I. Introduction

1.1 Motivation:

The large-scale structure of the universe, a vast cosmic web of galaxies, clusters, and voids, holds crucial information about the fundamental constituents and evolution of the cosmos. A key statistical tool for characterizing this structure is the matter power spectrum, P(k), which quantifies the distribution of matter density fluctuations as a function of scale. Precise measurements of the matter power spectrum, combined with theoretical models, allow cosmologists to constrain cosmological parameters, test theories of gravity, and probe the nature of dark matter and dark energy.

However, calculating the matter power spectrum from first principles typically involves computationally expensive cosmological simulations. These simulations, while powerful, can take significant time and resources, especially when exploring a wide range of cosmological parameters. This computational bottleneck hinders rapid exploration of the parameter space and efficient inference of cosmological parameters from observational data.

To overcome this challenge, the concept of *emulation* has emerged as a promising alternative. An emulator is a fast, surrogate model that approximates the output of a complex simulation or function. By training an emulator on a carefully selected set of simulation outputs, it can predict the results for new input parameters much faster than running the full simulation. Therefore, there is a pressing need for accurate and efficient emulators for the matter power spectrum, enabling rapid and reliable exploration of the cosmological parameter space. This project addresses this need by developing a machine learning-based emulator.

1.2 Objectives:

The primary objective of this project is to develop a Fully Connected Neural Network (FCNN) emulator for the matter power spectrum, P(k). The emulator will be trained on a large dataset of matter power spectra generated using the publicly available Code for Anisotropies in the Microwave Background (CAMB). The target is to achieve a mean accuracy of better than 1% across the relevant range of scales (k-values) and cosmological parameters. The emulator will be implemented using the Jax library for high-performance numerical computation and automatic differentiation. The key cosmological parameters considered in this project are Ω_m (matter density parameter), Ω_b (baryon density parameter), Ω_Λ (dark energy density parameter), h (Hubble constant), σ_8 (amplitude of matter fluctuations), and n_s (spectral index).

1.3 Project Overview:

This project involves several key stages. First, a large dataset of 1,000,000 matter power spectra is generated using CAMB, spanning a wide range of cosmological parameters sampled using a Latin Hypercube Sampling strategy. Second, a FCNN architecture with 4 hidden layers, each containing 1024 nodes, is designed and implemented using Jax. Third, a comprehensive hyperparameter search is conducted using Optuna to optimize the network's performance. Fourth, the trained network is evaluated on a held-out test set to assess its accuracy and generalization ability. The evaluation metrics include Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and the R-squared value. Finally, the emulator's performance is analyzed and compared with theoretical expectations. The remainder of this report is structured as follows: Section II provides the necessary theoretical background in cosmology and neural networks. Section III details the methodology used for dataset generation, network architecture, and hyperparameter tuning. Section IV describes the implementation details using Jax and the computational resources. Section V presents the experimental results and analysis. Section VI discusses the findings and limitations. Finally, Section VII concludes the report and outlines future research directions.

II. Theoretical Background

2.1 Cosmology:

2.1.1 The Expanding Universe:

The universe is expanding, a fact established by observations of the redshift of distant galaxies. This expansion is described by the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, which describes a homogeneous and isotropic universe:

$$ds^2 = -c^2dt^2 + a(t)^2[dr^2/(1-kr^2) + r^2(d\theta^2 + \sin^2\theta d\varphi^2)]$$

where *ds* is the spacetime interval, *c* is the speed of light, *t* is cosmic time, *a(t)* is the scale factor (describing the relative expansion of the universe), *r*, *θ*, and *ϕ* are comoving coordinates, and *k* is the curvature parameter (k = 0 for a flat universe, k > 0 for a closed universe, and k < 0 for an open universe). Hubble's Law, $v = H_0 d$, relates the recession velocity (*v*) of a galaxy to its distance (*d*) via the Hubble constant (H_0). Redshift (*z*) is defined as the fractional change in wavelength of light due to the expansion of the universe: $z = (\lambda_{observed} - \lambda_{emitted})/\lambda_{emitted} = a(t_0)/a(t_e) - 1$, where t_0 is the present time and t_e is the time of emission.

2.1.2 Cosmological Parameters:

The evolution of the universe and the formation of structure are governed by a set of cosmological parameters. The key parameters considered in this project are:

- **Ω_m (matter density parameter):** The ratio of the total matter density (including both baryonic and dark matter) to the critical density of the universe.
- **Ω_b (baryon density parameter):** The ratio of the density of baryonic matter (ordinary matter made of protons and neutrons) to the critical density.
- **Ω_Λ (dark energy density parameter):** The ratio of the dark energy density to the critical density. Dark energy is a mysterious component that drives the accelerated expansion of the universe.
- **h (Hubble constant):** The present-day expansion rate of the universe, expressed in units of 100 km/s/Mpc. $H_0 = 100h$ km/s/Mpc.
- **σ_8 (amplitude of matter fluctuations):** The root-mean-square (RMS) fluctuation of the matter density field smoothed on a scale of 8 h^{-1} Mpc. This parameter quantifies the overall amplitude of density perturbations.
- **n_s (spectral index):** The spectral index of the primordial power spectrum of density fluctuations. A value of $n_s = 1$ corresponds to a scale-invariant spectrum (Harrison-Zel'dovich spectrum).

2.1.3 Structure Formation:

The large-scale structure we observe today originated from small primordial density fluctuations in the early universe. These fluctuations, likely generated during a period of inflation, grew over time due to gravitational instability. Regions with slightly higher density attracted more matter, leading to the formation of galaxies, clusters, and superclusters.

2.1.4 Linear Perturbation Theory:

The evolution of small density perturbations can be described using linear perturbation theory. We define the density contrast, δ(x, t), as:

```
δ(x, t) = (ρ(x, t) - ρ̄(t)) / ρ̄(t)
```

where ρ(x, t) is the density at position x and time t, and ρ̄(t) is the mean density of the universe. In the linear regime (δ << 1), the evolution of the density contrast in a matter-dominated universe is governed by the following equation (derived from the continuity, Euler, and Poisson equations):

```
d²δ/dt² + 2H(t)dδ/dt - (3/2)H(t)²Ω<sub>m</sub>(t)δ = 0
```

This is a second-order differential equation. The growing mode solution describes how density perturbations grow with time. The Hubble parameter H(t) is given by the Friedmann equation:

$$H(t)^2 = (8\pi G/3)\rho - kc^2/a(t)^2$$

where G is Newton's gravitational constant.

2.1.5 Transfer Function:

The transfer function, T(k), describes the evolution of the density perturbations from the early universe to a later time (typically the epoch of matter-radiation equality). It accounts for the different growth rates of perturbations inside and outside the horizon. The transfer function relates the primordial power spectrum, $P_{prim}(k)$, to the matter power spectrum at a later time, P(k):

$$P(k) = P_{prim}(k) * |T(k)|^2$$

The primordial power spectrum is often assumed to be a power law: $P_{prim}(k) \propto k^{n_s}$.

2.2 The Matter Power Spectrum:

2.2.1 Definition:

The matter power spectrum, P(k), is a fundamental quantity in cosmology. It is defined as the Fourier transform of the two-point correlation function, ξ(r), of the density field:

$$\xi(r) = \langle\delta(x)\delta(x + r)\rangle$$

$$P(k) = \int \xi(r) \exp(-ik\cdot r)\, d^3r$$

where the angle brackets denote an ensemble average, *k* is the wavevector (with magnitude *k* representing the wavenumber, related to the scale by λ = 2π/k), and *r* is the separation between two points. P(k) represents the variance of the density fluctuations at different scales. It quantifies how much power is contained in fluctuations of different sizes.

2.2.2 Importance:

The matter power spectrum is a crucial tool for connecting cosmological theory with observations. By measuring the distribution of galaxies and other tracers of the matter density field, we can estimate the matter power spectrum. Comparing this observed power spectrum with theoretical predictions allows us to constrain cosmological parameters and test different cosmological models. For example, the shape and amplitude of P(k) are sensitive to the values of Ω_m, Ω_b, h, σ_8, and n_s.

2.2.3 Features:

The matter power spectrum exhibits several key features:

- **Turnover Scale:** At large scales (small k), P(k) follows the primordial power spectrum (approximately $P(k) \propto k^{n_s}$). At smaller scales (large k), the growth of perturbations is suppressed during the radiation-dominated era, leading to a turnover in the power spectrum. The location of this turnover is related to the size of the horizon at matter-radiation equality.
- **Baryon Acoustic Oscillations (BAO):** Before recombination, baryons and photons were tightly coupled, forming a plasma. Sound waves propagated through this plasma, leaving a characteristic imprint on the matter power spectrum in the form of oscillations. These BAO features provide a standard ruler for measuring cosmological distances.
- **Damping Tail:** At very small scales (large k), the power spectrum is damped due to various processes, such as Silk damping (photon diffusion) and free-streaming of neutrinos.

2.2.4 CAMB:

The Code for Anisotropies in the Microwave Background (CAMB) is a widely used Boltzmann code that calculates the matter power spectrum and the Cosmic Microwave Background (CMB) anisotropies given a set of cosmological parameters. It solves the coupled Boltzmann equations for the evolution of perturbations in the early universe. CAMB is a crucial tool for generating the theoretical predictions needed for this project.

2.3 Neural Networks:

2.3.1 Introduction to Artificial Neural Networks:

Artificial neural networks (ANNs) are computational models inspired by the structure and function of biological neural networks. They consist of interconnected nodes (neurons) organized in layers. ANNs can learn complex patterns and relationships from data, making them powerful tools for various tasks, including function approximation, classification, and pattern recognition.

2.3.2 Fully Connected Neural Networks (FCNNs):

In a Fully Connected Neural Network (FCNN), each node in one layer is connected to every node in the adjacent layers. The input layer receives the input data (in this case, the cosmological parameters). The hidden layers perform non-linear transformations of the data. The output layer produces the network's prediction (in this case, the matter power spectrum). Each connection between nodes has an associated weight, and each node has a bias. The output of a node is calculated by applying an activation function to the weighted sum of its inputs plus the bias:

```
output = activation(∑(weight * input) + bias)
```

Common activation functions include ReLU (Rectified Linear Unit), sigmoid, and tanh. The forward pass involves calculating the output of the network layer by layer, starting from the input layer and propagating through the hidden layers to the output layer.

2.3.3 Training Neural Networks:

Training a neural network involves adjusting the weights and biases to minimize a loss function, which quantifies the difference between the network's predictions and the true values. The Mean Squared Error (MSE) is a common loss function for regression problems:

```
MSE = (1/N) ∑(y<sub>true</sub> - y<sub>predicted</sub>)²
```

where N is the number of data points, y_{true} is the true value, and $y_{predicted}$ is the network's prediction. Optimization algorithms, such as stochastic gradient descent (SGD) and Adam, are used to iteratively update the weights and biases based on the gradient of the loss function. Backpropagation is an efficient algorithm for calculating the gradients.

2.3.4 Hyperparameter Optimization:

Hyperparameters are parameters that control the learning process and the network architecture, but are not learned during training. Examples include the learning rate, batch size, number of hidden layers, and number of nodes per layer. Hyperparameter optimization involves finding the best combination of hyperparameters that minimizes the validation loss. Optuna is a framework for automated hyperparameter optimization, using techniques like Bayesian optimization and tree-structured Parzen estimators.

2.3.5 Emulation with Neural Networks:

Neural networks can be used as emulators to approximate complex functions or simulations. By training a neural network on a set of input-output pairs from the original function or simulation, the network learns the underlying relationship and can predict the output for new inputs much faster than evaluating the original function or running the simulation. This is particularly useful for computationally expensive tasks, such as calculating the matter power spectrum.