National Aeronautics and Space Administration

Maximizing the Science return from Cosmic Microwave Background Anisotropy Observations

Principal Investigator: Graca Rocha (326); Co-Investigators: You Lu (398), Krzysztof Gorski (326), Kiri Wagstaff (398)

Program: FY22 R&TD Strategic Initiative

Strategic Focus Area: An Integrated Community of Practice for Scientific Understanding from Data Science (SUDS) - Strategic Initiative Leader: Susan E Owen

Objectives:

This project aims to maximize the science return from cosmic microwave background (CMB) observations and to ensure success in probing the ultimate frontiers of cosmology and physical science by searching for “B-modes” in polarized CMB data generated by relic gravitational waves - a signature of inflation. We aim to push the search for elusive “B-modes” by detecting the ratio of perturbation power in relic gravitational waves vs. matter at the level of 5×10^{-4} (5 sigma), which is 100 times below the current upper limits, and 5 times lower than limits forecast for any currently planned experiment. Importantly, a non-detection would exclude at 5 sigma those models for which the characteristic scale in the potential is at the Planck scale, a key threshold in inflation physics. While estimating the parameter r, a wide range of other parameters will be determined jointly. This adds intricacies and benefits: our work will enable lifting degeneracies amongst the cosmological parameters and increase the robustness and accuracy of all of the extracted parameters.

Background:

Precision and accuracy of measurements and model constraints are crucial to answering fundamental questions such as: How did our Universe begin and evolve? What are the laws of the Universe? How did we get here? How did galaxies, stars and planets come to be? These are key goals of NASA SMD in Astrophysics (PCOS, COR). This is so because only precision and accuracy enable us to properly/unequivocally constrain models of structure formation and offer a theory of cosmogenesis.

We explore the interface of physical (PS) and data science (DS) domains and leverage expert understanding of the science problem, the complexity of data and models, and both the potential and limitations of the data science approach. Our project stimulated the cooperation between PS and DS via direct, problem-oriented hands-on work, while consolidating joint ownership of the science question. We will disseminate the results of this project via seminars (internal to JPL, and external), participation in hackathons, and publications in high-impact journals. We see our project as a catalyst for integration of the community of practice for Science Understanding with Data Science at JPL.

Approach and Results:

We developed a graph-based Bayesian convolutional neural network (CNN) model to remove foregrounds due to our Galaxy and extragalactic sources. The CNN model operates on full-sky maps, predicting the cleaned CMB from observations at nine frequency bands. We use graph-based Bayesian CNN with U-Net architecture.

To train and evaluate the CNN model, we simulate sky observations as a sum of $C_i + F + N_j$, where: 1) C_i is a random realization of primordial CMB anisotropies consistent with Planck 2015 estimates of the cosmological model parameters; 2) F are the maps, at the nine Planck frequency bands, of the composite foreground emission provided in the 2015 Planck data release, 3) N_j are random, spatially modulated noise realizations consistent with Planck instrument sensitivity per frequency band. For the initial stage of the project, we reduced the angular resolution of the signal maps to FWHM=150 arcmin, and rendered them at HEALPix resolution of N_{side}=64.

We compare CNN model’s prediction against the outcome of the Internal Linear Combination (ILC) technique (Figure 1). We evaluate the r.m.s. error (rmse) and Pearson correlation coefficient, r, between the true and predicted pixel values outside a ± 30 degree sky cut. In a typical realization, ILC yields RMSE = 7.597 μK and r = 0.993 while graph-based Bayesian CNN model yield RMSE = 4.449 μK and r = 0.997, suggesting potential usefulness of the CNN based component separation approach.

But, CNN predicted maps exhibit an over-smoothing

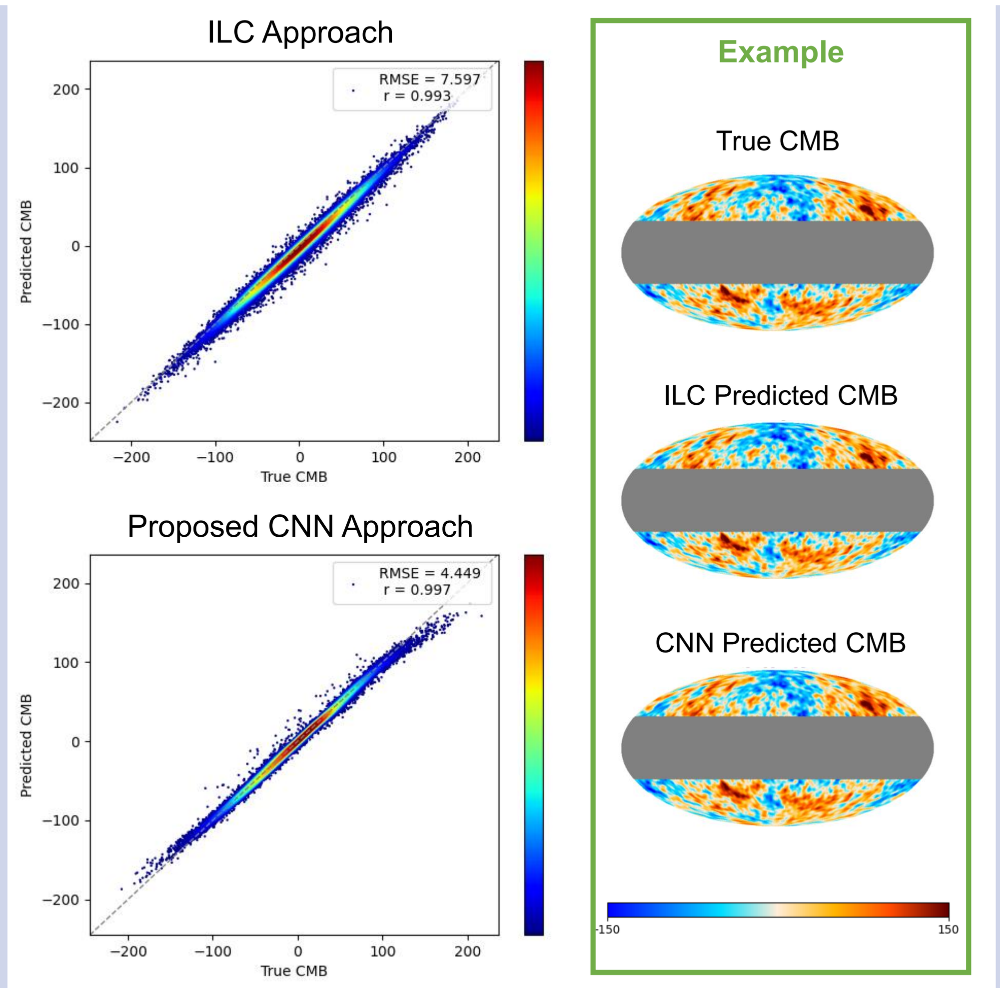

Figure 1. Comparison of the CMB input vs. CNN-predicted, and ILC sky maps, and statistical assessment of their quality.

effect in the signal dominated regime (Figure 2, Top plot). Uncorrected, this effect would result in biased cosmological parameter estimation from the CNN product maps. We investigated the plausible drivers of this effect (Figure 2. Middle and Bottom plots) to seek possible remedies. The discrepancy between the input and CNN-output CMB power spectra exhibits oscillatory behavior independent of the size of the Galaxy cut, but strongly dependent on the CNN depth. This CNN depth dependence implies links to implementation

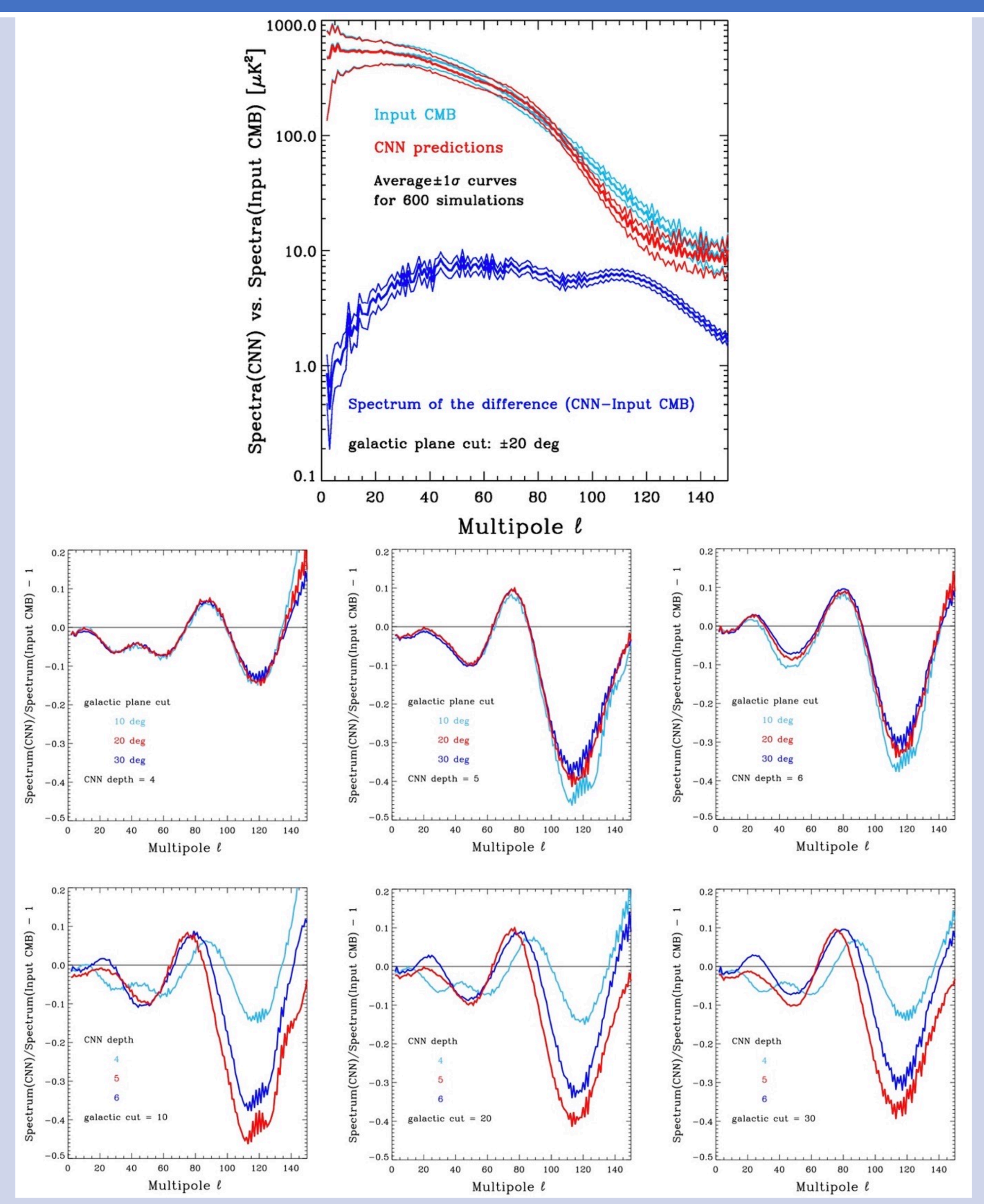

Figure 2. Investigation in search for the reasons for the undesirable “over-smoothing” effect discovered in the CNN-predicted CMB anisotropy maps.

of “graph convolution” (e.g. eigen-values and -modes of the graph Laplacian, and approximate evaluation of convolution via Chebyshev polynomials) and should be investigated further, understood, and remedied.

Significance/Benefits to JPL and NASA:

Year 1 results of this task are promising and support the expectation that upon completion the full 3 year program on the proposed CNN approach to CMB component separation will be helpful in synchronizing the science operations of future cosmological missions (e.g. CMB Probe) between instrument development and data analysis, and will support a more efficient optimization of instrumental specification of the future cosmological missions (viz. number and placement of frequency channels, and designation of the required sensitivities). Furthermore, the proposed CNN approach will also benefit JPL in competition for NASA ROSES proposal opportunities. Finally, the proposed CNN approach can be generalized to support similar problems in other disciplines, especially w.r.t. the global data sets, as in Earth, and planetary sciences.

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

Clearance Number: CL# 22-5001
Poster Number: RPC#R22003

Publications:

Jadie Adams, Steven Lu, Krzysztof M. Gorski, Graça Rocha, Kiri Wagstaff, “Cosmic Microwave Background Recovery: A Graph-Based Bayesian Convolutional Network Approach”, *35th Annual Conference on Innovative Applications of Artificial Intelligence*, 2023.

PI/Task Mgr. Contact Information:

Email: Graca.M.Rocha@jpl.nasa.gov