NeuralHMC: Accelerated Hamiltonian Monte Carlo with a Neural Network Surrogate

Linnea M. Wolniewicz[1], Peter Sadowski[1], Claudio Corti[2,3]

[1] Information and Computer Science Department, University of Hawaiʻi at Mānoa, HI, USA
[2] Physics and Astronomy Department, University of Hawaiʻi at Mānoa, HI, USA
[3] NASA Goddard Space Flight Center, Greenbelt, MD, USA

ICS 699, Spring 2024, Plan B Masters Project

Published at NeurIPS 2024 Workshop on Machine Learning for the Physical Sciences

SCAN FOR FULL PAPER

1. Introduction

- Hamiltonian Monte Carlo (HMC) is widely used to perform Bayesian inference in the sciences, and accelerates MCMC sampling by using gradients of the likelihood function, but can be computationally expensive [1] [2]
- In this work, we address the scenario where the likelihood function is calculated by a numerical solver that is incompatible with HMC for three reasons:
  1. It is non-differentiable
  2. It is too slow
  3. It suffers from numerical instabilities
- **Our solution is to train a neural network (NN) surrogate likelihood that solves all three of these problems**
- Previous work has used machine learning to accelerate HMC sampling, but none have used a surrogate likelihood function [3] [4] [5] [6]
- **We make the following contributions:**
  1. We demonstrate that surrogate model can be used for HMC
  2. We propose a strategy for keeping limiting the HMC to a trusted region in which the surrogate model is expected to generalize
  3. We use this approach to model the transport of galactic cosmic rays (GCRs) in the heliosphere, which constitute a major radiation hazard for deep-space human exploration

2. Data and methods

- We constrain the transport of GCRs with an HMC to estimate the posterior density function (PDF) over 5 variables that parameterize the heliospheric transport coefficients, given observations of the GCR flux at Earth

Observational data:

Experiment	Period	Position	Measurements
PAMELA	2006-2014	Earth LEO	H: 0.4-50 GV
AMS-02	2011-2019	Earth LEO	H, He: 1-100 GV

GCR transport is characterized by the Parker equation, which depends on the following heliospheric magnetic field (HMF) parameters:

Symbol	Description
I_{HMF}	Intensity at 1 AU
$\vec{V}_{SW}$	Solar wind speed
α	Solar dipole tilt angle
$k^0_{\parallel}$	Normalization constant of the diffusion tensor **K**
$a_{\parallel}$, $b_{\parallel}$	Rigidity slopes of **K** in the direction parallel to the HMF
$a_{\perp}$, $b_{\perp}$	Rigidity slopes of **K** in the direction perpendicular to the HMF

Neural Network

- Instead of solving the Parker equation for each set of parameters, we train a NN model as a surrogate of the numerical model of [7]
  - NN takes in 8 parameters above, and predicts GCR over 32 rigidity steps. Two NNs were trained separately on solar magnetic positive and negative polarity data
  - Numerical solutions were computed for 6 million parameter values in a grid over reasonable parameter values using the method of [7]
    - After removing solutions with numerical instabilities, 2,088,385 positive and 1,987,658 negative polarity samples remained (90/10 train/test split)

Hamiltonian Monte Carlo

- We perform HMC to sample from the posterior for 210 time intervals of interest
  - Sample 5 model parameters: $k^0_{\parallel}$, $a_{\parallel}$, $b_{\parallel}$, $a_{\perp}$, and $b_{\perp}$
  - I_{HMF}, $\vec{V}_{SW}$, and α are set to their 1–year backward average for each interval using data from OMNIWeb and Wilcox Solar Observatory
- Likelihood of a set of parameters z, given observed GCR fluxes (x) and NN-predicted GCR fluxes $\hat{x}$:

$$p(x|z) = \exp\left(\frac{-\chi^2(x,\hat{x})}{2}\right)$$

- We prevent the HMC from sampling outside the "trusted" domain with a prior distribution, $p(z)$, which is uniform in the training data domain and rapidly decays outside the domain

Bayesian Inference with Markov Chain Monte Carlo requires the computation of the likelihood function. For some scientific applications, such as modeling the transport of galactic cosmic rays, this is prohibitively expensive. Here, we **accelerate Hamiltonian Monte Carlo sampling with a surrogate likelihood function** implemented by a neural network

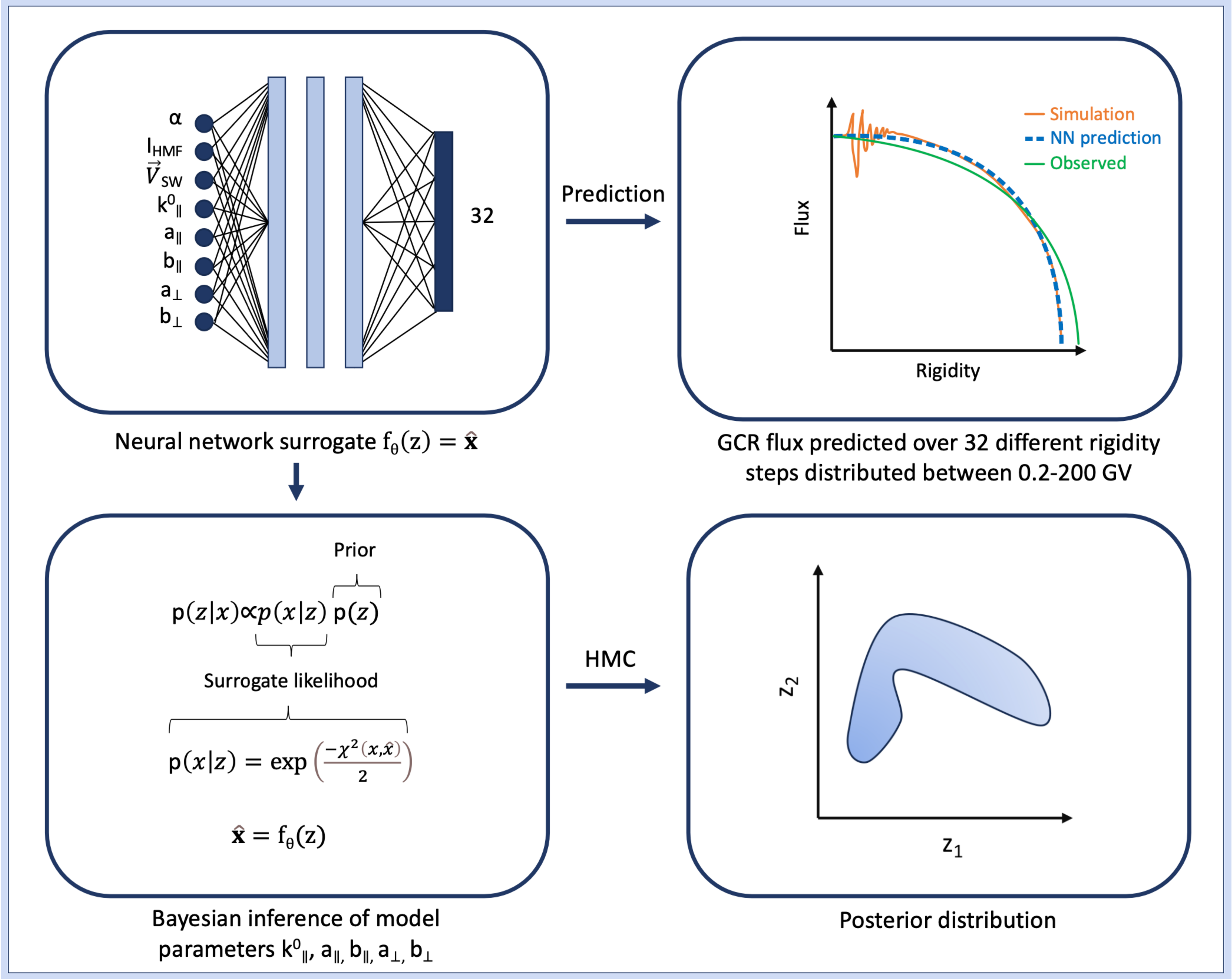

Figure 1: Diagram of the method. The NN takes in the 8 parameters detailed in Table 2 (top left) and predicts the GCR flux at 1 AU for 32 rigidity steps (top right). Surrogate likelihood is computed by comparing the NN outputs with observed fluxes (bottom left). This likelihood is used to sample from the posterior using HMC (bottom right)

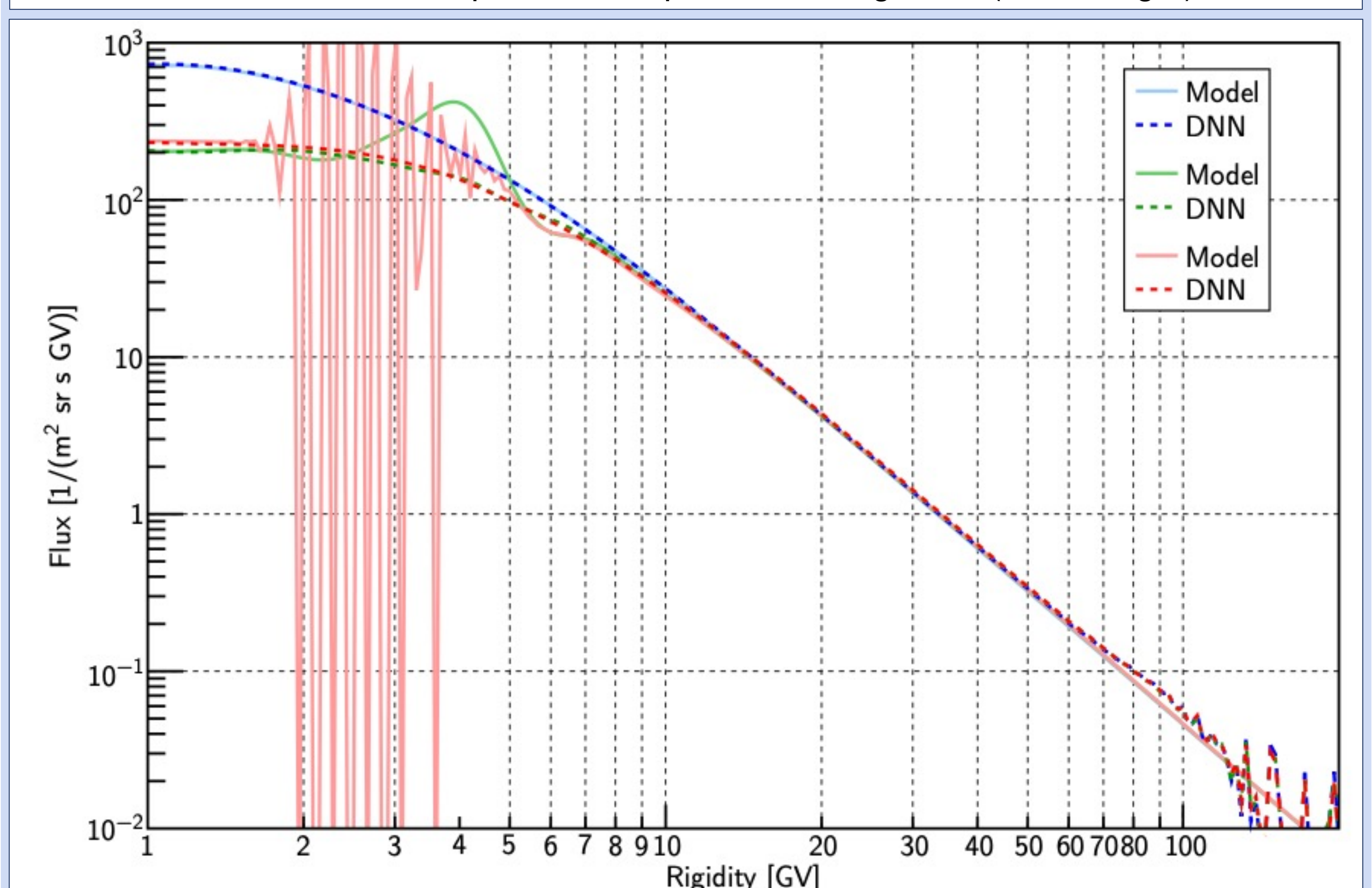

Figure 2: Modulated proton flux as a function of rigidity computed via the model from [7] and our NN for three typical examples with varying levels of numerical instability. Examples are drawn from the test set

3. Results

- HMC using the NN surrogate likelihood is fast and avoids the pitfalls of using the numerical solver for computing the likelihood, as seen by the smooth NN curve of Figure 2
- Our strategy of constraining the HMC to a trusted region was successful, as shown by the well-constrained parameter values of Figure 3
- Our results agree with previous attempts to sample from the posterior in [7], except with a significant speed-up (~480 days to ~1 day)

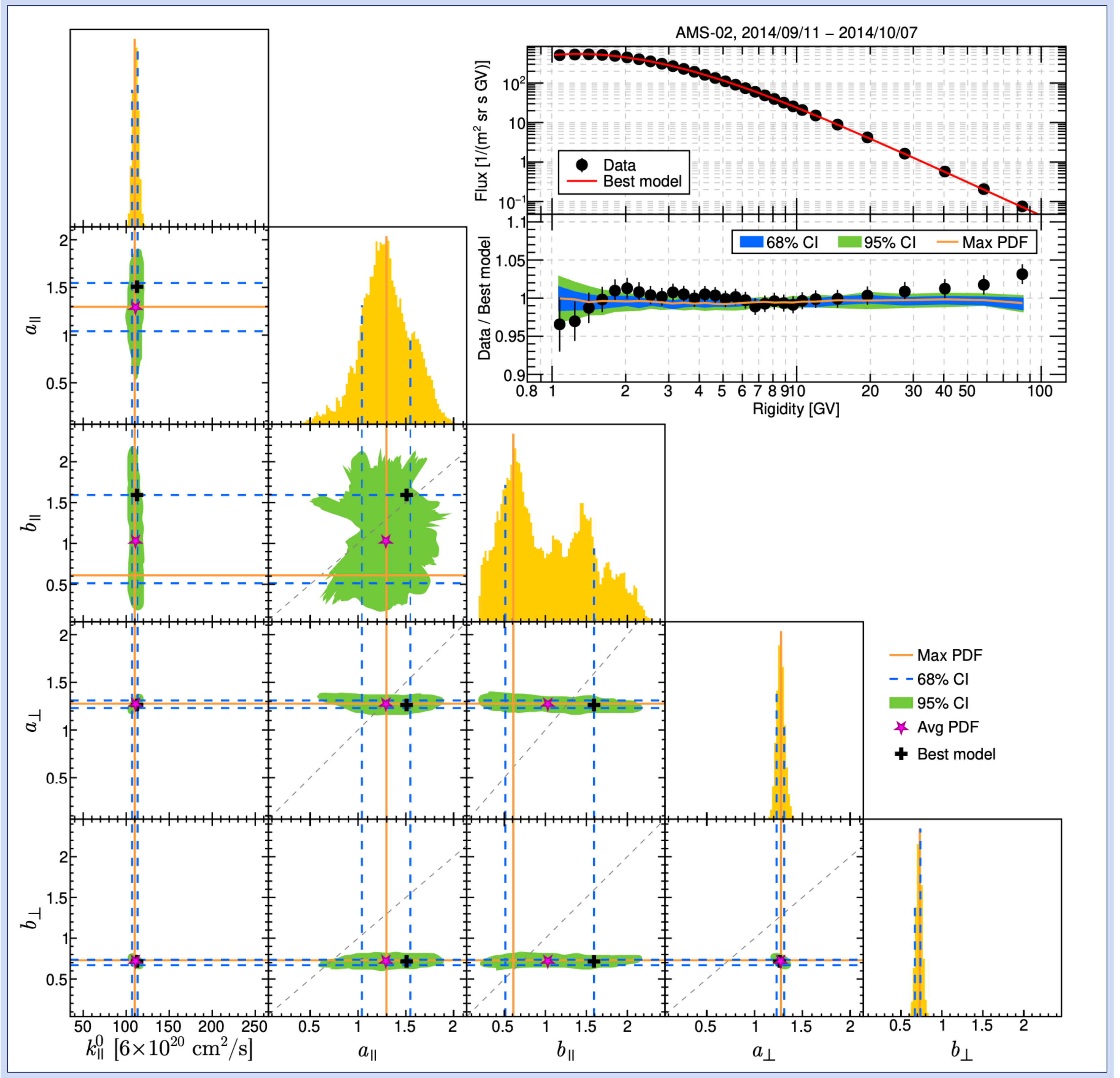

Figure 3: 2D and 1D PDFs of the HMC-sampled parameters for GCR flux, measured by AMS-02 in the positive polarity time interval 2014/09/11-2014/10/07. Top panel compares the maximum likelihood NN model (red) with observations, with the 68% and 95% credible intervals

4. Acknowledgments and references

This work was funded by NSF GRFP Grant No. 2236415 and NASA Living with the Star Grant No. LWS80NSSC20K1819.

[1] Kyle Cranmer, Johann Brehmer, Gilles Louppe. The frontier of simulation-based inference. Proceedings of the National Academy of Sciences, 117(48):30055–30062, 2020.
[2] Simon Duane, Anthony D Kennedy, Brian J Pendleton, Duncan Roweth. Hybrid monte carlo. Physics letters B, 195(2):216–222, 1987
[3] Cheng Zhang, Babak Shahbaba, Hongkai Zhao. Hamiltonian monte carlo acceleration using surrogate functions with random bases. Statistics and computing, 27:1473–1490, 2017
[4] Lingge Li, Andrew Holbrook, Babak Shahbaba, Pierre Baldi. Neural network gradient hamiltonian monte carlo. Computational statistics, 34:281–299, 2019
[5] Daniel Levy, Matthew D. Hoffman, Jascha Sohl-Dickstein. Generalizing Hamiltonian Monte Carlo with neural networks, 2018
[6] Sam Foreman, Xiao-Yong Jin,, James C. Osborn. Deep learning hamiltonian monte carlo, 2021b.
[7] Claudio Corti, Marius S. Potgieter, Veronica Bindi, Cristina Consolandi, Christopher Light, Matteo Palermo, and Alexis Popkow. Numerical modeling of galactic cosmic-ray proton and helium observed by AMS-02 during the solar maximum of solar cycle 24. The Astrophysical Journal, 871 (2):253, 2019.