Bayesian Stokes inversion with Normalizing flows

***C. J. Díaz Baso**[1], A. Asensio Ramos[2,3], J. de la Cruz Rodríguez[1]

[1] Institute for Solar Physics, Dept. of Astronomy, Stockholm University, AlbaNova University Centre, SE-10691 Stockholm, Sweden
[2] Instituto de Astrofísica de Canarias, C/Vía Láctea s/n, E-38205 La Laguna, Tenerife, Spain
[3] Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain
*Email : carlos.diaz@astro.su.se

1. Introduction

In spectropolarimetry, we call inversion to the process of finding the best fit to a given spectra by modifying an atmospheric model with physical quantities like temperature, velocity, magnetic field and comparing the emergent spectra. This method is too slow when is applied to many pixels. Recently, we have seen that standard artificial neural networks are much faster by learning the “average” mapping between spectra and physical quantities; but this is a problem if degeneracies or multimodal solutions exist. On the other hand, Bayesian inference allows us to obtain the full probability distribution including uncertainties, correlations, and if our distribution is (or not) multimodal, but it implies exploring the parameter space by running the forward model several thousand times. So, is there any way to perform fast Bayesian inference?

2. Normalizing flows

The novel tecnique that allow us to perform fast Bayesian inference is known as normalizing flows (NFlows), and they are a set of invertible and parameterized transformations that convert a simple distribution into an approximation of any other complex distribution. With this new technique, we approximate the probability distribution of our target by a transformation from a simple probability distribution. If these transformations are conditioned on observations (CNFlows, see right panel of Fig. 1), we can train normalizing flows to return Bayesian posterior probabilities for any observation. For comparison, there is a sketch of a standard artificial neural network on the left side of Fig. 1 which outputs the average mapping.

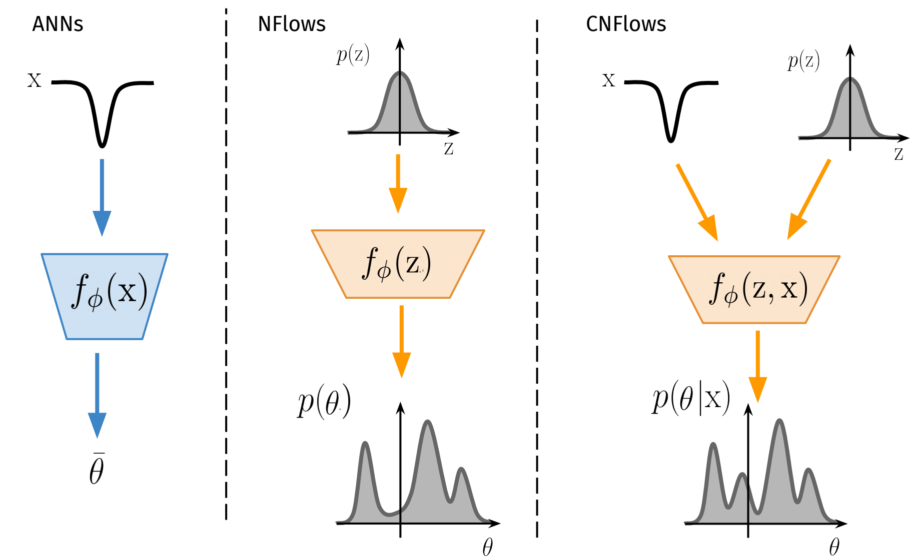

Fig. 1: Comparison between an artificial neural network and a conditional normalizing flow with parameters **φ**.

3. Milne-Eddington model

As a first example, we illustrate here the capabilities of the method in a case where the forward modelling is fast enough to allow a comparison with the exact solution obtained with a Markov Chain Monte Carlo (MCMC) method. For that we have chosen a simple Milne-Eddington model with five parameters that controls the intensity profile of the spectral line.

We have created a database of 10^6 pairs of examples (parameters vs spectra). We optimize the transformations of the normalizing flow, like in classical neural networks, but in this case to reproduce the distribution of the data. Once trained, the **NFlow** can produce the distribution for any given observation as accurate as the **MCMC** sampling method, with the corresponding uncertainties and degeneracies like the ones between the absorption of the line, the source function and the Doppler width with a banana-shape.

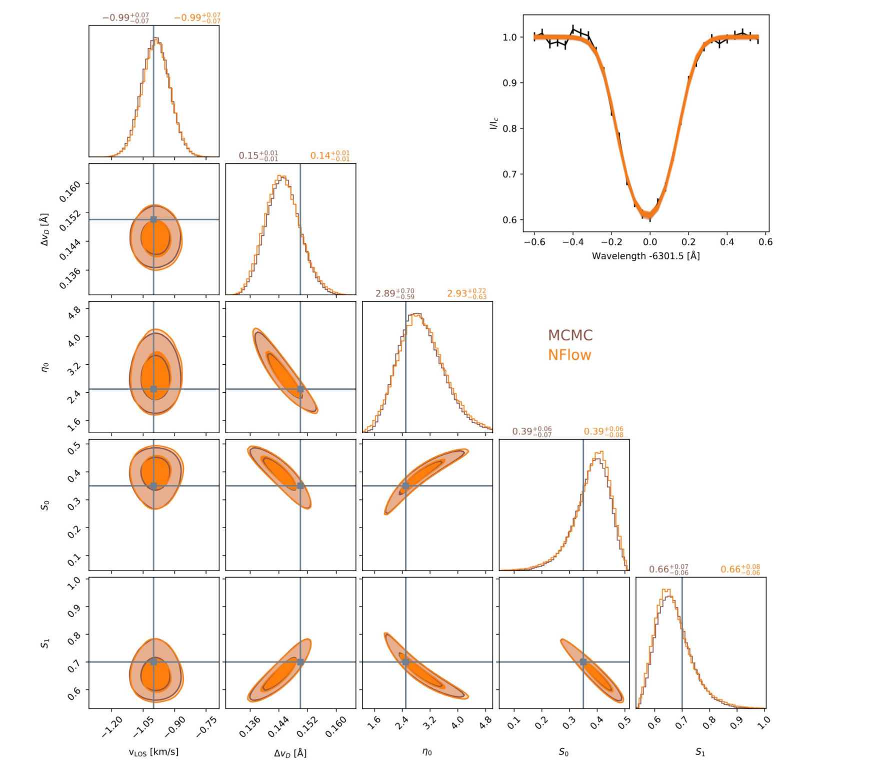

Fig. 2: Joint and marginal posterior distributions for the physical parameters involved in the Milne-Eddington model.

4. N-LTE inversion

The real improvement comes when the new method is applied to non-LTE inversions, which are **more computationally demanding** than the previous example. Following the same procedure, we create a large dataset with synthetic profiles of a photospheric iron line and the chromospheric CaII line at 8542Å. Figure 3 shows two profiles and their stratification of temperature, velocity, and microturbulent velocity.

To illustrate the performance of this technique with different spectral lines, we have trained two normalizing flows: one only using the FeI line which gives the orange solution and another which also **uses the CaII profile and produces the brown solution**. From the width of the solutions (here the bands of 1 sigma of the distribution), we see that just looking to the database, the normalizing flow learns the range of sensitivity of each spectral line. This inference takes around 1 second (producing 10^4 samples) while an MCMC would take many hours or even days.

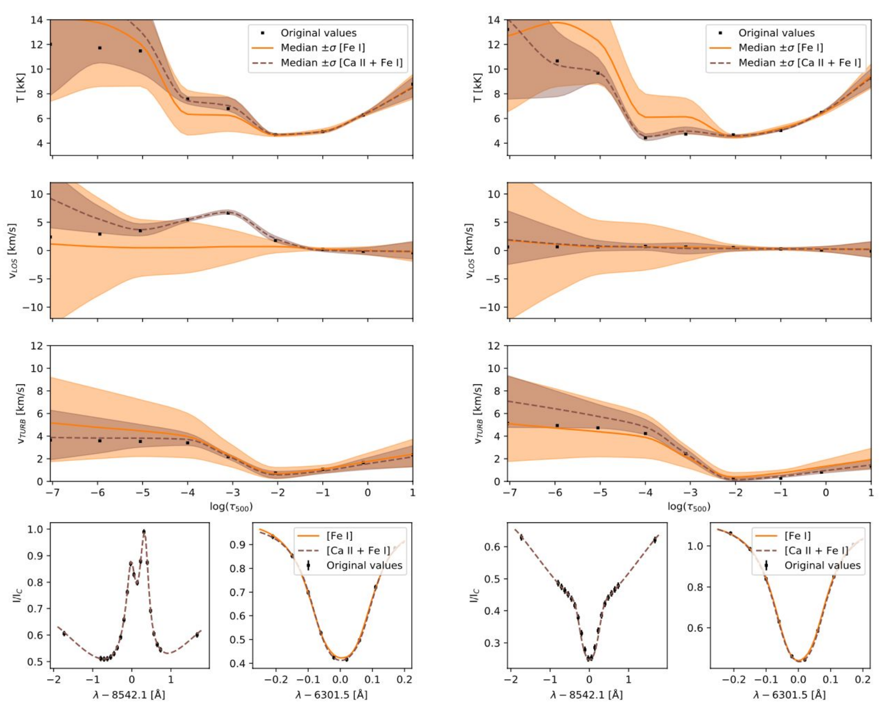

Fig. 3: Atmospheric stratification for two examples. The colored bands of each curve indicate the standard deviation of each distribution.

5. N-LTE correlations

Joint distributions give interesting information about how different parameters are correlated in the inference, clearly pointing out the presence of ambiguities. To obtain an approximate insight, the joint distribution can be summarized by showing the correlation matrix in Fig. 4:

The checkerboard pattern found in the temperature indicates that reductions in the temperature at some locations can be compensated by increases in other locations.

We also found expected correlations between the broadening effect of the microturbulent velocity and the temperature to the shape of the profile.

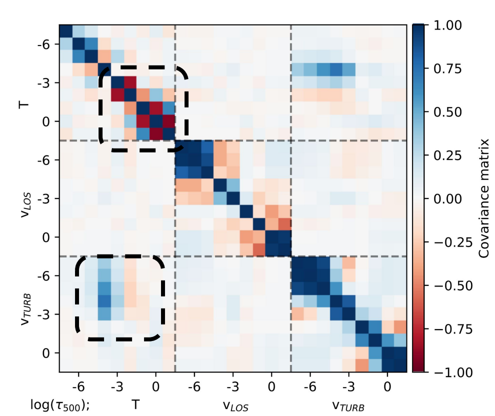

Fig. 4: Correlation matrices calculated for the inferred atmospheric stratification. Blue/red indicates positive/negative correlations, respectively.

6. Summary

We have explored the usage of normalizing flows to accurately infer the posterior distribution of a solar model atmosphere (parameters, correlations, and uncertainties) from the interpretation of observed photospheric and chromospheric lines.

A natural extension of this work would be to include the four Stokes parameters to infer the magnetic properties of our target of interest, while also setting more constraints in the rest of the physical parameters.

More info here: Díaz Baso et al, submitted to A&A (https://arxiv.org/abs/2012.06229)

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 759548)