Iris-ML: Simulation-Based Inference for the Spectral Energy Distribution fitting.

Mateusz Kapusta
Astronomical Observatory, Warsaw, Poland

Astronomical Observatory of the University of Warsaw
Founded in 1825

Introduction

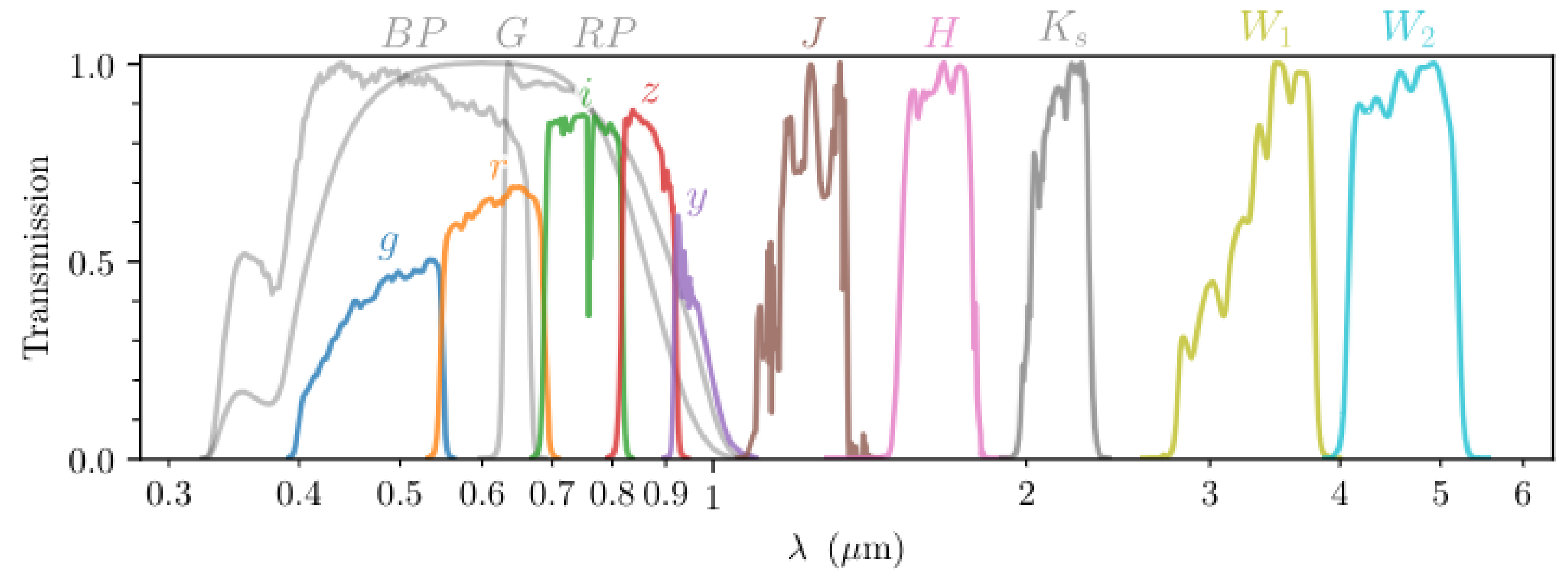
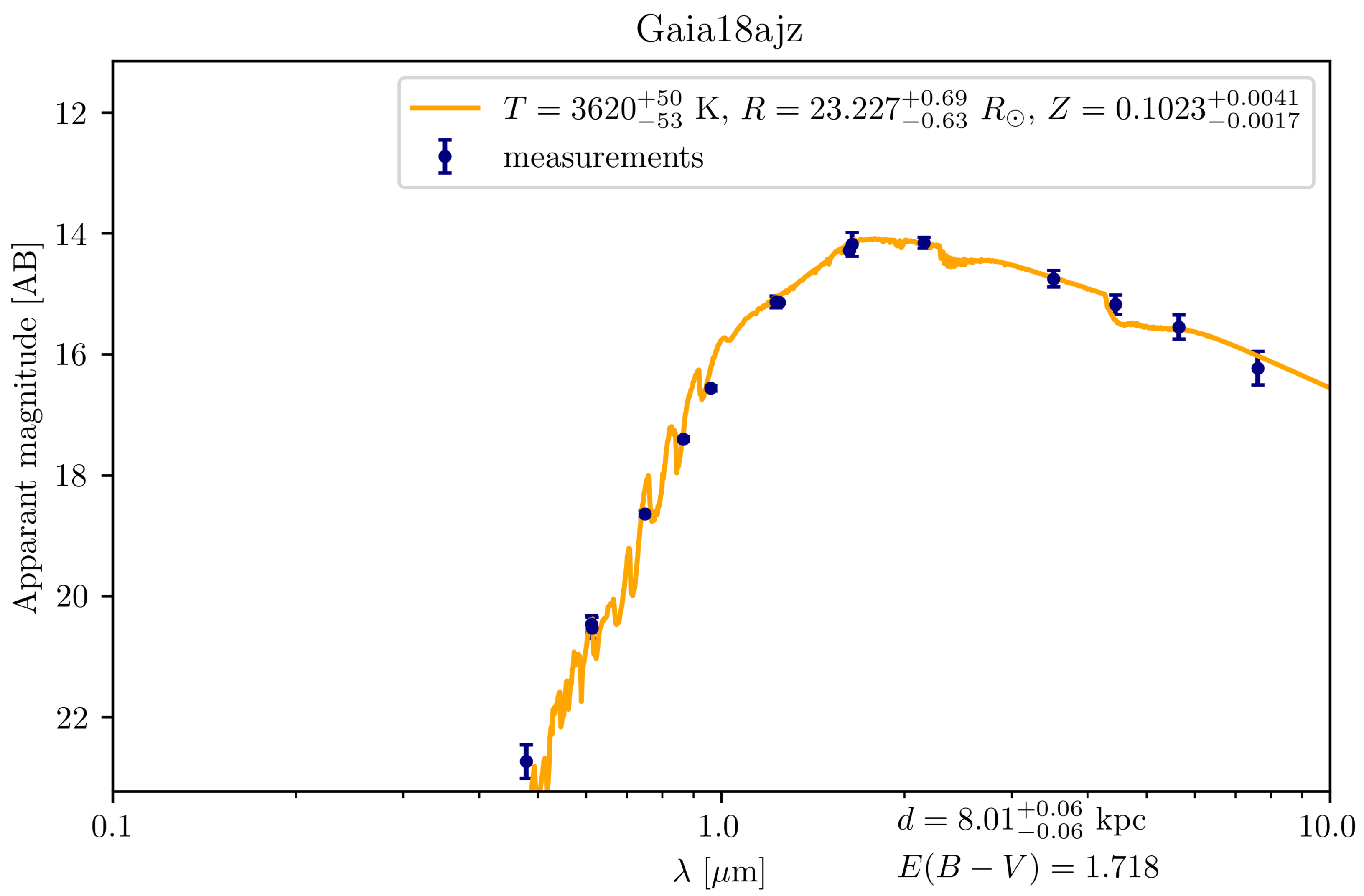


Figure 1: Different filters used in the study (image from Green et al. [2021]).

Photometric observations allow one to take measurements of thousands of objects at the same time. Contrary to the spectroscopic data they allow one to observe much larger portions of the sky. Hence, there is a huge interest in the inference strategies that allow to utilize photometric measurements for parameter inference.



Usually one proceeds with following procedure:

- Given parameters $\theta = (T, logg, [M/H], A_v, R_v)$ compute brightness in filter G,BP,RP.
- Compare predicted values and true values with true and compute likelihood.
- Use likelihood to perform MCMC sampling step.

However, MCMC sampling takes enormous amount of time.
(Filter, Brightness, Error)

(G,15.818, 0.002)	$T = 4500 \pm 100$
(BP, 16.899, 0.004)	$logg = 2.4 \pm 1$
(RP, 14.792, 0.003)	$[M/H] = -0.32 \pm 0.51$
(H, 14.32, 0.04)	$A_V = 2.27 \pm 0.17$
(J, 12.13, 0.02)	$R_V = 3.61 \pm 0.14$
(K, 13.11, 0.03)	

In order to obtain samples from the posterior distribution one needs an architecture that:

- takes N different observations in format: name of the filter, brightness, brightness error
- Produces samples from the posterior distribution (in Bayesian sense), which are invariant with respect of the permutation of observation.

Problem: How to allow for a variable-size input and return samples from the model?

Transformer + Normalizing Flow (Masked Autoregressive flow)

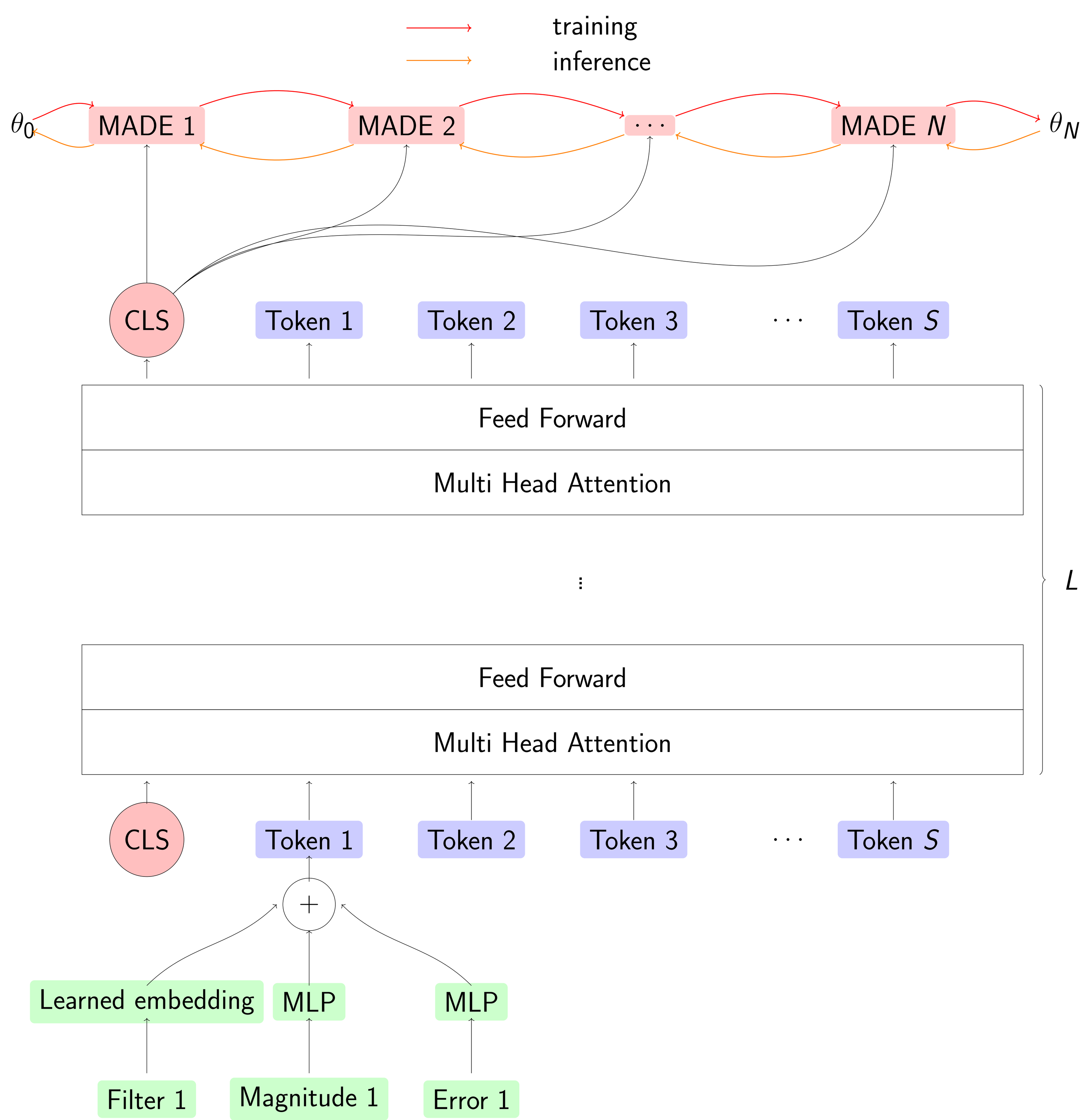
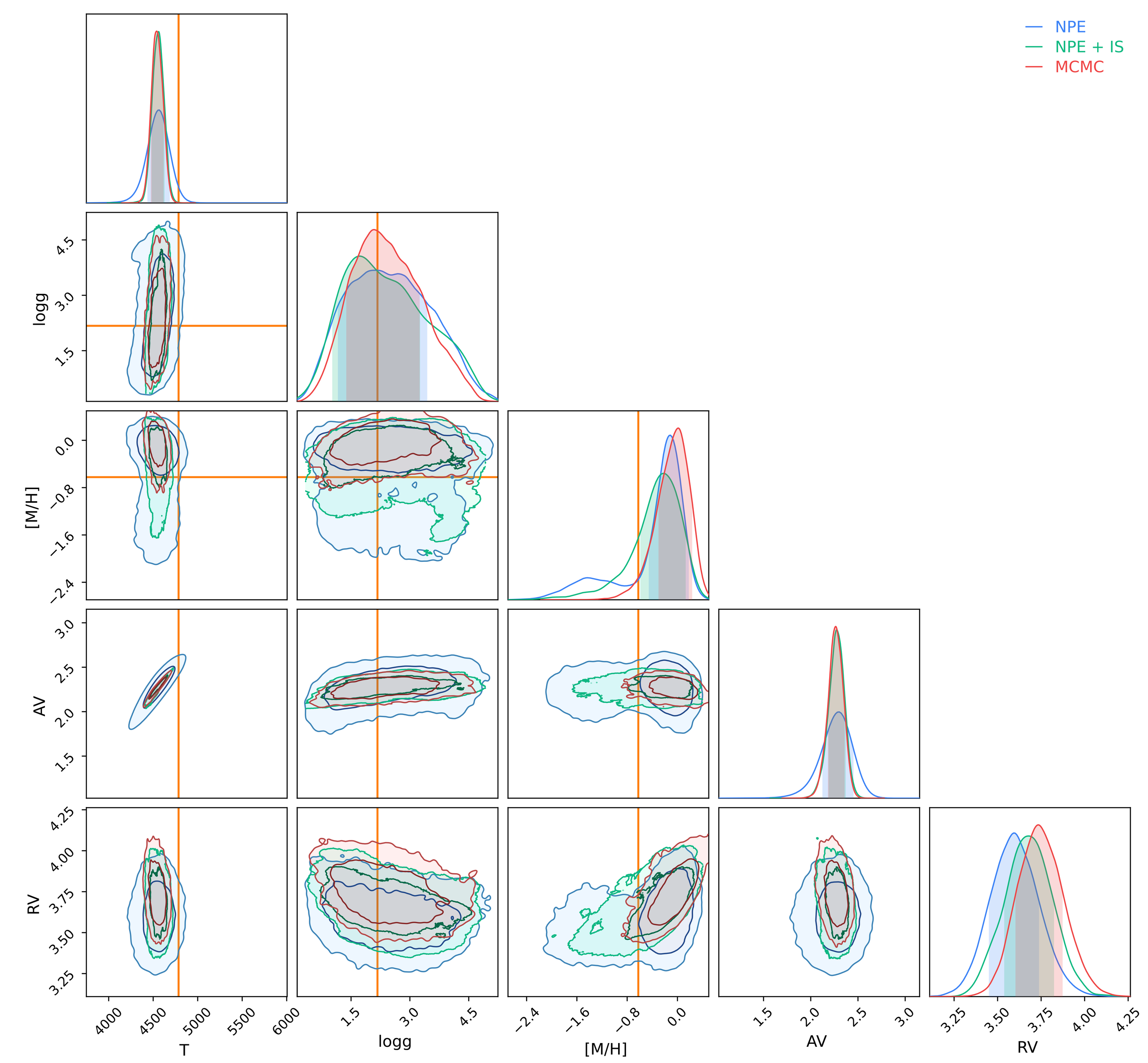


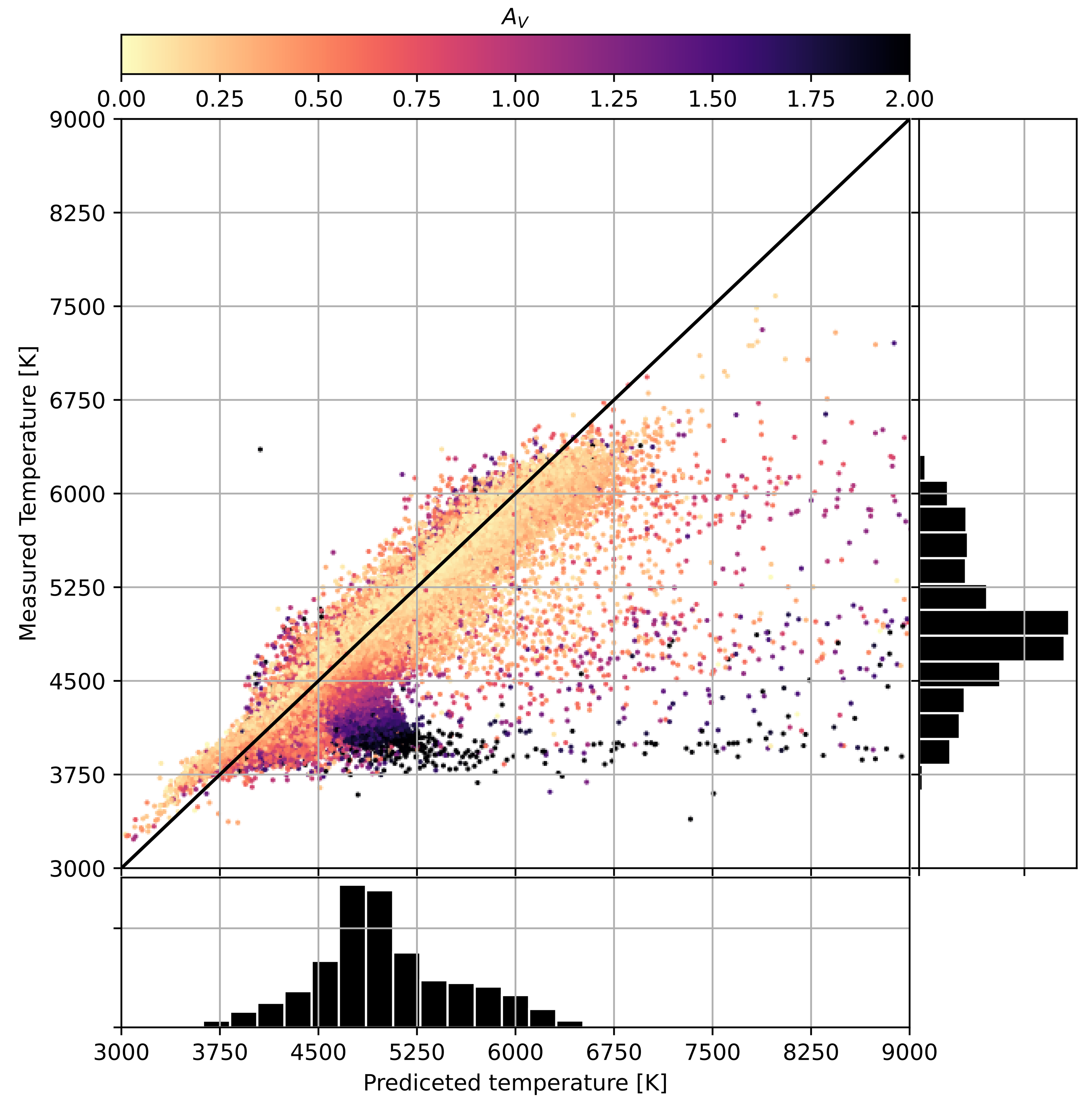
Figure 2: Dataflow of the SED Transformer. MAF architecture is presented in Papamakarios et al. [2017].

- Filter $\rightarrow$ embedding vector (learned embedding),
- Brightness $\rightarrow$ MLP embedding,
- Brightness error $\rightarrow$ MLP embedding,
- Final vector is composed of three aforementioned vectors added together.



Results

Contrary to the Variational Autoencoders or Diffusion models, Normalizing Flows return the likelihood of the associated sample. This allows to use **Importance Sampling procedure** to refine the results of the sampling as pointed out in Dax et al. [2023]. Hence, total sampling procedure is very robust and allows to detect outlier data.



MAD error on temperature $\sim 150K$ with ~ 250000 stars, MCMC posterior computation would be impossible on such huge dataset without computer cluster, with NDE it takes 4 minutes.

Conclusions

- Great accuracy for different stars - even with only few observations.
- Very easy integration of other filters.
- A great sampling efficiency, ~ 0.1 **mln samples per second**, nearly three orders of magnitude faster than the MCMC sampling on the CPU.
- Easy to integrate (as it returns probability associated with each sample), possibility to refine the results using importance sampling.
- Next generations of optical observatories (like LSST) are supposed to produce enormous amounts of data, Neural-based methods will play important role in data post-processing.

M. Dax, S. R. Green, J. Gair, M. Pürrer, J. Wildberger, J. H. Macke, A. Buonanno, and B. Schölkopf. Neural Importance Sampling for Rapid and Reliable Gravitational-Wave Inference. , 130(17):171403, Apr. 2023. doi: 10.1103/PhysRevLett.130.171403.

G. M. Green, H.-W. Rix, L. Tschesche, D. Finkbeiner, C. Zucker, E. F. Schlafly, J. Rybizki, M. Fouesneau, R. Andrae, and J. Speagle. Data-driven Stellar Models. *ApJ*, 907(1):57, Jan. 2021. doi: 10.3847/1538-4357/abd1dd.

G. Papamakarios, T. Pavlakou, and I. Murray. Masked autoregressive flow for density estimation. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/6c1da886822c67822bcf3679d04369fa-Paper.pdf.