International Conference on Machine Learning for Astrophysics
ML4ASTRO2 8-12 July, 2024

Conditional Variational Autoencoders for the analysis of the Gaia spectrophotometric data of pre-main sequence stars

This work is being carried out by the Extended Stellar Parametrizer - Cool Stars (ESP-CS) group within DPAC's Coordination Unit 8 (CU8) as part of the Gaia mission.

Speaker: Dr. Cristina Paola Marcellino

INAF - Osservatorio Astrosico di Catania

T Tauri stars

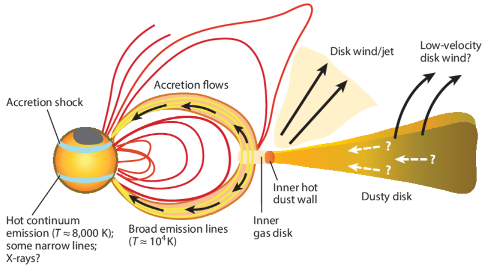


Figure 1: Magnetospheric accretion sketch for a low mass T Tauri star. From Hartmann et al. (2016).

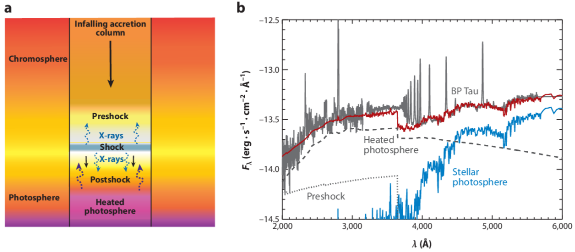


Figure 2: **a**: schematic view of an accretion column. **b**: a typical spectrum of an accreting star with the various flux contributions. From Hartmann et al. (2016)

- presence of a circumstellar disk
- magnetospheric accretion model (Hartmann 1994)
- accretion diagnostic: continuum excess

Goal: inference of stellar and accretion parameters from Gaia BP/RP spectra

Grid of $\sim$ 160 000 precomputed model BP/RP spectra parametrized by

- effective temperature, $T_{\rm eff}$: (3000, 7000) K
- surface gravity, $\log g$: (2.5, 5.0) dex
- Metallicity, $[M/H] = \log_{10}(\frac{N_{Fe}}{N_H})_{star} - \log_{10}(\frac{N_{Fe}}{N_H})_{Sun}$: (-0.5, 1.0) dex
- Mass accretion rate, $\dot{M}$: (-11, -4.52) $\log M_\odot / yr$
- filling factor, ff: (0.0, 0.3)
- Mass, M: (0.25, 1.3) $M_\odot$
- Extinction, A_0: (0.0, 10.0)

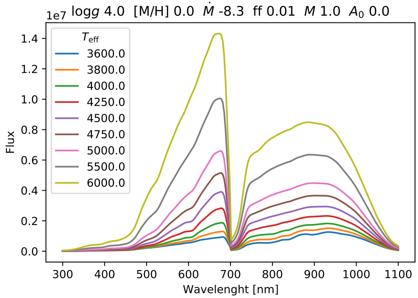


Figure 3: From Lanzafame et al. (In preparation). Gaia XP model spectra at different $T_{\rm eff}$ values.

Goal: inference of stellar and accretion parameters from Gaia BP/RP spectra

Grid of $\sim$ 160 000 precomputed model BP/RP spectra parametrized by

- effective temperature, T_{eff}: (3000, 7000) K
- surface gravity, $\log g$: (2.5, 5.0) dex
- Metallicity, $[M/H] = \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{star} - \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{Sun}$: (-0.5, 1.0) dex
- Mass accretion rate, $\dot{M}$: (-11, -4.52) $\log M_\odot/yr$
- filling factor, *ff*: (0.0, 0.3)
- Mass, M: (0.25, 1.3) $M_\odot$
- Extinction, A_0: (0.0, 10.0)

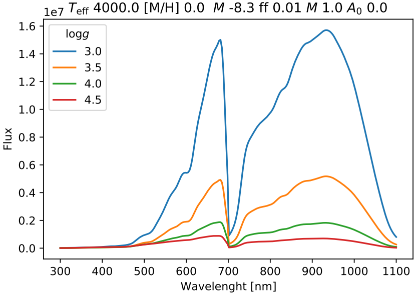


Figure 3: From Lanzafame et al. (In preparation). Gaia XP model spectra at different $\log g$ values.

Goal: inference of stellar and accretion parameters from Gaia BP/RP spectra

Grid of $\sim$ 160 000 precomputed model BP/RP spectra parametrized by

- effective temperature, T_{eff}: (3000, 7000) K
- surface gravity, $\log g$: (2.5, 5.0) dex
- Metallicity, $[M/H] = \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{star} - \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{Sun}$: (-0.5, 1.0) dex
- Mass accretion rate, $\dot{M}$: (-11, -4.52) $\log M_{\odot}/yr$
- filling factor, ff: (0.0, 0.3)
- Mass, M: (0.25, 1.3) $M_{\odot}$
- Extinction, A_0: (0.0, 10.0)

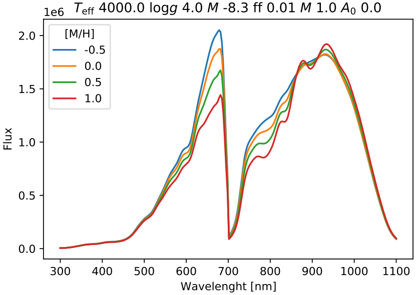


Figure 3: From Lanzafame et al. (In preparation). Gaia XP model spectra at different [M/H] values.

Goal: inference of stellar and accretion parameters from Gaia BP/RP spectra

Grid of ∼ 160 000 precomputed model BP/RP spectra parametrized by

- effective temperature, $T_{\rm eff}$: (3000, 7000) K
- surface gravity, $\log g$: (2.5, 5.0) dex
- Metallicity, $[M/H] = \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{star} - \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{Sun}$: (-0.5, 1.0) dex
- Mass accretion rate, $\dot{M}$: (-11, -4.52) $\log M_\odot/yr$
- filling factor, ff: (0.0, 0.3)
- Mass, M: (0.25, 1.3) $M_\odot$
- Extinction, A_0: (0.0, 10.0)

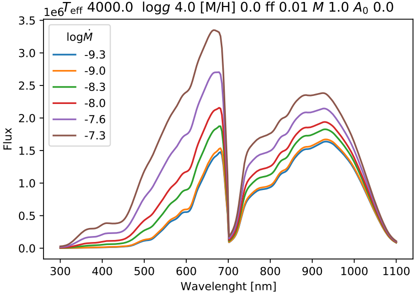


Figure 3: From Lanzafame et al. (In preparation). Gaia XP model spectra at different $\dot{M}$ values.

Goal: inference of stellar and accretion parameters from Gaia BP/RP spectra

Grid of $\sim$ 160 000 precomputed model BP/RP spectra parametrized by

- effective temperature, $T_{\rm eff}$: (3000, 7000) K
- surface gravity, $\log g$: (2.5, 5.0) dex
- Metallicity, $[M/H] = \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{star} - \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{Sun}$: (-0.5, 1.0) dex
- Mass accretion rate, $\dot{M}$: (-11, -4.52) $\log M_{\odot}/yr$
- filling factor, ff: (0.0, 0.3)
- Mass, M: (0.25, 1.3) $M_{\odot}$
- Extinction, A_0: (0.0, 10.0)

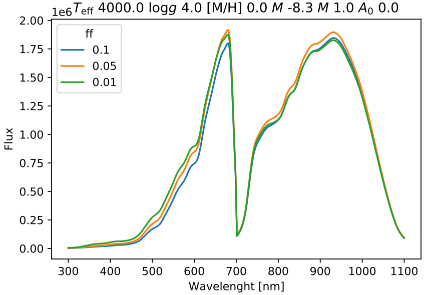


Figure 3: From Lanzafame et al. (In preparation). Gaia XP model spectra at different ff values.

Goal: inference of stellar and accretion parameters from Gaia BP/RP spectra

Grid of $\sim$ 160 000 precomputed model BP/RP spectra parametrized by

- effective temperature, T_{eff}: (3000, 7000) K
- surface gravity, $\log g$: (2.5, 5.0) dex
- Metallicity, $[M/H] = \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{star} - \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{Sun}$: (-0.5, 1.0) dex
- Mass accretion rate, $\dot{M}$: (-11, -4.52) $\log M_\odot/yr$
- filling factor, ff: (0.0, 0.3)
- Mass, M: (0.25, 1.3) $M_\odot$
- Extinction, A_0: (0.0, 10.0)

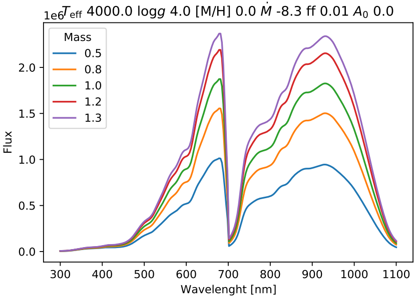


Figure 3: From Lanzafame et al. (In preparation). Gaia XP model spectra at different M values.

Goal: inference of stellar and accretion parameters from Gaia BP/RP spectra

Grid of $\sim$ 160 000 precomputed model BP/RP spectra parametrized by

- effective temperature, T_{eff}: (3000, 7000) K
- surface gravity, $\log g$: (2.5, 5.0) dex
- Metallicity, $[M/H] = \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{star} - \log_{10}\left(\frac{N_{Fe}}{N_H}\right)_{Sun}$: (-0.5, 1.0) dex
- Mass accretion rate, $\dot{M}$: (-11, -4.52) $\log M_{\odot}/yr$
- filling factor, ff: (0.0, 0.3)
- Mass, M: (0.25, 1.3) $M_{\odot}$
- Extinction, A_0: (0.0, 10.0)

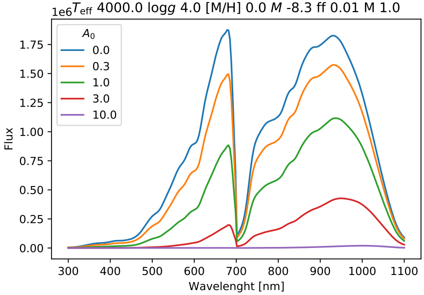


Figure 3: From Lanzafame et al. (In preparation). Gaia XP model spectra at different A_0 values.

Conditional Variational Autoencoders (CVAE)

- Autoencoders: encoder-decoder structure $\rightarrow$ compression of input data into a lower dimensional space (latent space)
- Variational: the encoder and decoder networks parametrize statistical distributions $\rightarrow$ samples from the latent space go into the decoder $\rightarrow$ range of possible reconstructed input
- Conditional: the conditional information is embedded in the encoding and decoding process

$\downarrow$

CVAE applicable to approximate the Bayesian posterior

$$p(x|y) \propto \underbrace{p(y|x)}_{likelihood}\underbrace{p(x)}_{prior}, \tag{1}$$

where

- x, stellar and accretion parameters (T_{eff}, $\log g$, [M/H], $\dot{M}$, ff, M, A_0)
- y, the BP/RP spectra

Network architecture from Gabbard et al. (2022)

Cost Function $$H(p,r) = \int dx \, \underbrace{p(x|y)}_{\text{true posterior}} \log \underbrace{r_\theta(x|y)}_{\text{parametric distribution}} \quad (2)$$

$$r_\theta(x|y) = \int dz \, \underbrace{r_{\theta_1}(z|y)}_{\text{encoder distribution}} \underbrace{r_{\theta_2}(x|y,z)}_{\text{decoder distribution}} \quad (3)$$

From (2) $\rightarrow$ Evidence Lower Bound optimization

$$\downarrow$$

$$H \lesssim \frac{1}{N} \sum_{n=1}^{N_b} \left(\underbrace{-\log r_{\theta_2}(x_n|z_n|y_n)}_{\text{reconstruction loss L}} \right) + \underbrace{KL\left(q_\phi(z|x_n, y_n)||r_{\theta_1}(z|y_n)\right)}_{\text{Kullback-Leibler (KL) divergence}} \quad (4)$$

where $q_\phi(z|x_n, y_n)$, 'recognition' encoder network

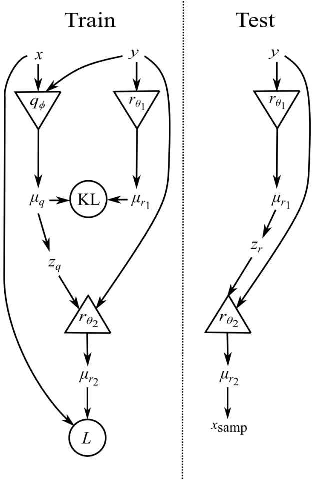
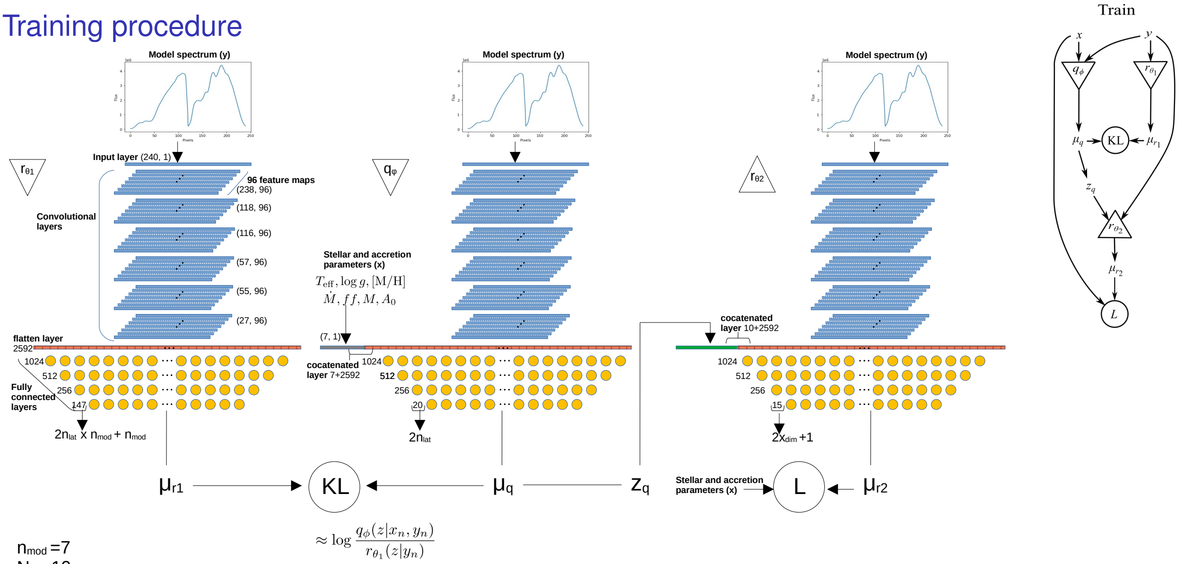


Figure 4: Network architecture. From Gabbard et al. (2022).

Training procedure



$n_{mod}=7$
$N_{lat}=10$
$x_{dim}=7$

The cost function as a function of the training epochs

- annealing procdure: KL contribution ignored for the first 50 epochs, then logarithmically increased from 0 to 1 between epoch 50 and 150.
- Adam optimizer, learning rate 10^{-5}
- Total elapsed time $\sim$ 5.5 days (HPC resources: Pleiadi Cluster, OACT)

No overfitting to the training data

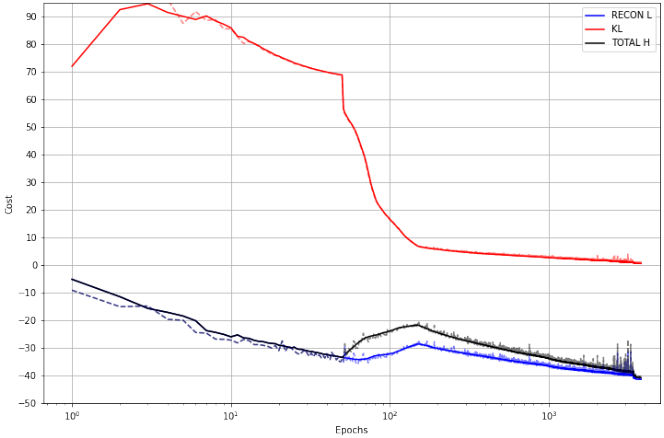
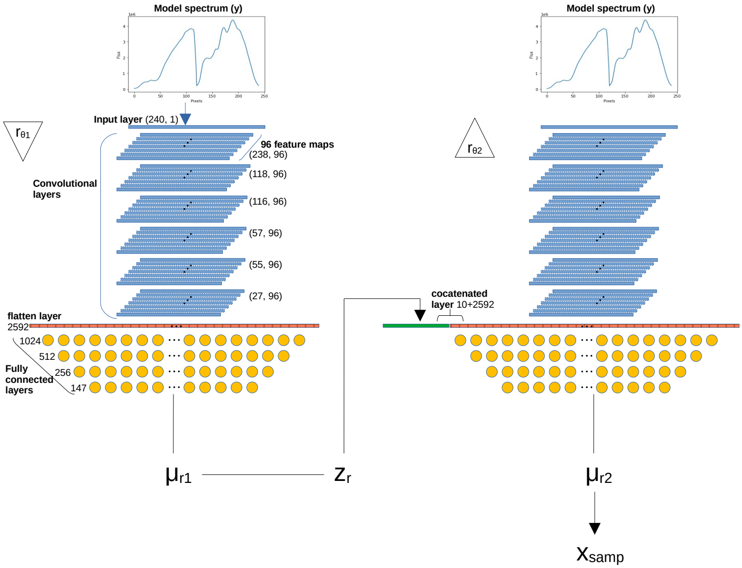
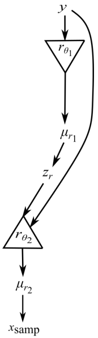


Figure 6: The total cost function H, the KL-divergence contribution and the reconstruction component L as the training proceeds.

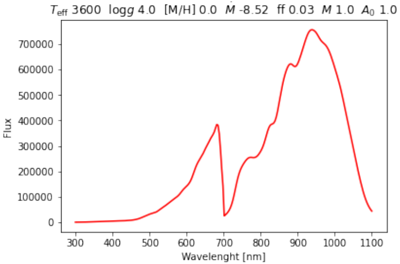
Test procedure

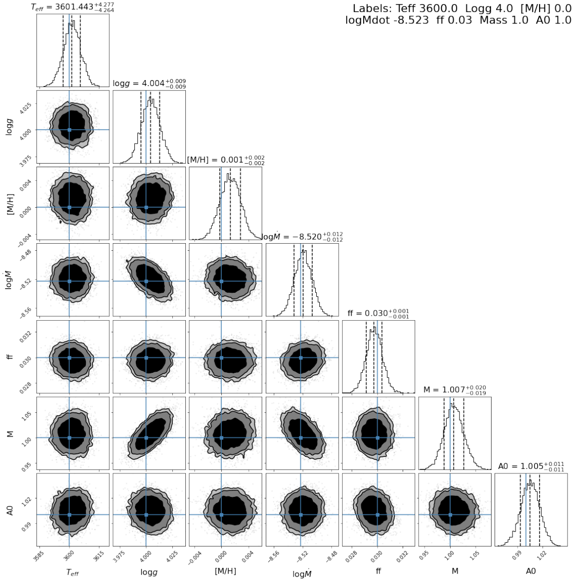


Test



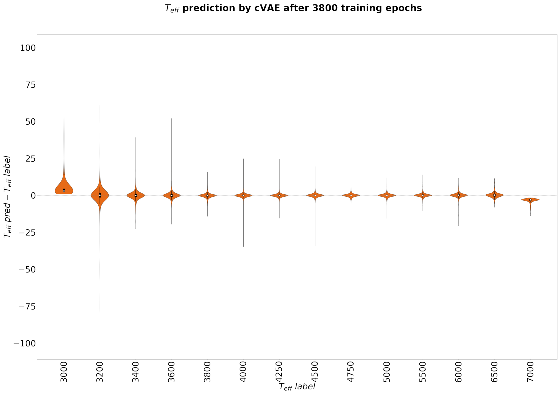
Recovering stellar and accretion parameters of a model spectrum from the test dataset

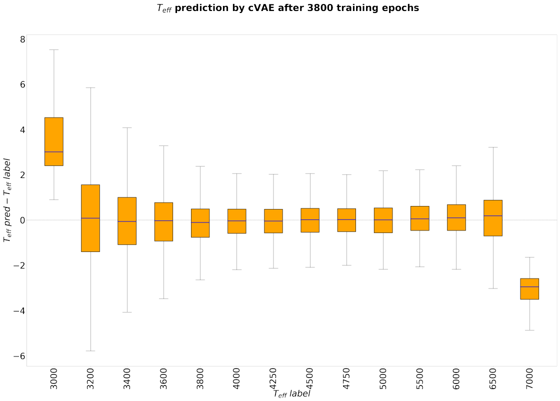




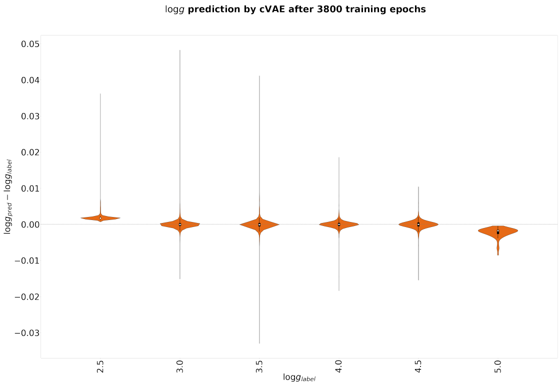
- x, stellar and accretion parameters ($T_{\rm eff}$, $\log g$, [M/H], $\dot{M}$, ff, M, A_0)
- y, the BP/RP spectra

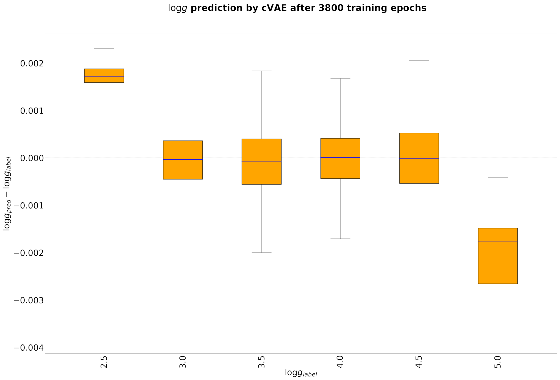
Accuracy at recovering stellar and accretion parameters of the whole test dataset



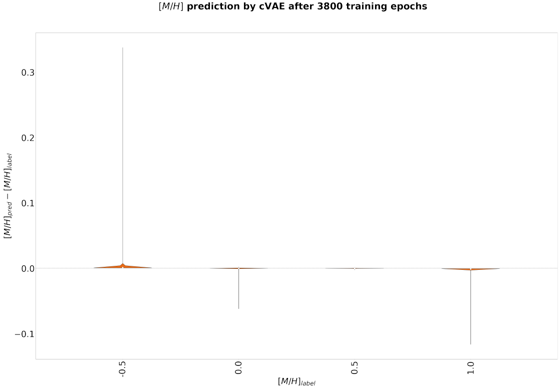


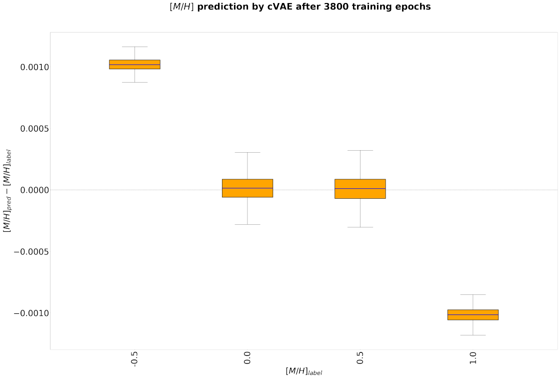
Accuracy at recovering stellar and accretion parameters of the whole test dataset



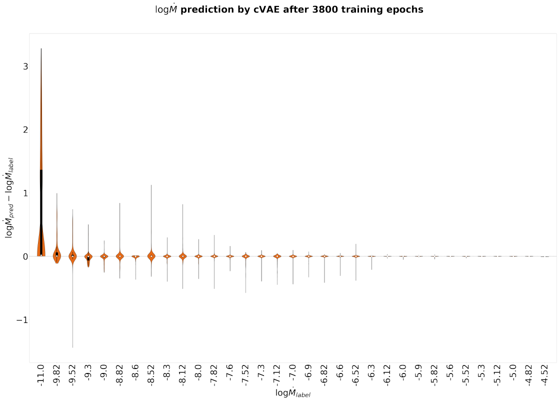


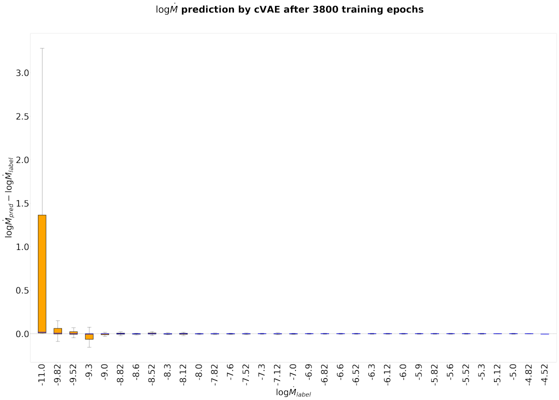
Accuracy at recovering stellar and accretion parameters of the whole test dataset



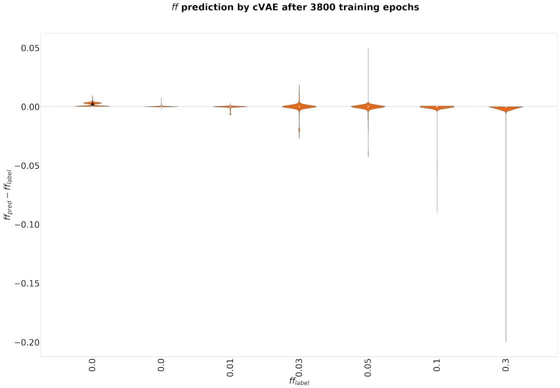


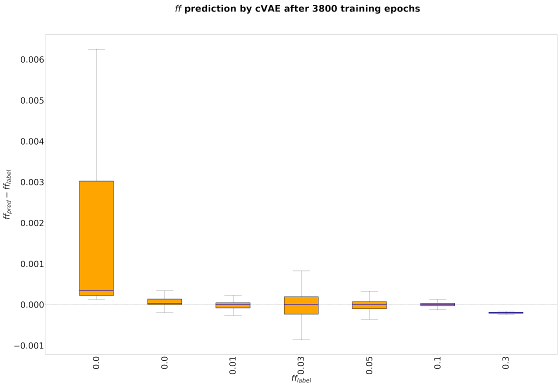
Accuracy at recovering stellar and accretion parameters of the whole test dataset



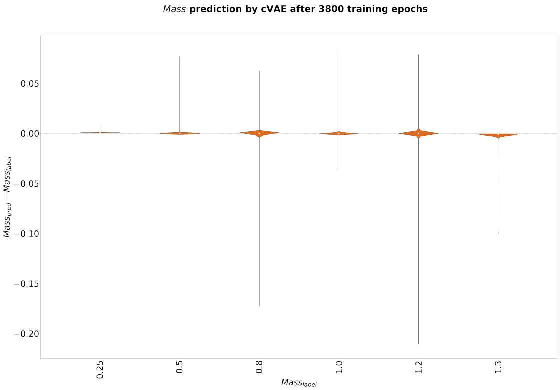


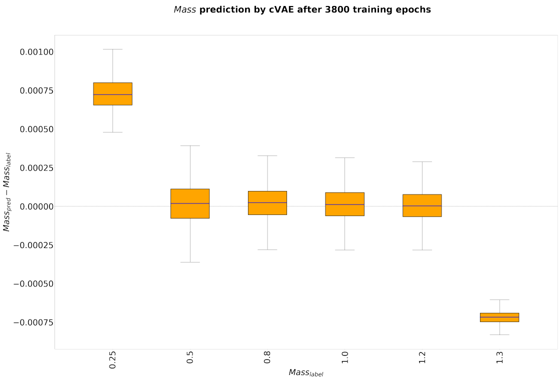
Accuracy at recovering stellar and accretion parameters of the whole test dataset



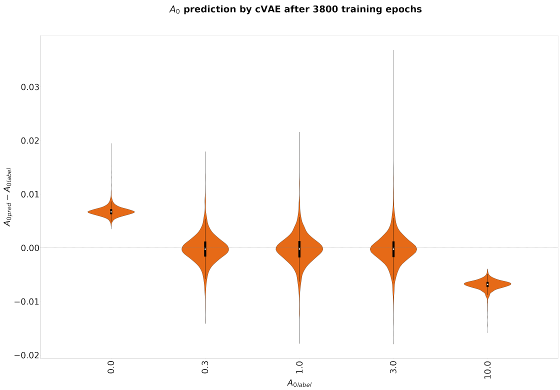


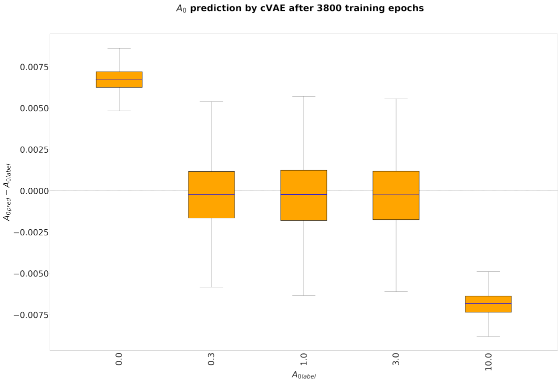
Accuracy at recovering stellar and accretion parameters of the whole test dataset



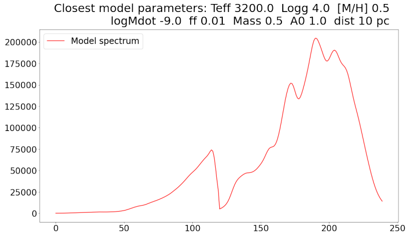


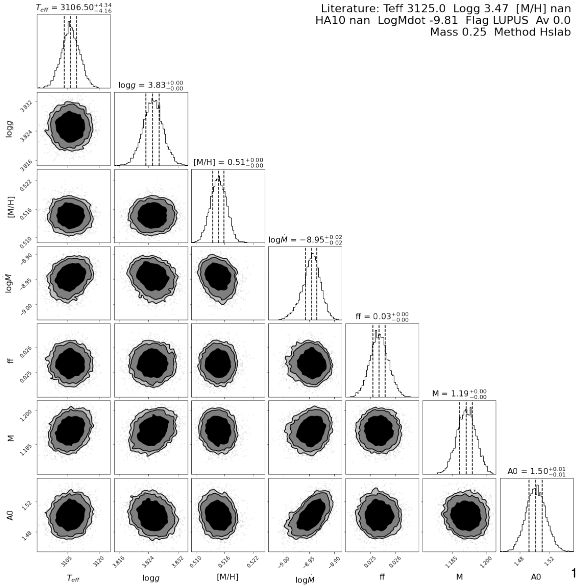
Accuracy at recovering stellar and accretion parameters of the whole test dataset



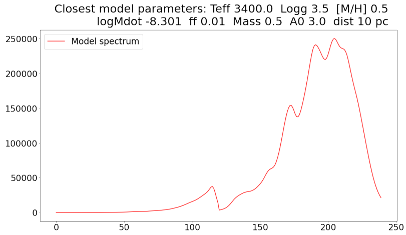


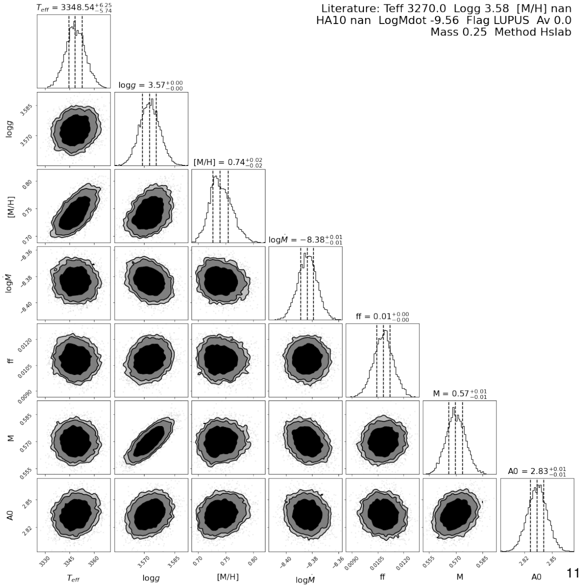
Inferring the posterior of stellar parameters of an anonymous source



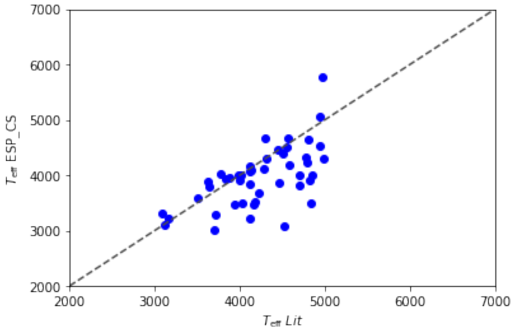


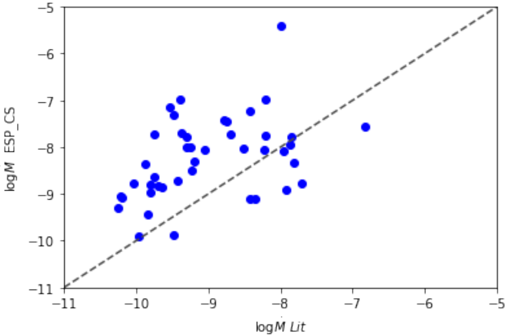
Inferring the posterior of stellar parameters of an anonymous source





Comparison with the literature for 40 anonymous sources





Conclusions

- CVAE allow to approximate a multidimensional bayesian posterior $\rightarrow$ uncertainties on the parameters of interest
- Inference with CVAE is faster than by traditional sampling methods (0.5 s per source with Java)
- Our CVAE are not data-driven: they are conditioned on BP/RP model spectra generated by taking the magnetospheric accretion model into account. The CVAE model is good as our training set.

References

Gabbard, H., Messenger, C., Heng, I. S., Tonolini, F., & Murray-Smith, R. 2022, Nature Physics

Hartmann, L. 1994, in Theory of Accretion Disks — 2, ed. W. J. Duschl, J. Frank, F. Meyer, E. Meyer-Hofmeister, & W. M. Tscharnuter (Dordrecht: Springer Netherlands), 19–33

Hartmann, L., Herczeg, G., & Calvet, N. 2016, ARA&A, 54, 135

Lanzafame, A. C., Marcellino, C. P., Brugaletta, E., Licata, E., & et al. In preparation