Exploiting large multiplex spectroscopic surveys of young clusters with neural networks

Leonardo Testi, Da Eun Kang (University of Bologna)

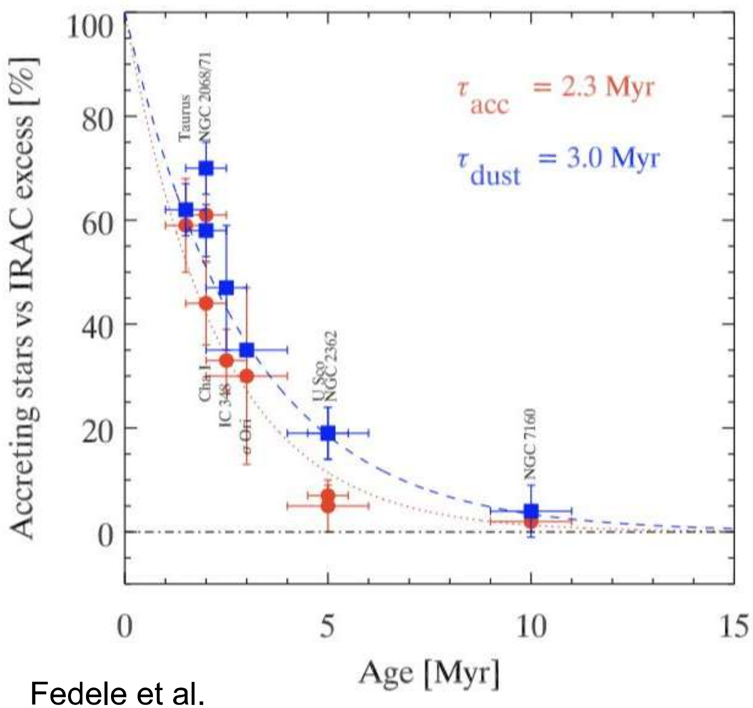
Dominika Itrich, Victor Ksoll, Ralf Klessen, Megan Reiter

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

WST workshop, March 10-13th, 2025, Naples, Italy

Young stars, planets and cluster environme

Planet formation in protoplanetary disk



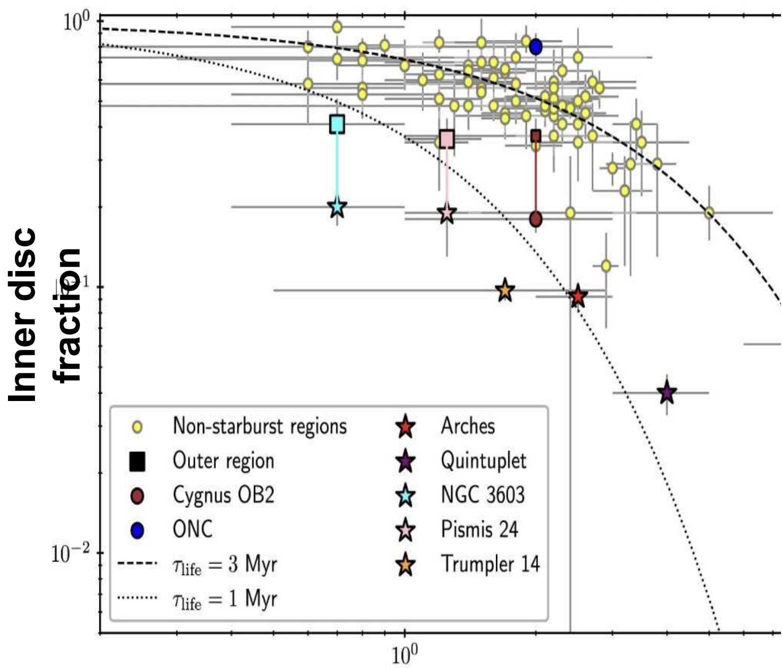
Fedele et al. 2010

Disk lifetime (e-folding): ~2-3 Myr

External photoevaporation intense UV radiation from massive s

Proply d

Aru et al. 2024



Reduced disk-life time affects forn Jupiter-like giant planets

Young stars, planets and cluster environme

Stellar and disk properties

Many young stars for statistics

Diverse clu environm

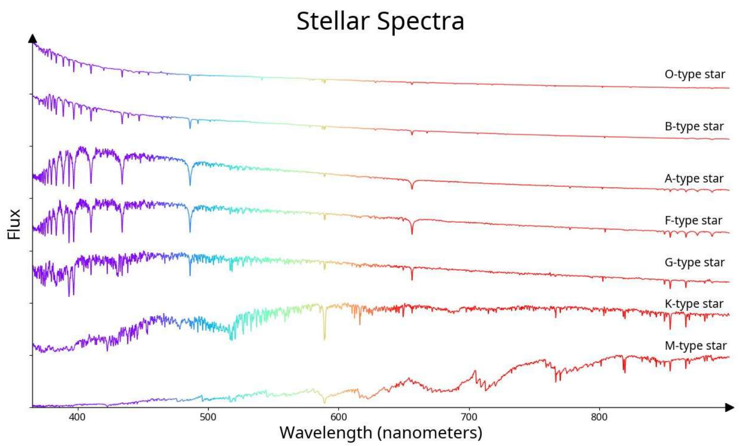
Young stars, planets and cluster environme

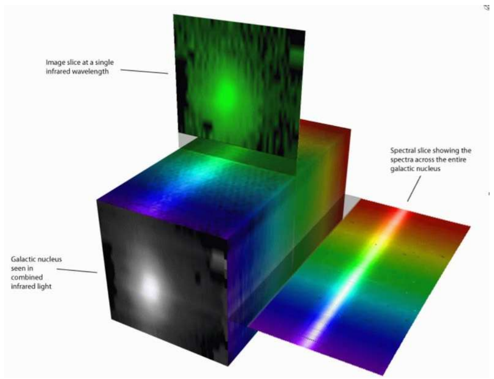
Stellar and disk properties

Many young stars for statistics

Diverse clu environm

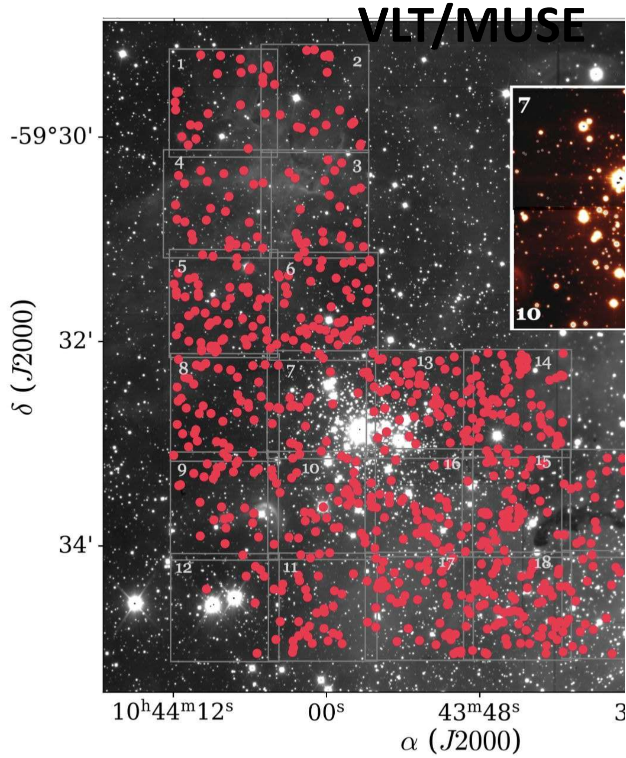
Optical stellar spectrum & Integral Field Spectroscopy (IFS)





muse
multi unit spectroscopic explorer

Spectra of ~700 low-ma
Trumpler 14 observe
VLT/MUSE

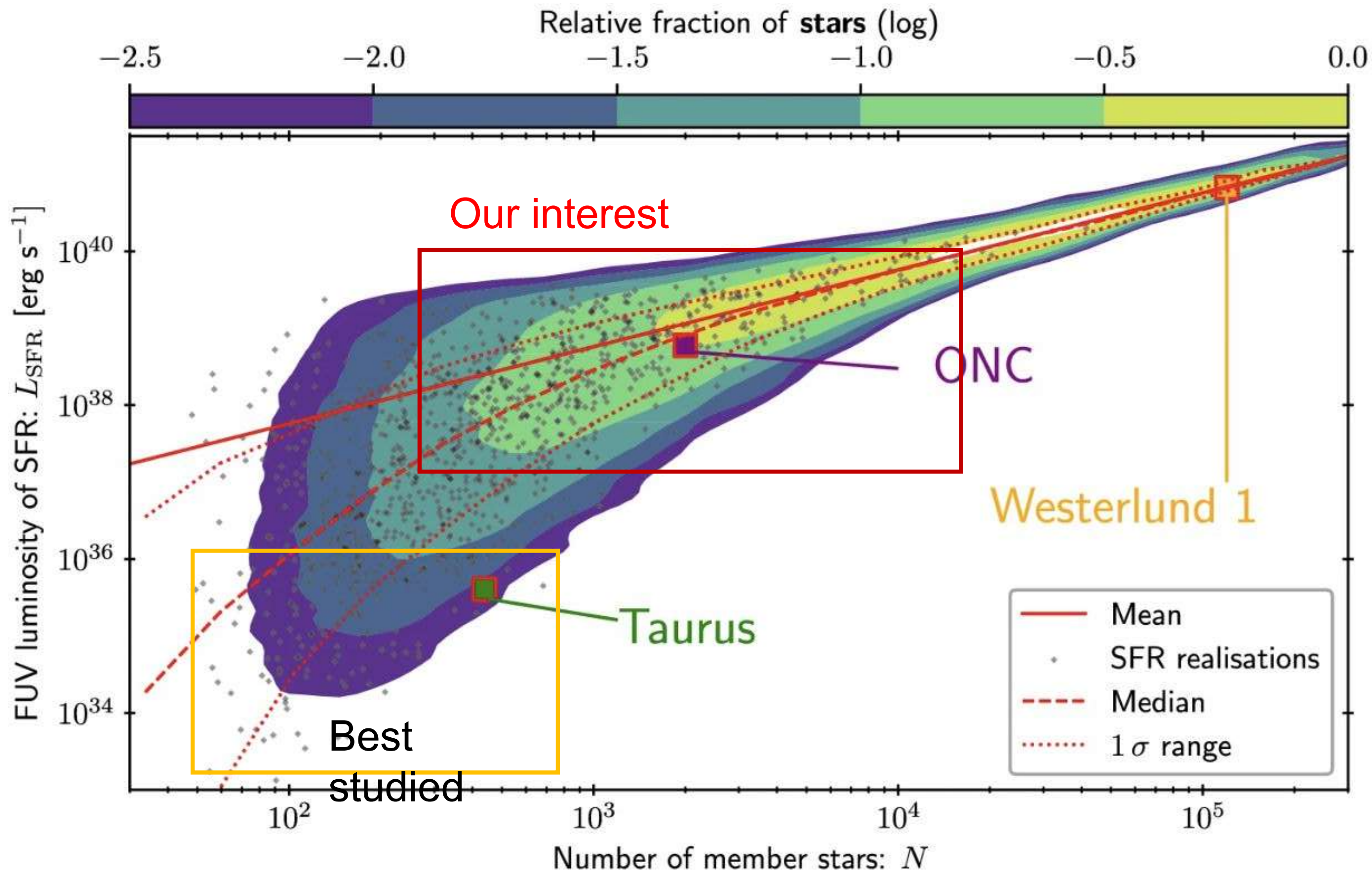


Young stars, planets and cluster environme

Stellar and disk properties

Many young stars for statistics

Diverse clu environm



Efficient, large surv young clusters, co diverse cluste environments

Focus on typical, middle range e

Ongoing MUSE survey (PI: M. Re

Wide-field Spectroscopic Telescope (W

Stellar and disk properties

Many young stars for statistics

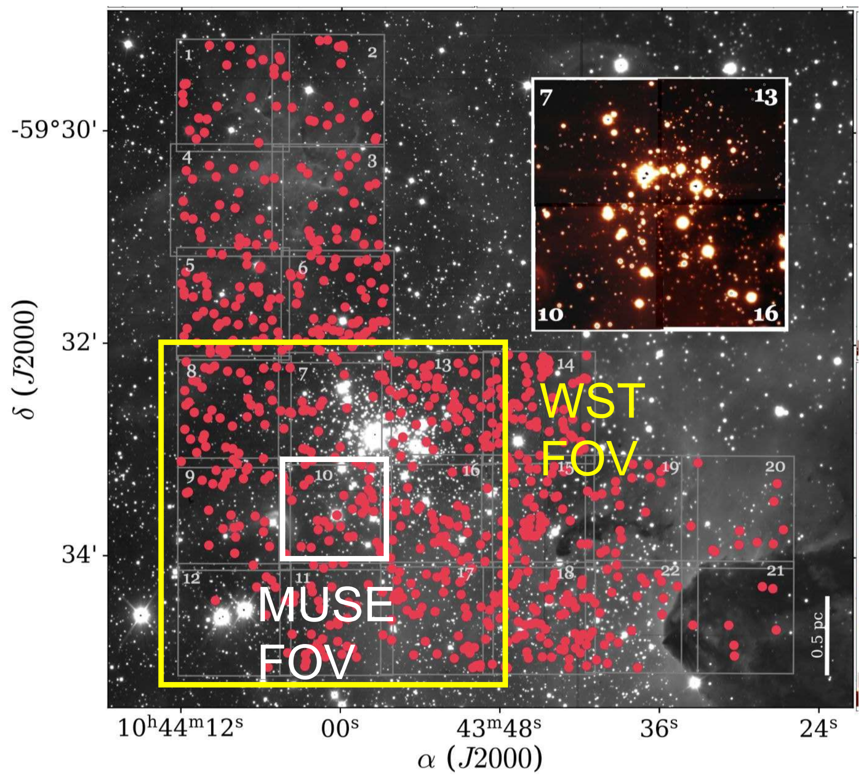
Diverse cl environm

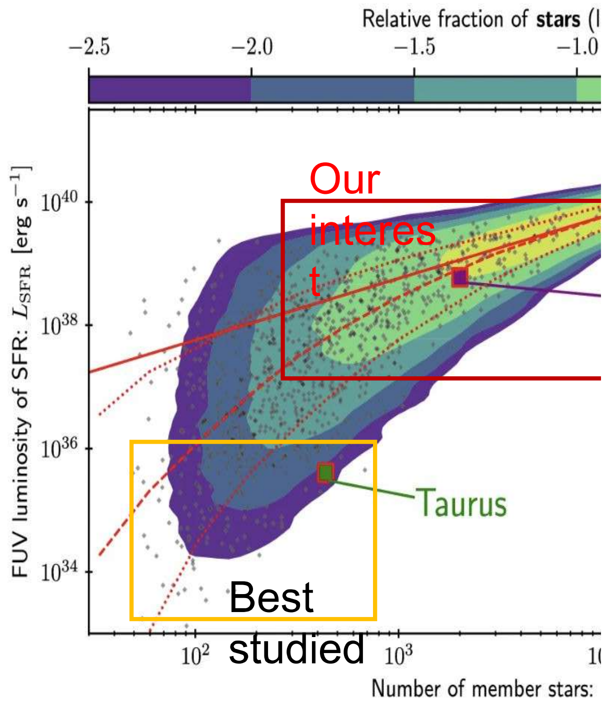
VLT/MUSE: 4600—9300Å

(R: 1770—3590), 0.2" spaxel

WST (IFS): 3700—9700Å

(R: ~3500), 0.25" spaxel





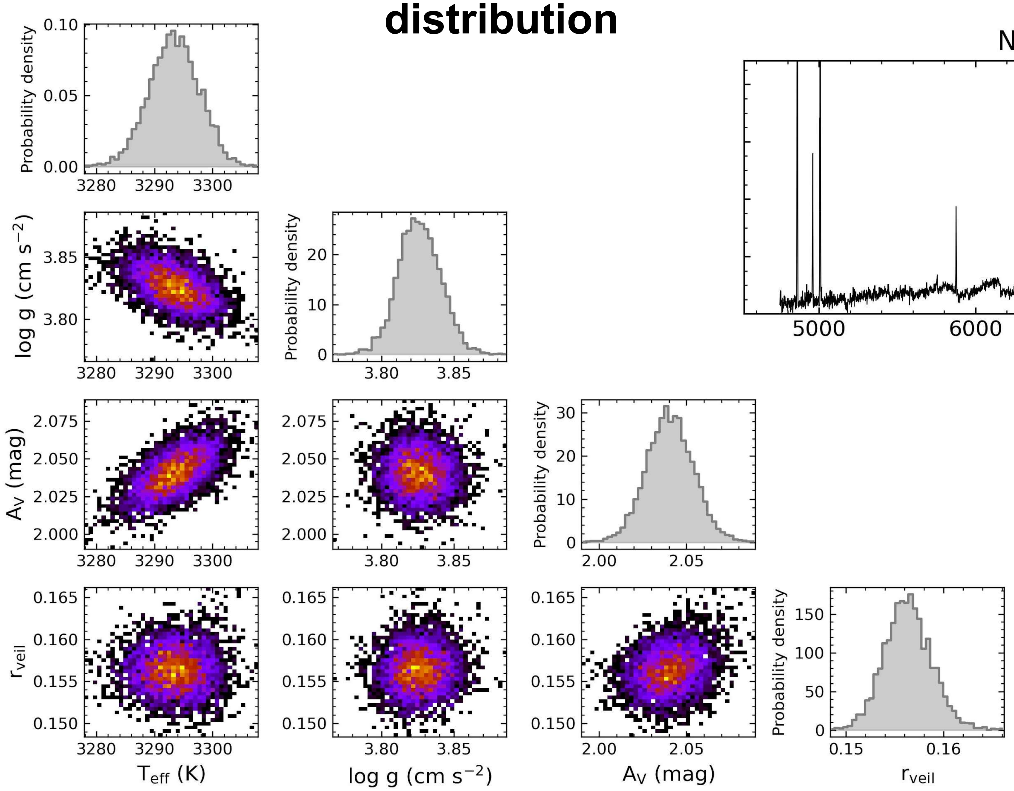
Need efficient method to analyze large data volum

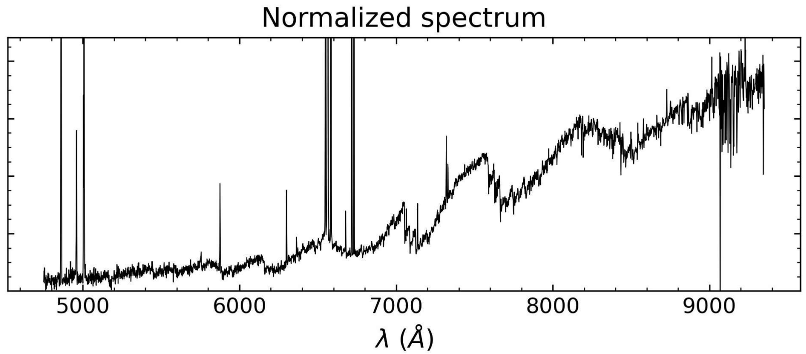
Neural network for young low-mass sta

conditional invertible neural network (cINN) for fast spectral classificatio

Stellar parameters (T_{eff}, log g, A_V, r_{veil}) ← cINN ← **Stellar spectrum (F_λ) (VLT/MUSE)**

Full posterior distribution





cINN

Kang+23: applicability test

Kang+25 (in revision): app

Speed

~10 obs/s (M1 pro CPU)

~13 obs/s (M2 pro CPU)

~100 obs/s (GPU 2080Ti)

~ 250 obs/s (GPU H100)

To analyze 10^5 stars:

- 2.7 h (M1 pro), 17 min 2080Ti)
- 2.14h (M2 pro), 7 min (

Training data

Phoenix stellar atmosphere model (Allard et al. 2012, Husser et al. 2013, Baraffe et al. 2015)

Stellar photosphere spectrum (flux) from given $\mathbf{T_{eff}}$ and **log g**

- Interpolate publicly available library: **BT-Settl**, NextGen, **Dusty**, etc
- $\mathbf{T_{eff}}$ **: 2600 – 7000K, log g: 2.5 – 5**
- **MUSE** spectrum: **4700 – 9300Å (R~4000**, dλ~1.25Å)

Veiling: extra emission from mass accretion or disk emission (constant veiling model)

$$F_{\lambda,\mathrm{veiled}} = F_\lambda + F_{7500} \times r_{\mathrm{veil}}$$ or at 7500, $\mathbf{r_{veil}}$**: 0—2**

Extinction: (A_V, R_V, Cardelli et al. 1989)

Rv=4.4 fixed for Trumpler 14 (Hur+2012), **Av: 0—10 mag**

Generate 60,000 ~ 100,000 models (use 80% for training and 20% for tests)

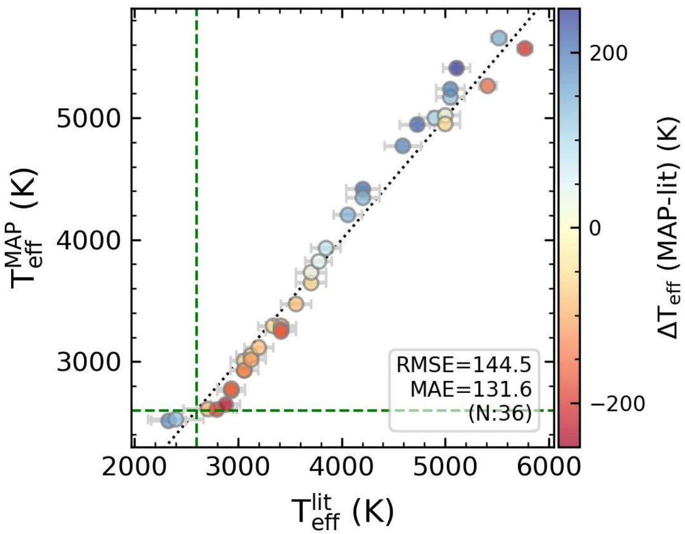
Performance tests and Application to Tr14

36 class III template stars

- Observed with VLT/X-shooter
- Downgrade spectral resolution

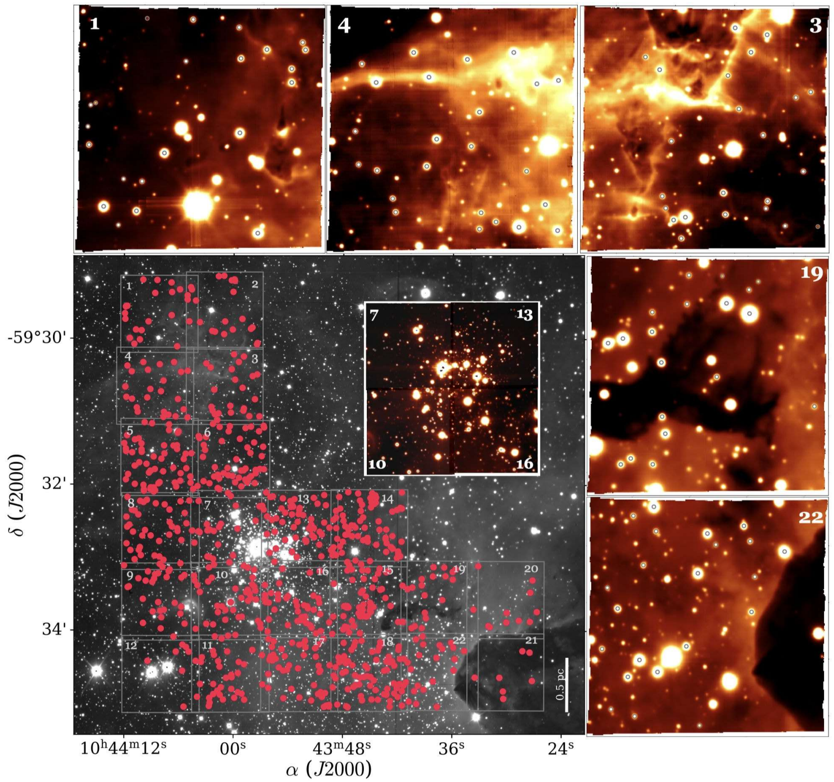
Average accuracy

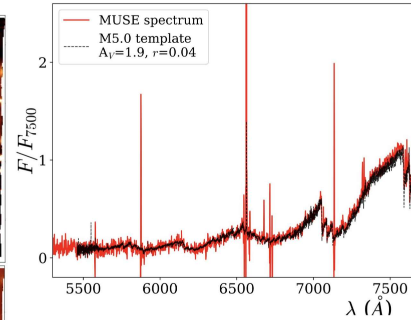
- T_{eff} : 3.5%, 0.95-subclass
- log g: 0.3 dex
- Av: 0.2 – 1 mag
- r : 0.13



2051 stars observed with VLT/MUSE

- Tr14: a massive cluster in the Carina Nebula Compl
- ~20 O-type stars. Intense UV field
- T_{eff}, A_V and r_{veil} by template fitting (Itirch+2024)



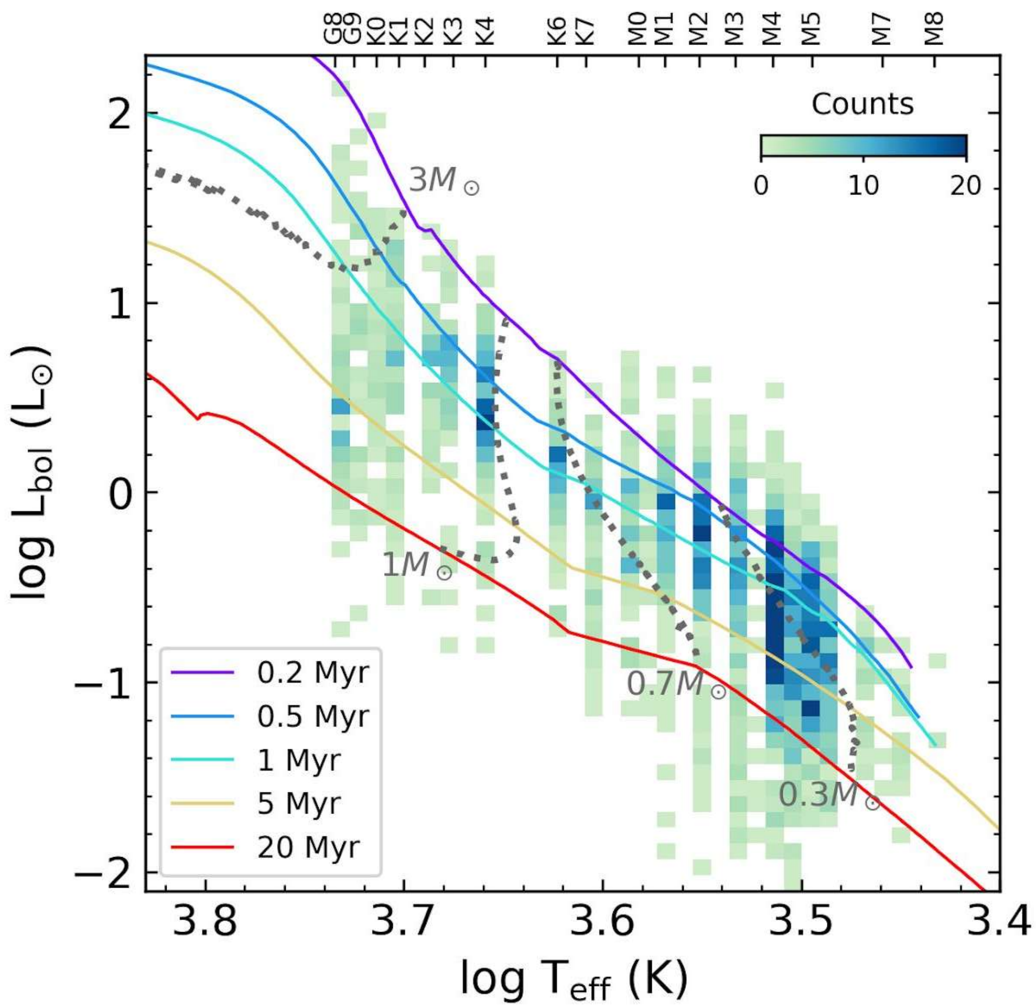


Itrich et al. 2024

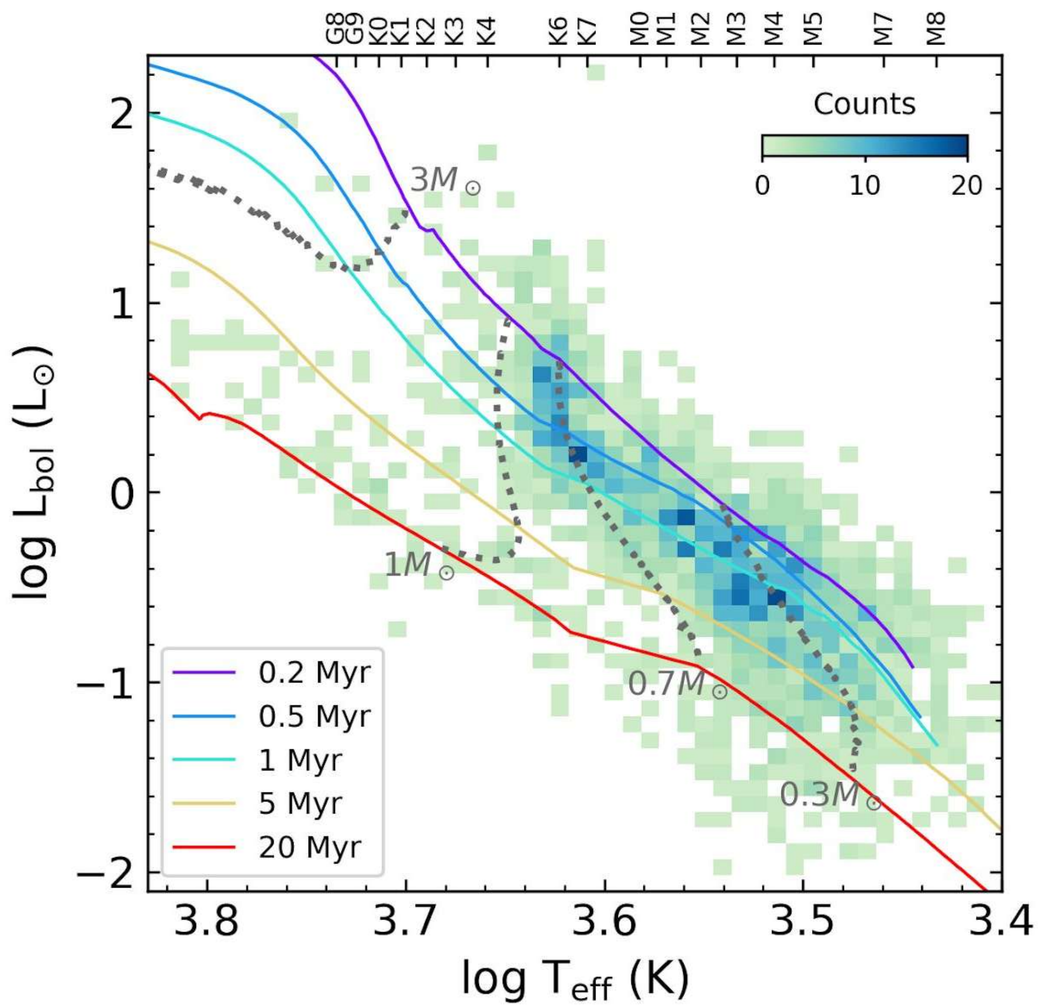
Temp
fitti
vs. c

Template fitting vs. cINN

Template fitting



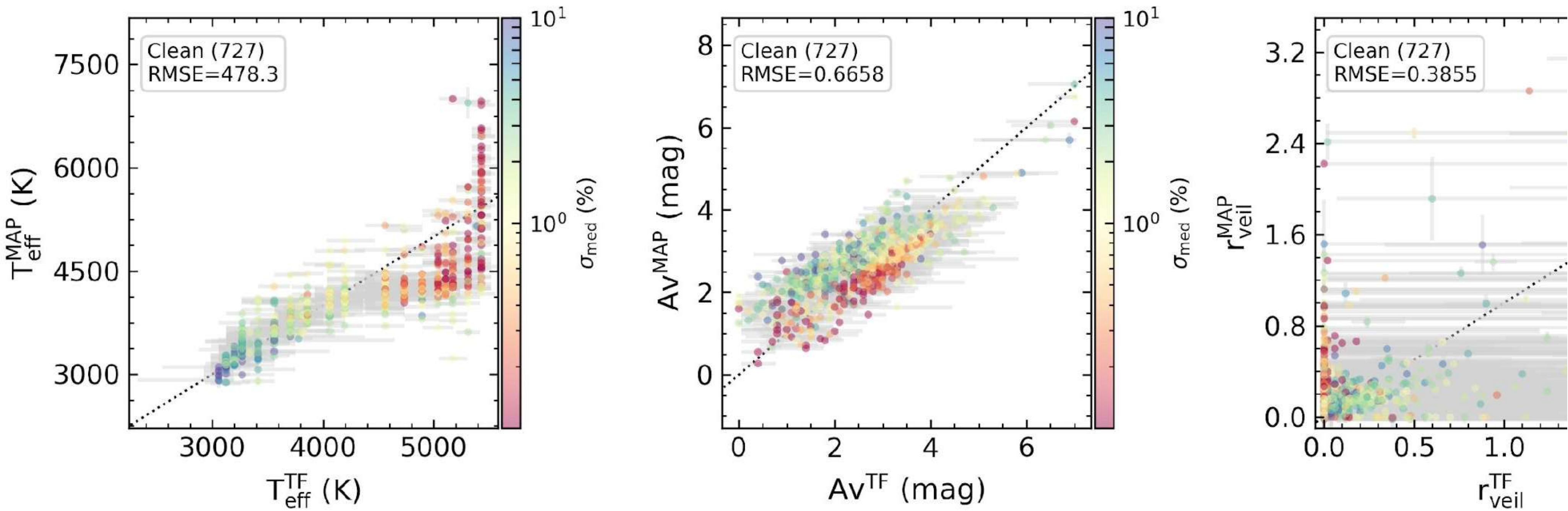
cINN



A
s
0

Template fitting vs. cINN

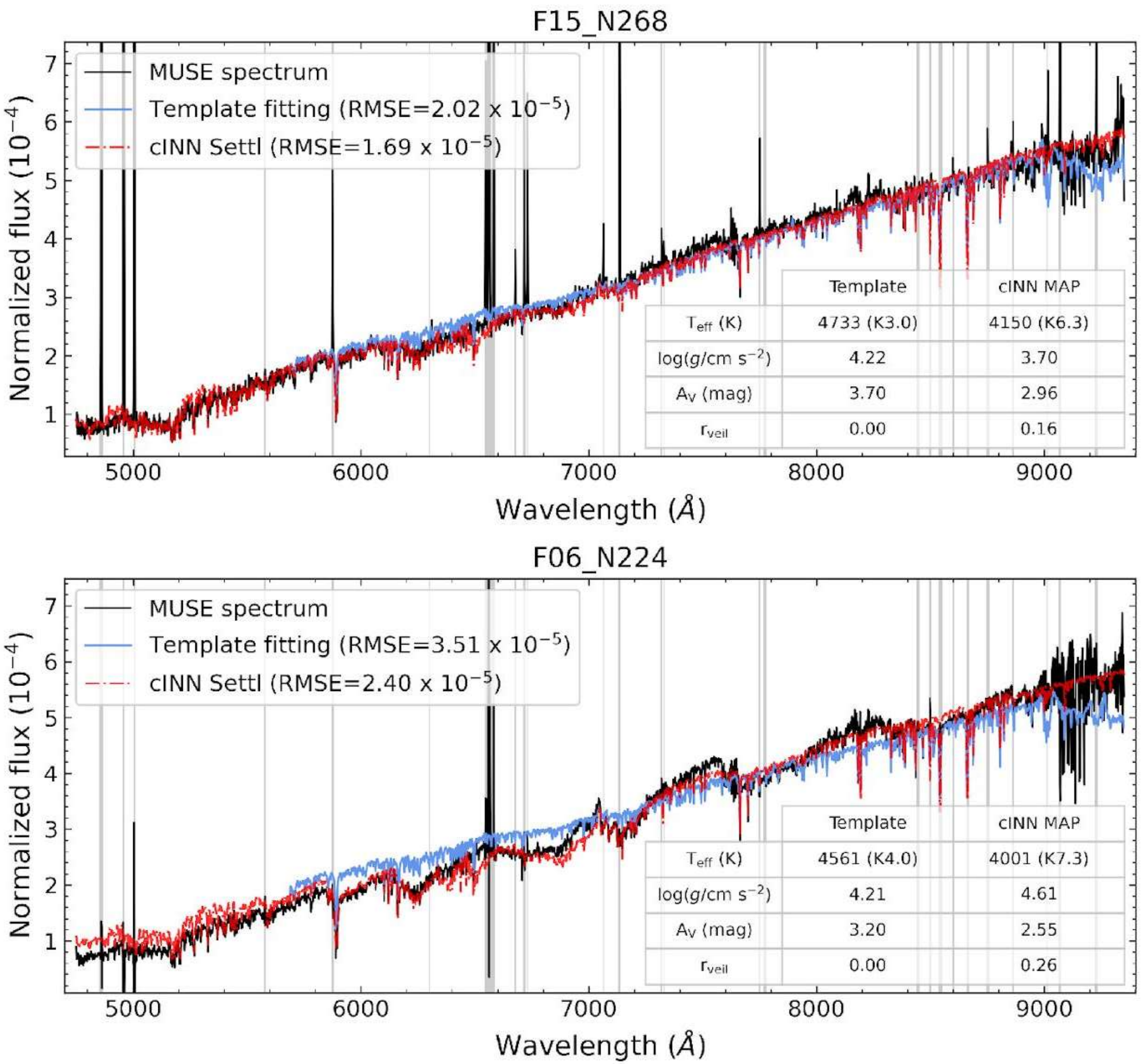
cINN



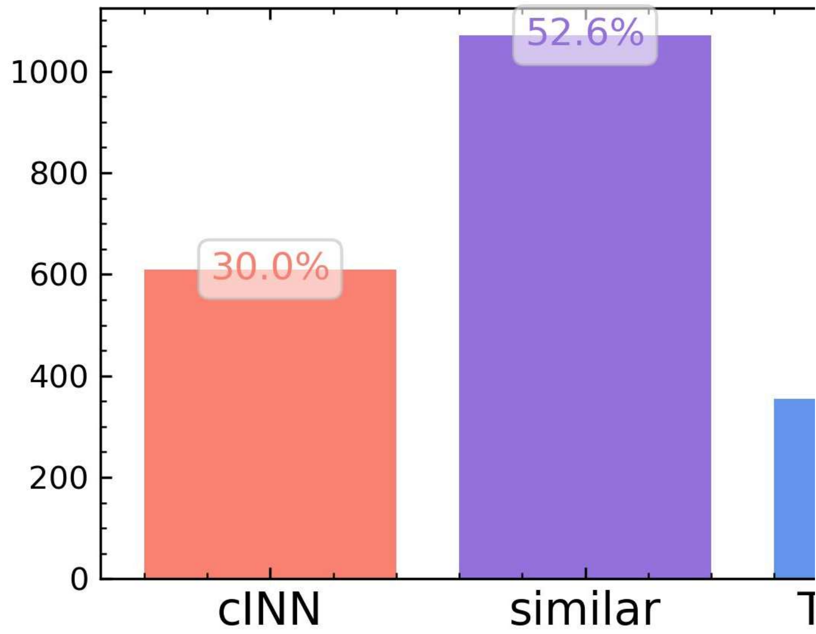
Template fitting

- Good agreements for M-type stars, large differences (T and veil) for K/G types
    - Lack of templates (No K5 type, upper limit of G8)
    - Degeneracy between T_{eff} and r_{veil}

Template fitting vs. cINN



Blue: Template fitting
Red: reproduced Phoenix spectrum



Planned updates

1. Additional parameters:
    - extinction law, Rv (currently fixed for Tr14, Rv=4.4)
    - bolometric correction
2. Improve veiling
    - use a hydrogen slab model (Manara+14, Acala+2017, etc)
3. Increase Teff range: ~ A types ($\lesssim$10,000K)
4. Improve simulation gap: domain adaptation, include real spect
5. Improved consideration of flux error

Summary

WST for studying young stellar population and cluster environment

- Need a demography of young stars in diverse cluster environments to study tl
  influence of environment on the stellar evolution and planet formation
- Need efficient large survey in optical spectroscopy
- WST can provide unprecedented large data of young stars from various cluste

Neural network to handle large data from WST

- Fast, accurate and efficient method is required to handle large data from WST
- We developed cINN that predicts 4 stellar parameters from optical spectrum
- Our cINN is very fast and shows comparable performance as template fitting
- Our works on VLT/MUSE data are preparation for WST

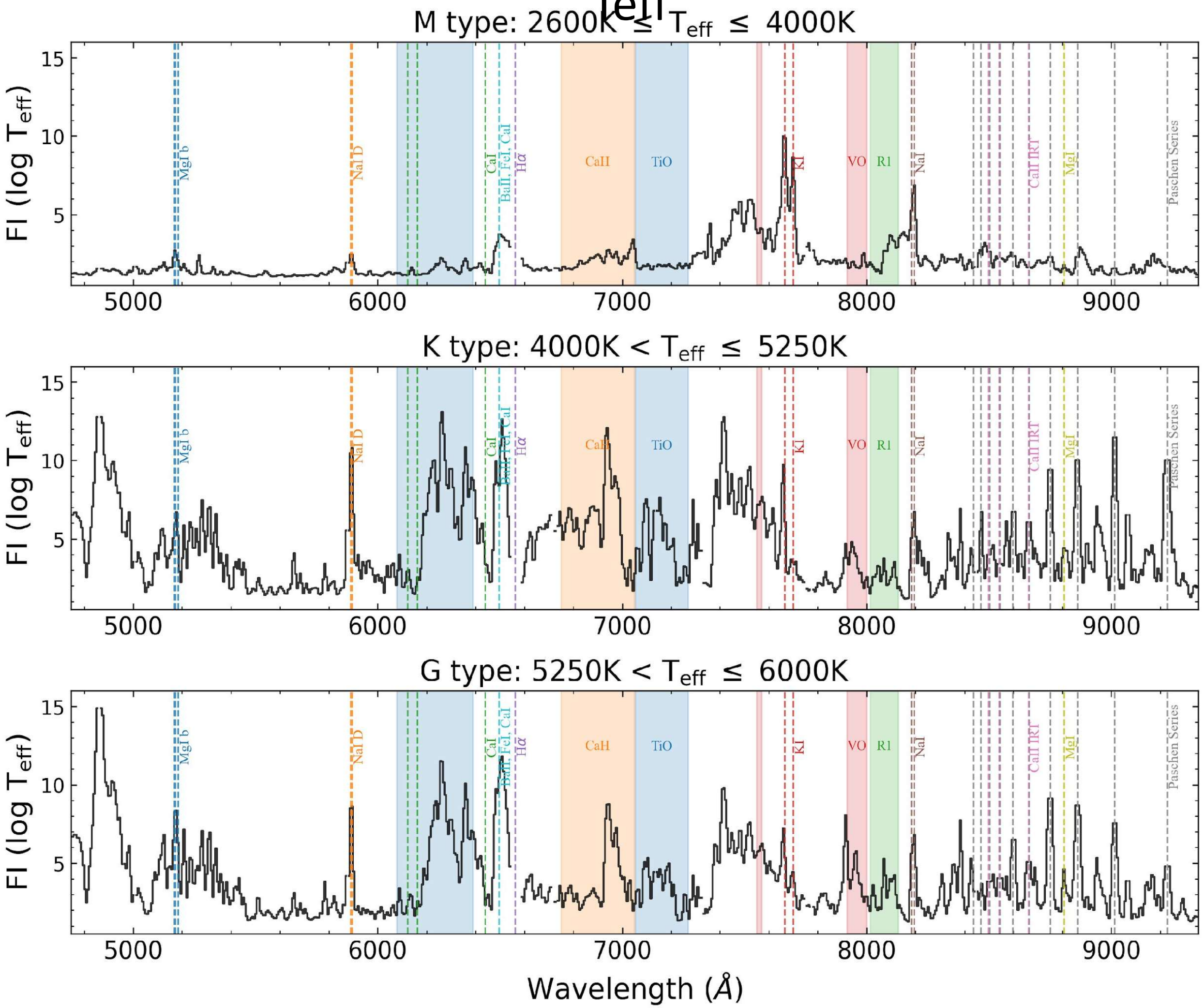
Thank you!

Q &
A

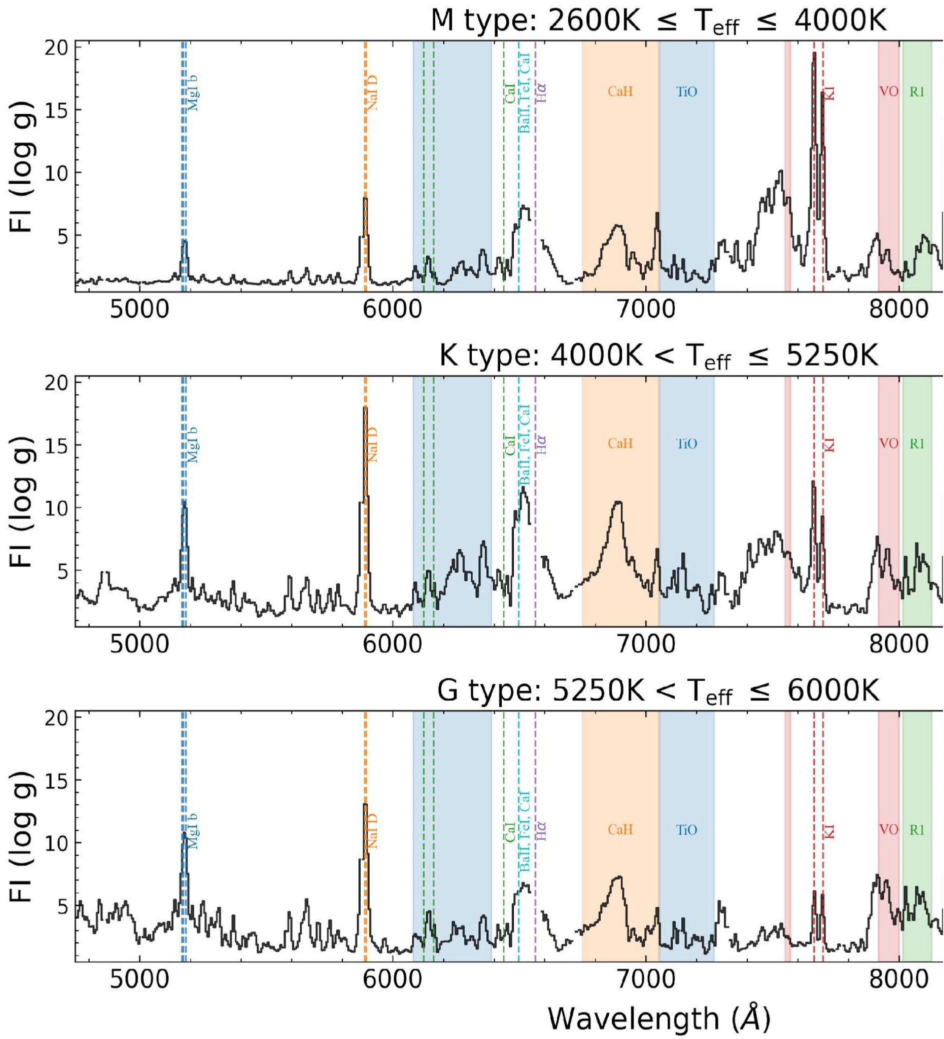
How does cINN estimate parameters?

Feature importance tests

log Teff

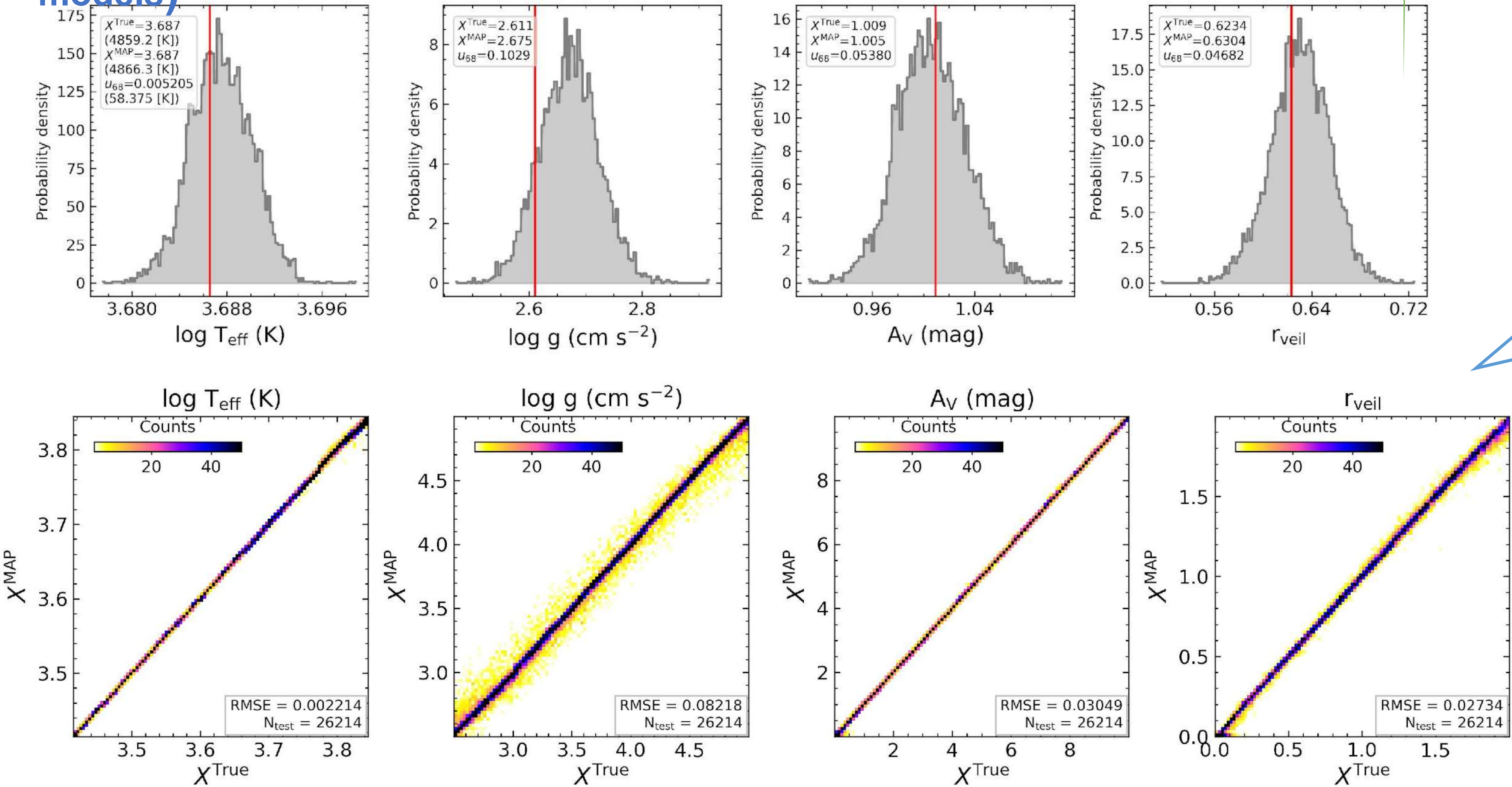


log g



Generalization on test models

Synthetic test models (~26,000 test models)



1D posterior dis
for one test

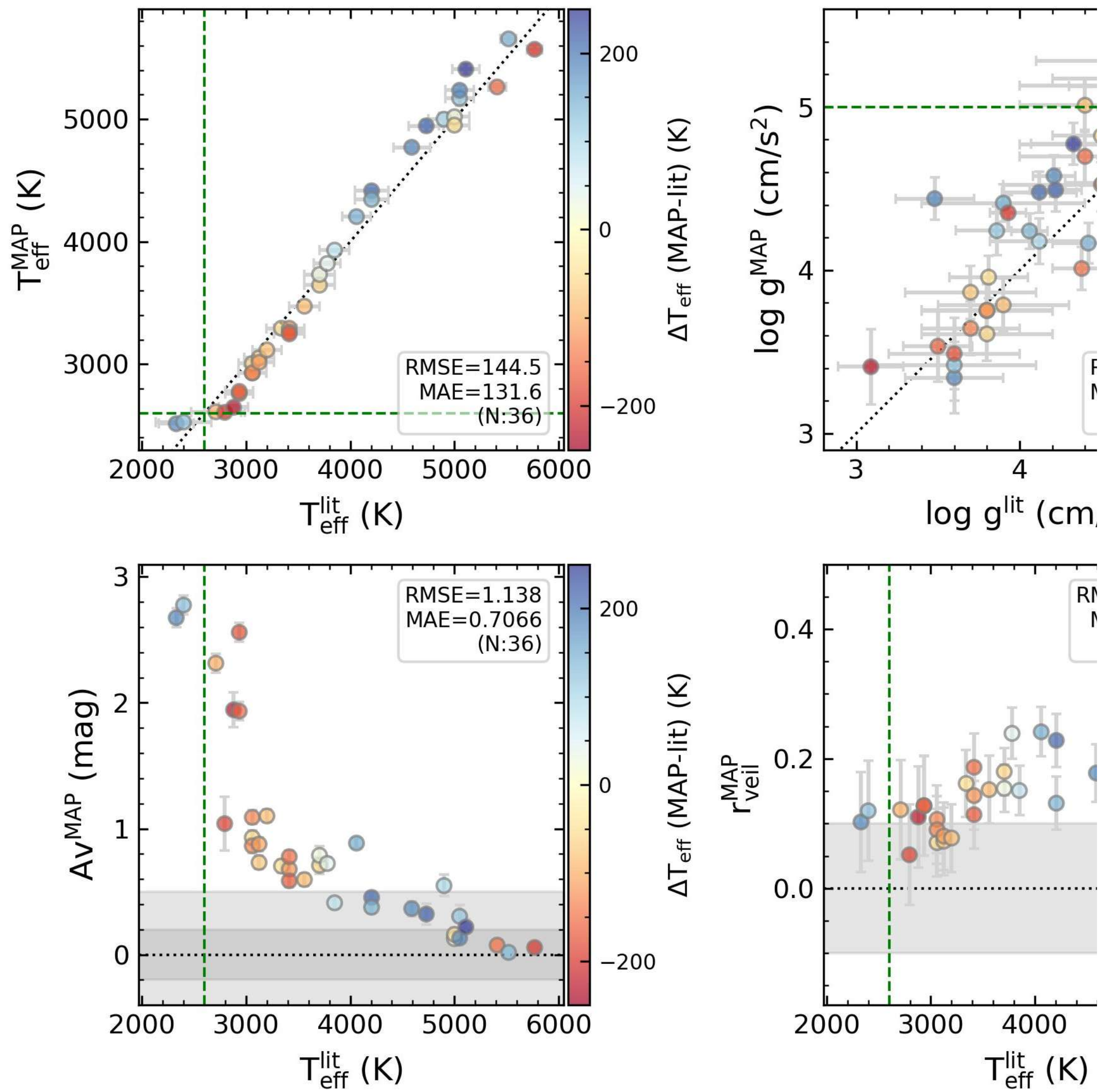
Peak va
(All te

T_{eff} (K)
log g (cm s
Av (mag)
r_{veil}

Generalization on real observations

36 class III template stars observed with VLT/X-shooter

- Wider wavelength range & higher spectral resolution combined UV+VIS arms and downgraded the spectral resolution
- Well-analyzed by Manra+2013, 2017 and Stelzer+2013 (SpT, log g)
- Zero extinction, zero veiling

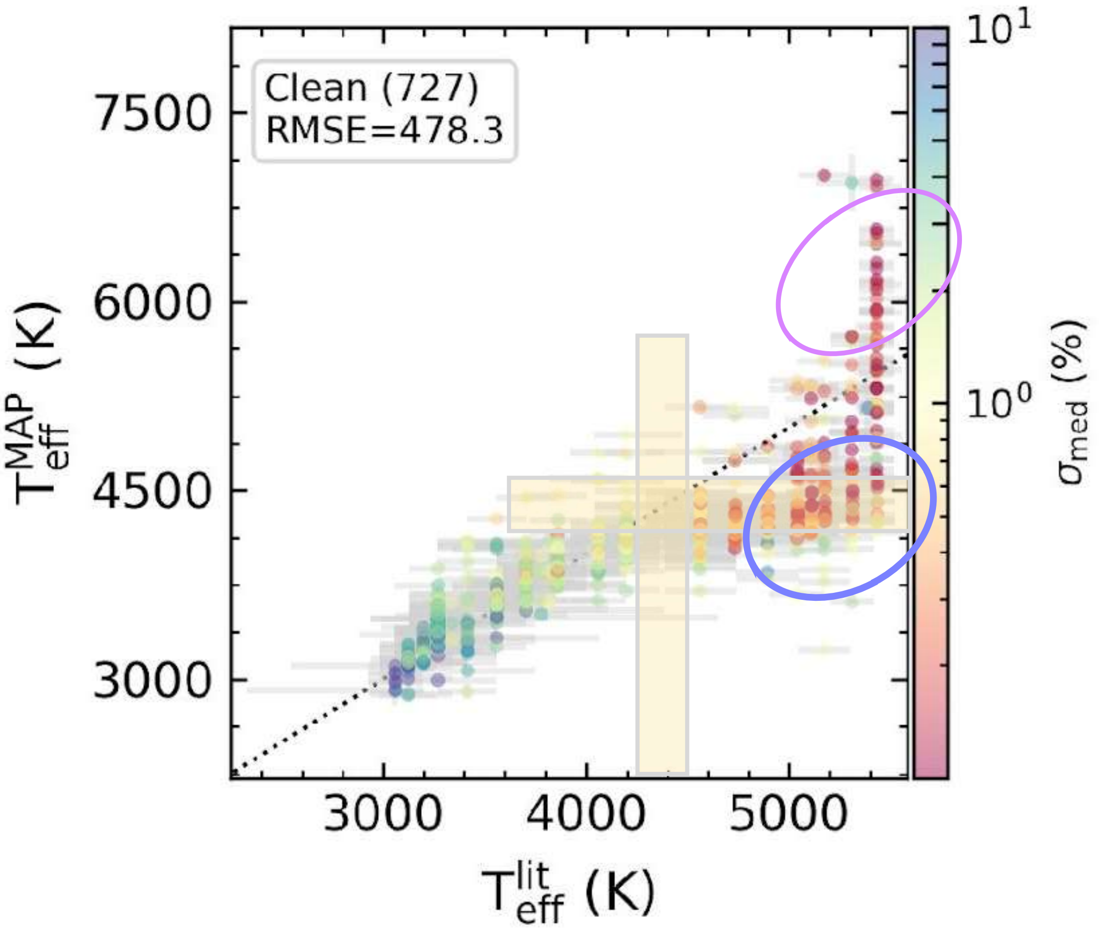


Average accuracy

- T_{eff} : 3.5%, 0.95-subclass
- log g: 0.3 dex
- Av: 0.2 – 1 mag
- r_{veil}: 0.13

cINN (MAP) vs. Template fitting (lit)

cINN



Literatur
e

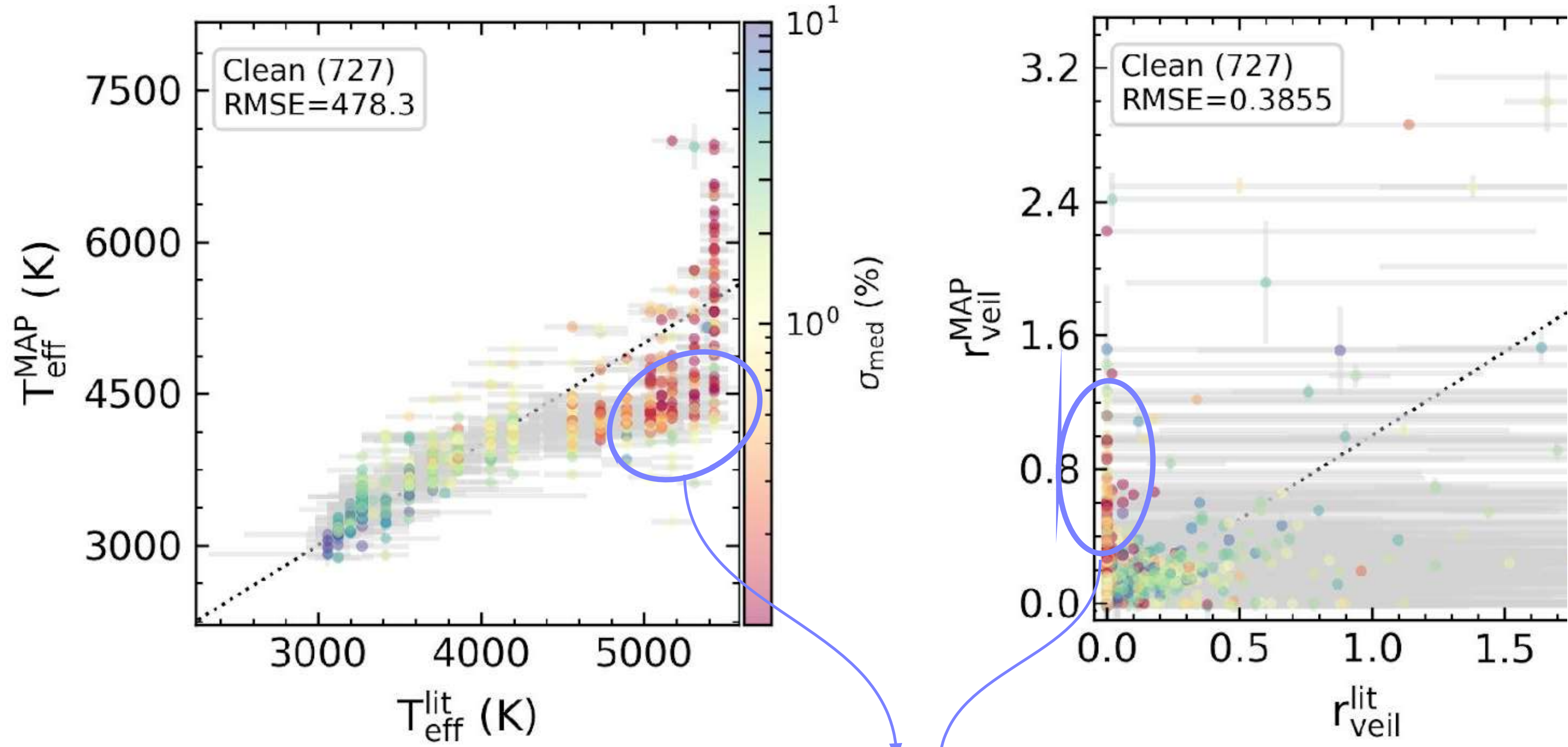
Templates used in literature:

- No templates higher than C
- No templates between K6
  (ΔT ~360K)

cINN training range: 2600K –

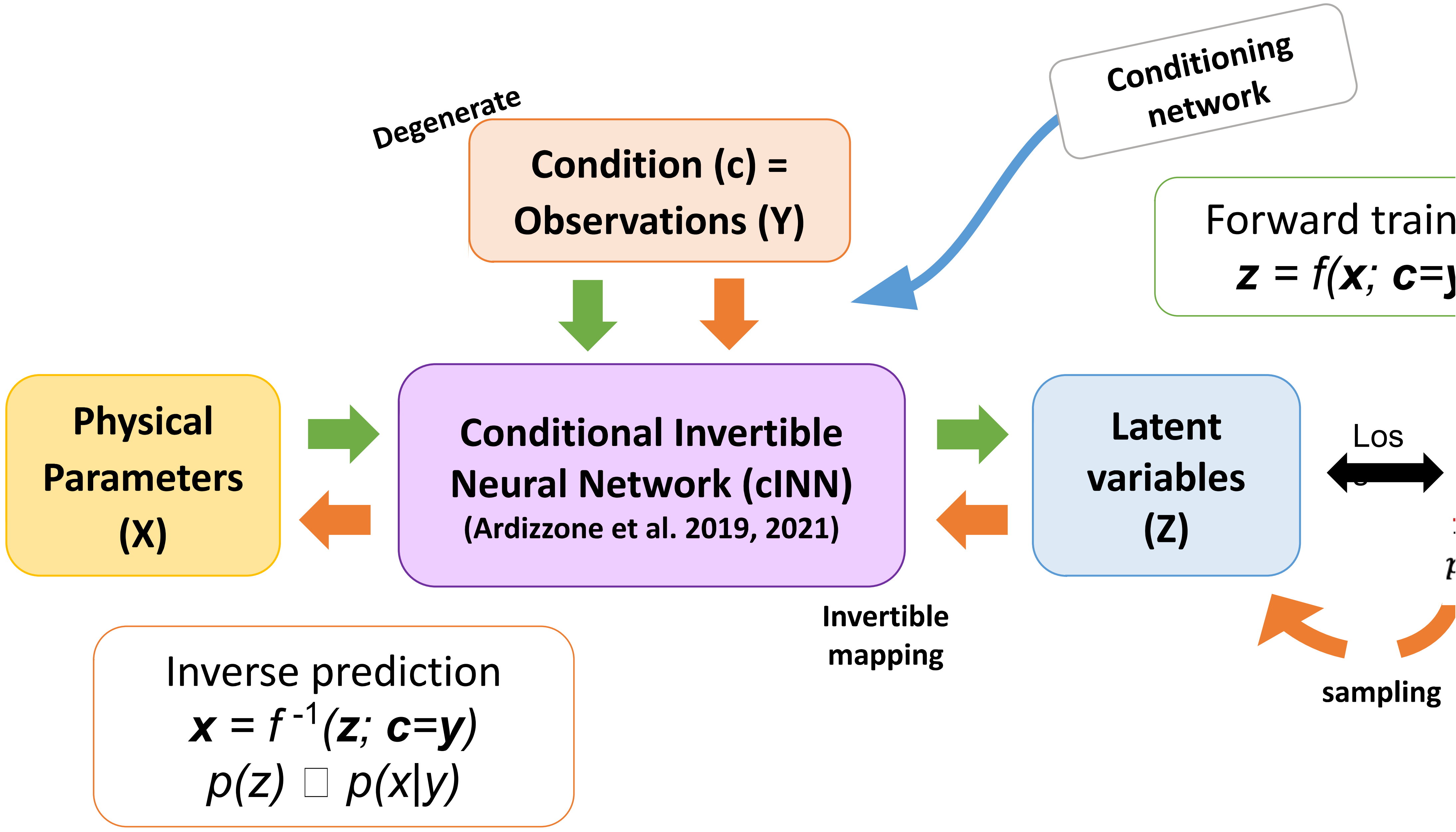
cINN (MAP) vs. Template fitting (lit)

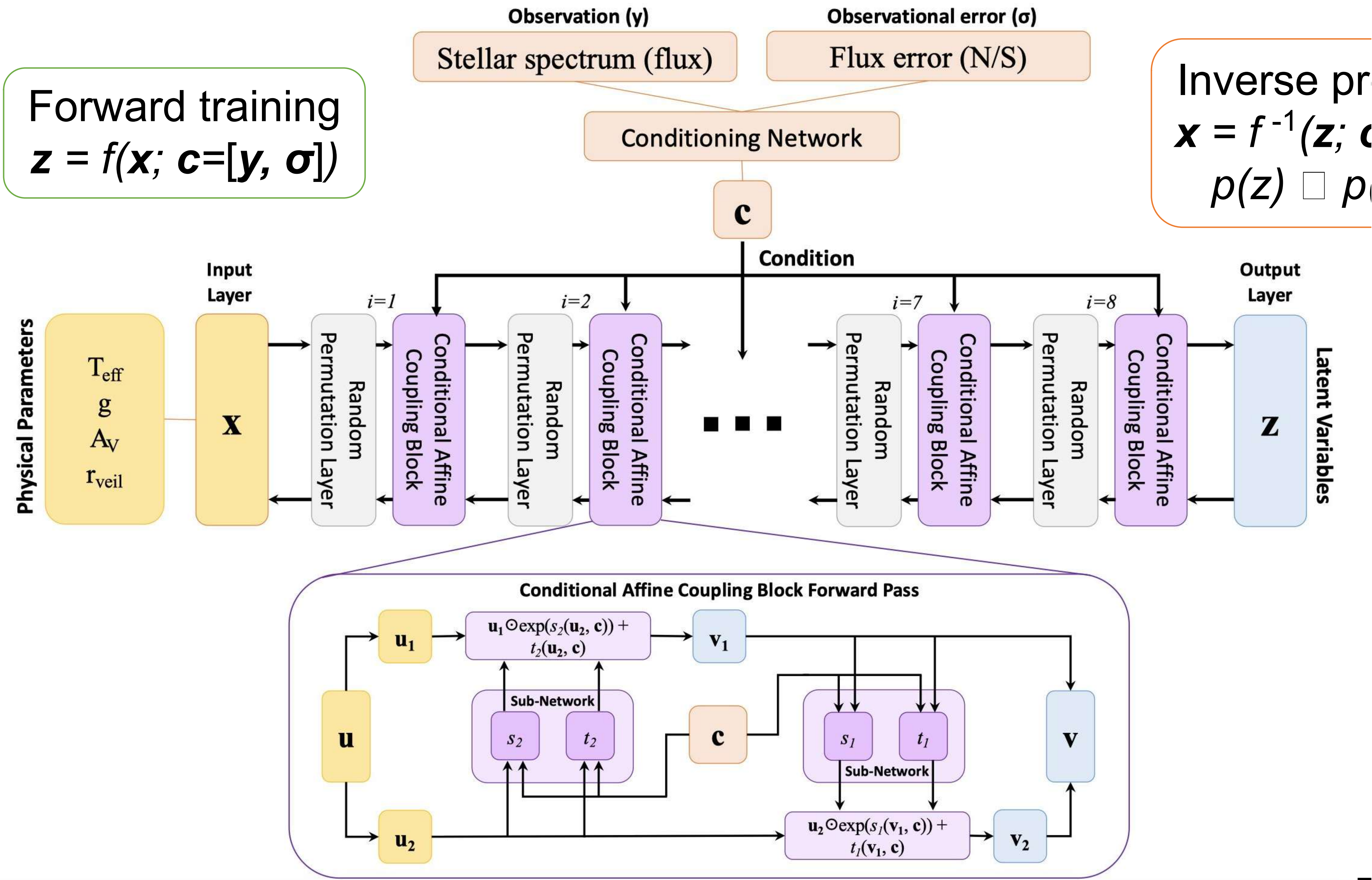
cINN



Literature

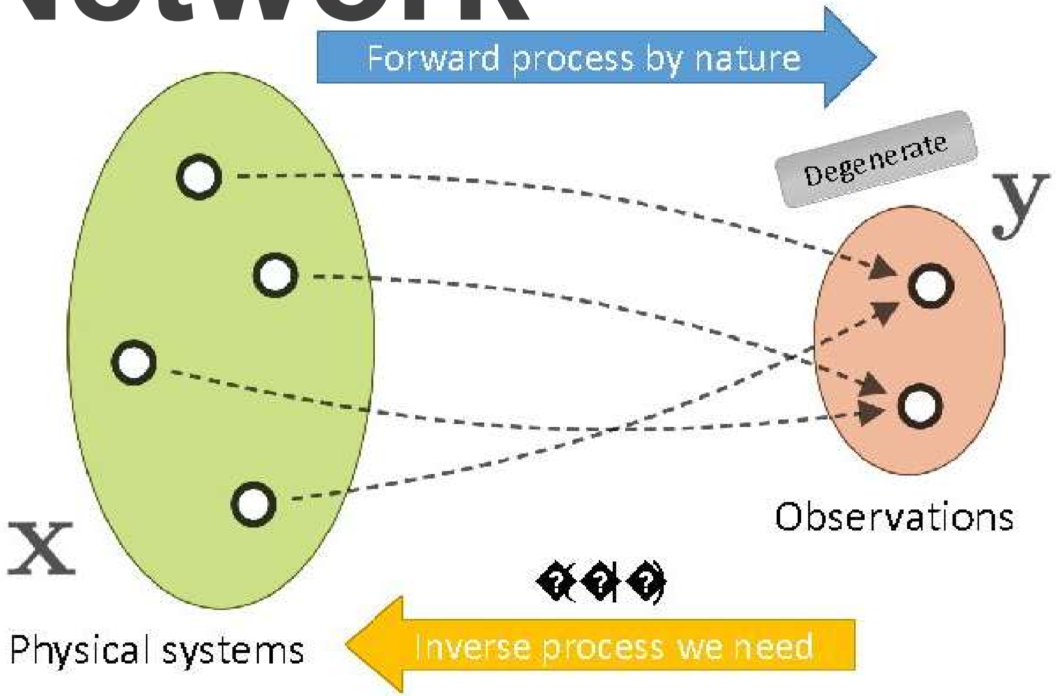
Degeneracy between T_{eff} and r_{veil} is not considered


Degenerate
Condition (c) =
Observations (Y)
Conditioning
network
Forward train
z = f(x; c=y
Physical
Parameters
(X)
Conditional Invertible
Neural Network (cINN)
(Ardizzone et al. 2019, 2021)
Latent
variables
(Z)
Los
Invertible
mapping
sampling
Inverse prediction
x = f -1(z; c=y)
p(z) □ p(x|y)

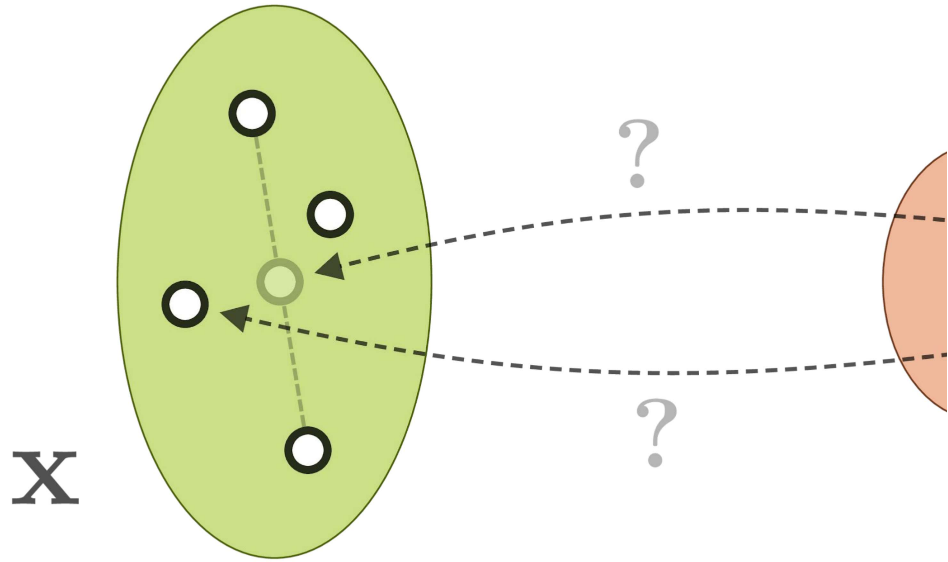

Forward training
$z = f(x; c=[y, \sigma])$
Inverse pr
$x = f^{-1}(z; c$
$p(z)$ □ p
Observation (y)
Stellar spectrum (flux)
Observational error (σ)
Flux error (N/S)
Conditioning Network
c
Condition
Physical Parameters
T_{eff}
g
A_V
r_{veil}
Input Layer
x
i=1
Random Permutation Layer
Conditional Affine Coupling Block
i=2
Random Permutation Layer
Conditional Affine Coupling Block
i=7
Random Permutation Layer
Conditional Affine Coupling Block
i=8
Random Permutation Layer
Conditional Affine Coupling Block
Output Layer
z
Latent Variables
Conditional Affine Coupling Block Forward Pass
u
u_1
u_2
$u_1 \odot \exp(s_2(u_2, c)) + t_2(u_2, c)$
Sub-Network
s_2
t_2
c
v_1
s_1
t_1
Sub-Network
$u_2 \odot \exp(s_1(v_1, c)) + t_1(v_1, c)$
v_2
v

Invertible Neural Network

(Ardizzone+2019a,b, 2021)



Information loss during the forward process makes observations degenerate

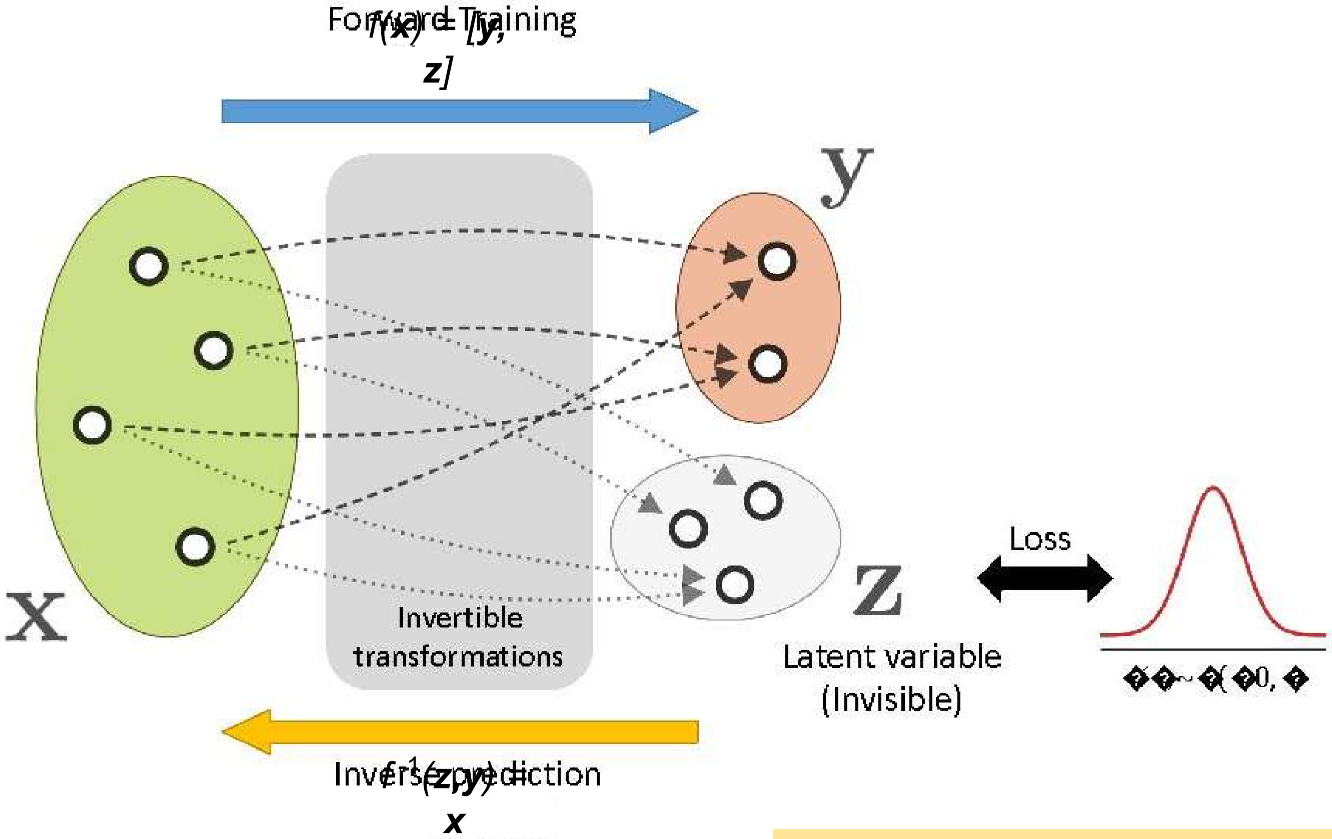


If we use a standard NN trained on the inv process, it will either pick only one of the e even worse, will form an average.

Images from Visual Learning Lab, Heidelberg University

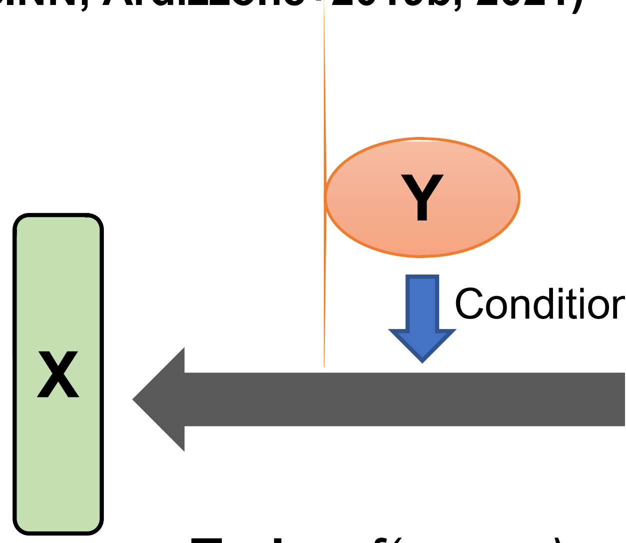
Invertible Neural Network & cINN

Invertible Neural Network (Ardizzone+2019a)



conditional Invertible Neur

(cINN; Ardizzone+2019b, 2021)



Train: $f(\boldsymbol{x}; \boldsymbol{c}=\boldsymbol{y}) =$

Usage: $\boldsymbol{x} = f^{-1}(\boldsymbol{z}; c$

INN and cINN give a full posterior distribution, p(x|y), without any additional calculations like MCMC.