Neural Networks for NLTE

Proof of concept

Gilles KLUTH[1]
K. D. Humbird, B. K. Spears, L. Peterson
H. A. Scott, M. V. Patel, J. Koning, M. Marinak,
L. Divol, C. Young[2].

1 CEA-DAM. VSP at LLNL.
2 Lawrence Livermore National Laboratory

Lawrence Livermore National Laboratory

LLNL-TR-769700

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Outline

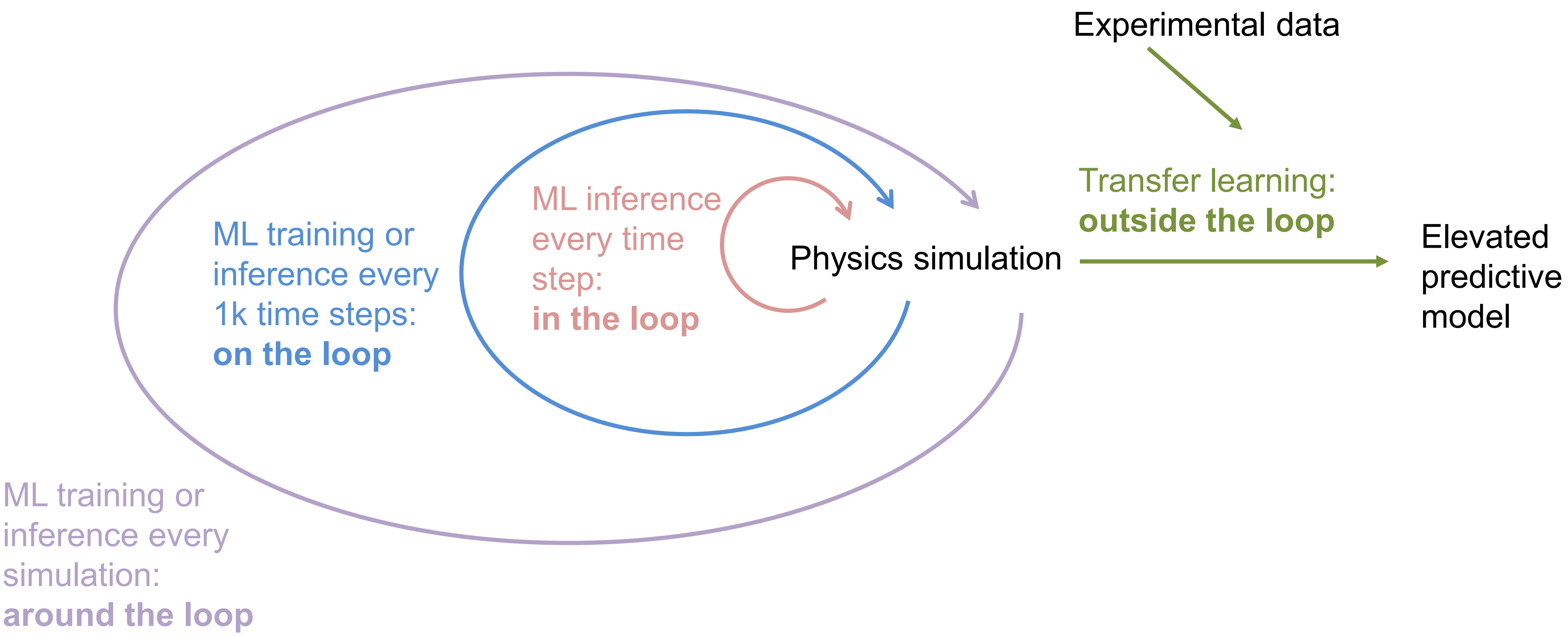
- We use deep neural networks to obtain NLTE absorption and emissivity spectra in radiation transport

  1) To accelerate ICF simulations,

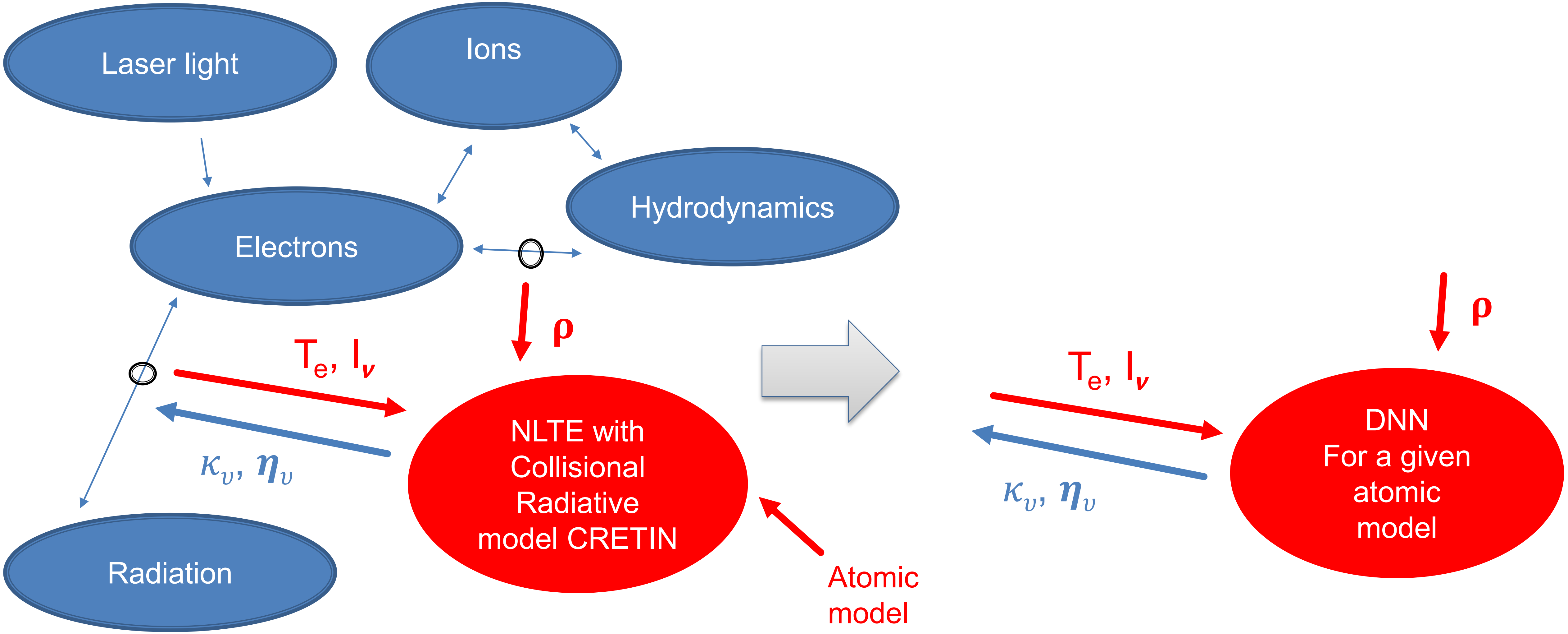
  2) To allow the use of **a new physics and a more accurate numeric**, too expensive now.

- We show the feasibility of the first point on a ICF representative test-case.

Deep learning & simulations at LLNL:
The global picture
Wrap simulation in multiple layers of Machine Learning



Deep neural network in-the-loop:
Physics-informed DNN surrogate model in HYDRA



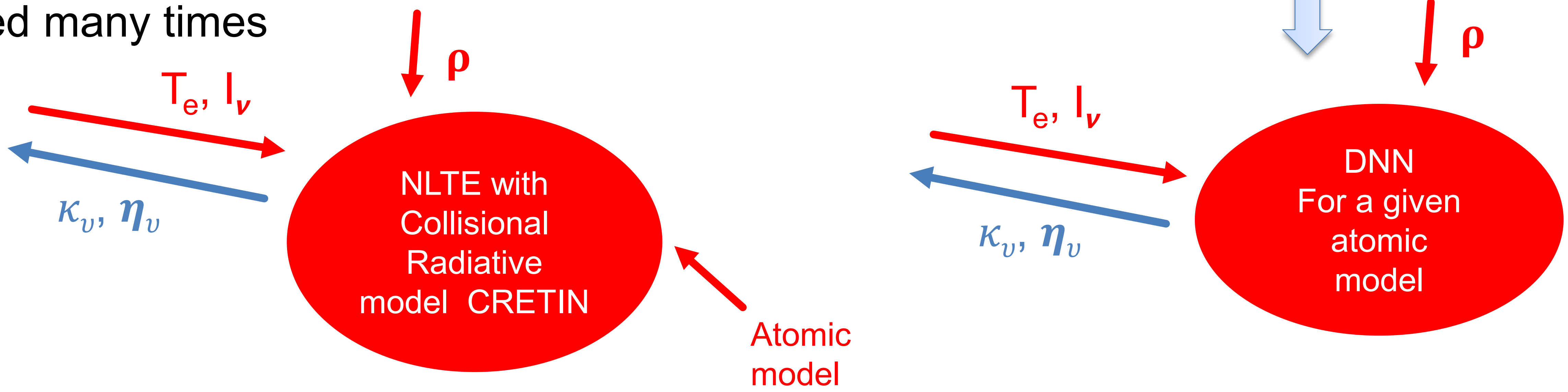
Deep neural network in-the-loop:
Physics-informed DNN surrogate model in HYDRA

In-line CRETIN
- Expensive
- Called many times

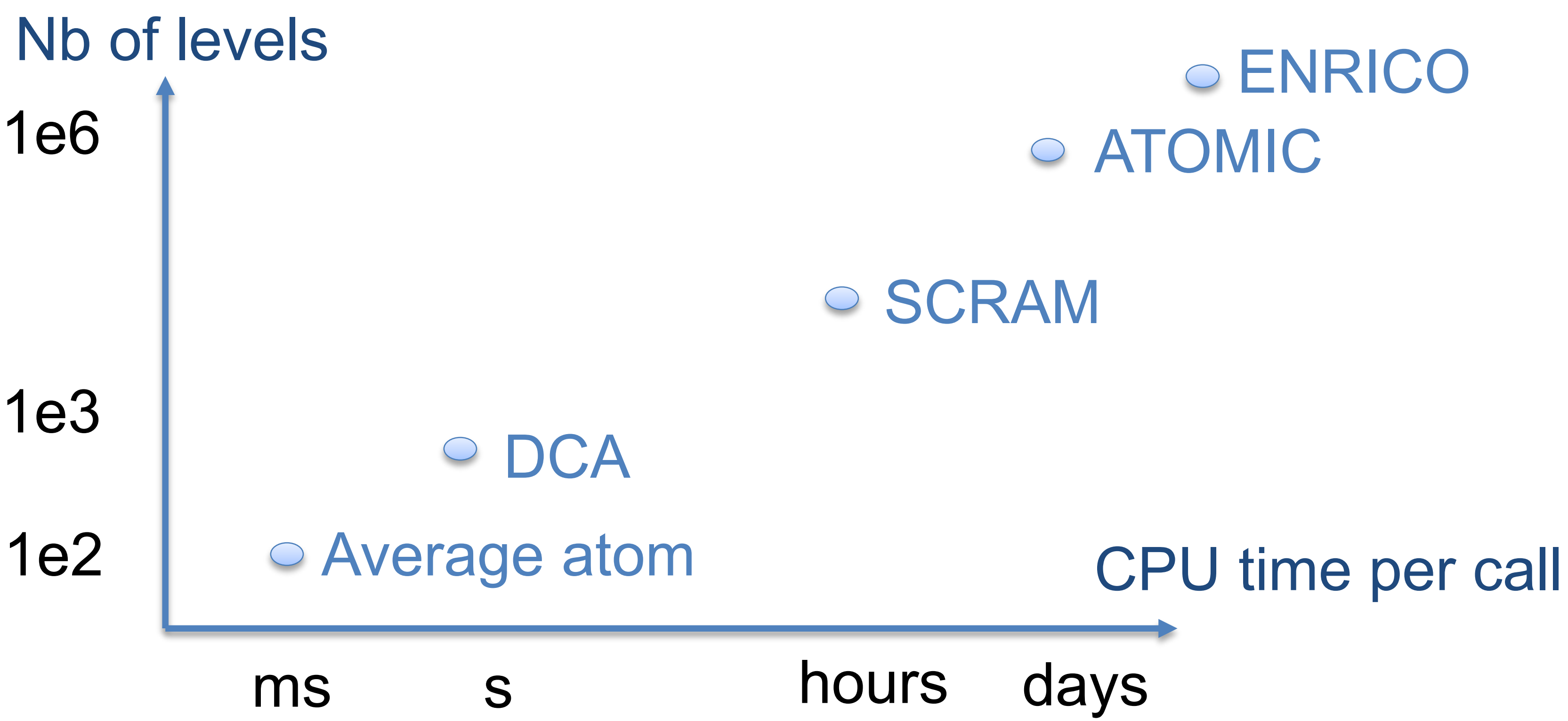
Create a dataset with CRETIN
- Apart from HYDRA
- Expensive one time

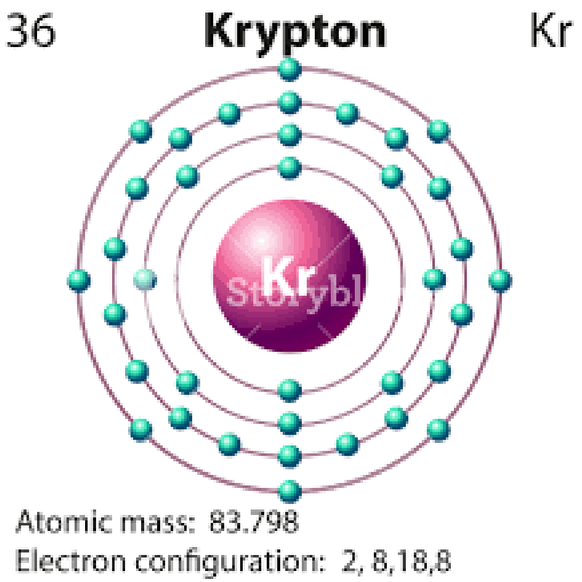
Train a DNN
- Apart from HYDRA
- Expensive one time

In-line DNN
- Fast



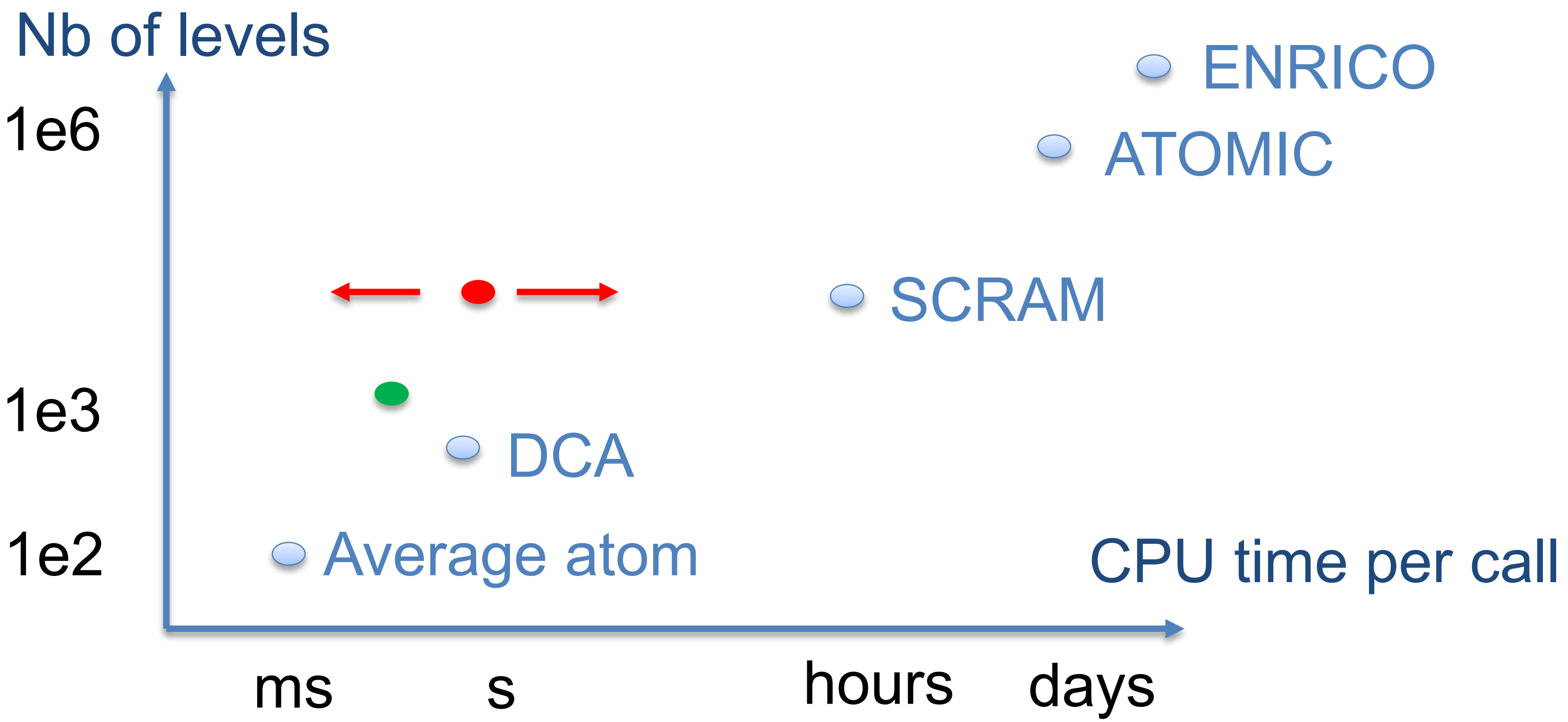
The CPU conundrum of NLTE codes





NLTE can be tens of percent of computational time in hohlraum simulations.

Replaced with fast deep neural networks



In this talk:

1808 levels
→ Speed-up

261 229 levels
On-going work
→ New physics in ICF simulations

What is the steady-state collisional-radiative model

- For a given atomic structure:

Levels, transitions with associated atomic cross sections.

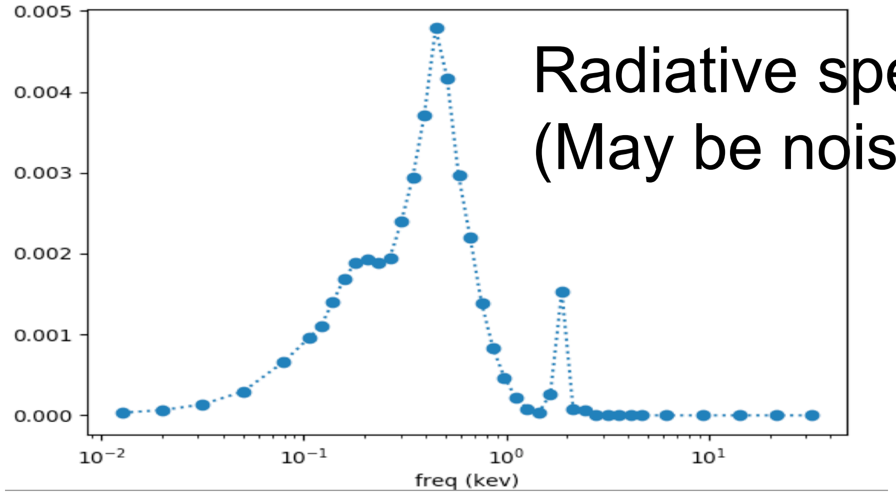
- For a given plasma and radiation:

Mass density, temperature, radiative field.

- For a given frequency binning.

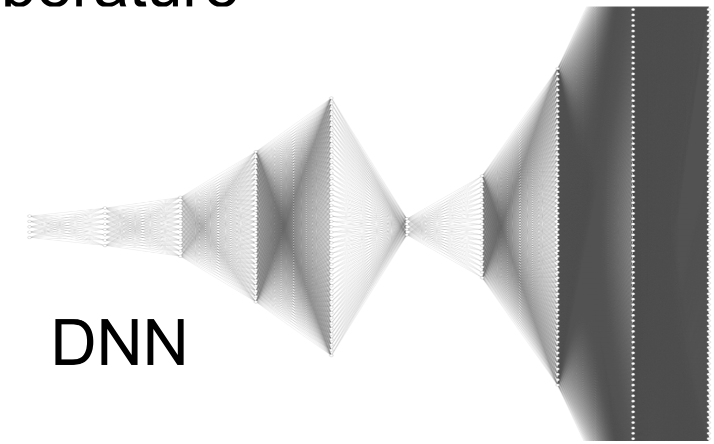
- Calculates the rates between levels,
- Calculates densities of each ion populations (Linear system whose size is the numbers of levels).
- Calculates ionization, absorption and emissivity spectra.

To replace the CR model, we solve a regression problem in high dimensions ~100.



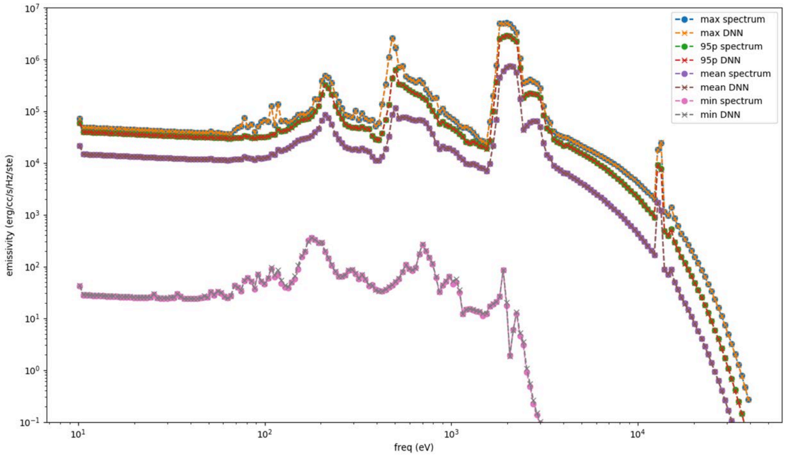
Radiative spectrum
(May be noisy with IMC)

+ Mass Density
+ Electronic temperature

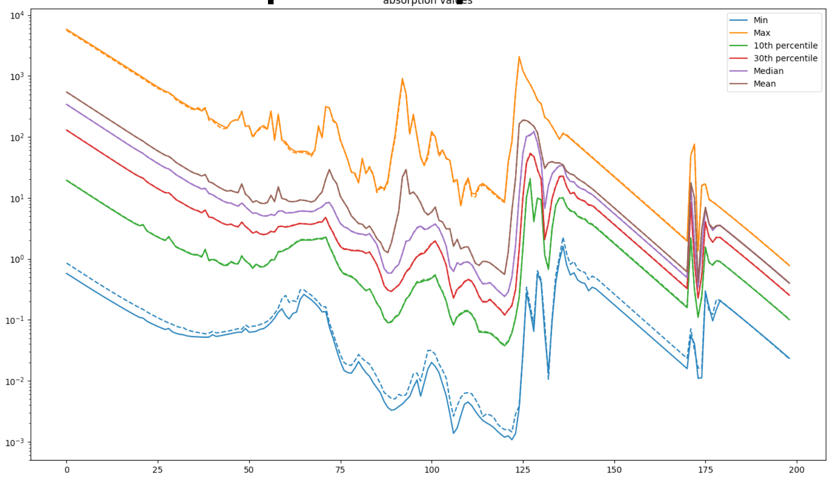


DNN

Emissivity spectrum

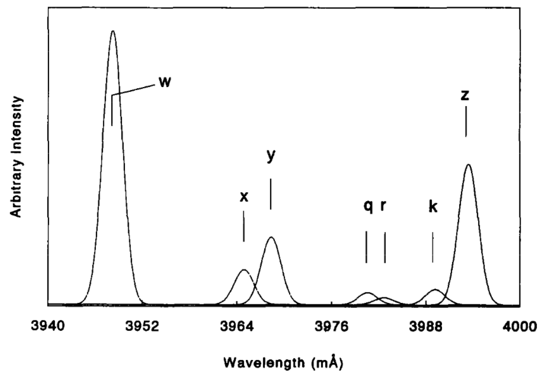


Absorption spectrum



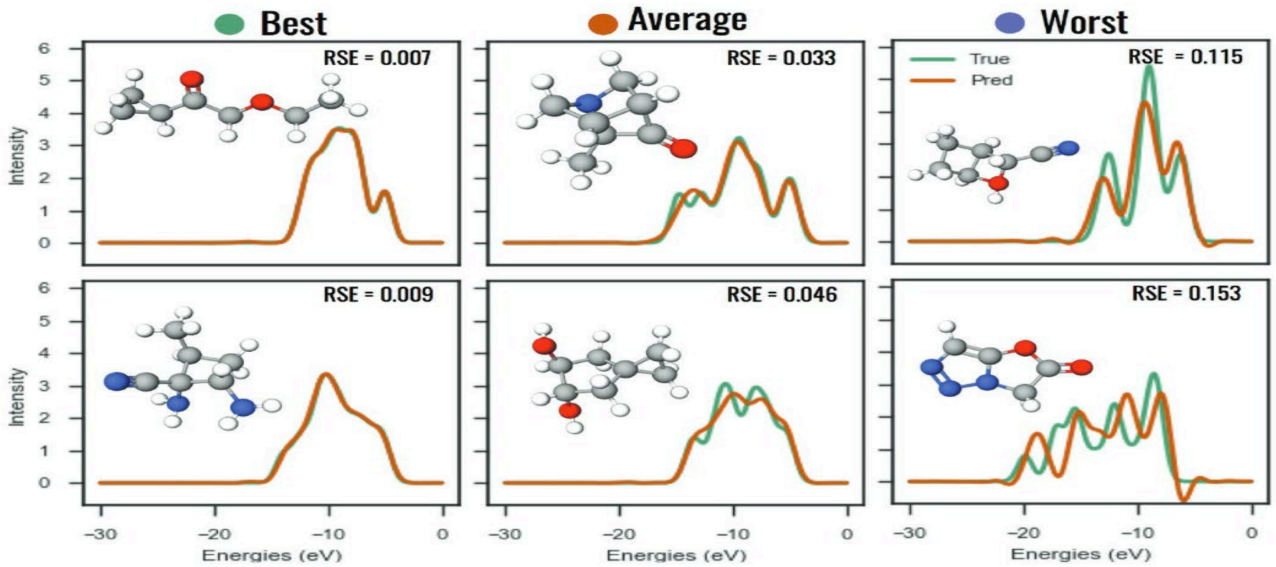
Neural networks in spectroscopy.

- MOSTLY CLASSIFICATION (type of astronomic objects, type of material)
- OR SCALAR REGRESSION



9 lines intensity → Ne, Te

- MOLECULAR SPECTRA begins to be created on a very narrow frequency.



Here, plasma conditions and embedded radiation create spectra on a large frequency range.

We will focus on 2 problems

Problem 1

Encapsulation of CRETIN alone, using an analytical radiative field.

To study the accuracy we may obtain on dataset.

We have here infinite data, as regards the input dimensional space (=4).

Problem 2

In-lining a DNN in a HYDRA ICF test-case.

To study the speed-up we may achieve, and relies the accuracy of the training to the final accuracy in HYDRA.

We have less data, given by close HYDRA test-cases, on a bigger input dimensional space (=42).

Problem 1: Cretin data with analytic radiative field

Inputs D=4 Outputs D=400

- 200 bins: log-spaced between 10eV and 40kev
- Krypton: Z=36, 1808 levels, 98 531 transitions
- Radiative field given by Tr and $\boldsymbol{\alpha}$.

$$I(\nu) = aT_r^4 \, [\, (1- \boldsymbol{\alpha}) \, b(\nu, Tr) + \boldsymbol{\alpha} \, g(\nu) \,]$$

With $b(\nu, Tr)$ the reduced Planckian

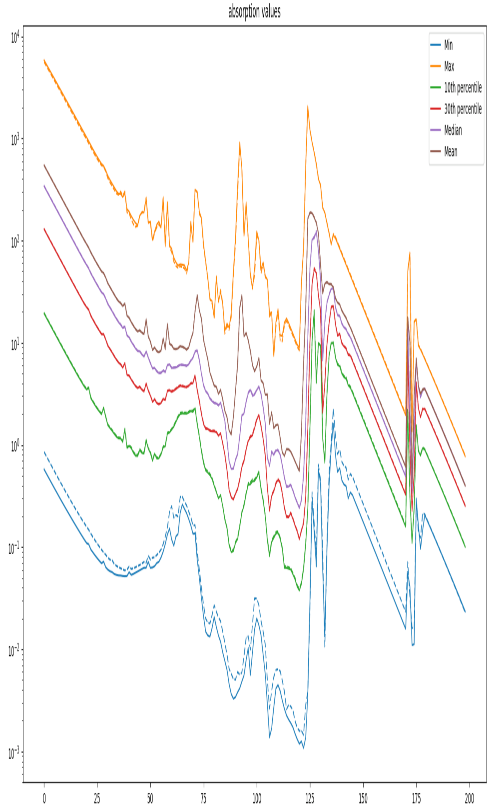
$g(\nu)$ the reduced gaussian of mean 3 keV and FWHM 1 keV

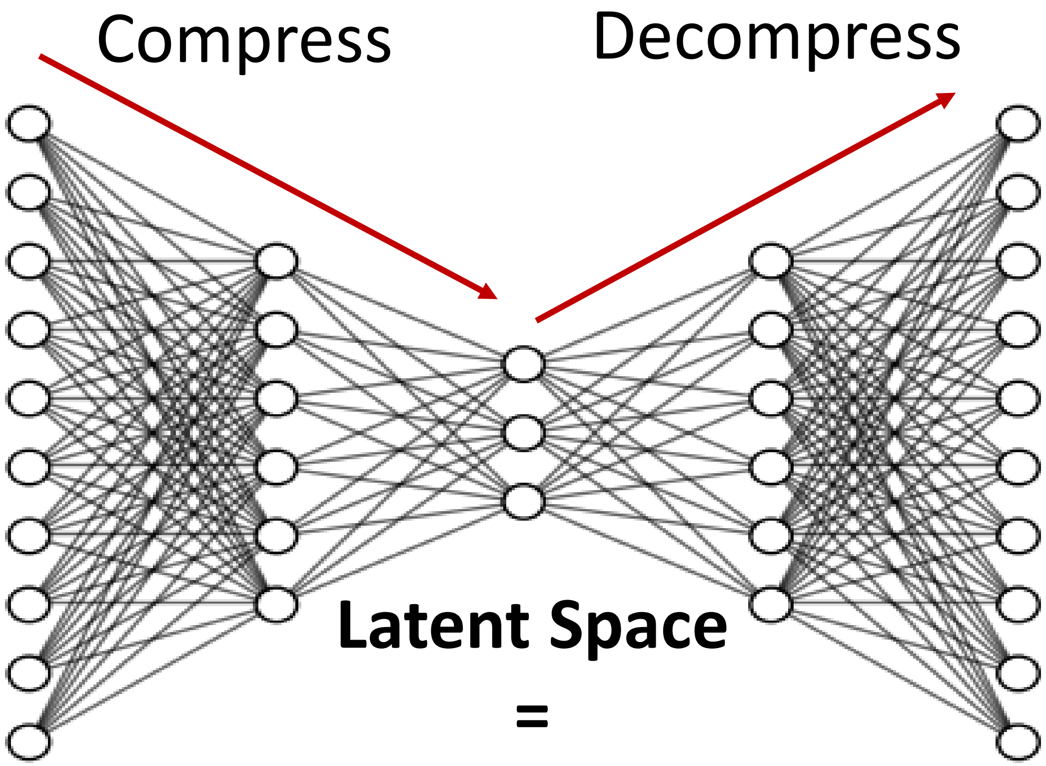
- Training dataset: 30K samples (= CRETIN simulations), uniform.

$3 < \boldsymbol{\rho} < 100$mg/cc , $300 < Te < 3000$eV , $30 < Tr < 300$eV , $0. < \boldsymbol{\alpha} < 0.3$

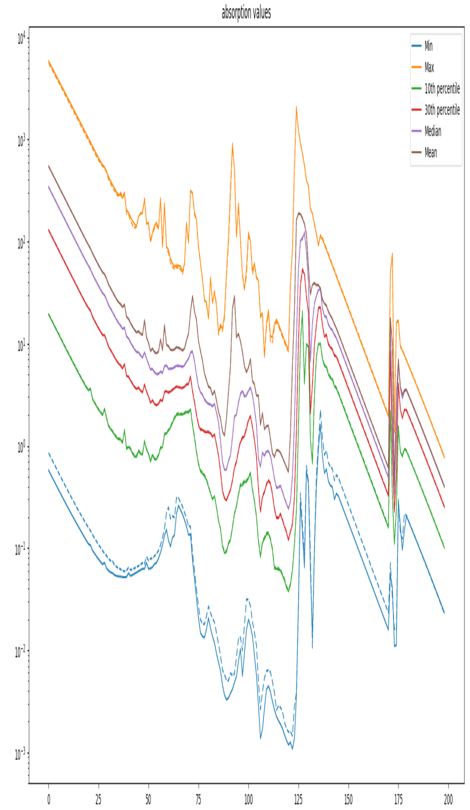
Auto-encoders enable us to reduce spectra dimensions.
Here from 200 bins to D neurons between 3 and 7.

200-bins spectra





200-bins spectra



Latent Space

=

F_i(w_i , 200-bins spectra)

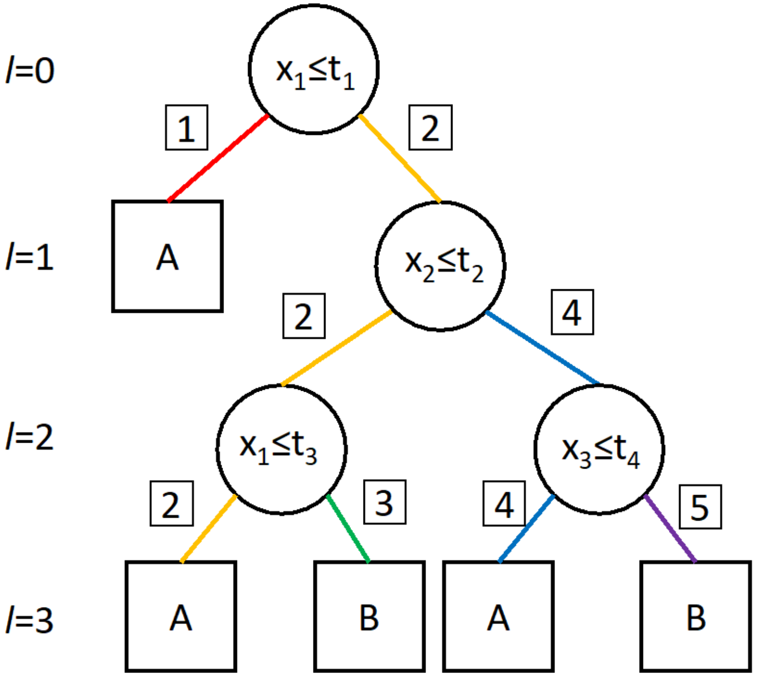
w_i fixed after training,

$1<i<D$.

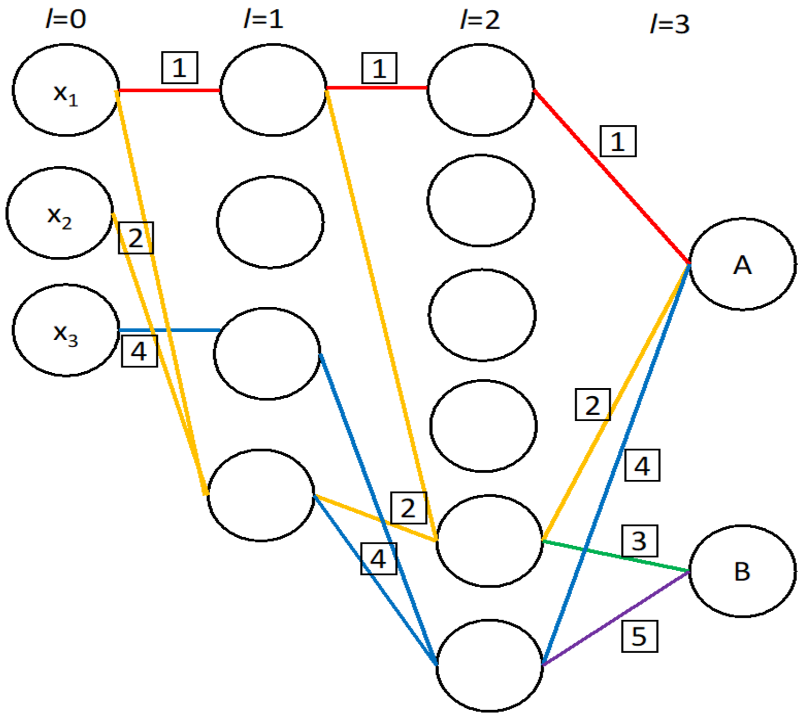
Then DJINN connects the inputs to the latent space.

DJINN maps decision trees to initialized deep feed-forward neural networks

Decision tree



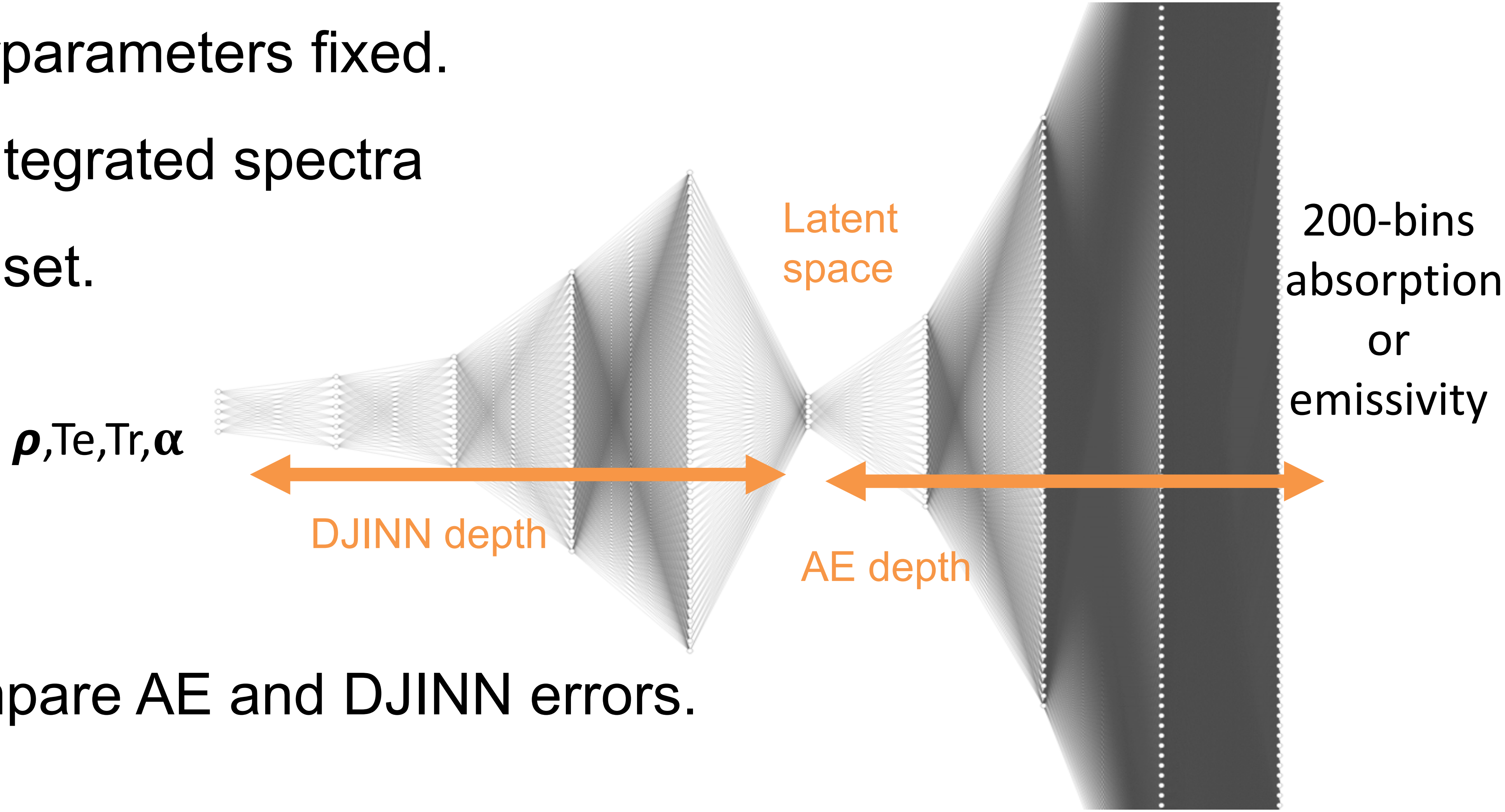
Initialized neural network



We optimize architecture with ~3 parameters

- Many hyperparameters fixed.
- Maximize integrated spectra

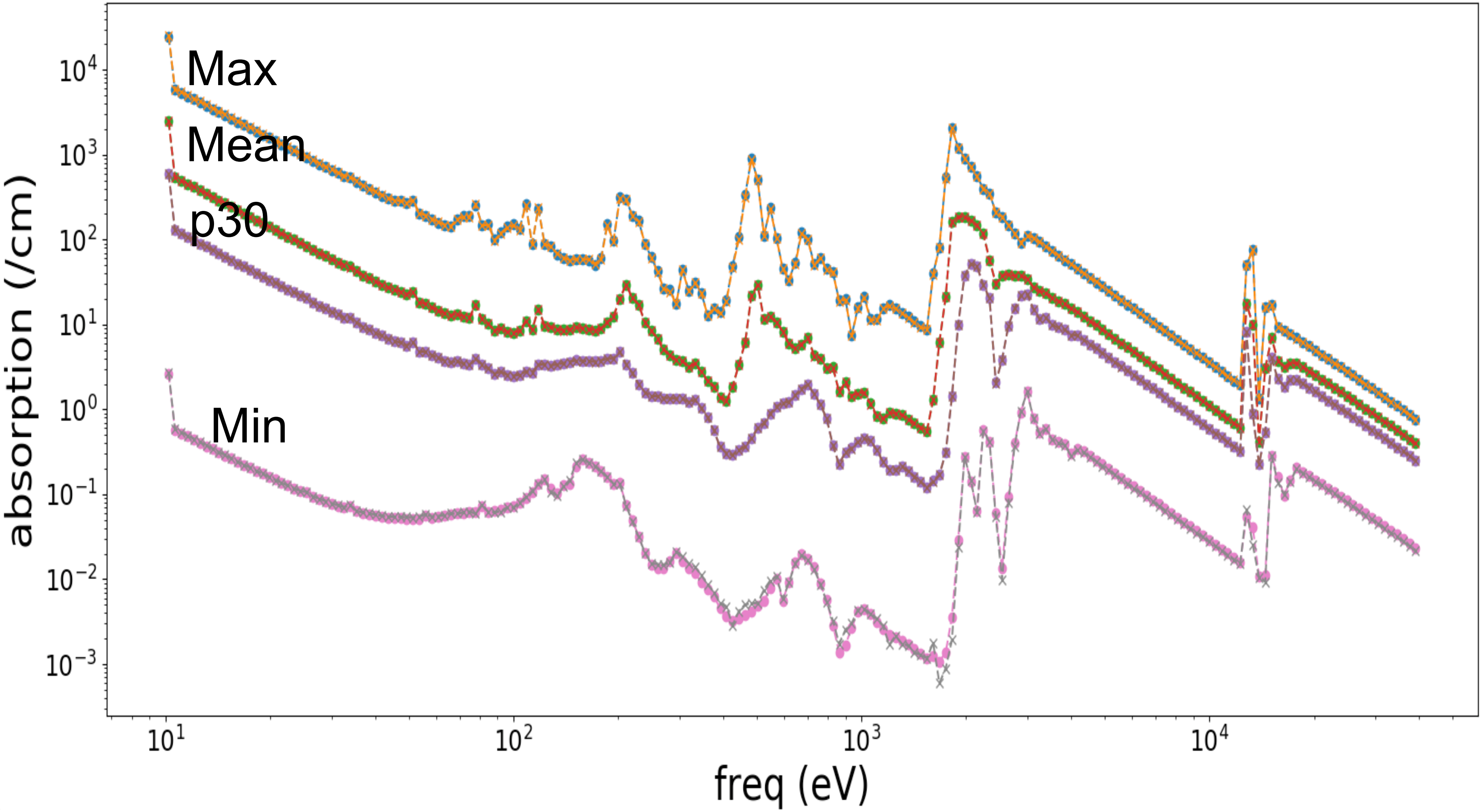
on a test dataset.



- Always compare AE and DJINN errors.

Results DNN predicts absorption spectra with accuracy.

O Cretin X DNN, over 30k test dataset.



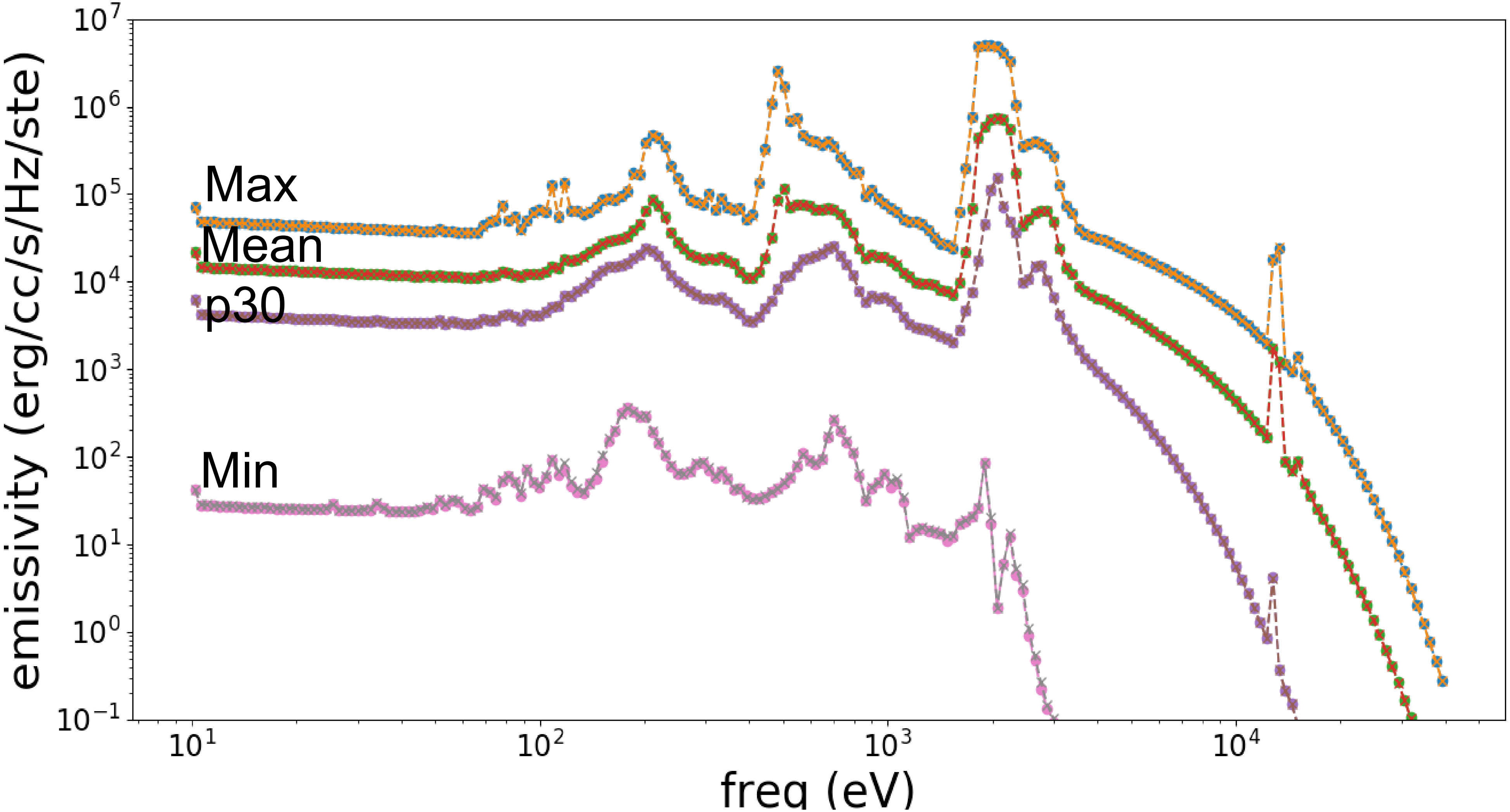
<u>Relative errors</u>

PLANCK
Mean 0.16%
Max 6.06%

ROSSELAND
Mean 0.19%
Max 8.77%

DNN predicts emissivity spectra with accuracy.

O Cretin X DNN, over 30k test dataset.



<u>Relative errors</u>

INTEGRAL
Filtered over percentile 30.

Mean 0.24%
Max 3.89%

Problem 2: CRETIN in-lined in HYDRA

Inputs D=42 Outputs D=80

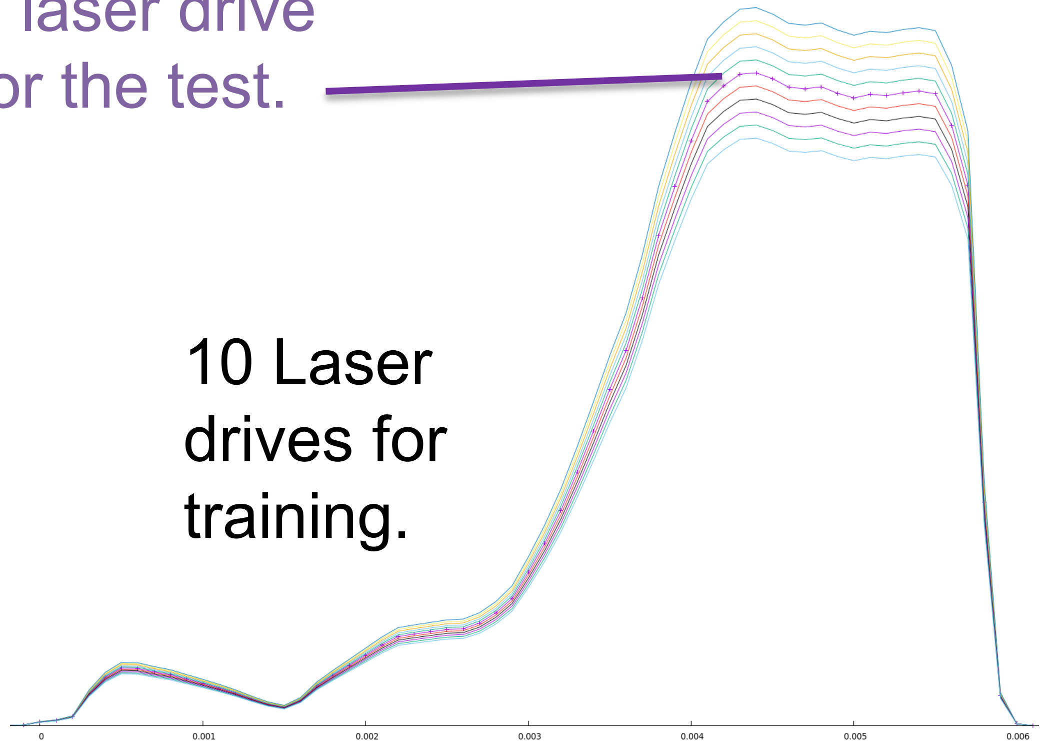
- Spherical Kr hohlraum with internal laser source, He gas and a capsule (DT cryo , Be/Cu).
- Te-Ti-Tr multigroup diffusion solver
- Conduction with flux limiter of 15%.

- 40 bins: uniform adapted to L, Kα, Kβ between 10eV and 40kev
- Krypton: Z=36, 1808 levels, 98 531 transitions

Problem2

Datasets for training and test.

1 laser drive
for the test.



10 Laser
drives for
training.

11 HYDRA
simulations.
Dumps every 500ps.
63 Krypton cells

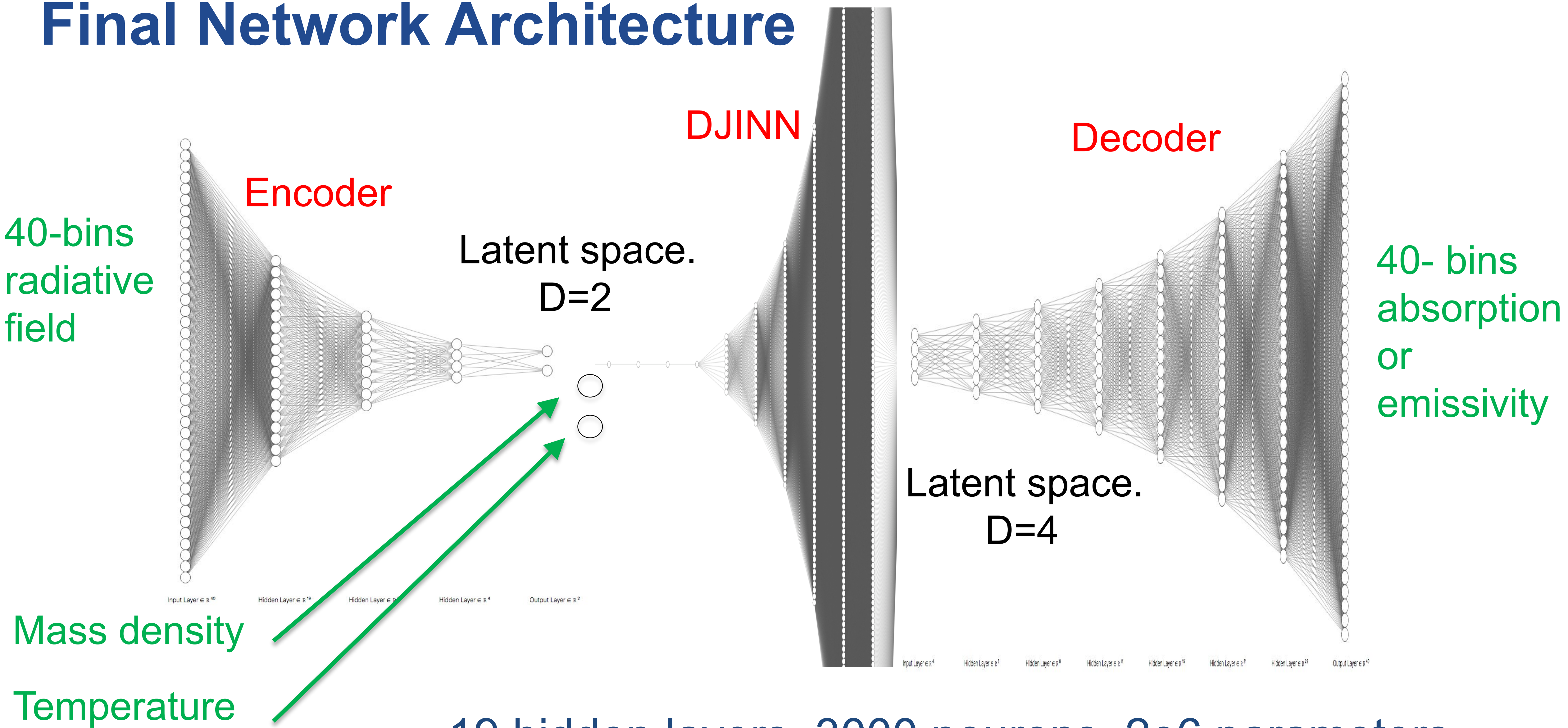
Radiative fields
dataset.
~ 10K

Density
Te.

66K CRETIN

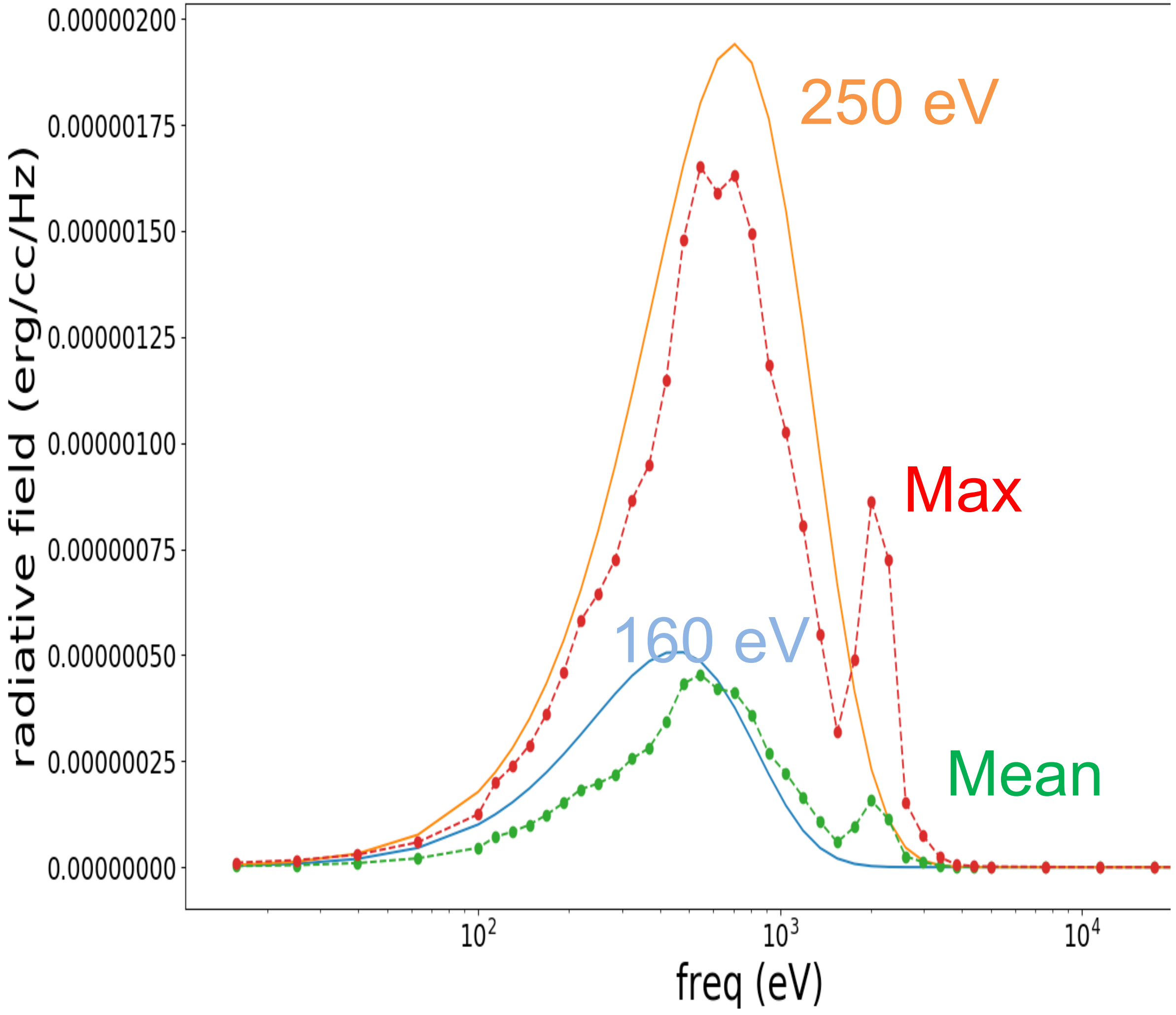
Absorption.
Emissivity.
66K

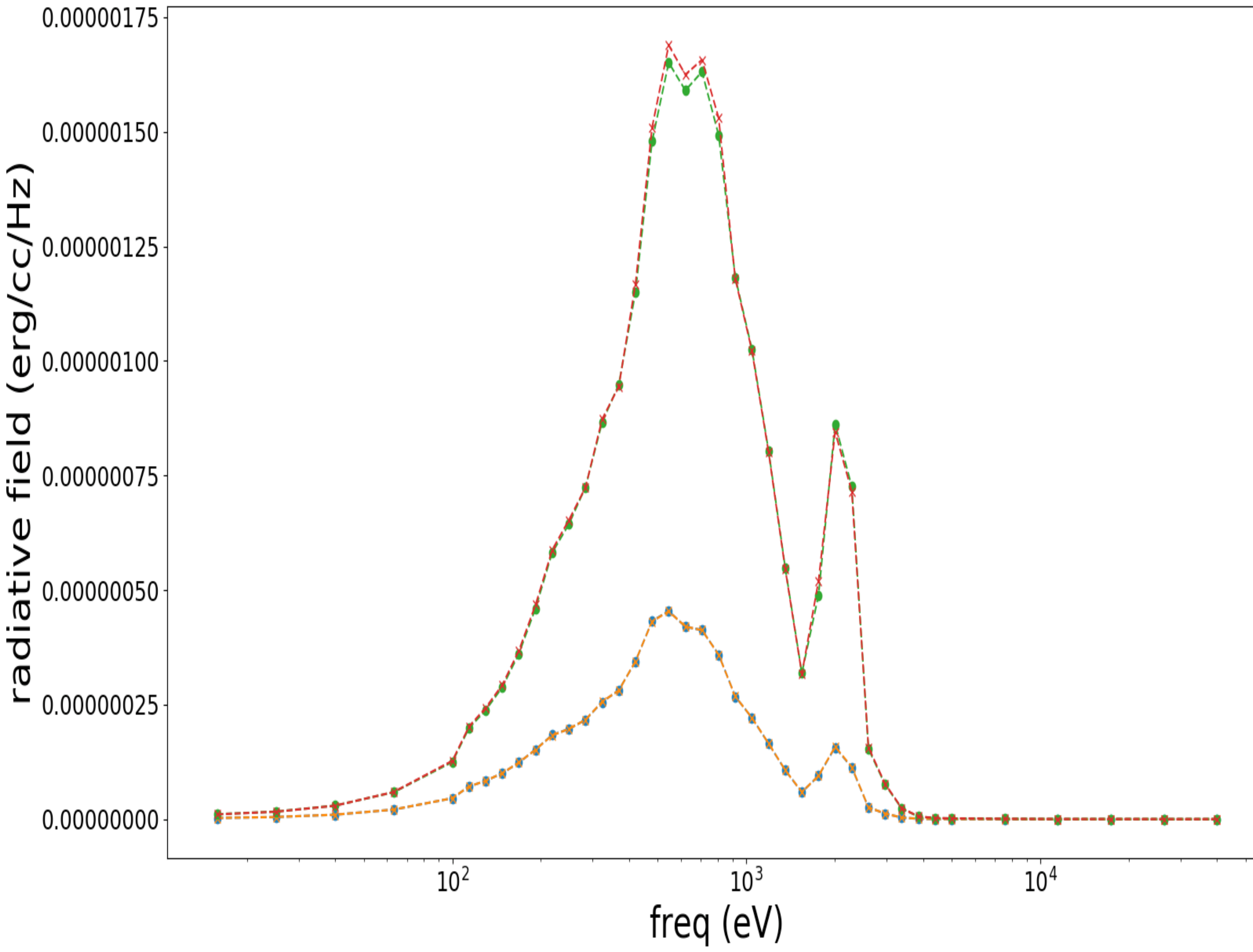
Final Network Architecture



19 hidden layers, 3000 neurons, 2e6 parameters.

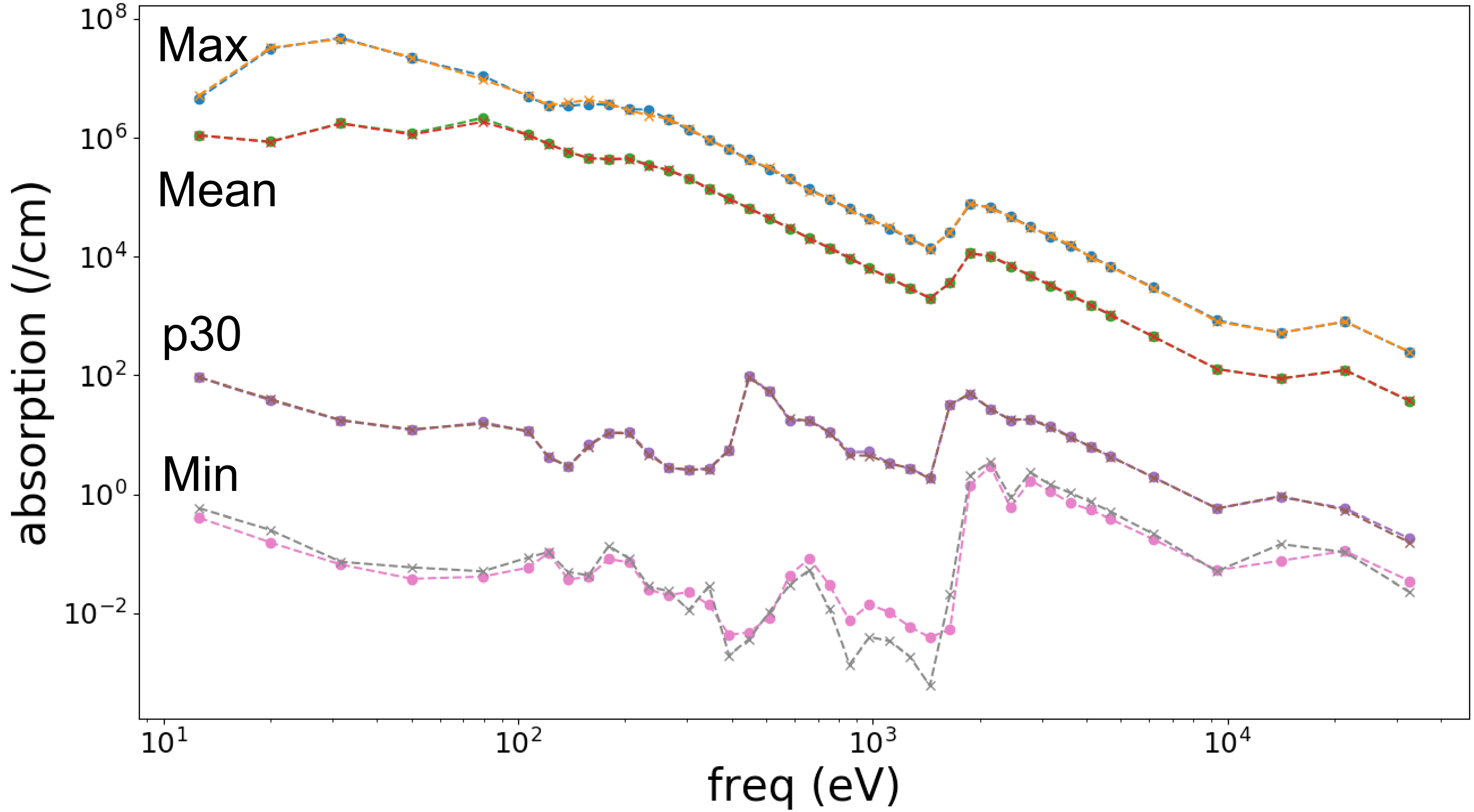
Auto-encoders manages to encode the radiative field on 2 dimensions latent space.





DNN match less with hydra rad fields.

O Cretin X DNN, over 832 test dataset.



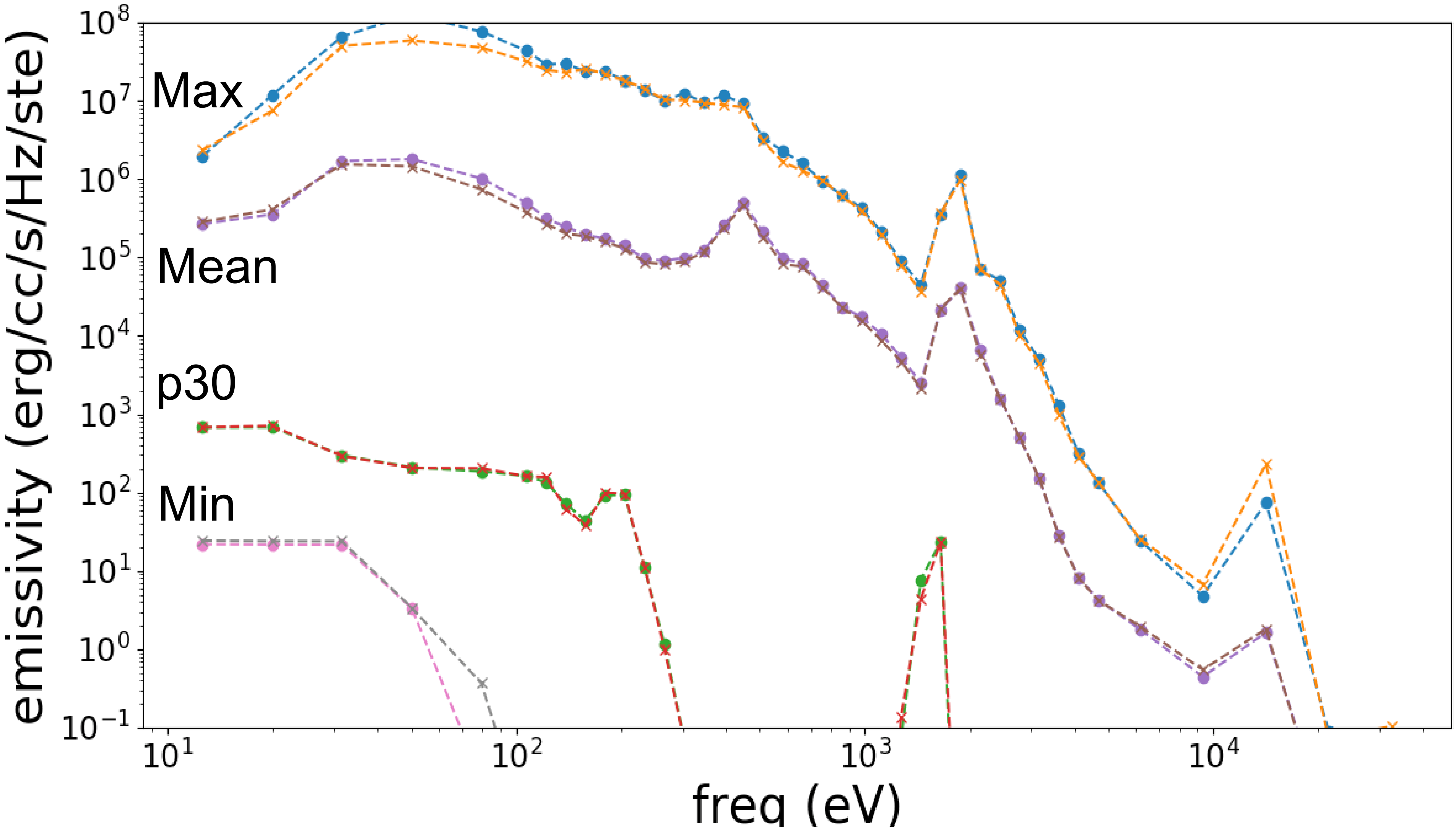
Relative errors

PLANCK
Mean 6.9%
Max 33%

ROSSELAND
Mean 9.9%
Max 54%

DNN match less with hydra rad fields.

O Cretin X DNN, over 832 test dataset.



<u>Relative errors</u>

FREQ-integrated
Filtered over
percentile 30.

Mean 8.1%
Max 43%

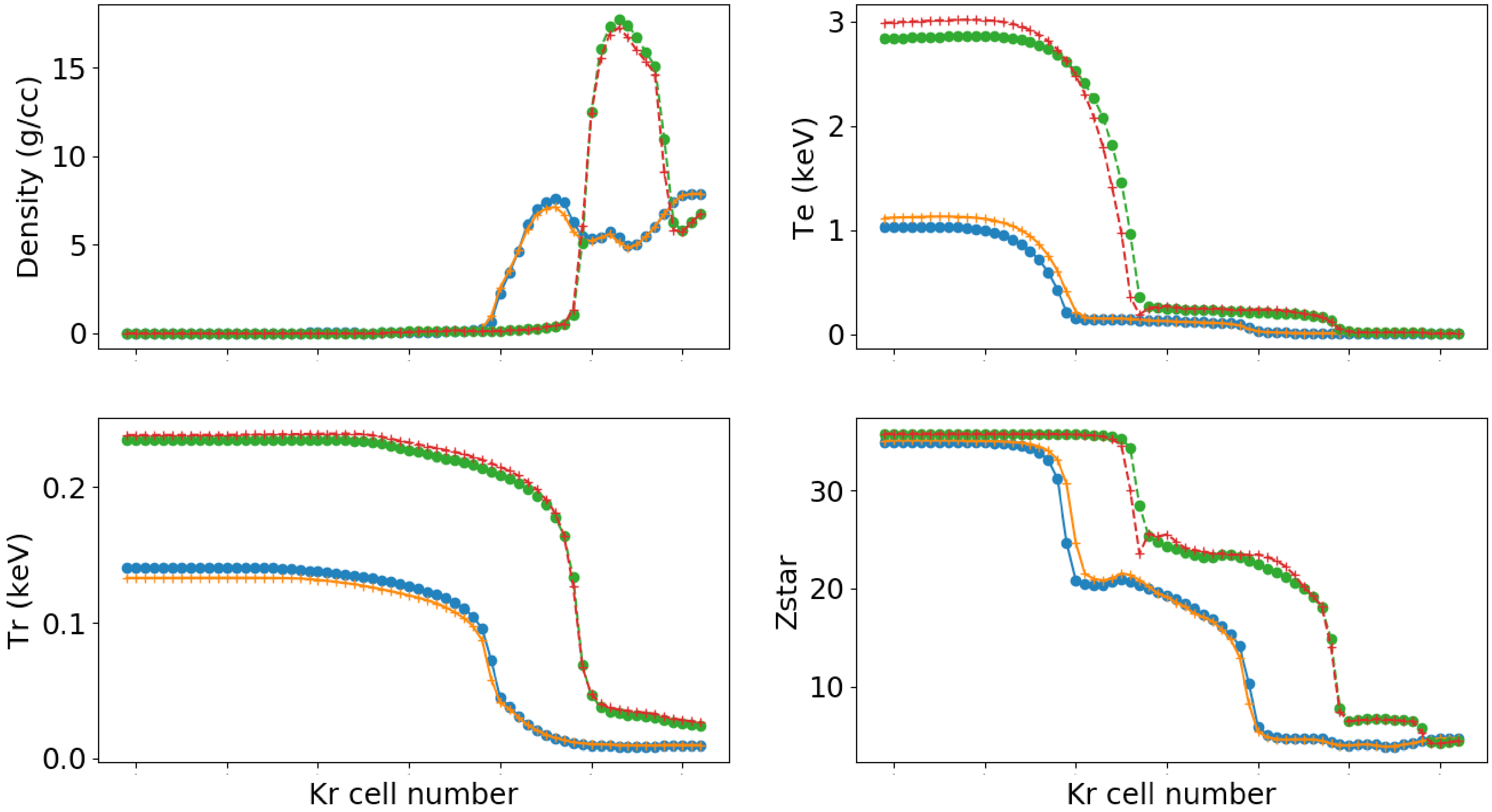
Summary of results on both problems

	PROBLEM 1 mean	max	PROBLEM 2 mean	max
Absorption Planck	0.16%	6.06 %	6.9 %	33%
Absorption Rosseland	0.19 %	8.77 %	9.9 %	54 %
Emissivity	0.24 %	3.89 %	8.1 %	43%

Analytical rad. field (Tr , α) $\not\approx$ Real rad. field (40 independents bins)
30K training dataset $\not\approx$ 10K rad. field dataset
Smaller range in ρ , Te $\not\approx$ Broader range in ρ , Te

Does it matters in the HYDRA test-case?

Hydra comparisons on the problem 2.
DNN and CRETIN results are similar.



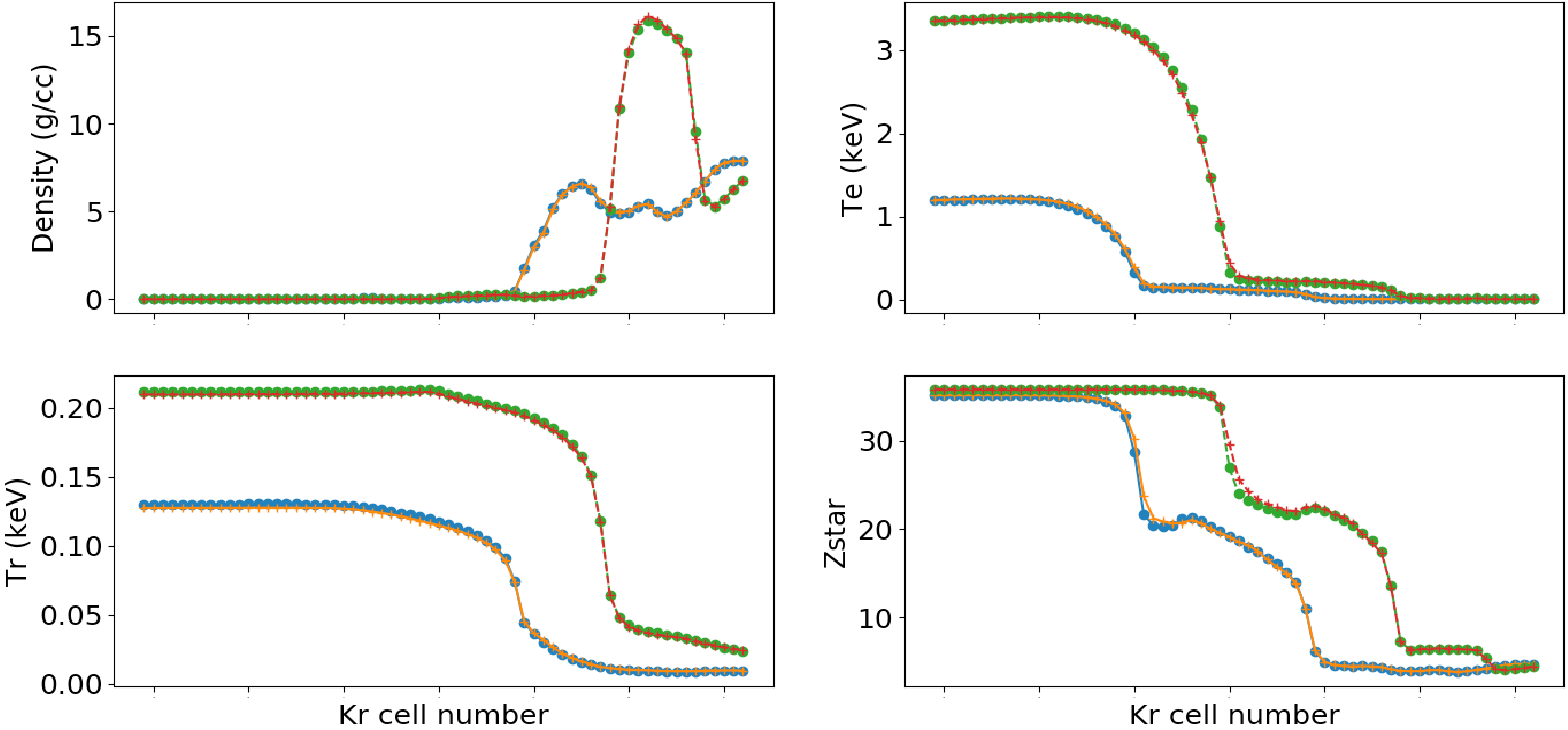
Relative errors in the bubble, at peak flux, on a batch of simulations:

At 5 ns,

	Mean	Max
Te:	5.3%	6.8%
Tr:	1.7%	7.3%

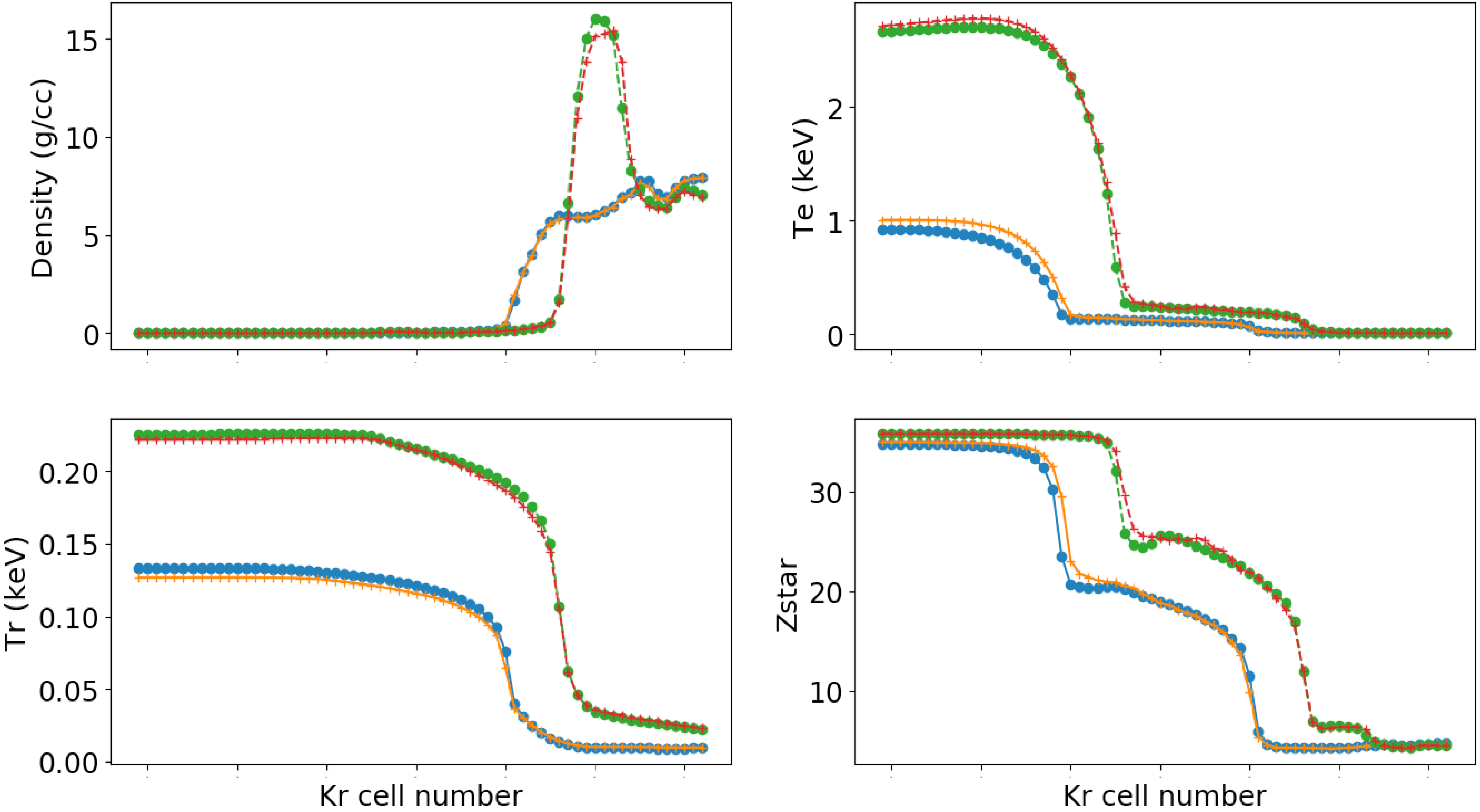
Without radiative field.

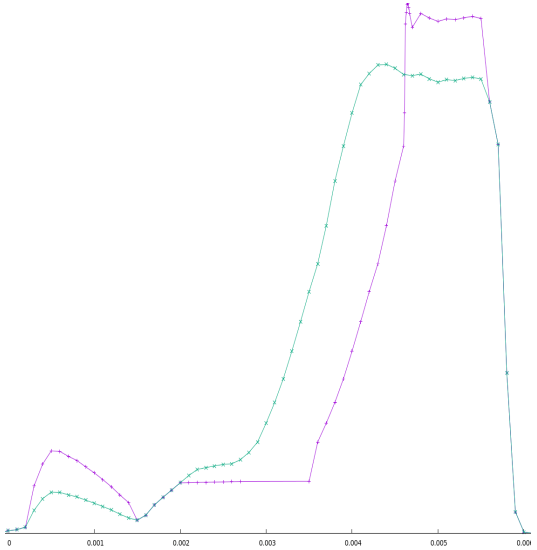
DNN and CRETIN results are identic.



With an other drive.

DNN and CRETIN results are similar in “Extrapolation”.





Modified drive
Initial drive

Conclusion

- We showed here that we can gain CPU time:

  on problem 2, with 1 CPU and 1 thread:

  DCA 434s DNN 65s,

- We will figure out:
  - Why we degrade accuracy from problem 1 to problem 2 ?

    → Better scan of the radiative field input.

  - Why DCA and in-line CRETIN results are different?

    (Not showed here)

Future works

- Machine Learning: improve architecture, transfer learning.
- UQ: use efficient tools to analyze and propagates errors in networks.
- Physics: use it in DNN (free-free part, important lines…)
- HPC work: accelerate training and predictions (CPU, GPU, NN accelerators)
- ICF hohlraum simulations: Au, 2d-3d, radiative fields, w ionization and derivatives coming from NN, w IMC…
- Capsule simulations: non steady-state collisional-radiative model

Try on more accurate atomic model.

Results to answer questions asked during the seminar:

- Accuracy for Problem 2depends on the dataset of radiative fields,

- Visualization of the latent space for the radiative field

Lawrence Livermore National Laboratory

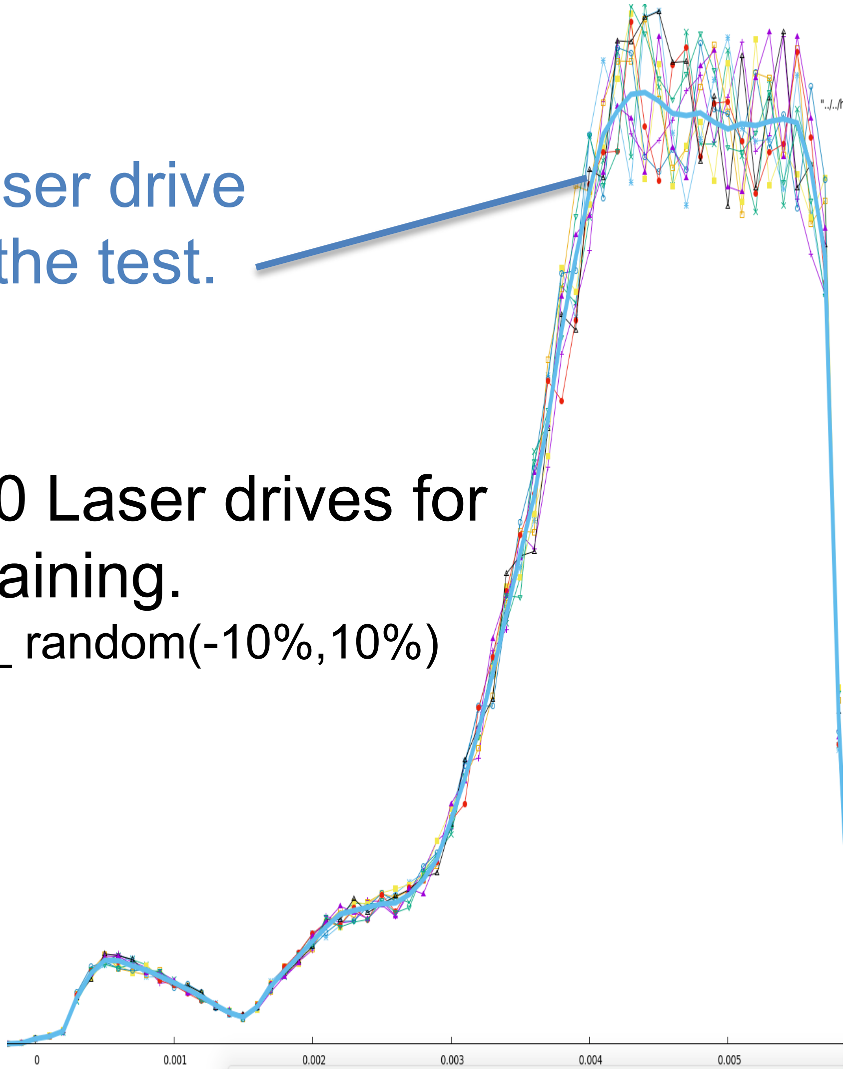
Actualization on Problem 2

- We show that we can obtain a good accuracy on Problem 2 by a large enough dataset for the radiative fields.

- To focus on important data only, we will call:
  - The same NN as before when $T<300$ eV
  - A new NN when $T>300$eV.

Problem2 new dataset

1 laser drive
for the test.



10 Laser drives for
training.
+_ random(-10%,10%)

11 HYDRA
simulations.
Dumps every 50ps.
63 Krypton cells

Radiative fields
dataset.
~ 78K

Density
Te>300.

120K CRETIN

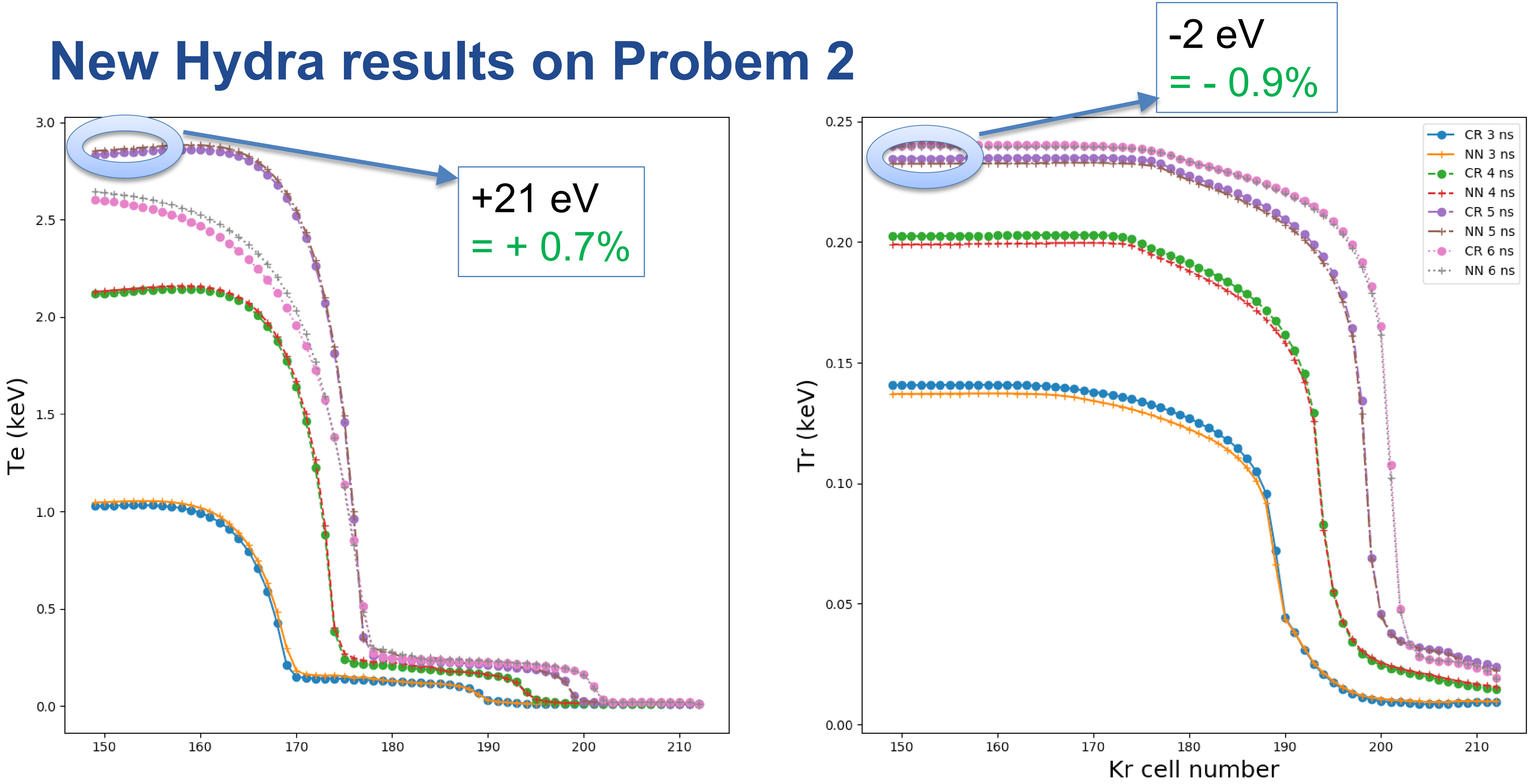
Absorption.
Emissivity.
120K

New results for problem 2

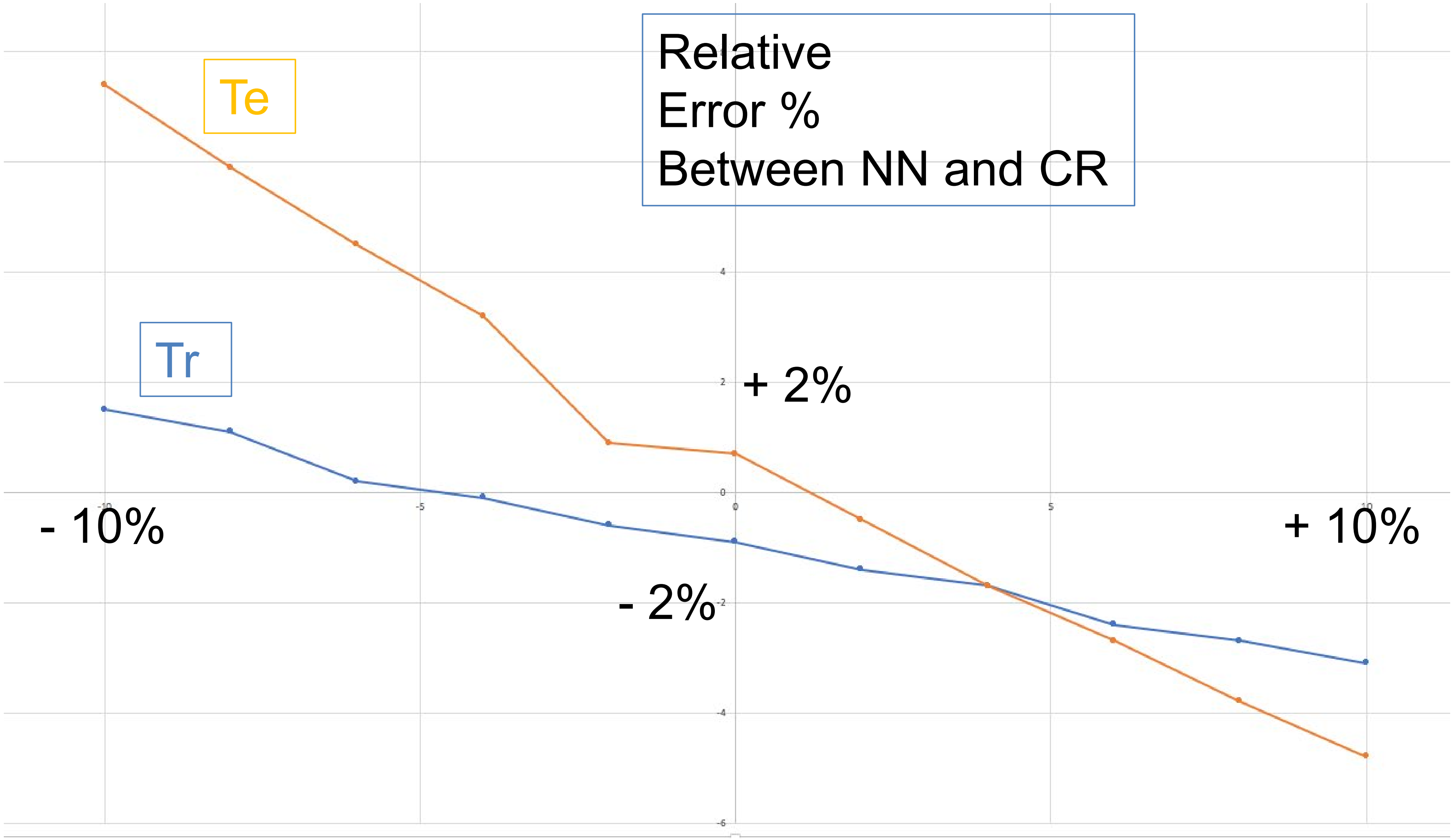
	PROBLEM 1 mean	max	PROBLEM 2 mean	max
Absorption Planck	0.16%	6.06 %	1.07 %	3.56 %
Absorption Rosseland	0.19 %	8.77 %	3.31 %	7.42 %
Emissivity	0.24 %	3.89 %	1.26 %	12.70%

- With these new dataset, we obtain again a good accuracy (for data over 300 eV).

New Hydra results on Probem 2

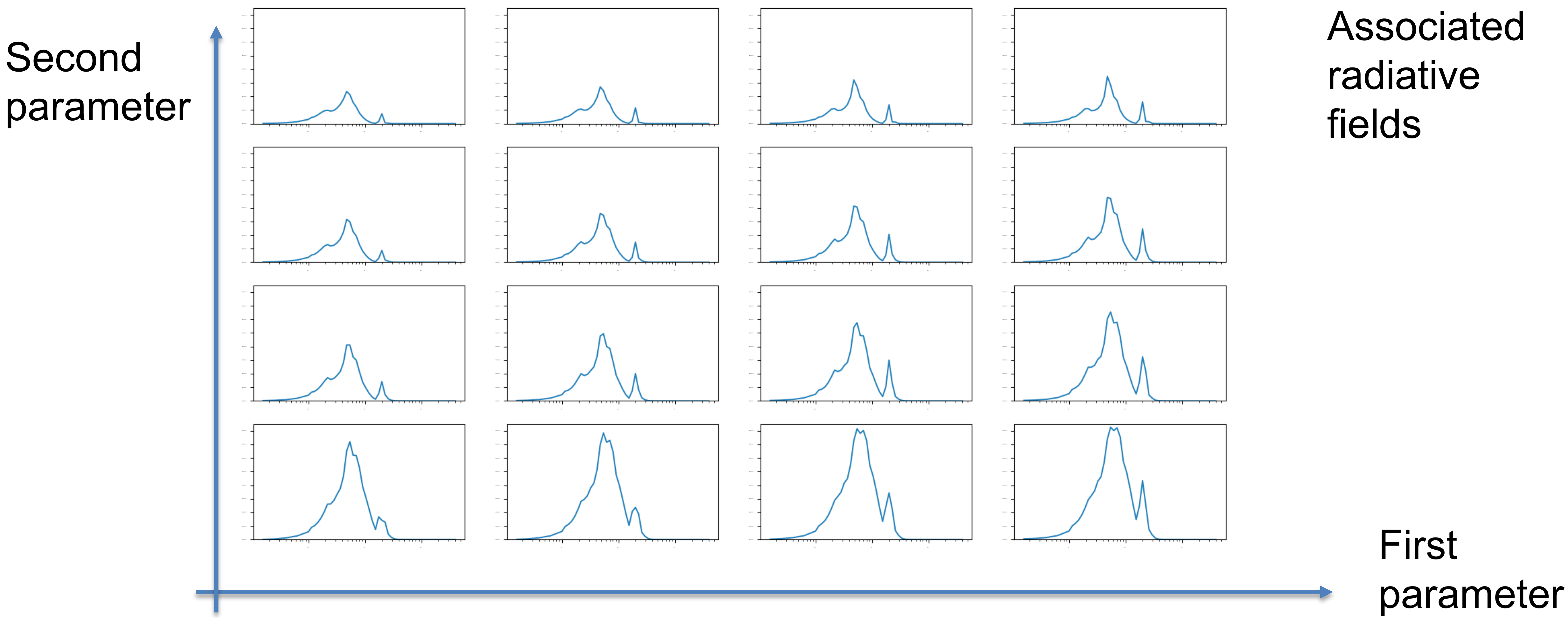


Extrapolation on other laser drives



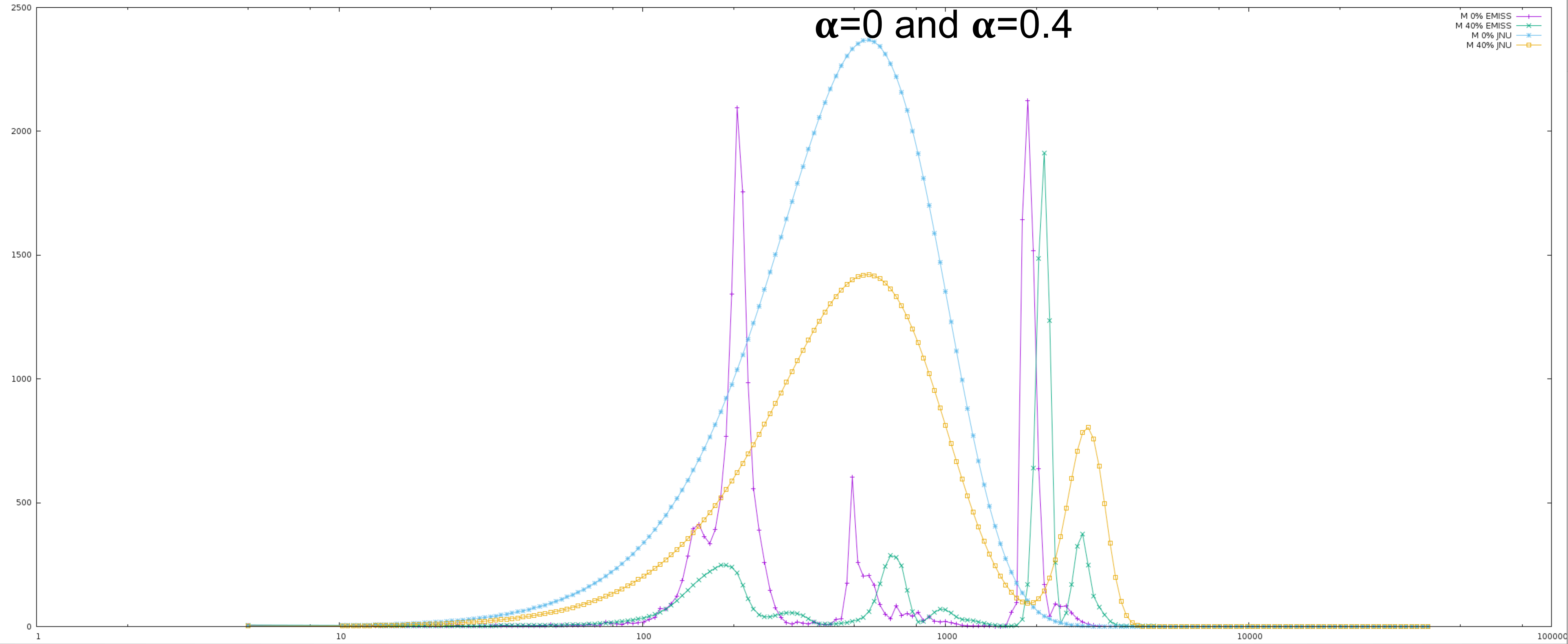
10 new calculations:
laser Drive
=
Drive 0
+- 2%
+- 4%
+- 6%
+- 8%
+- 10%

Visualization of the latent space for the radiative field.



ρ=1mg/cc   Tr=200eV
T=2 keV

Effect of the "M-band" on emissivity
$\boldsymbol{\alpha}$=0 and $\boldsymbol{\alpha}$=0.4



Bubble metric

- 21 calculations (from -10% to +10% on the laser drive),

  with Cretin , and with DNN.
- Mean of Te and Tr on the first 10 Krypton cells,
- Err[calculation k]

  = |mean_DNN[k] – mean_CR[k]|/ mean_CR[k]*100
- Mean and Max over the calculations.