Neural Networks for NLTE

Proof of concept

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

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

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

- ML training or inference every 1k time steps: on the loop
- ML inference every time step: in the loop
- Transfer learning: outside the loop
- Elevated predictive model
Deep neural network in-the-loop: Physics-informed DNN surrogate model in HYDRA

- Laser light
- Ions
- Electrons
- Radiation
- Hydrodynamics

NLTE with Collisional Radiative model CRETIN

DNN For a given atomic model

For a given atomic model $Te, lv$, $k_v, \eta_U$

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

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

Train a DNN
- Apart from HYDRA
- Expensive one time

In-line CRETIN
- Expensive
- Called many times

In-line DNN
- Fast

NLTE with Collisional Radiative model CRETIN

DNN For a given atomic model

\[ \rho, T_e, I_v, \kappa_u, \eta_u \]
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

---

Diagram:
- Nb of levels:
  - $10^6$
  - $10^3$
  - $10^2$

- CPU time per call:
  - ms
  - s
  - hours
  - days

- Points:
  - Average atom
  - DCA
  - SCRAM
  - ENRICO
  - ATOMIC

---

Lawrence Livermore National Laboratory
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

Emissivity spectrum

Absorption spectrum

DNN
Neural networks in spectroscopy.

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

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 $T_r$ and $\alpha$.

$$I(\nu) = aT_r^4 [ (1- \alpha) b(\nu,T_r) + \alpha g(\nu) ]$$

With $b(\nu,T_r)$ the reduced Planckian
$g(\nu)$ the reduced gaussian of mean 3 keV and FWHM 1 keV

- Training dataset: 30K samples (= CRETIN simulations), uniform.
  $3 < \rho < 100\text{mg/cc} \ , \ 300 < T_e < 3000\text{eV} \ , \ 30 < T_r < 300\text{eV} \ , \ 0. < \alpha < 0.3$
Auto-encoders enable us to reduce spectra dimensions. Here from 200 bins to D neurons between 3 and 7.

Latent Space = $F_i(w_i, \text{200-bins spectra})$

$w_i$ 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.
We optimize architecture with \( \sim 3 \) 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.

Relative errors

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.

Relative errors
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
Problem 2

Datasets for training and test.

10 Laser drives for training.

1 laser drive for the test.

Radiative fields dataset. ~ 10K

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

66K CRETIN

Absorption. Emissivity. 66K

Density Te.
Final Network Architecture

40-bins radiative field

Mass density
Temperature

Latent space. D=2
Latent space. D=4

19 hidden layers, 3000 neurons, 2e6 parameters.

40- bins absorption or emissivity

Encoder

Decoder

DJINN

Lawrence Livermore National Laboratory
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%
- **ROSSLAND**
  - Mean: 9.9%
  - Max: 54%
DNN match less with hydra rad fields.
O Cretin X DNN, over 832 test dataset.

Relative errors
FREQ-integrated
Filtered over percentile 30.
Mean 8.1%
Max 43%
Summary of results on both problems

<table>
<thead>
<tr>
<th></th>
<th>PROBLEM 1</th>
<th></th>
<th>PROBLEM 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>max</td>
<td>mean</td>
<td>max</td>
</tr>
<tr>
<td>Absorption</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planck</td>
<td>0.16%</td>
<td>6.06%</td>
<td>6.9 %</td>
<td>33%</td>
</tr>
<tr>
<td>Absorption</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosseland</td>
<td>0.19 %</td>
<td>8.77%</td>
<td>9.9 %</td>
<td>54%</td>
</tr>
<tr>
<td>Emissivity</td>
<td>0.24 %</td>
<td>3.89%</td>
<td>8.1 %</td>
<td>43%</td>
</tr>
</tbody>
</table>

Analytical rad. field (Tr, α) ≈ Real rad. field (40 independents bins)
30K training dataset ≈ 10K rad. field dataset
Smaller range in ρ, Te ≈ 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 Te: 5.3%
- Max Te: 6.8%
- Mean Tr: 1.7%
- Max Tr: 7.3%
Without radiative field.

DNN and CRETIN results are identical.
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 2 depends on the dataset of radiative fields,
- Visualization of the latent space for the radiative field
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>300eV.
Problem 2 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

<table>
<thead>
<tr>
<th></th>
<th>PROBLEM 1 mean</th>
<th>PROBLEM 1 max</th>
<th>PROBLEM 2 mean</th>
<th>PROBLEM 2 max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absorption Planck</td>
<td>0.16%</td>
<td>6.06%</td>
<td>1.07%</td>
<td>3.56%</td>
</tr>
<tr>
<td>Absorption Rosseland</td>
<td>0.19%</td>
<td>8.77%</td>
<td>3.31%</td>
<td>7.42%</td>
</tr>
<tr>
<td>Emissivity</td>
<td>0.24%</td>
<td>3.89%</td>
<td>1.26%</td>
<td>12.70%</td>
</tr>
</tbody>
</table>

- With these new dataset, we obtain again a good accuracy (for data over 300 eV).
New Hydra results on Problem 2

+21 eV
= + 0.7%

-2 eV
= - 0.9%
Extrapolation on other laser drives

<table>
<thead>
<tr>
<th>Laser Drive</th>
<th>Relative Error %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive 0</td>
<td>+10%</td>
</tr>
<tr>
<td>+2%</td>
<td></td>
</tr>
<tr>
<td>+4%</td>
<td></td>
</tr>
<tr>
<td>+6%</td>
<td></td>
</tr>
<tr>
<td>+8%</td>
<td></td>
</tr>
<tr>
<td>+10%</td>
<td></td>
</tr>
<tr>
<td>-2%</td>
<td></td>
</tr>
<tr>
<td>-4%</td>
<td></td>
</tr>
<tr>
<td>-6%</td>
<td></td>
</tr>
<tr>
<td>-8%</td>
<td></td>
</tr>
<tr>
<td>-10%</td>
<td></td>
</tr>
</tbody>
</table>
Visualization of the latent space for the radiative field.
\( \rho = 1 \text{mg/cc} \quad \text{Tr} = 200\text{eV} \quad \) Effect of the “M-band” on emissivity
\( \alpha = 0 \) and \( \alpha = 0.4 \)
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]
  \[= \left| \text{mean}_{\text{DNN}}[k] - \text{mean}_{\text{CR}}[k] \right| / \text{mean}_{\text{CR}}[k] \times 100\]
- Mean and Max over the calculations.