Analysis of Extended X-ray Absorption Fine Structure data using Deep Neural Networks

F. Iesari[1], H. Setoyama[2], T. Okajima[1]
[1]Aichi Synchrotron Radiation Center
[2]Kyushu Synchrotron Light Research Center
fabio.iesari@aichisr.jp

Over the last decade, the use of neural networks (NNs) for scientific applications has been steadily increasing. Consisting of numerous "neurons" stacked into layers, they are able to distinguish patterns or understand relationships between different quantities after appropriate training. Extended X-ray Absorption Fine Structure (EXAFS) data contain structural information in terms of the *n*-body distribution functions [1]:

$$\langle\chi(k)\rangle = \int_0^\infty dr\ 4\pi r^2 \rho g_2(\mathrm{r})\gamma^{(2)}(\mathrm{r},\mathrm{k}) + \int dr_1 dr_2\, d\phi 8\pi^2 r_1^2 r_2^2 sin(\phi)\rho^2 g_3(r_1, r_2, \phi)\gamma^{(3)}(r_1, r_2, \phi, k) + \mathcal{O}(g_4).$$

The inversion of this equation is an ill-posed problem and among the various strategies to obtain a solution, recently NNs have been used [2]. We wanted to investigate whether the same methodology could be applied to disordered phases and whether it would be possible to obtain information beyond the pair distribution function.

The critical point of any NN is the dataset used for the training process, that should be sufficiently large and heterogeneous. For this purpose, we used MD simulations of mono-atomic nickel in different structural configurations and at various temperature. The temperature was increased past the melting point to also include liquid configurations. From each configuration, we calculated the radial distribution function, bond-angle distribution of the nearest neighbors and the EXAFS signal, using GNXAS suite of programs. The created dataset was then used to optimize and train a set of deep NNs to estimate radial and bond-angle distribution functions from a given EXAFS signal.

We used the NNs to analyze data of nickel at different temperatures. Obtained results show that the NNs are able to distinguish between ordered and disordered configurations and are also able to detect small changes in the local ordering of liquid structure, comparable with previously published results [3].

This work was supported by JST CREST JPMJCR1861.

[1] A. Filipponi and A. Di Cicco, *Phys. Rev. B* **52**, 15135 (1995)
[2] J. Timoshenko, A. Anspoks, et al., *Phys. Rev. Lett.* **120**, 225502 (2018)
[3] A. Di Cicco, F. Iesari, et al., *Phys. Rev.* B **89**, 060102 (2014)

Analysis of Extended X-ray Absorption Fine Structure data using Deep Neural Networks

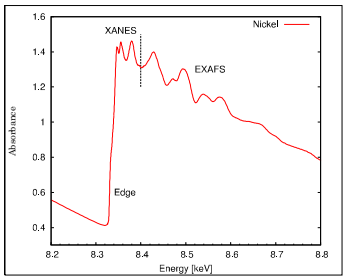
F. Iesari[1], H. Setoyama[2], T. Okajima[1]

[1]Aichi Synchrotron Radiation Center (Japan),
[2]Kyushu Synchrotron Light Research Center (Japan)

fabio.iesari@aichisr.jp

第15回九州シンクロトロン光研究センター研究成果報告会
2021 August 23

Introduction: X-ray absorption spectroscopy



Absorbance: $\alpha(E) = \ln(I_0/I_1)$
EXAFS structural signal:

$$\chi(k) = \frac{\alpha(k) - \alpha_0(k) - \alpha_{exc}(k)}{\alpha_0(k)}$$

where $k = \sqrt{2m(E - E_e)}/\hbar$
n-body distribution expansion:

$$\langle\chi(k)\rangle = \int_0^\infty dr\, 4\pi r^2 \rho\, g_2(r)\, \gamma^{(2)}(r,k) + \int dr_1 dr_2 d\phi\, 8\pi^2 r_1^2 r_2^2 \sin(\phi)\rho^2 g_3(r_1, r_2, \phi) \times \gamma^{(3)}(r_1, r_2, \phi, k) + \mathcal{O}(g_4)$$

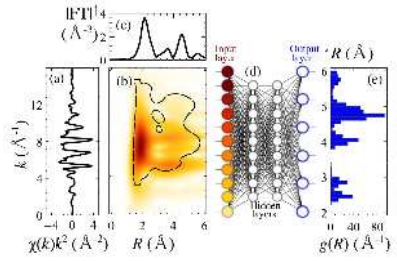
[A. Filipponi, A. Di Cicco, C.R. Natoli. Phys. Rev. B 52, 15122 (1995)]

n-body distribution represents the probability of finding *n* atoms in a given configuration.
The more known equation of EXAFS:

$$\chi(k) = \sum_j \frac{N_j}{kR_j^2} e^{-2R_j/\lambda_j(k)} |f_j(k,\pi)| \sin(2kR_j + 2\delta_1(k) + \phi_j(k)) e^{-2\sigma_j^2 k^2} dr.$$

is a single scattering approximation where the shell is approximated by a Gaussian distribution. This is adequate for some cases, but can lead to ambiguous results and mistakes in presence of distortions or for amorphous systems.

Inversion of the EXAFS equation is known to be an ill-posed problem. Recently, Neural Networks (NNs) have been used to solve this issue for crystals.



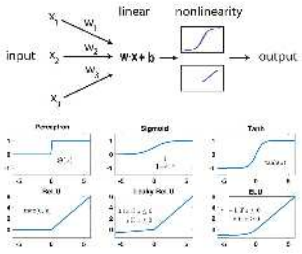
[J. Timoshenko et al., Phys. Rev. Lett. 120, 225502 (2018)]

1. Can we use it also for disordered configurations?
2. Is it possible to obtain more than pair distribution function?

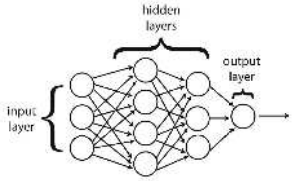
What is a Neural Network?

Nonlinear models for supervised learning, extension of linear and logistic regression.

Basic unit is a "neuron" which takes a vector $\vec{x} = (x_1, x_2, \cdots x_d)$ and output a scalar $a_i(\vec{x})$.
a_i consists in a linear transformation $z^{(i)} = \vec{w}^{(i)} \cdot \vec{x} + b^{(i)}$ and an activation function (non-linear) $\sigma_i(z)$.



Neural Network (NN) consists in many neurons stacked into layers, with the output of one layer acting as the input for the next.



(this is feed-forward network, but many other architectures are possible)
Universal approximation theorem: a neural network with a single hidden layer can approximate any continuous, multi-input/multi-output function with arbitrary accuracy.

[P. Mehta et al., Physics Reports 810, 1-124 (2019)]

During training the weights and bias of the neurons are adjusted, using backpropagation algorithm.

1. Activation at input layer
2. Feedforward
3. Error at top layer
4. "Backpropagate" the error
5. Calculate gradient of the error respect to the parameters

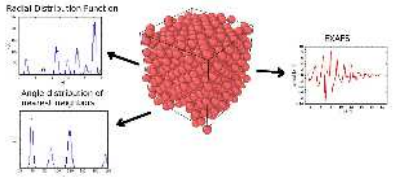
Additional practices are commonly used to avoid typical pitfalls in using NNs (early stopping, dropout, batch normalization).

NN Flowchart:

1. Load and process the data
2. Define the model and its architecture
3. Choose the optimizer and the cost function
4. Train the model
5. Evaluate the model performance on unseen test data
6. Modify the hyperparameters and architecture to optimize performance for the specific problem

Constructing the dataset for training

We use Molecular Dynamics (MD) simulations through LAMMPS for creating various structures at different temperatures using Modified Embedded-Atom Method potential. For each configuration we calculate radial distribution function (RDF), bond angle distribution of nearest neighbors and EXAFS signal.



$$n(r) = \langle \sum_{\substack{j,j \\ i \neq j}} \delta(|\vec{r}_i - \vec{r}_j| - r) \rangle = 4\pi r^2 \rho g(r)$$

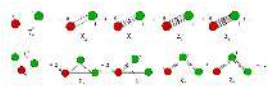
Created configurations (fcc, hcp, bcc, diamond and liquids) using NVT ensemble of Ni mono-atomic system.
Atoms: ∼1000 (for crystals), 4000 (for liquids)
Temperature: 60 - 1500 K
1500 - 2000 K ($\Delta T = 20K$)
Volume variation: $\Delta a = 0.02$ Å (±10 times)

MD configurations are only used to establish the relationship between structure and EXAFS signal, so it's not required that they represent real configurations.

EXAFS calculation

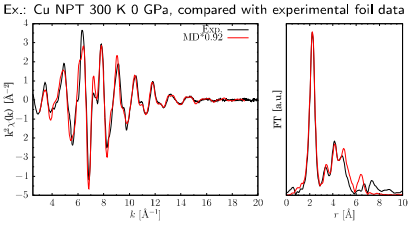
EXAFS signal is calculated averaged over all atoms in the configuration (ensemble average) considering 2-body and 3-body terms using GNXAS.

$$\langle \chi(k) \rangle = \left\langle \sum_i \gamma^{(2)}(0,i) + \sum_{(i,j)} \gamma^{(3)}(0,i,j) \right\rangle \text{ over all atoms}$$



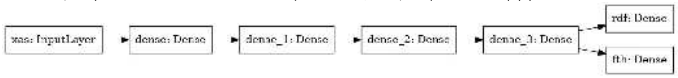
due to mean free path effects, only paths up to a given cut-off are considered.
The energy difference parameter ΔE (difference between theoretical and experimental energy scale) is chosen randomly between -5 eV to +10 eV for every configuration.
This way NN becomes independent of this parameter.

Ex.: Cu NPT 300 K 0 GPa, compared with experimental foil data



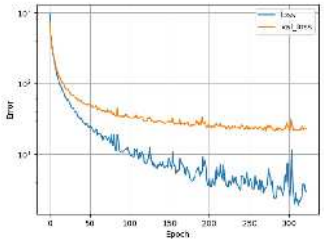
NN Architecture

Python3.6 (TensorFlow, Keras)
Optimized by grid search
4 dense layers (1000 neurons, Activation: ReLU) and 2 output layers (RDF and $N(\theta)$)

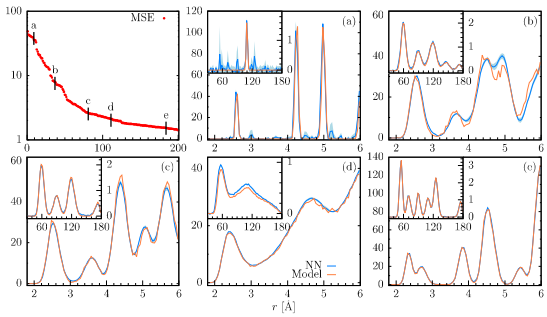


Optimizer = Adam (stochastic gradient descent, learning rate = 1.0E-4)
Cost function = Mean Squared Error of RDF and $N(\theta)$

20% of the dataset (chosen randomly) was used as test set, the rest was divided into 5 equal parts: 1 part was used as validation set to implement early stopping and the other 4 as training set. The part used as validation set was rotated and each time a new NN was trained, for a total of 5. Predictions are then averaged and the standard deviation used as error.



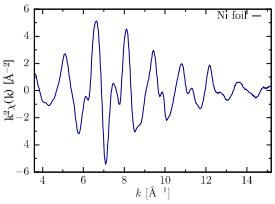
Test Dataset



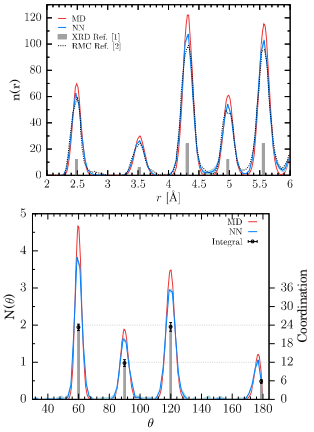
Performance of the trained NN on the test dataset. On the top left, the Mean Square Error (MSE) of the worst 200 cases are shown in descending order. Predicted RDF and BAD (insets) from selected configurations are shown in blue and the shadowed area represent standard deviation of the predictions. Orange lines are the same quantities calculated from the models, which are: (a) dia, (b) hcp, (c) fcc, (d) liquid and (e) bcc.

Experimental Ni at RT

Ni foil @ RT measured at BL11 of Saga Light Source
Comparison MD simulations of Ni NVT 300 K

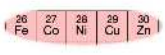


[1] H. M. Otte, J. Appl. Phys. 32, 1536–1546 (1961)
[2] J. Timoshenko et al., Phys. Rev. Lett. 120, 225502 (2018)

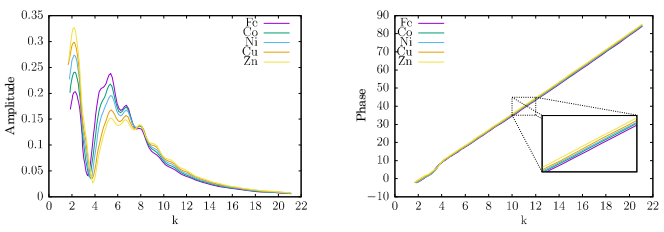


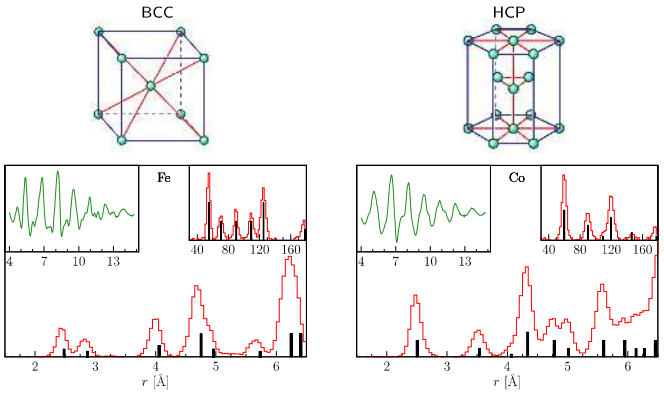
Neighboring elements

Due to the fact that scattering properties are similar for neighboring atoms in the periodic table, NN can be used also used to analyze spectra of surrounding elements.

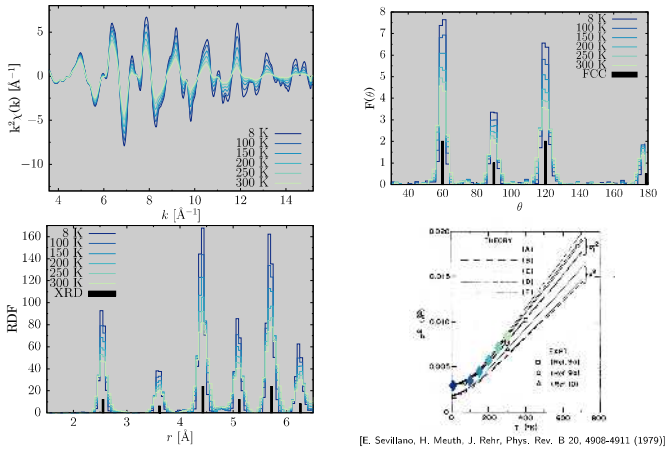


$$\chi(k) = A(k)\sin\Psi(k)$$





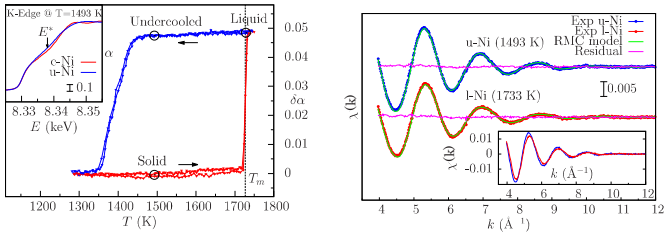
Cu at low temperatures



[E. Sevillano, H. Meuth, J. Rehr, Phys. Rev. B 20, 4908-4911 (1979)]

Liquid and Undercooled Ni

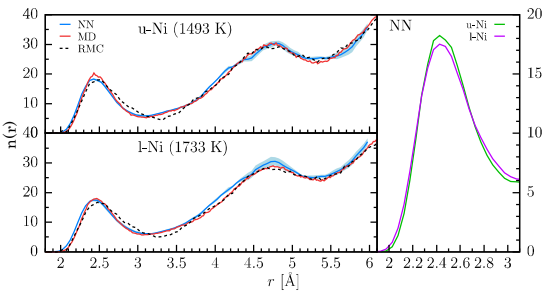
After melting, small metal particles can remain in the liquid phase for temperature below the melting point.



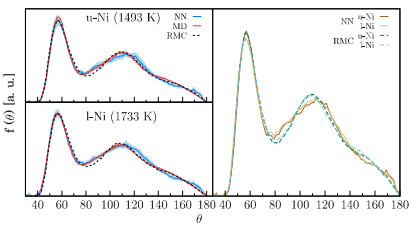
[A. Di Cicco et al., Phys. Rev. B 89, 060102 (2014)]

Radial distribution function

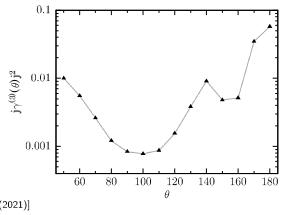
Comparison of NNs prediction and RMC analysis on the same data and MD simulations



Bond-angle distribution



Amplitude of EXAFS signal as a function of the bond-angle



[F. Iesari et al., Symmetry 13, 1070 (2021)]

Conclusions

1. Can we use it also for disordered systems?

   We could obtain RDF and BDA for liquid structures
   NNs have been able to distinguish small differences between liquid and undercooled phase

2. Is it possible to obtain more than pair distribution function?

   Correctly obtained bond angle distributions of various structure at different temperatures
   Reconstruction may be challenging when 3-body EXAFS signal becomes small

Acknowledgements
Financial support by JST CREST JPMJCR1861.
Part of the calculations have been carried out using Supercomputer System ITO (R.I.I.T., Kyushu University).

Thank you for your kind attention.