The Journal of Open Source Software

# reflectorch: a deep learning package for X-ray and neutron reflectometry

**Valentin Munteanu** [1*], **Vladimir Starostin** [2*], **Alexander Hinderhofer** [1], **Alexander Gerlach** [1], **Dmitry Lapkin** [1], and **Frank Schreiber** [1¶]

**1** University of Tübingen, Institute of Applied Physics, Auf der Morgenstelle 10, 72076 Tübingen, Germany **2** University of Tübingen, Cluster of Excellence "Machine learning - new perspectives for science", Maria-von-Linden-Straße 6, 72076 Tübingen, Germany ¶ Corresponding author * These authors contributed equally.

**DOI:** 10.xxxxxx/draft

**Software**

- Review
- Repository
- Archive

**Editor:** Johanna Bayer

**Reviewers:**

- @turnerjennings
- @thomashopkins32
- @rachitanaik-lyft

**Submitted:** 08 August 2024
**Published:** unpublished



## Summary

We introduce `reflectorch`, a Python package which facilitates the full machine learning pipeline for the data domain of X-ray and neutron reflectometry. Firstly, the package allows the choice of different parameterizations of the scattering length density profile of a thin film, or generally, a layered structure, and the sampling of the ground truth physical parameters from user-defined ranges. Secondly, the package provides functionality for the fast simulation of reflectometry curves on the GPU using a vectorized implementation of the Abelès matrix formalism (Abelès, 1950) and the augmentation of the theoretical curves with noise informed by experimental considerations. The architecture of the neural network as well as the training callbacks and hyperparameters can be easily customized from YAML configuration files. Notably, our implementation makes use of a special training procedure introduced in our publication (Munteanu et al., 2024), in which prior boundaries for the target parameters are provided alongside the reflectivity curve as an additional input to the neural network.

## Statement of need

X-ray and neutron reflectometry (XRR and NR) are widely-used and indispensable experimental techniques for elucidating the structure of interfaces and thin films, i.e. layered systems. Experimentalists, having measured a reflectometry curve, are faced with the task of obtaining the physical parameters corresponding to an assumed parameterization of the scattering length density (SLD) profile along the depth of the investigated sample. The inverse problem is non-trivial and generally ambiguous due to the lack of phase information and experimental limitations. Recent years have seen an increased interest in the fast analysis of reflectometry data using machine learning techniques, as such methods could be adapted for real-time investigations at large scale facilities (i.e. synchrotron and neutron sources), potentially enabling closed loop experiments (Pithan et al., 2023).

Our contribution to the open-source scientific software community is `reflectorch`, a deep learning package built in the Pytorch framework, which posseses a modular, object-oriented and customizable design. Reflectorch enables the user to simulate reflectivity curves in a fast and vectorized manner which takes advantage of the computing power of modern GPUs. The neural network architecture as well as the training callbacks and hyperparameters can be easily customized by editing YAML configuration files. As a result of the special training procedure, which incorporates minimum and maximum prior bounds for the parameters as described in (Munteanu et al., 2024), the user is able to make use of prior knowledge about the investigated sample at inference time.

## General workflow

Different types of parameterizations of the SLD profile can be represented as subclasses of the `ParametricModel` class. The parameters are represented as an instance of the `BasicParams` class, which encapsulates the parameter values and their prior bounds, also taking care of scaling these values to neural-network friendly ranges. The prior sampler is responsible for sampling the values of the parameters and their prior bounds from their predefined ranges, in which a subclass of `SamplingStrategy` can be used to further restrict the values of some parameters with respect to others. Based on the sampled parameters and on the momentum transfer (q) values generated by a subclass of `QGenerator`, batches of reflectivity curves are simulated. The reflectivity curves are further augmented with experimentally-informed noise provided by a subclass of `IntensityNoiseGenerator` and scaled to a neural network-friendly range by a subclass of `CurvesScaler`.

The trainer encapsulates the data loader, the Pytorch optimizer, the neural network, and other training hyperparameters, allowing the seamless training of the model and the saving of the resulting model weights and of the history of losses and learning rates. The neural network architecture consists of an embedding network for the reflectivity curves and a multilayer perceptron, components which can be flexibly customized. Callback objects can be used to control the training process, such as scheduling the learning rate or periodically saving the model weights.

Finally, the `EasyInferenceModel` class serves as a wrapper around trained models, simplifying the inference step. It also provides functionality for automatically downloading model weights and configuration files not locally available from a remote Huggingface repository.

## Related Work

There are several well-established packages designed for the conventional analysis of X-ray and neutron reflectometry data such as `GenX` (Glavic & Björck, 2022), `refnx` (Nelson & Prescott, 2019) and `refl1d` (Maranville et al., 2020). While several machine learning approaches pertaining to X-ray or neutron reflectometry have been proposed in various publications, `mlreflect` (Greco et al., 2022) is the only other properly packaged and documented software so far to the best of our knowledge. Previously developed in our group, `mlreflect` is built in the Tensorflow deep learning framework and has limited functionality compared to `reflectorch` (specifically, the SLD profile is limited to a single film). Still, it has been successfully adopted for use in the scattering community such as in the publication (Schumi-Mareček et al., 2024).


## Acknowledgements

The research was part of a project (VIPR 05D23VT1 ERUMDATA) funded by the German Federal Ministry for Science and Education (BMBF). This work was partly supported by the consortium DAPHNE4NFDI in the context of the work of the NFDI e.V., funded by the German Research Foundation. Supported by the Cluster of Excellence "Machine Learning – New Perspectives for Science" funded by the German Research Foundation under Germany's Excellence Strategy – reference Number EXC 2064/1 - project number 390727645.


## References


Abelès, F. (1950). La théorie générale des couches minces. *Journal de Physique Et Le Radium*, *11*(7), 307–309. https://doi.org/10.1051/jphysrad:01950001107030700

Glavic, A., & Björck, M. (2022). *GenX 3*: the latest generation of an established tool. *Journal of Applied Crystallography*, *55*(4), 1063–1071. https://doi.org/10.1107/S1600576722006653

Greco, A., Starostin, V., Edel, E., Munteanu, V., Russegger, N., Dax, I., Shen, C., Bertram, F., Hinderhofer, A., Gerlach, A., & Schreiber, F. (2022). Neural network analysis of neutron and x-ray reflectivity data: Automated analysis using mlreflect, experimental errors and feature engineering. *J. Appl. Crystallogr.*, *55*(2), 362–369. https://doi.org/10.1107/S1600576722002230

Maranville, B. B., Kienzle, P., Richardsheridan, Doucet, M., Nelson, A., Hoogerheide, D. P., Book, A., & Ghose, R. (2020). *Reflectometry/refl1d: v0.8.13*. Zenodo. https://doi.org/10.5281/ZENODO.4329338

Munteanu, V., Starostin, V., Greco, A., Pithan, L., Gerlach, A., Hinderhofer, A., Kowarik, S., & Schreiber, F. (2024). Neural network analysis of neutron and x-ray reflectivity data incorporating prior knowledge. *Journal of Applied Crystallography*, *57*(2). https://doi.org/10.1107/s1600576724002115

Nelson, A. R. J., & Prescott, S. W. (2019). *refnx*: neutron and X-ray reflectometry analysis in Python. *Journal of Applied Crystallography*, *52*(1), 193–200. https://doi.org/10.1107/S1600576718017296

Pithan, L., Starostin, V., Mareček, D., Petersdorf, L., Völter, C., Munteanu, V., Jankowski, M., Konovalov, O., Gerlach, A., Hinderhofer, A., Murphy, B., Kowarik, S., & Schreiber, F. (2023). Closing the loop: autonomous experiments enabled by machine-learning-based online data analysis in synchrotron beamline environments. *Journal of Synchrotron Radiation*, *30*(6), 1064–1075. https://doi.org/10.1107/S160057752300749X

Schumi-Mareček, D., Bertram, F., Mikulík, P., Varshney, D., Novák, J., & Kowarik, S. (2024). Millisecond X-ray reflectometry and neural network analysis: unveiling fast processes in spin coating. *Journal of Applied Crystallography*, *57*(2), 314–323. https://doi.org/10.1107/S1600576724001171