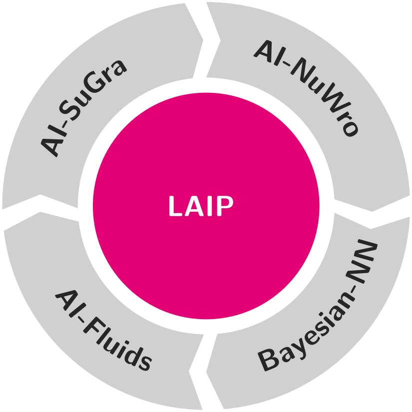
Deep learning in modeling lepton-nucleus scattering

Krzysztof M. Graczyk

Institute for Theoretical Physics
Neutrino Physics Division
NuWro Group

September 12, 2024

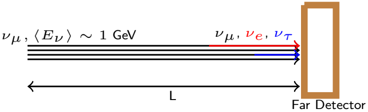


https://kgraczyk.github.io/laip/

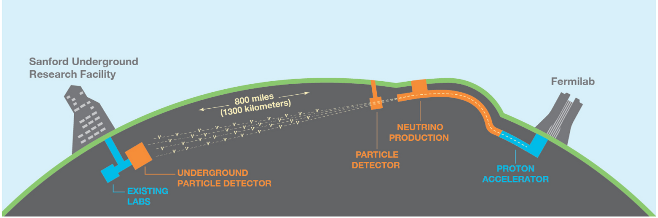
- AI-NuWro: (anti)neutrino-nucleon scattering physics
  - current NuWro team: **Jan Sobczyk**, Artur Ankowski, Rwik Banerjee, Luis Bonilla, Krzysztof Graczyk, Beata Kowal, Hemant Prasad
- Development of:
  - MC code of NuWr
  - nuclear and hadronic models
  - deep learning techniques for NuWro
  - → **Can we teach neural networks physics?**

Neutrino-Matter Interactions

- neutrinos ν_e, ν_μ, ν_τ,..., fundamental particles
  - weakly interacting, neutral, difficult to detect...
- neutrino oscillations:



  - CP-violation phase $\rightarrow$ the Matter-Antimatter asymmetry
  - mass hierarchy



- **A huge experimental and theoretical effort in studies of neutrino properties**
- Deep Underground Neutrino Experiment (DUNE), HyperKamiokande and T2K experiments

Neutrino-Matter Interactions

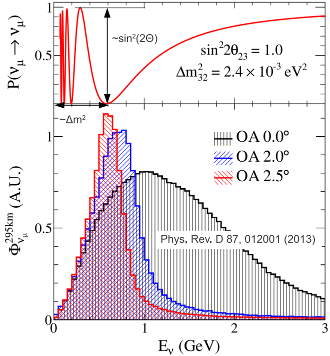
- Accelerator neutrinos:

$\rightarrow$ 1 GeV neutrinos interact with Oxygen (HyperKamiokande), Argon (DUNE), ...

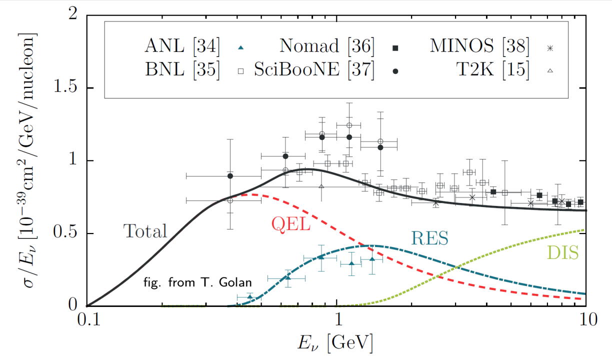
- From the precision of the order of **20-30%** in the knowledge of ν-nuclei scattering cross sections **to the order of percent**

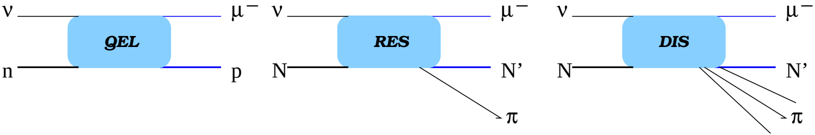
$$P(\nu_\mu \to \nu_\tau) = \sin^2\theta_{23}\left(\frac{\Delta m^2_{32} L}{4E_\nu}\right)$$

To investigate oscillations, we need to know the energy of incoming neutrinos



- **neutrino energy**, E_ν, given by some distribution, one must reconstruct energy event-by-event





- Neutrino energy reconstructed mainly from the analysis of QuasiElastic (QE) scattering events!

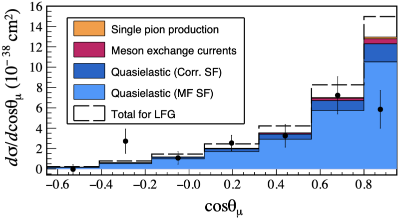
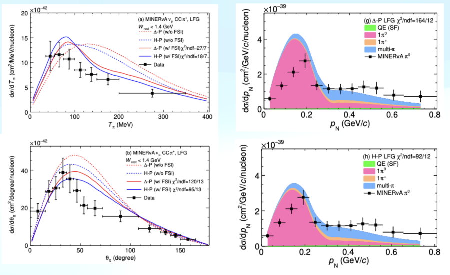


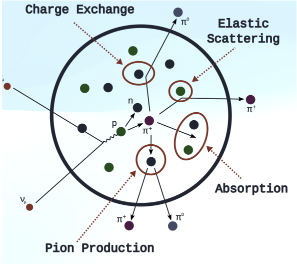
FIG. 2. MicroBooNE CC1p0π differential cross section as

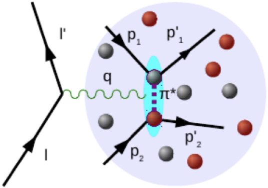


Data from: B. Eberly et al. (MINERvA), Phys. Rev. D92 092008 (2015), arXiv:1406.6415 [hep-ex].

Data from: D. Coplowe et al. (MINERvA), Phys. Rev. D 102 072007 (2020), arXiv:2002.05812 [hep-ex].

Figs. from Banerjee, Ankowski, Graczyk, Kowal, Prasad, Sobczyk, Phys.Rev.D 109 (2024) 073004

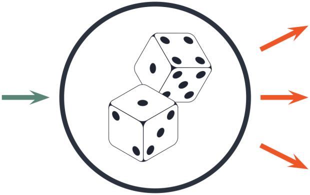


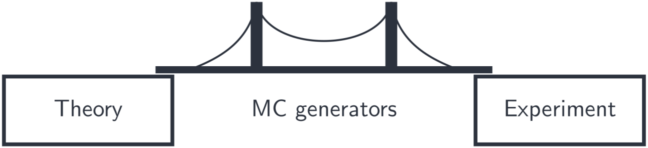


figs. from J. Sobczyk

- We need to simulate ν-Nucleus in realistic conditions
  - $\rightarrow$ Monte Carlo Generator of Neutrino Interactions

- Monte Carlo Generator of neutrino interactions, written in C++
- From 2005 (University of Wroclaw, project lead by Jan Sobczyk)
- Optimized for neutrino energy 1 GeV
- Handle all kinds of targets, and neutrino fluxes, equipped with detector interface
- Output files in ROOT format
- PYTHIA6 used for harmonization and Deep Inelastic Scattering
- open source code, repository: `https://github.com/NuWro/nuwro`





figs. from T. Golan

NuWro

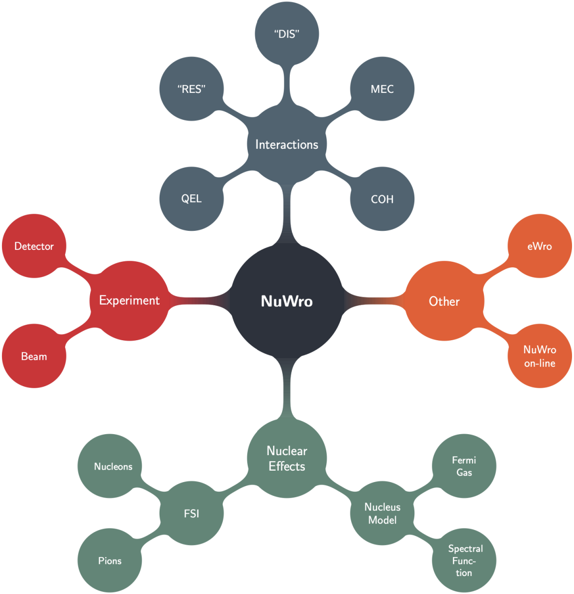
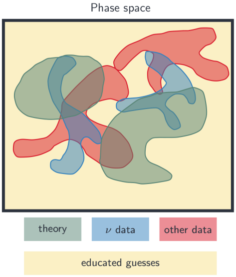


fig. from T. Golan

INGREDIENTS:



RECIPE:

figs. T. Golan

- One must combine various theoretical/phenomenological models with different data types in different kinematic regimes and reaction scenarios.

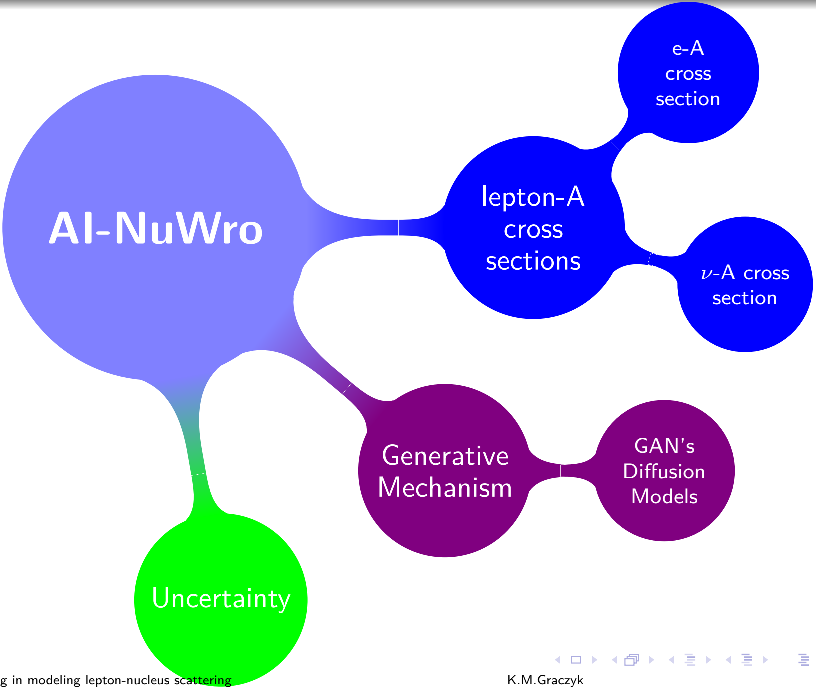
Problems to solve

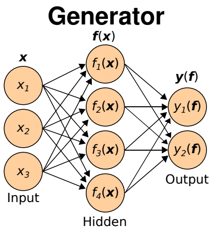
General Problems

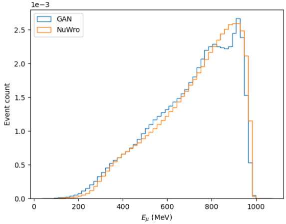
- Limited neutrino scattering data
- Limited knowledge of neutrino-nucleus scattering cross sections
- How to transform the knowledge of physics in one kinematic domain to the other where there are no measurements?
- How to transfer a knowledge from electron scattering physics?

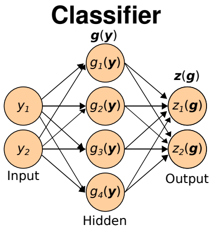
Monte Carlo Generator goals

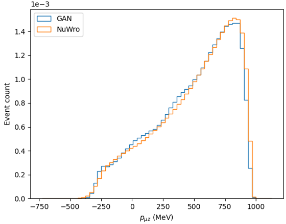
- Obtain the system that automatically and objectively updates its knowledge of physics when new data and new theoretical constraints are delivered
- Obtain the set of methods that allows one to access how the system is uncertain in the predictions


AI-NuWro
lepton-A cross sections
e-A cross section
ν-A cross section
Generative Mechanism
GAN's Diffusion Models
Uncertainty







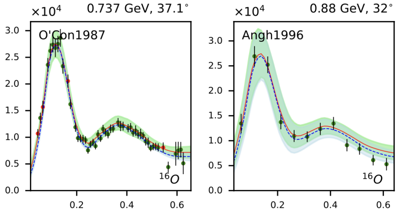


Sucesfull test for QE, paper in November 24

- **Similarities between electron and neutrino interactions with nuclear targets**
  - vector-axial contribution
  - the same nuclear physics
  - similarities in mechanism for final state interactions ...
  - accurate data in eA scattering

    $\longrightarrow$ test neutrino interactions models

    $\longrightarrow$ transfer nuclear mechanisms knowledge from electron to neutrino scattering physics



QE, dip, RES peaks are clearly distinguished! Not the case in neutrino interactions!

- The deep neural network (DNN) can predict electron-nuclei scattering cross sections
  - see achievements of Martini et al. (Phys.Rev.C 107 (2023) 6, 065501), Lovato et al. (2406.06292) and our group (Phys.Rev.C 110 (2024) 2, 025501)
- but can DNN learn basic nuclear properties?

→ Yes it can, but: **representational learning** and **transfer learning**

STEP I: Teach deep neural networks nuclear physics

Teach DNN nuclear physics: electron-carbon scattering

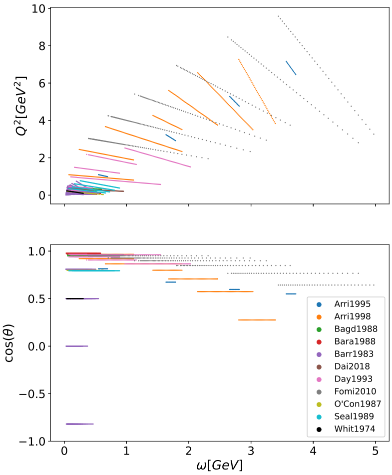
$\rightarrow$ Kowal, Graczyk, Ankowski, Banerjee, Prasad, and Sobczyk, Phys.Rev.C 110 (2024) 2, 025501

- **inclusive electron-carbon cross-sections**

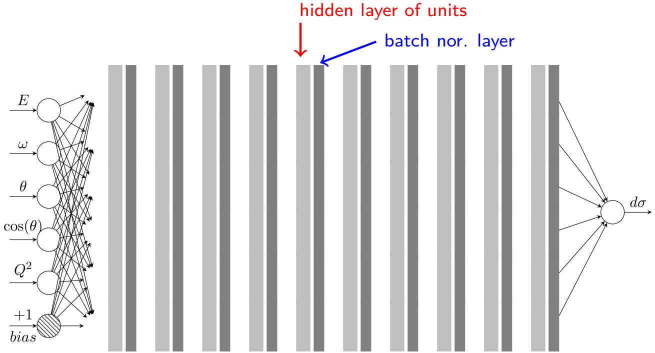
$$DNN(E, \theta, \omega, \cos\theta, Q^2) \rightarrow \frac{d^2\sigma}{d\cos\theta d\omega}$$

E = Energy, θ = scattering angle, ω =transfer of energy

- 11 independent datasets and 3265 points
- stat., sys. and nor. sys. uncertainties
- a broad kinematic region: **quasielastic scattering**, **pion production**, and the onset of **deep-inelastic scattering**
- **we removed the lowest ω data**



Data from http://discovery.phys.virginia.edu/research/groups/qes-archive/notes.html



- 10 blocks, each consists of 300 fully connected units and following batch normalization layer
- Batch Normalization (Ioffe and Szegedy, arxiv:1502.03167)

- We need a model that generalizes well

$\rightarrow$ deep neural networks tend to generalize well

* see Zhang, *et al.*, arXiv:1611.03530

- How uncertain are the network's predictions?

* Open problem in DL, see Gawlikowski *et al.*, arXiv:2107.03342

$\rightarrow$ We follow

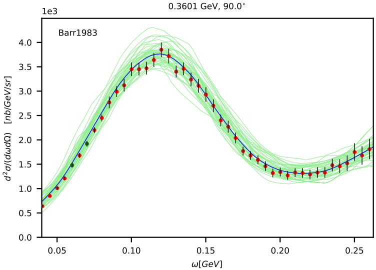
  - Ensemble methods $\rightarrow$ bootstrap approach (model A)
  - Bayesian methods $\rightarrow$ Variational Inference $\rightarrow$ MC Dropout (model B)

- Efron (1979): bootstrap parametric and non-parametric adapted for neural networks by Tibshirani (1996) and Breiman (1996).

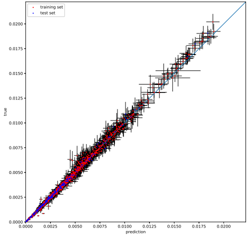
* Model's prediction = average over the ensemble of models

** Augmentation-like technique

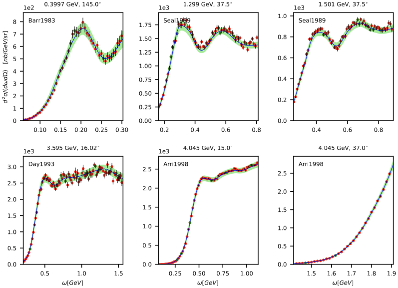
- We split the dataset into training and test datasets, with a proportion of 9:1.



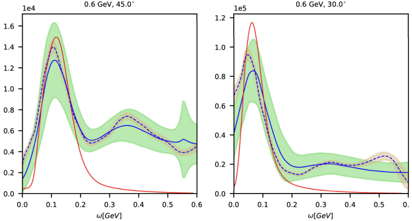
Model A



DNN predictions



Spectral function (QE scattering) vs DNN



Model A and model B

training, test data points

NEXT:1 electron - any target cross section model

NEXT:2 neutrino - any target cross section model

STEP II: Did neural networks learn nuclear physics? If yes, let us take profit from that

Graczyk, Kowal, Ankowski, Banerjee, Bonilla, Prasad, Sobczyk, arXiv:2408.09936
Electron-nucleus cross sections from transfer learning

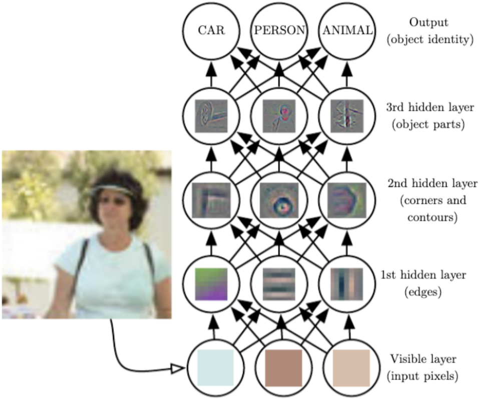
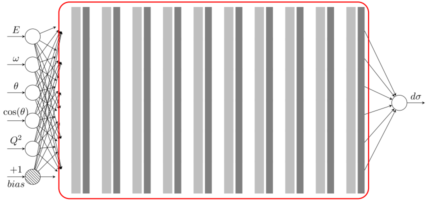


fig from: Deep Learning, Goodfellow, Bengio and Courville

- fundamental concept of deep learning, Bengio, Courville, and Vincent, IEEE Transactions on Pattern Analysis and Machine Intelligence 35, 1798 (2013).
- A graph connecting basic with abstract features
- Transfer learning known in psychology and education. It refers to the ability of a person who has learned skills in one specific field to easily acquire skills needed in related areas of life.
- used in deep learning: DNN, trained on non-medical data, after fine-tuning, is used to detect cancer in medical photos.

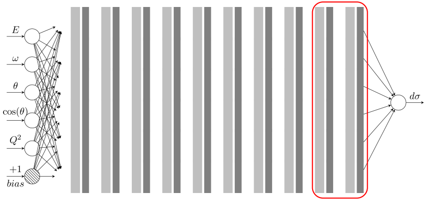
The First Scenario

all layers fine-tuned



The Second Scenario

the last two blocks fine-tuned

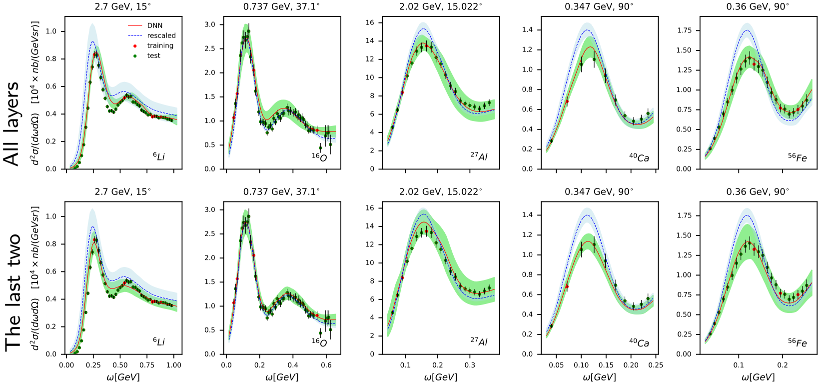


- Consider electron scattering on lithium, oxygen, aluminium, calcium and iron
- For each target consider its own fine-tuning procedure
- To tests transfer learning minimize as much as possible training dataset:

→ **training:test = 1:9**

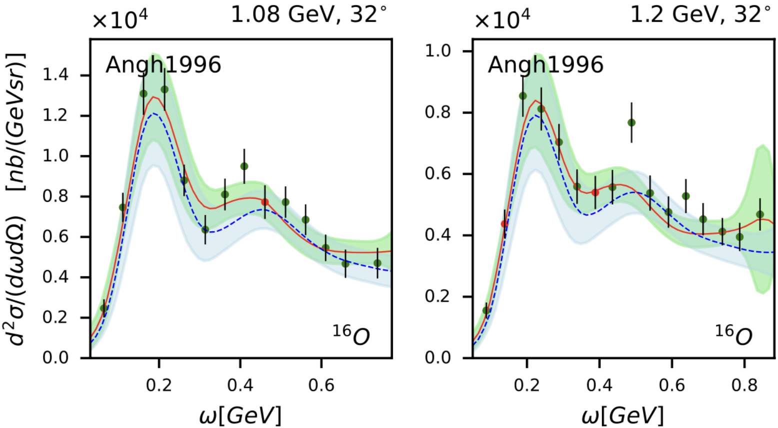
∗ Fine-tuning: from 450 to 1,200 epochs compared to 40,000 epochs to train the pre-trained model

Fine-Tuning: training (10%), test (90%)

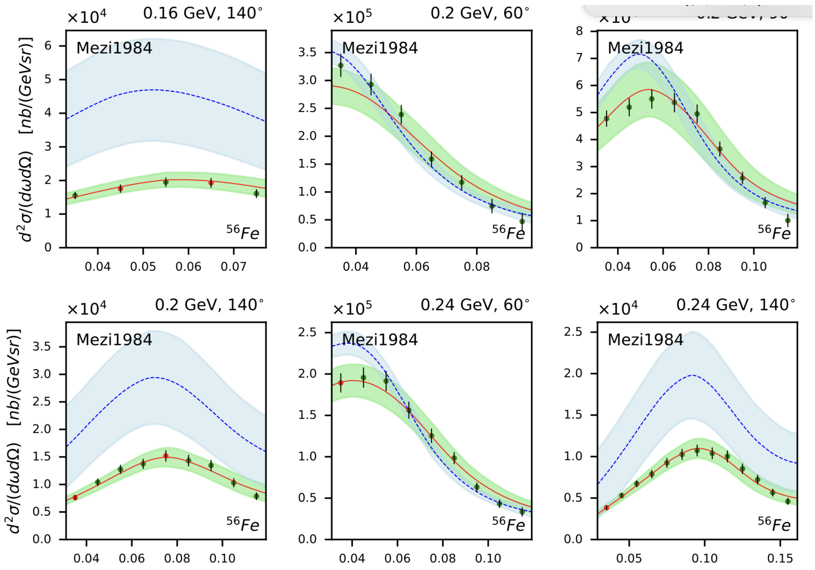


rescaled = (A/12) carbon cross section

training (10%), test (90%): Oxygen, all layers fine-tuned

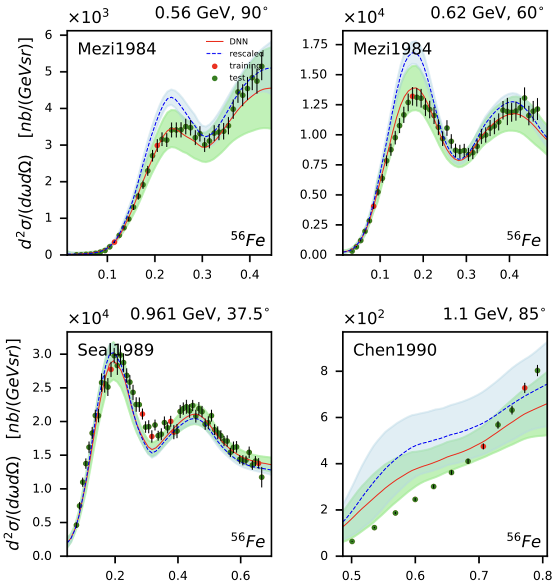


training (10%), test (90%): Iron, all layers fine-tuned



Note that relative normalization parameters (due to nor. sys. uncert.) were taken into account

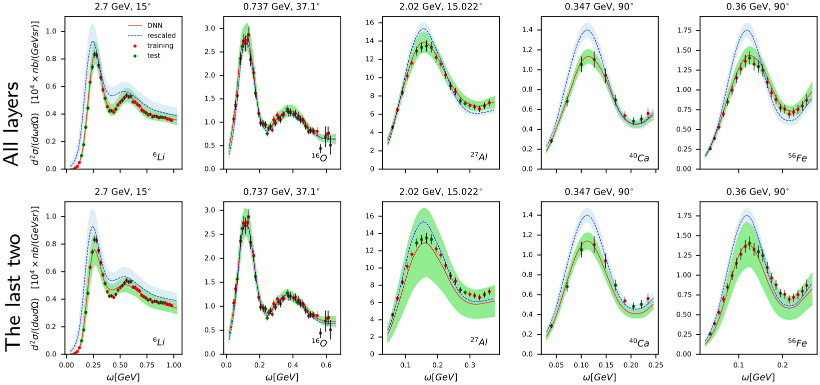
training (10%), test (90%): Iron, two last layers fine-tuned



Summary

- Transfer learning techniques can help reconstruct the cross sections for processes for which the experimental information is limited, such as electron-argon and neutrino-argon scattering.
- Transfer learning is a universal technique

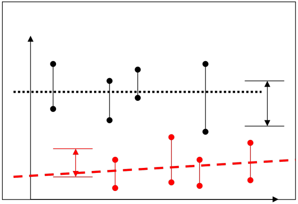
Backup



rescaled = (A/12) carbon cross section

Likelihood and Systematic Normalization

$$\chi_{\text{tot}} = \sum_{k=1}^{11}\left[\chi_k^2(\lambda_k) + \frac{1}{2}\left(\frac{1-\lambda_k}{\Delta\lambda_k}\right)^2\right], \quad \chi_k^2(\lambda_k) = \frac{1}{2}\sum_{i=1}^{N_k}\left(\frac{d\sigma_k^i - \lambda_k d\sigma_i^{\text{fit}}(E_k^i, \theta_k^i)}{\Delta d\sigma_k^i}\right)^2$$



see D'Agostini, NIMPR A 346 (1994) 306

- elastic ep scattering, see e.g. PRC79 (2009) 065204
- C_5^A-axial form factor and consistency of ANL and BNL data: PRD80 (2009) 093001
- DNN tends to lose proper normalization, Graczyk *et al.* Self-Normalized Density Map (SNDM) for Counting Microbiological Obejcts, Sci Rep 12, 10583 (2022)

Abbrev.	$\Delta\lambda_k$
Arri1995	4.0%
Arri1998	4.0%
Bagd1988	10.0%
Bara1988	3.7%
Barr1983	2.0%
Dai2018	2.2%
Day1993	3.4%
Fomi2010	4.0%
O'Con1987	5.0%
Seal1989	2.5%
Whit1974	3.0%

- λ_k's are hyperparameters

On the test data set, dropout p=0.01