Refining the Syllabus:

Can Physically Motivated Descriptors Improve Predictions

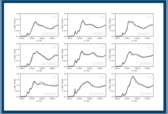
Thomas Pope

June 13, 2024

Newcastle University

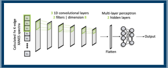
What is XANESNET?

Structure to Spectrum

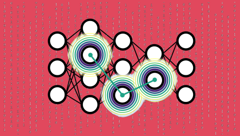


C Middleton et al. PCCP 25:13325 (2023)
L Watson et al. PCCP 24:9156 (2022)
C Rankine et al. JCP. 156:164102 (2022)
C Rankine et al. JPCA. 124:4263 (2020)

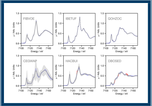
Spectrum to Structure



T David et al. Digital Discoveries 2:1461 (2023)

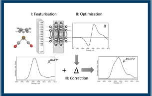


XANESNET



Quantifying Uncertainty

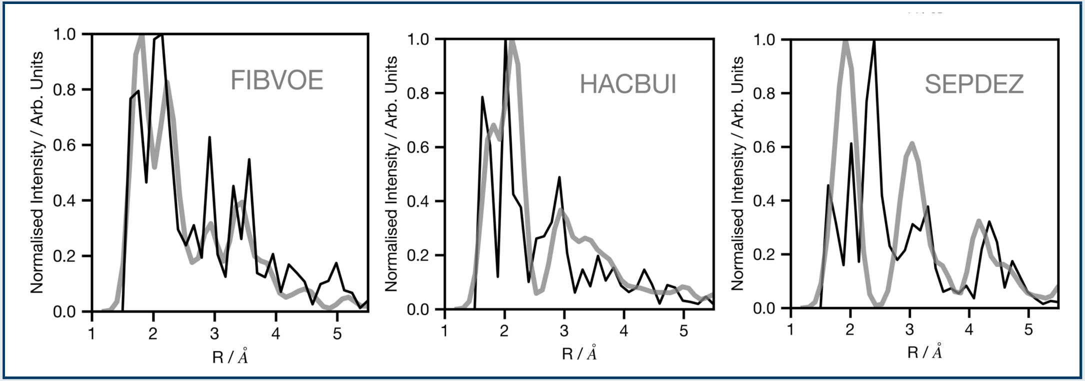
S Verma et al. ChemComm 59:7100 (2023)



Δ-Learning

L Watson et al. Structural Dynamics 10:064101 (2023)

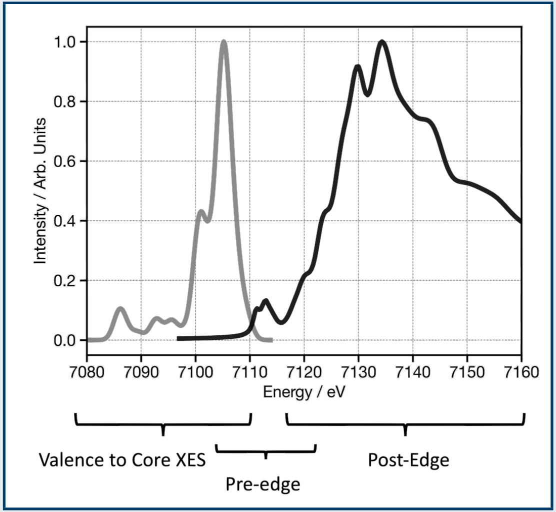
Shapley analysis: The Physics is Important



- Gray: G^2 wACSF feature importance from the Shapley analysis
- Black: Curved waved amplitude of 2 body multiple scattering pathways extracted from FEFF simulations

[1] T Penfold *et al* (2023) Mach. Learn.: Sci. Technol. 5, 021001

Physics in the Spectra



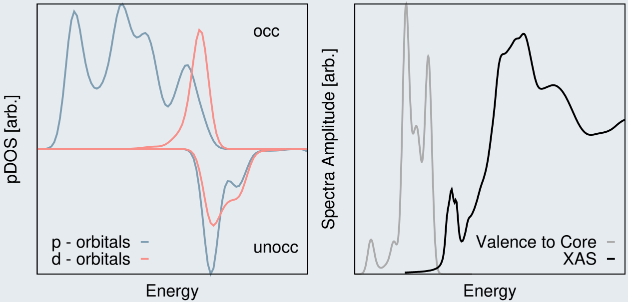
- Post-Edge:
  - Core electron is excited into the continuum
  - Spectrum contain information about the geometric structure around the absorbing atom
- Pre-Edge:
  - Core electron is excited into an unnoccupied orbital
  - Spectrum contains information about the electronic structure of the molecule
- Valence to Core:
  - Core electron is gone and a valence electron drops down to take its place
  - Spectrum contains information about the electronic structure of the molecule

Matching Input Descriptors

- Geometric Information
  - wACSF
- Electronic Information [1]
  - The probability of electronic transition is given by Fermi's golden rule:

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar} |\langle f|H|i\rangle|^2 \, \delta\left(E_i - E_f + \omega\right)$$

  - The 1s orbital of the absorbing atom is involved, so we only need info on the other orbitals
  - Partial density of states (pDOS) of occupied and unoccupied orbits is generated with GFN-xTB [2]

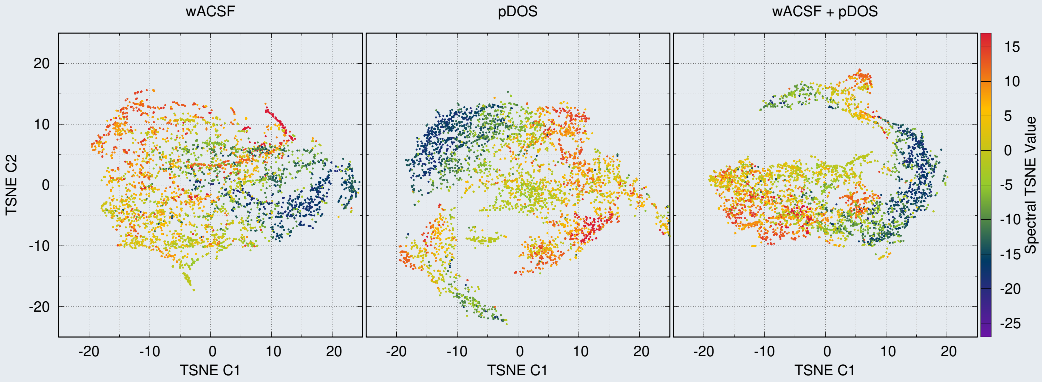


[1] C. Middleton *et al* (2024) ChemRxiv

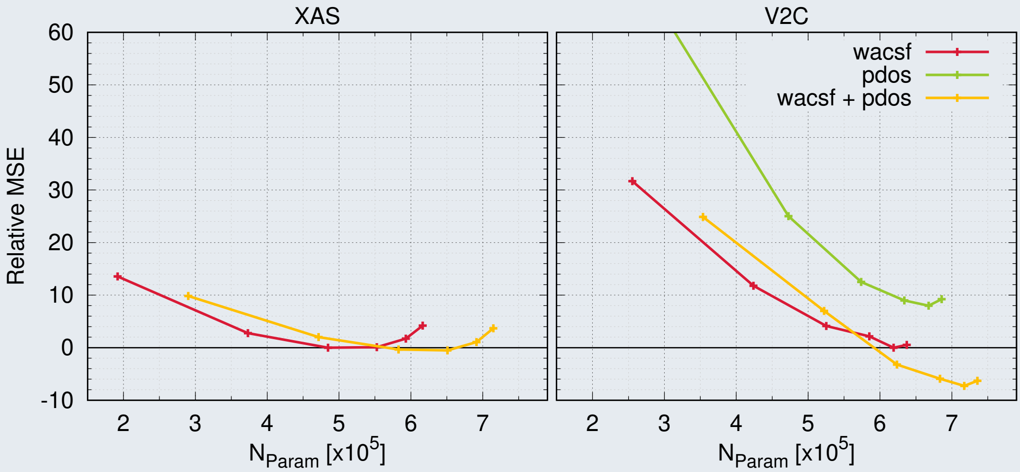
[2] Grimme *et al* (2017) Journal of chemical theory and computation, 13(5), 1989-2009.

Does it work?

Valence to Core TSNE

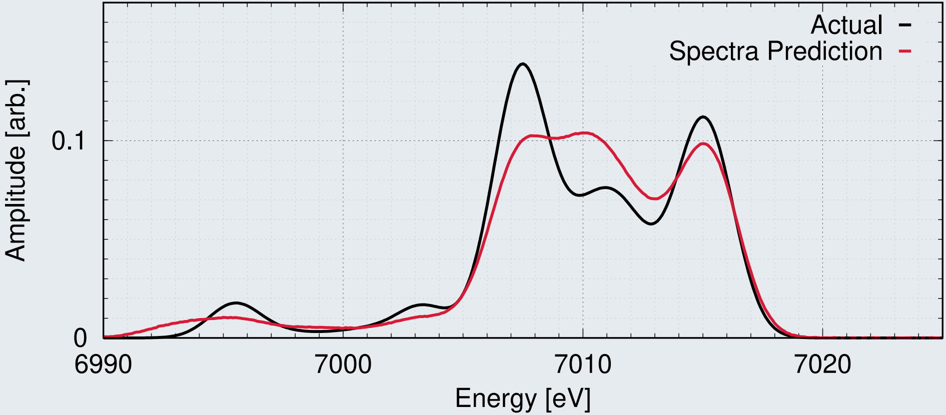


Predictions with and without the pDOS modification



Can we do better?

Cosine Transform Concatenation



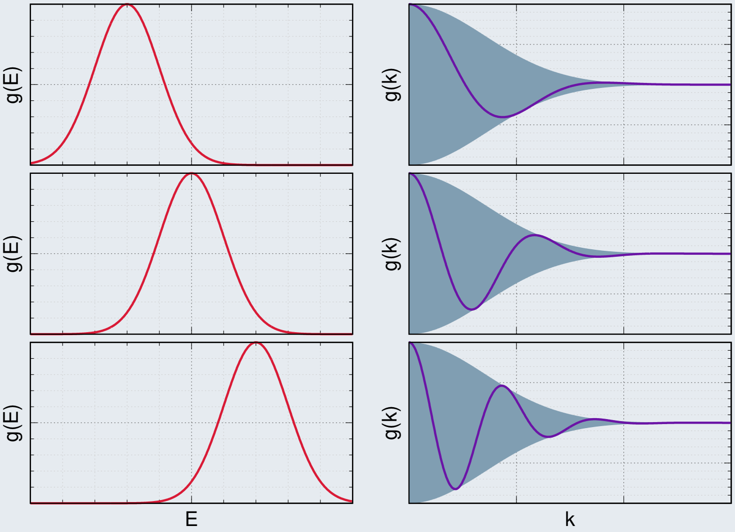
Cosine Transform Concatenation

Gaussian Function:

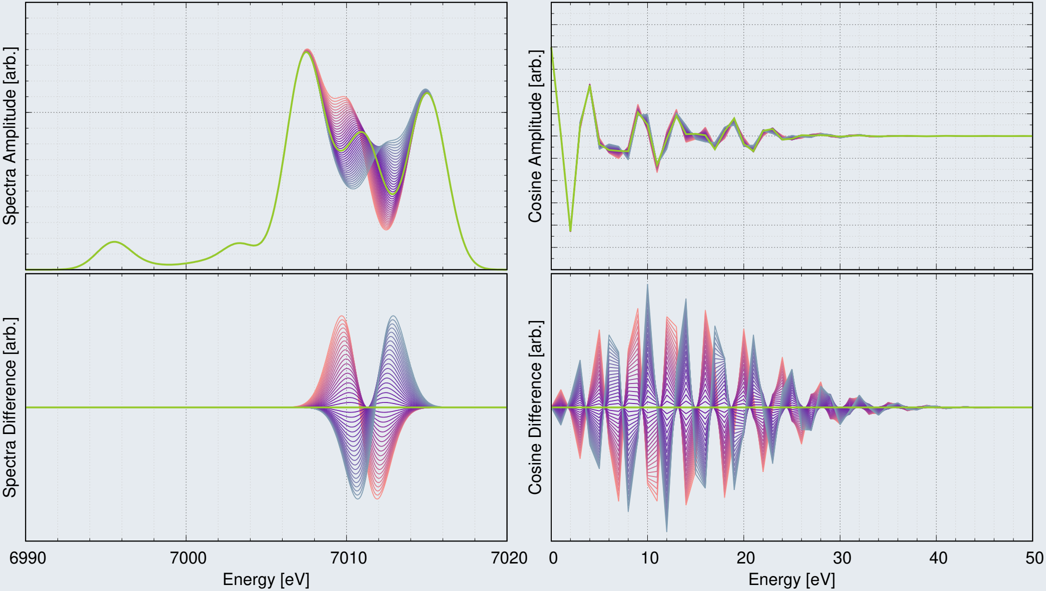
$$g(E) = e^{-(E-E_0)^2/2\sigma^2}$$

Cosine Transform:

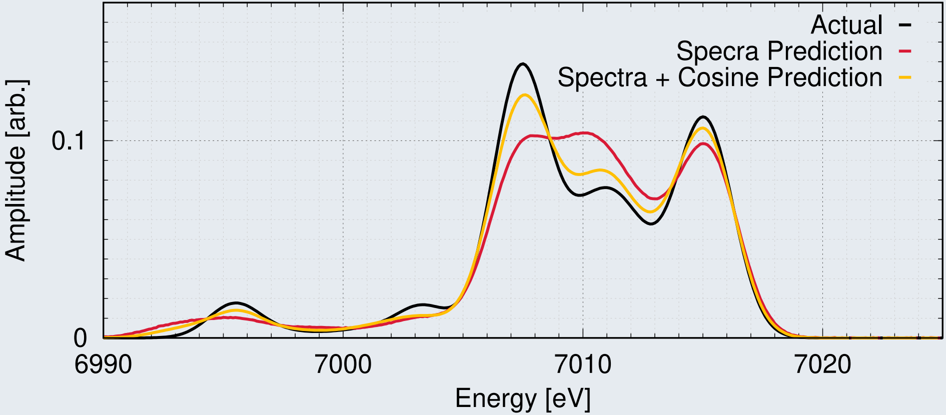
$$g(k) = \sigma e^{-(\sigma k)^2} \cos(E_0 k)$$



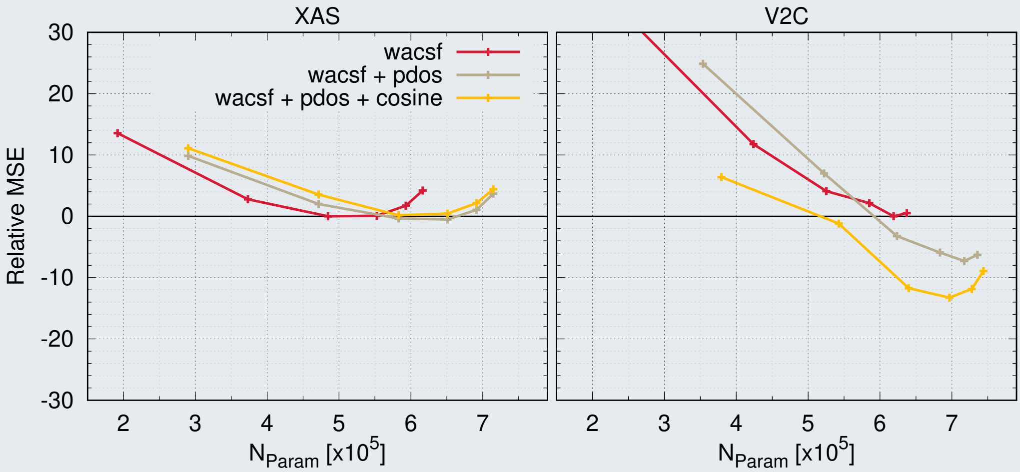
Cosine Transform Concatenation



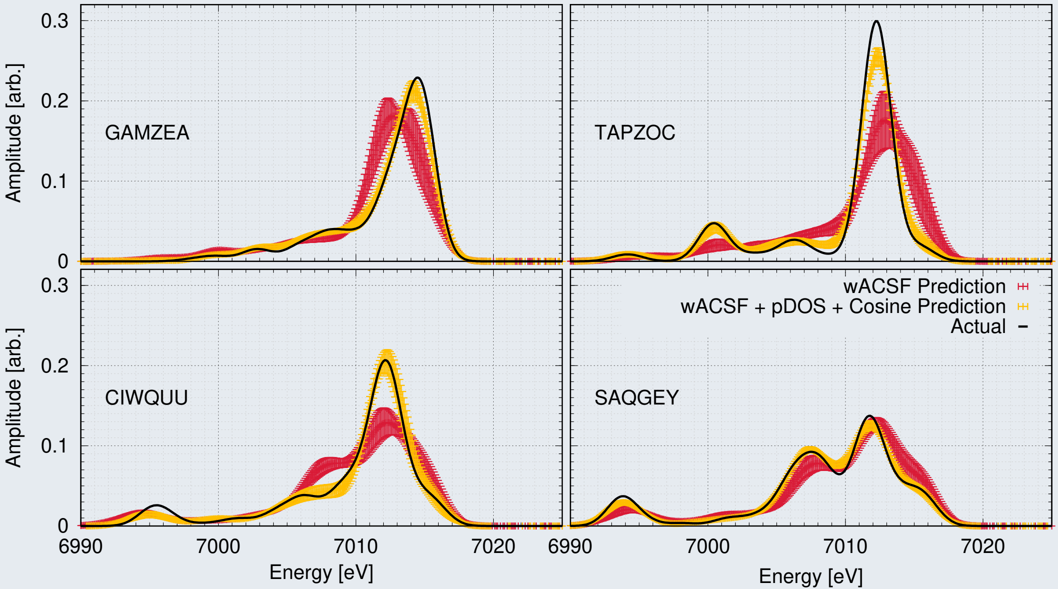
Does it work?


Actual
Specra Prediction
Spectra + Cosine Prediction
Amplitude [arb.]
0.1
0
6990
7000
7010
7020
Energy [eV]

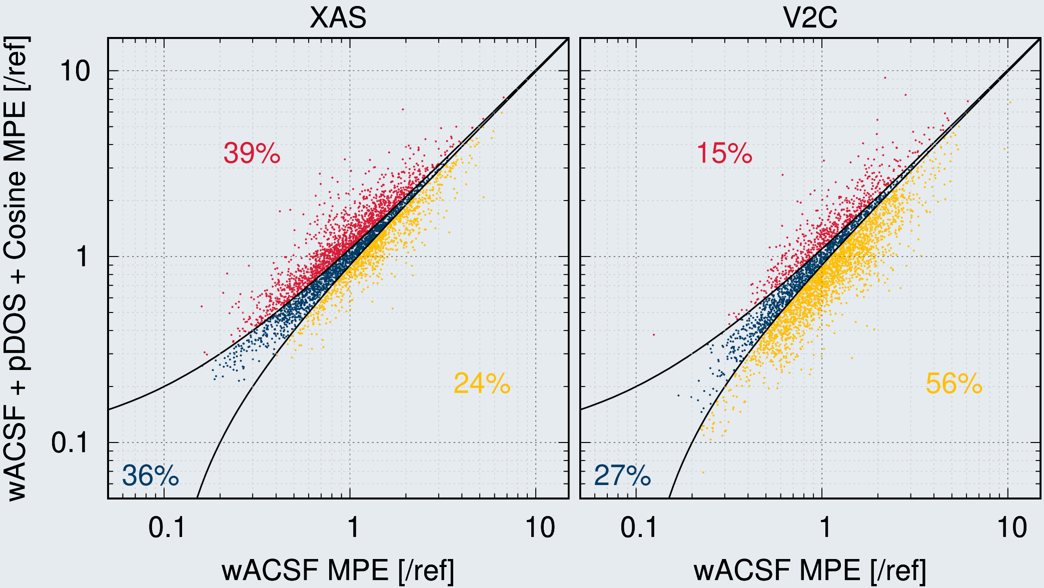
Predictions with and without the pDOS and cosine modifications



Predictions with and without the pDOS and cosine modifications



Predictions with and without the pDOS and cosine modifications



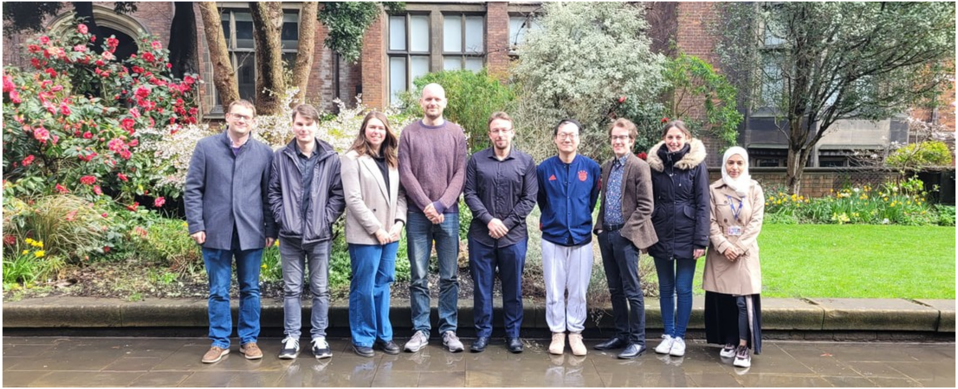
Conclusions:

Adding the right physics can improve predictions

But, adding irrelevant physics does not help and can hinder predictions

Including extra information in the predicted vectors can also improve predictions

Aknowledgements



Code: gitlab.com/team-xnet/xanesnet
Datasets: gitlab.com/team-xnet/training-sets
Tutorials: gitlab.com/team-xnet/tutorials
Online Interface: Coming Soon

EPSRC
Engineering and Physical Sciences Research Council

LEVERHULME TRUST

Thank you

Any Questions?