Neural Thermal Scattering (NeTS) Modules for Graphite & Beryllium

J.P.W. Crozier, A.I. Hawari

LEIP Laboratories
Department of Nuclear Engineering
North Carolina State University
Raleigh, North Carolina, USA

CSEWG Meeting
November 1st, 2022 BNL

NC STATE UNIVERSITY

Overview

A. Motivations
  - Compact formulation for TSL data
    - Extend AI knowledge
      - Advanced reactor simulation framework

B. Accounting for Beryllium and Graphite Complexity: Atomistic, Dynamical and Neural

C. New Material-informed Neural Thermal Scattering (NeTS) Modules

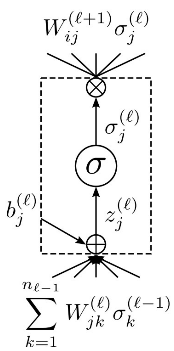
D. Implications and Conclusions

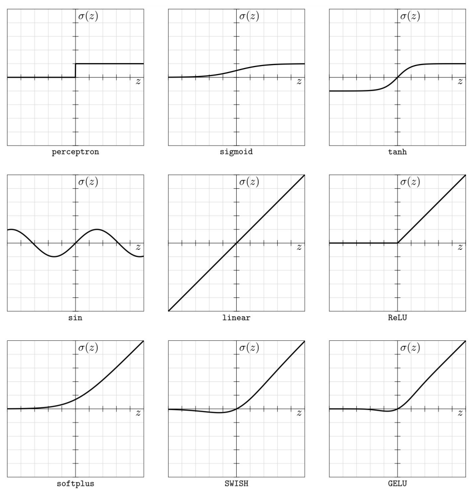
LEIP LABORATORIES

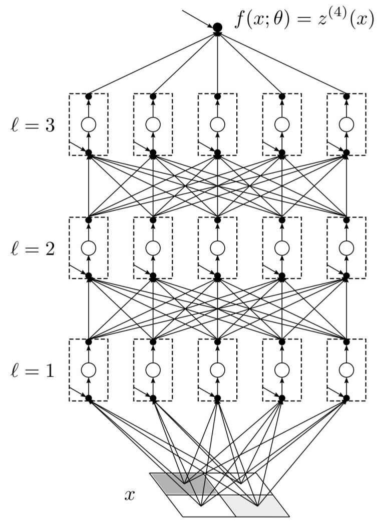
NC STATE UNIVERSITY

Deep Learning and Artificial Neural Networks

A mathematical abstraction inspired by biological neurons







$$S_s(\alpha,\beta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\beta t} e^{-\gamma(t)}\,dt$$

$$\gamma(t) = \alpha\int_{-\infty}^{\infty} \frac{\rho(\beta)}{2\beta\sinh\left(\frac{\beta}{2}\right)}\left[1 - e^{-i\beta t}\right] e^{-\frac{\beta}{2}}\,d\beta$$

$$\alpha = \frac{E' - E - 2\mu\sqrt{E'E}}{AkT}$$

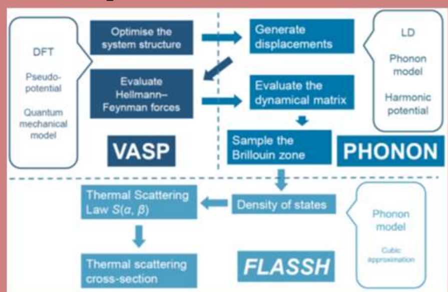
$$\beta = \frac{E' - E}{kT}$$

ENDF/B-VIII Material	File 7	Temps
Crystalline Graphite	9 MB	10
Beryllium Metal	3.7 MB	8
NeTS	< 200 KB	Continuous

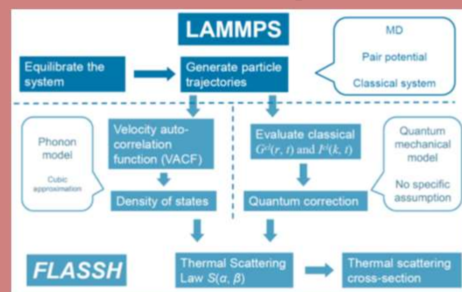
LEIP LABORATORIES

Existing and Emerging Frameworks

Density Functional Theory



Molecular Dynamics

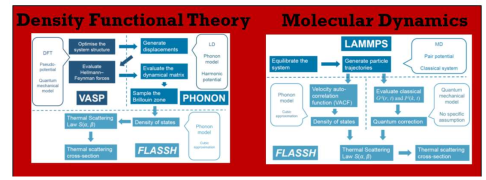



```
                                                                    1 0  0
1.480000+2 2.360058+2         -1          0          0          0  48 1451
0.000000+0 0.000000+0          0          0          0          6  48 1451
1.000000+0 5.000000+0          0          0         12          8  48 1451
0.000000+0 0.000000+0          0          0         33          3  48 1451
U in UC    LEIP LAB    EVAL-Jun21 J.P.W. Crozier, A.I. Hawari         48 1451
                       DIST-                                          48 1451
----ENDF/B-VIII        MATERIAL  48                                   48 1451
-----THERMAL NEUTRON SCATTERING                                       48 1451
------ENDF-6 FORMAT                                                   48 1451
                                                                      48 1451
 Temperatures = 296 400 500 600 700 800 1000 1200 1600 2000 K         48 1451
                                                                      48 1451
HISTORY                                                               48 1451
-------                                                               48 1451
This library was produced by the Low Energy Interaction Physics       48 1451
(LEIP) group at North Carolina State University, USA. The             48 1451
thermal scattering law data for Uranium in Uranium Carbide was        48 1451
developed using ab initio lattice dynamics (AILD)[1,2]. Ten           48 1451
temperatures are available in this library. The Full Law              48 1451
Analysis Scattering System Hub (FLASSH) system was used to            48 1451
produce File 7 MT = 2, 4 data for Uranium in UC [3]. MAT=48 and       48 1451
ZA=148 are used for U in UC.                                          48 1451
                                                                      48 1451
REFERENCES                                                            48 1451
```

ENDF File 7

Current TSL Evaluation Methodology

1. TSL Calculation

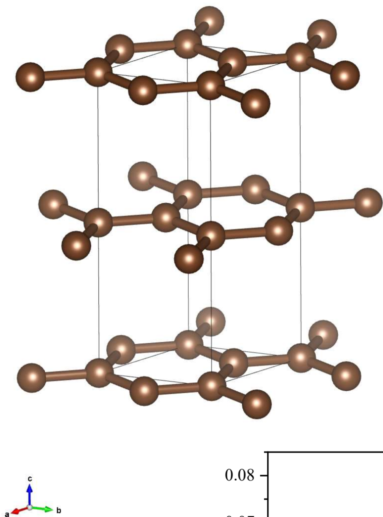


2. TSL Preprocessing → Train/Val/Test Data
3. Neural Network + Data → Trained W_{ij} / b_j
4. Network / Data Evaluation
5. A Priori $S(\alpha,\beta,T)$

Architecture, Weights, Biases

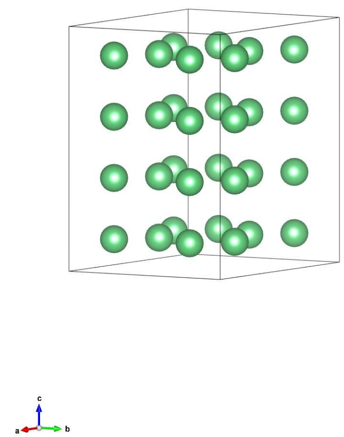
Emerging TSL Framework

Ab Initio Simulations and Lattice Dynamics



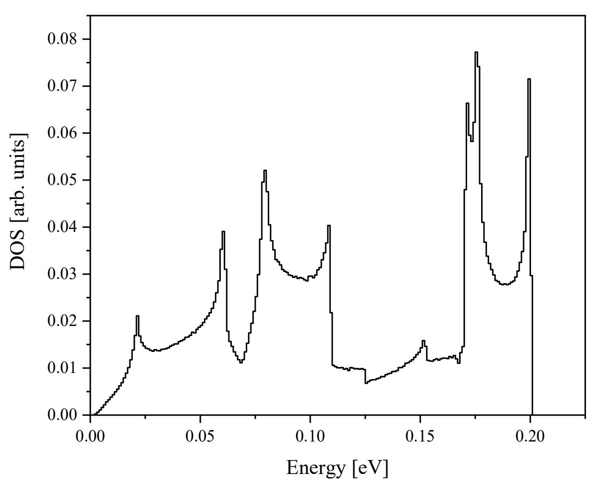
Crystalline Graphite

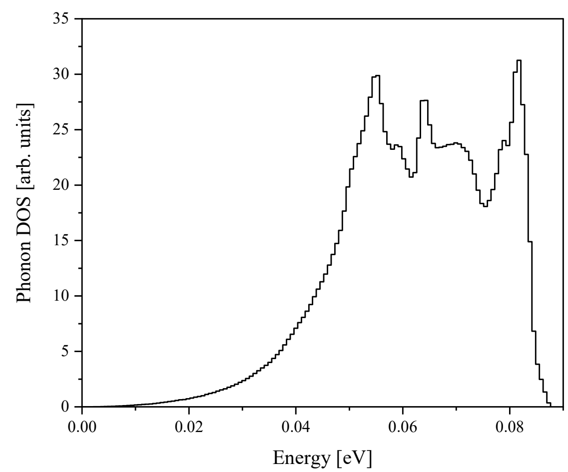
- Trigonal Planar, s.g. 186
- $1s^2 2s^2 2p^2$
- VASP: GGA-PBE
- 4x4x4 Supercell
- VDW radius parametrization
- Highly asymmetric, anharmonic forces between layers



Beryllium Metal

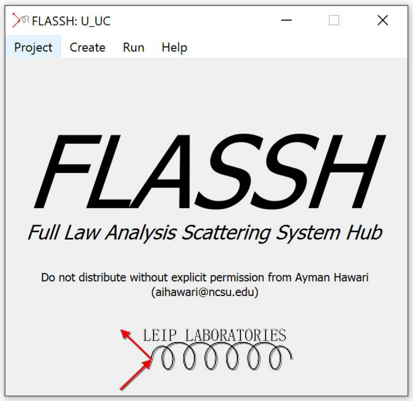
- HCP, s.g.192
- $1s^2 2s^2$
- VASP: PAW-GGA
- 5x5x5 Supercell
- High symmetry

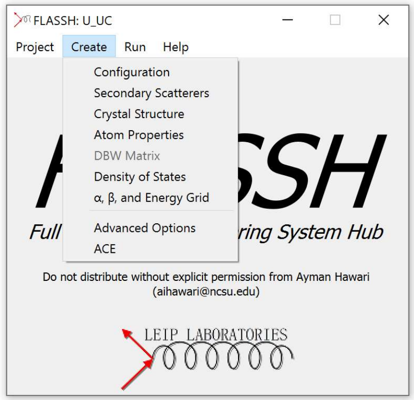


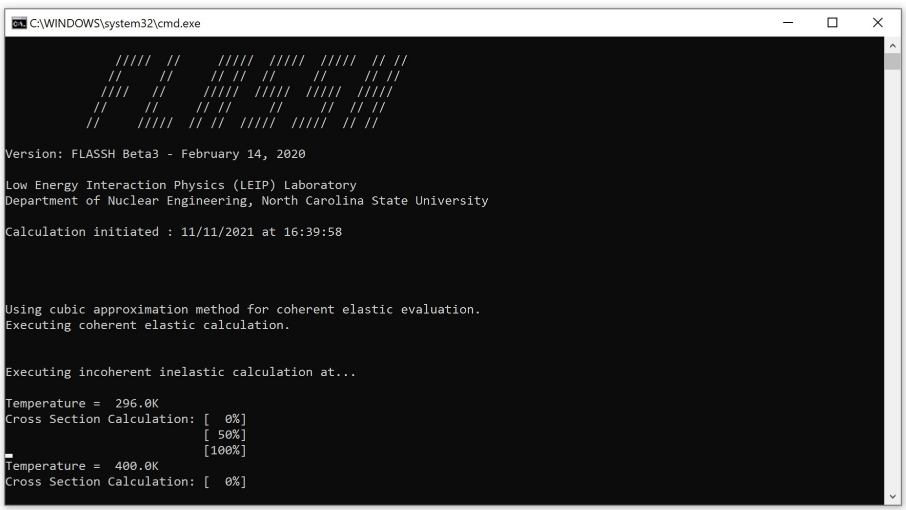


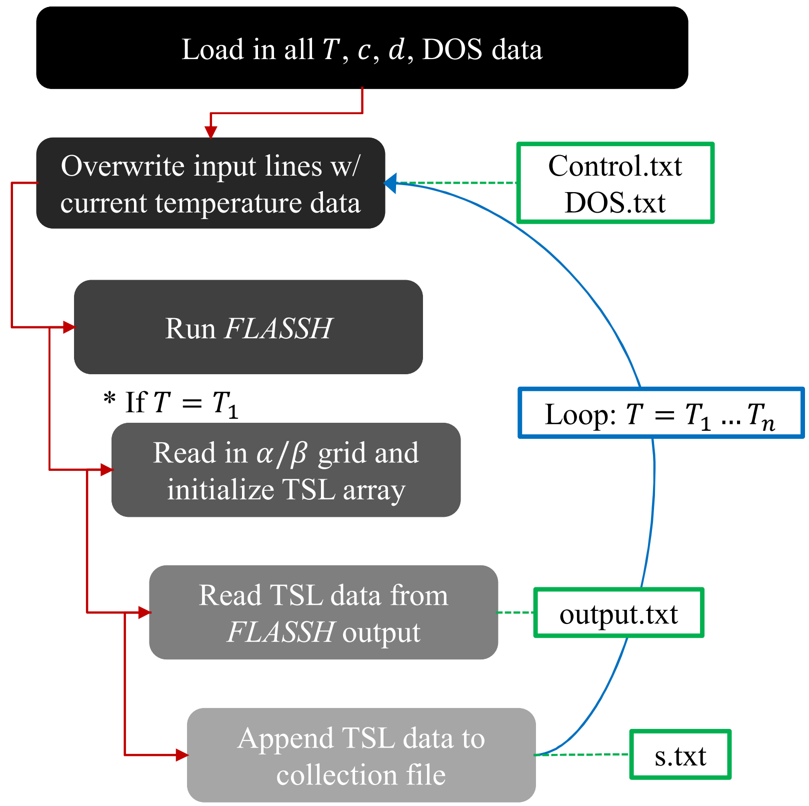
LEIP LABORATORIES

TSL Datasets via *FLASSH* Loop



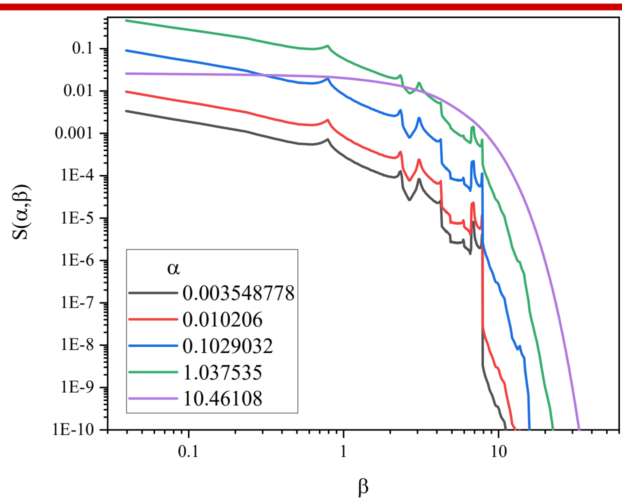


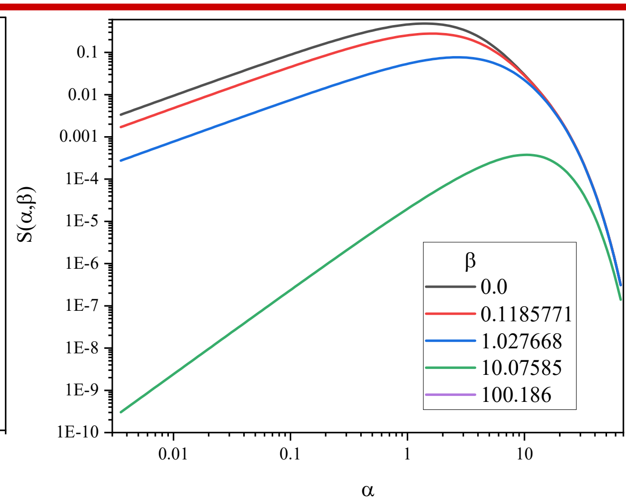




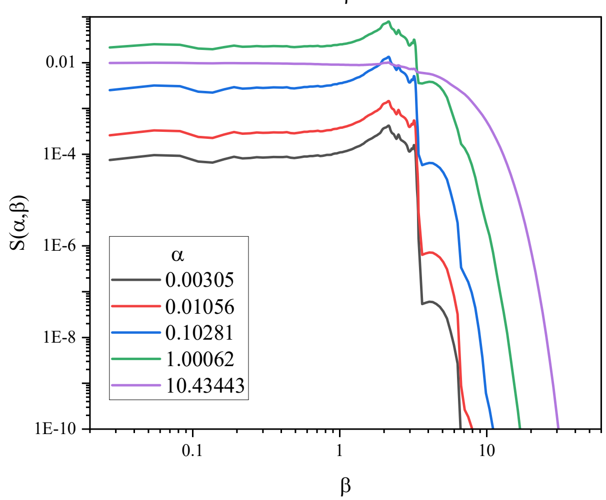
Thermal Scattering Law Data

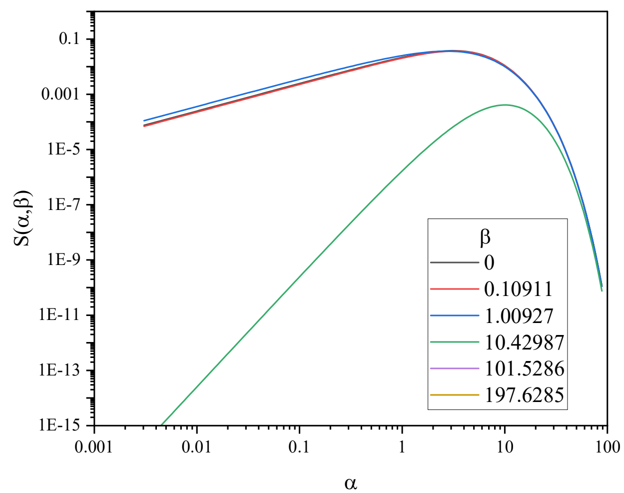
Graphite





Beryllium



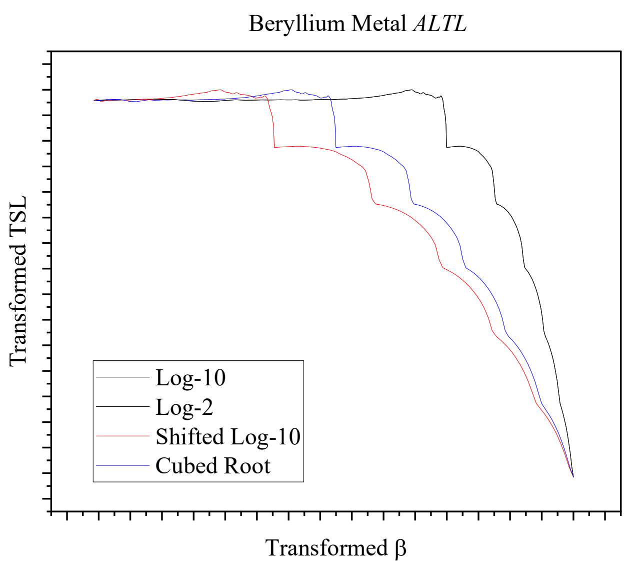


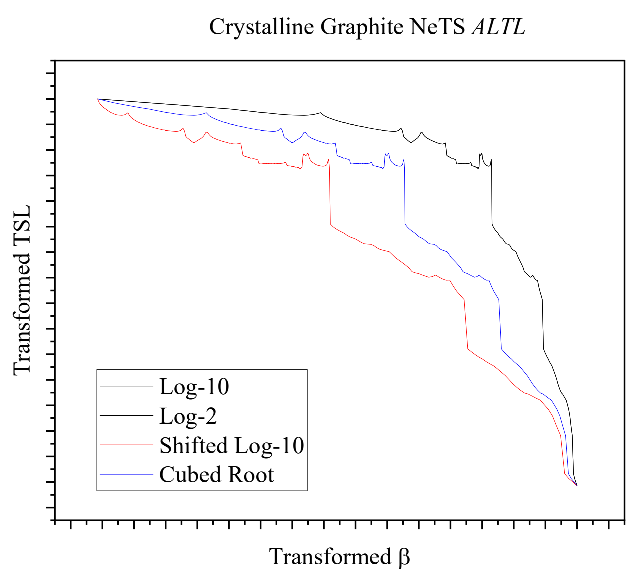
FLASSH Data Preprocessing

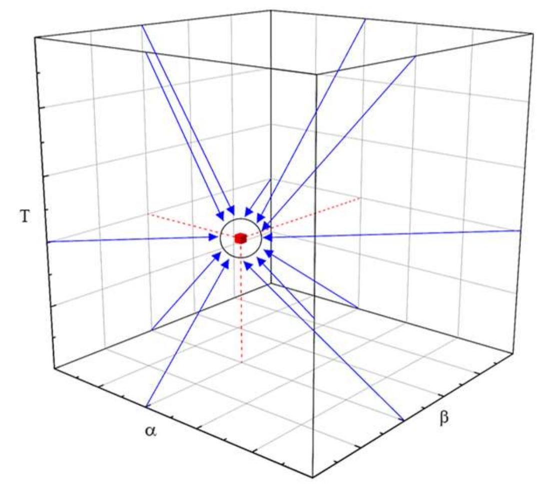
1. Transform alpha, beta, temperature, and TSL

2. Generate input "edges" and identify important beta features (peaks / troughs)

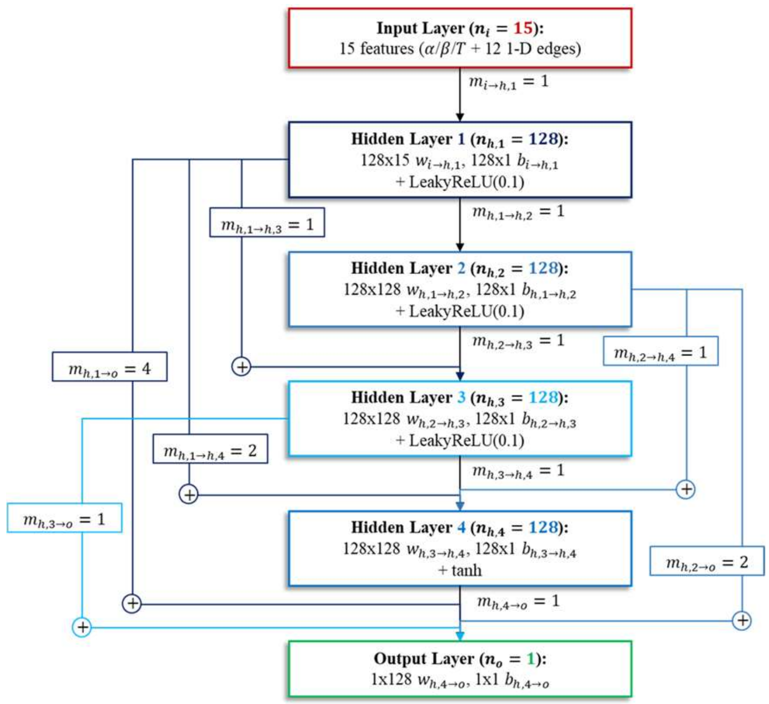
3. Dataset split into training, validation, test (99%,.5%,.5%)







Original NeTS Architecture



Setting	Selection
n_{hidden_layers}	4
n_h	128
cascading	none
skip connections	dense, weighted
skip weight scheme	multiplicative, 1:2:4
activation function	LeakyReLU(0.1) in hidden layers 1-3, tanh in hidden layer 4
weight initialization	Xavier uniform
optimizer	AdamW
regularization (explicit)	Weight decay ($\lambda \to$ 1e-6, as implemented in AdamW)
batch size	1024
training epochs	400
learning rate schedule	triangular, cyclic, w/ exponentially decaying amplitude (see parameters in Table 4.3)
loss function	MSE

C.A. Manring, A.I. Hawari, "Design of a Neural Thermal Scattering (NeTS) Module for Hydrogen in Light Water", PHYSOR 2022

LEIP LABORATORIES

Is this architecture excessive? Computationally unfeasible for many systems? Prohibits fast network usage?

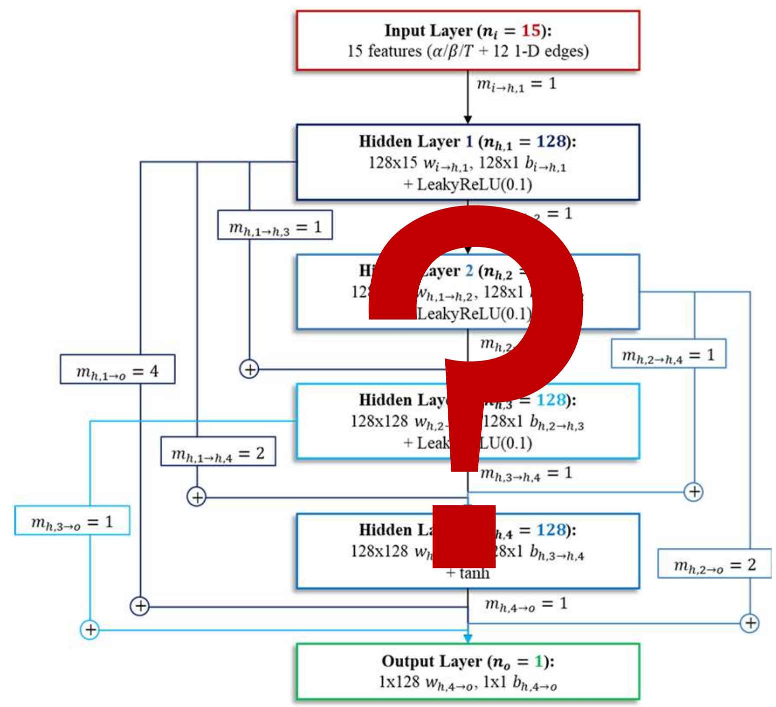
NC STATE UNIVERSITY

Optimizing NeTS Architecture

Potential areas for reduction / improvement:

- Number of hidden layers
- Neurons per hidden layer
- Cascading neural configuration
- Skip connections
- Activation functions
- Weight initialization
- Weight decay regularization
- Batch size / Learning Rate

Iterative tests done with 2D TSL data at 296 K



Objectives: reduce network size, training time, and computational footprint

LEIP LABORATORIES

NC STATE UNIVERSITY

Initial NeTS Search Space

Dataset Preparation:						
Materials (TSL dataset)	Beryllium Metal	Crystalline Graphite				
Input Transformations (for β)	$Log_{10}(\beta)$	$Log_{2}(\beta)$	$Log_{2}(\beta + \beta_{Max\ DOS})$	$\beta^{1/3}$	$\beta^{1/2}$	
Architecture:						
Width (Neurons / hidden layer)	512	256	128	64	32	16
Depth (Number of layers)	2	4	6	8	12	16
Skip Connections (Skip Weights)	None	Uniform	Cascading			
Hidden/final layer Activation	Tanh	ReLU	LeakyReLU(.1)	LeakyReLU(.2)		
Training Parameters:						
Learning Rate	1e-6	1e-5	1e-4	1e-3	1e-2	
Batch-size	4096	2048	1024	512	256	128
Epochs	100	400	800	1600		

Total Possibilities: **518,400**

Engineering best judgement → identify plausible combinations to train with 2D data at 296 K

LEIP LABORATORIES

Initial Architecture Search Highlights

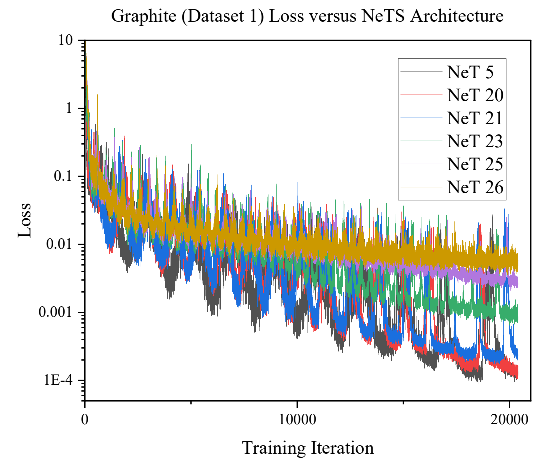
NeTS :	5	20	21	23	25	26
W_{ij} / b_i [KB]	1296	462	190	75	47	28
Input Dimensions :	7	7	7	7	7	7
neurons / layer 1 :	256	128	64	32	24	16
neurons / layer 2 :	256	128	64	32	24	16
neurons / layer 3 :	256	128	64	32	24	16
neurons / layer 4 :	256	128	64	32	24	16
neurons / layer 5 :	256	128	64	32	24	16
neurons / layer 6 :	256	128	64	32	24	16
neurons / layer 7 :		128	64	32	24	16
neurons / layer 8 :		128	64	32	24	16
neurons / layer 9 :			64	32	24	16
neurons / layer 10 :			64	32	24	16
neurons / layer 11 :			64	32	24	16
neurons / layer 12 :			64	32	24	16
neurons / layer 13:				32	24	16
neurons / layer 14 :				32	24	16
neurons / layer 15 :				32	24	16
neurons / layer 16 :				32	24	16
Output :	1	1	1	1	1	1

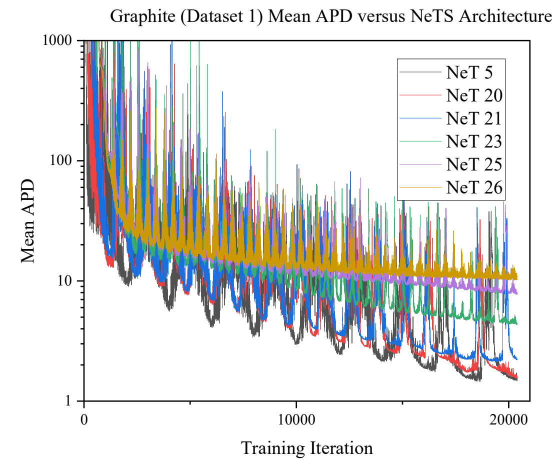
Initial Architecture Search

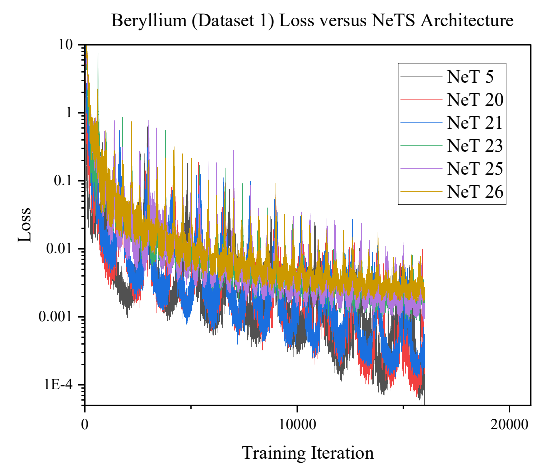
$$Loss_{MSE}(\, z(x_{\delta},\theta), y_{\delta} \,) = \frac{1}{2}\, [z(x_{\delta},\theta) - y_{\delta}\,]^2$$

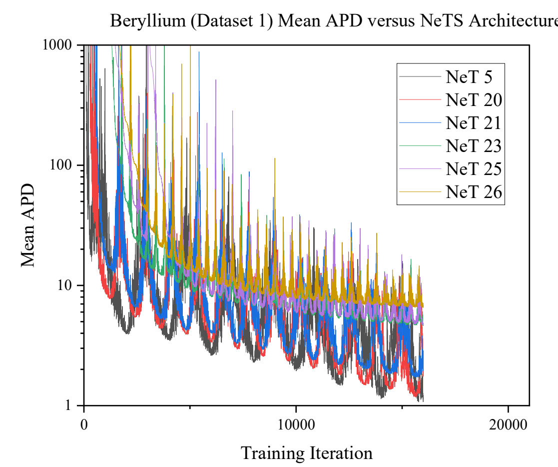
	5	20	21	23	25	26
	1296	462	190	75	47	28
	7	7	7	7	7	7
	256	128	64	32	24	16
	256	128	64	32	24	16
	256	128	64	32	24	16
	256	128	64	32	24	16
	256	128	64	32	24	16
	256	128	64	32	24	16
		128	64	32	24	16
		128	64	32	24	16
			64	32	24	16
			64	32	24	16
			64	32	24	16
			64	32	24	16
				32	24	16
				32	24	16
				32	24	16
				32	24	16
	1	1	1	1	1	1

$$APD = \frac{|y_{\delta} - \hat{y}\,|}{y_{\delta}} * 100$$









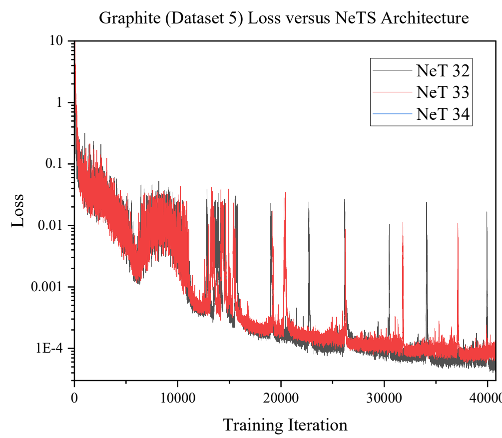
NC STATE
UNIVERSITY

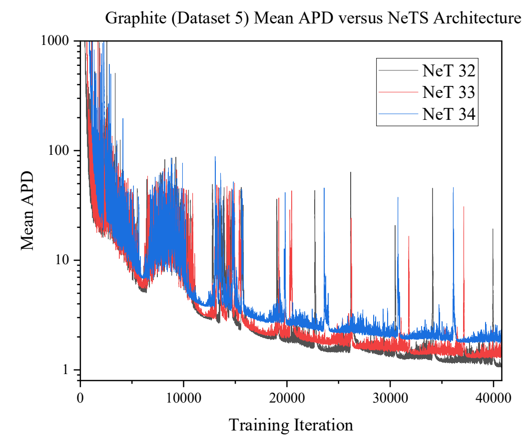
Further Architecture Search

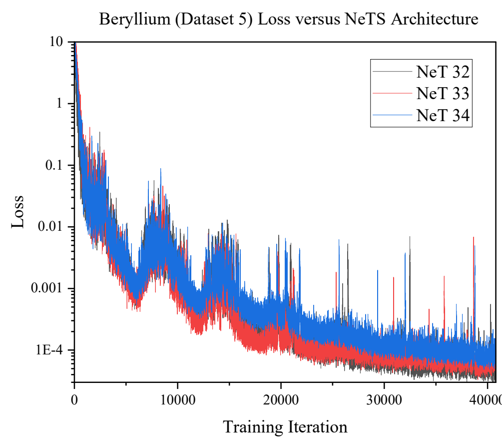
NeTS :	32	33	34
W_{ij} / b_i [KB]	190	147	111
Input Dimensions :	7	7	7
neurons / layer 1 :	64	56	48
neurons / layer 2 :	64	56	48
neurons / layer 3 …	64	56	48
neurons / layer 11 :	64	56	48
neurons / layer 12 :	64	56	48
Output :	1	1	1

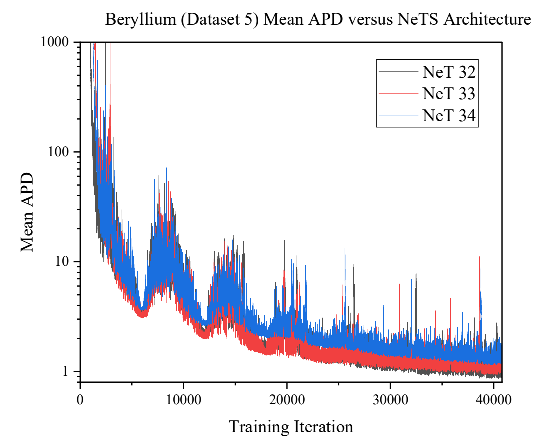
NeTS-2D '32' (64 x 12, 190 kb) → GRAPHITE		
	Mean APD	Median APD
train	1.0950	0.7154
valid	1.1620	0.7714
test	1.1882	0.7773

NeTS-2D '33' (56 x 12, 147 kb) → BERYLLIUM		
	Mean APD	Median APD
train	1.1574	0.7735
valid	1.2405	0.7926
test	1.4313	0.8506







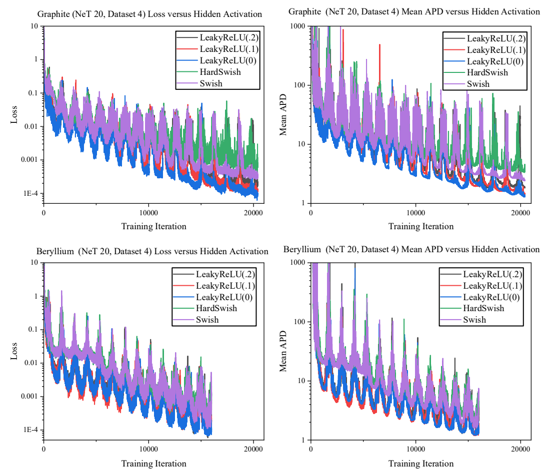


LEIP LABORATORIES

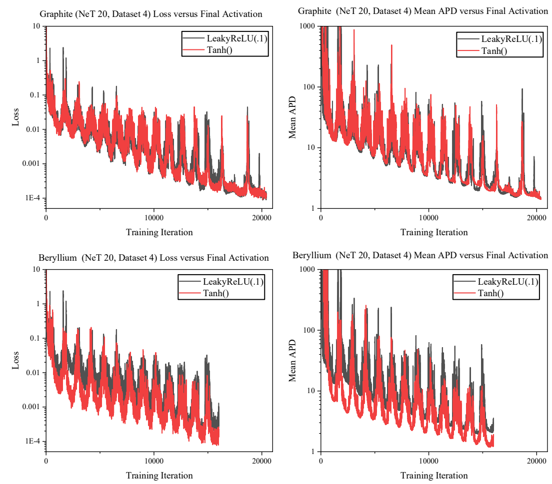
NC STATE UNIVERSITY

Hyper-parameter Search

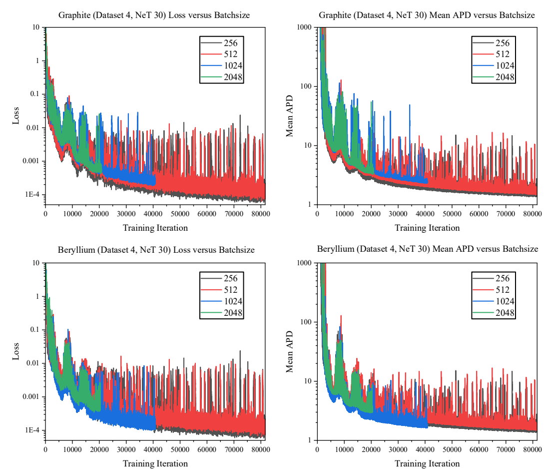
Hidden Layer Activation Function



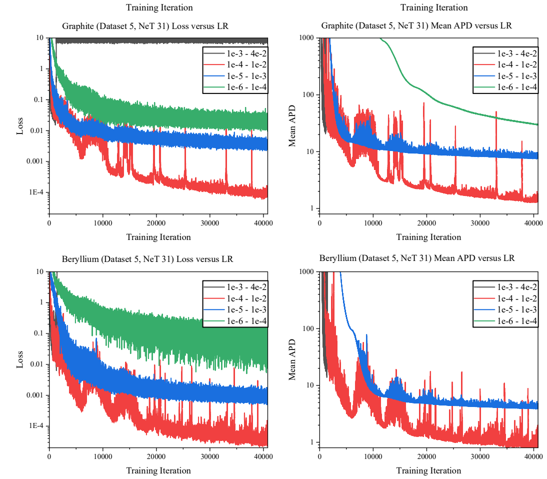
Final Layer Activation Function



Batch size



Learning Rate

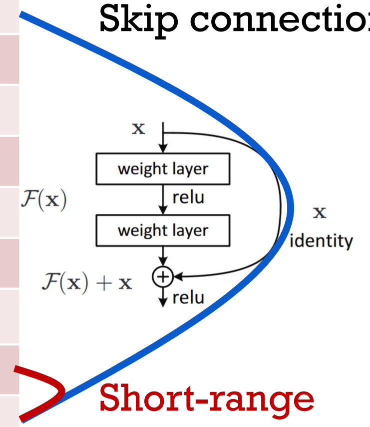


LEIP LABORATORIES

Final Architecture Search Results

NeTS :	70 – NeTS-Graphite	71 – NeTS-Beryllium
W_{ij} / b_i [KB]	**178**	**143**
Input Dimensions :	7	7
neurons / layer 1 :	96	96
neurons / layer 2 :	64	56
neurons / layer 3 :	64	56
neurons / layer 4 :	64	56
neurons / layer 5 :	64	56
neurons / layer 6 :	64	56
neurons / layer 7 :	64	56
neurons / layer 8 :	64	56
neurons / layer 9 :	64	56
neurons / layer 10 :	64	56
neurons / layer 11 :	32	32
neurons / layer 12 :	32	32
Output :	1	1

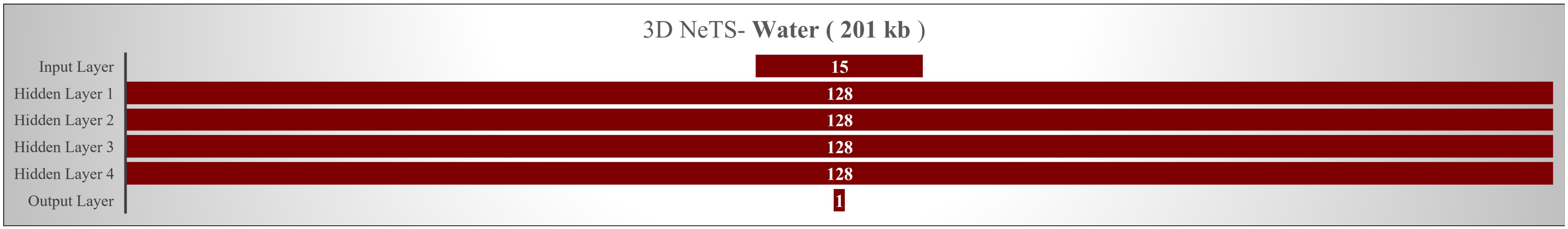
Long-range
Skip connection

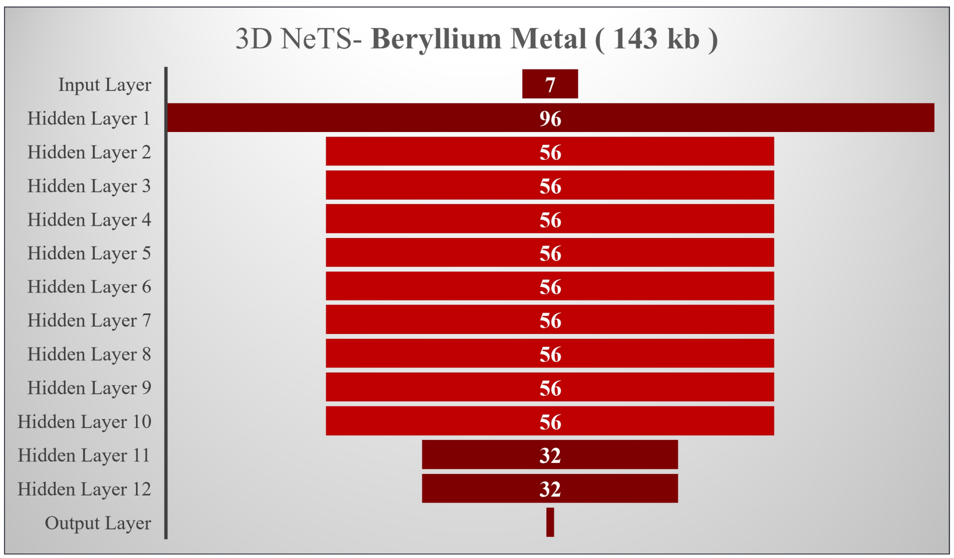


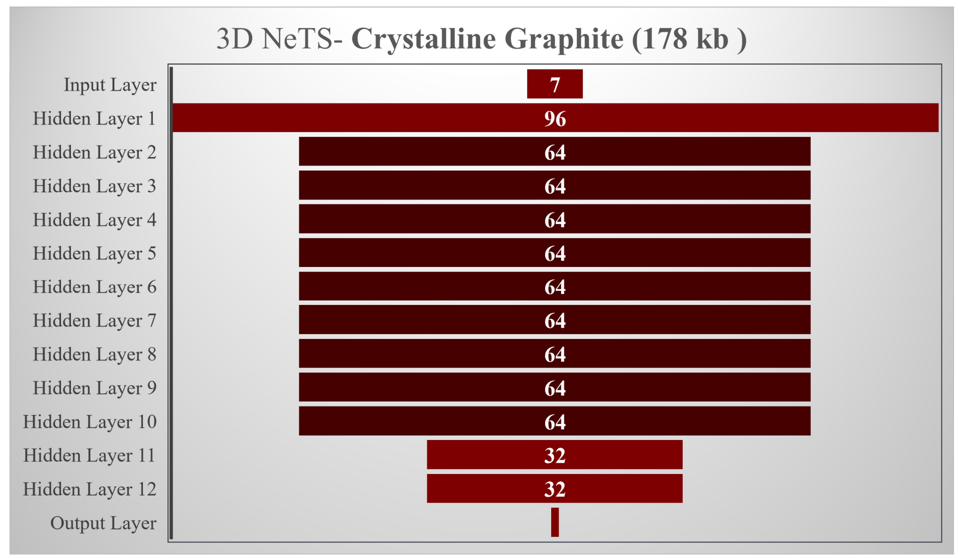
Short-range
Skip connection

ResNeT Original Paper: He K et al., Deep Residual Learning for Image Recognition, 2015

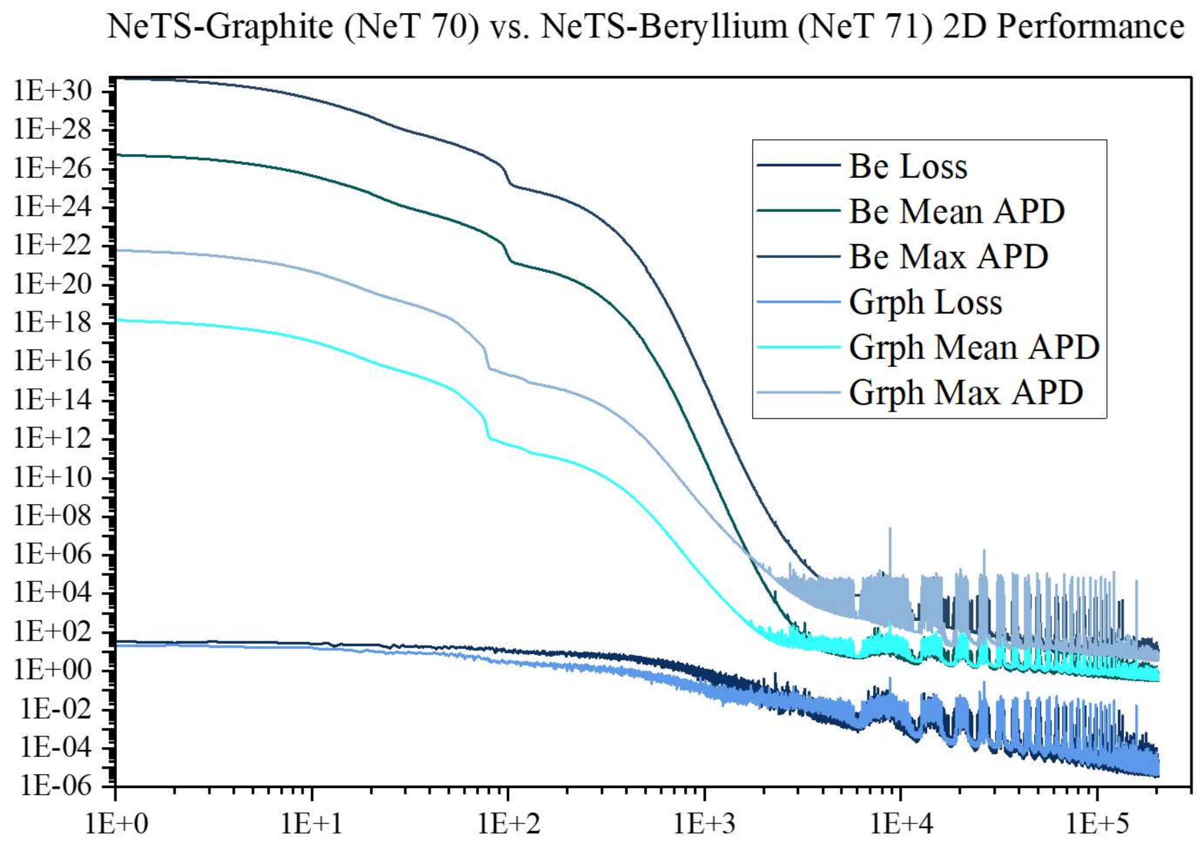
Architecture Visualization / Comparison







2D NeTS Training



NC STATE UNIVERSITY

2D NeTS Performance Statistics !

Crystalline Graphite, res(.4)NeTS-70 [96-64-32 x 12], 178 kb

Dataset	Mean APD [%]	Med. APD [%]	Max APD [%]	< 2%	< 1%
Train	**0.3700**	0.2603	**3.1055**	.9948	.9351
Validation	**0.4229**	0.3015	**4.1002**	.9908	.9160
Test	**0.4214**	0.3098	**3.0302**	.9941	.9160

Beryllium Metal, res(.4)NeTS-71 [96-56-32 x 12], 143 kb

Dataset	Mean APD [%]	Med. APD [%]	Max APD [%]	< 2%	< 1%
Train	**0.3066**	0.2086	**2.5763**	.9984	.9566
Validation	**0.4130**	0.2756	**3.8964**	.9849	.9129
Test	**0.4062**	0.2577	**2.9965**	.9892	.9118

LEIP LABORATORIES

3D NeTS Performance Statistics !

Crystalline Graphite, res(.4)NeTS-70 [96-64-32 x 12], 178 kb

Dataset	Mean APD [%]	Med. APD [%]	Max APD [%]	< 2% [%]	< 1% [%]
Train	0.1878	0.1464	2.9173	99.9979	99.5816
Validation	0.1792	0.1461	1.4321	1.00000	99.9695
Test	0.1818	0.1433	1.4687	1.00000	99.9085

Beryllium Metal, res(.4)NeTS-71 [96-56-32 x 12], 143 kb

Dataset	Mean APD [%]	Med. APD [%]	Max APD [%]	< 2% [%]	< 1% [%]
Train					
Validation					
Test					

NC STATE UNIVERSITY

LEIP LABORATORIES

New NeTS and Future Work

Feature	New NeTS-Beryllium	New NeTS-Graphite	Old NeTS	Improvement
β **input scaling**	$\beta^{\frac{1}{3}}$	$\beta^{\frac{1}{2}}$	$log_{10}\,\beta$	DOS linearization, include $\boldsymbol{\beta = 0}$
S(α, β, T) features	4 (altl,ahtl,bltl,bhtl)	4 (altl,ahtl,bltl,bhtl)	12 (all)	Reduced input redundancy
Network Architecture	[96,56,32] x 12, **143 kb** Cascading	[96,64,32] x 12, **178 kb** Cascading	128 x 4**, 201 kb** Uniform	Smaller, material-specific 10-30% size reduction
Total training time			~4-6 days	Comparable
Mean / Max APD			.1-.2 / 2-3 %	Comparable

1. Include one-phonon correction and train highly structured S(a,b,T) surface
2. Extend dimensionality: additional properties (i.e., porosity, burnup, pressure), new materials, extreme temperatures (cryogenic, advanced reactor conditions)
3. Couple NeTS to reactor physics framework

Thank You!