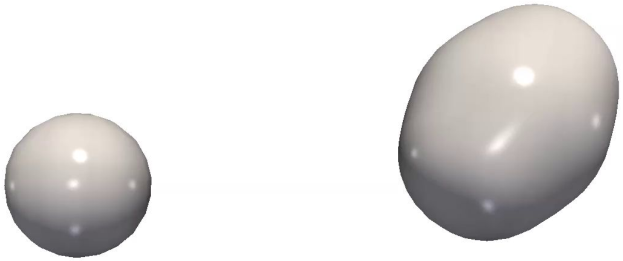
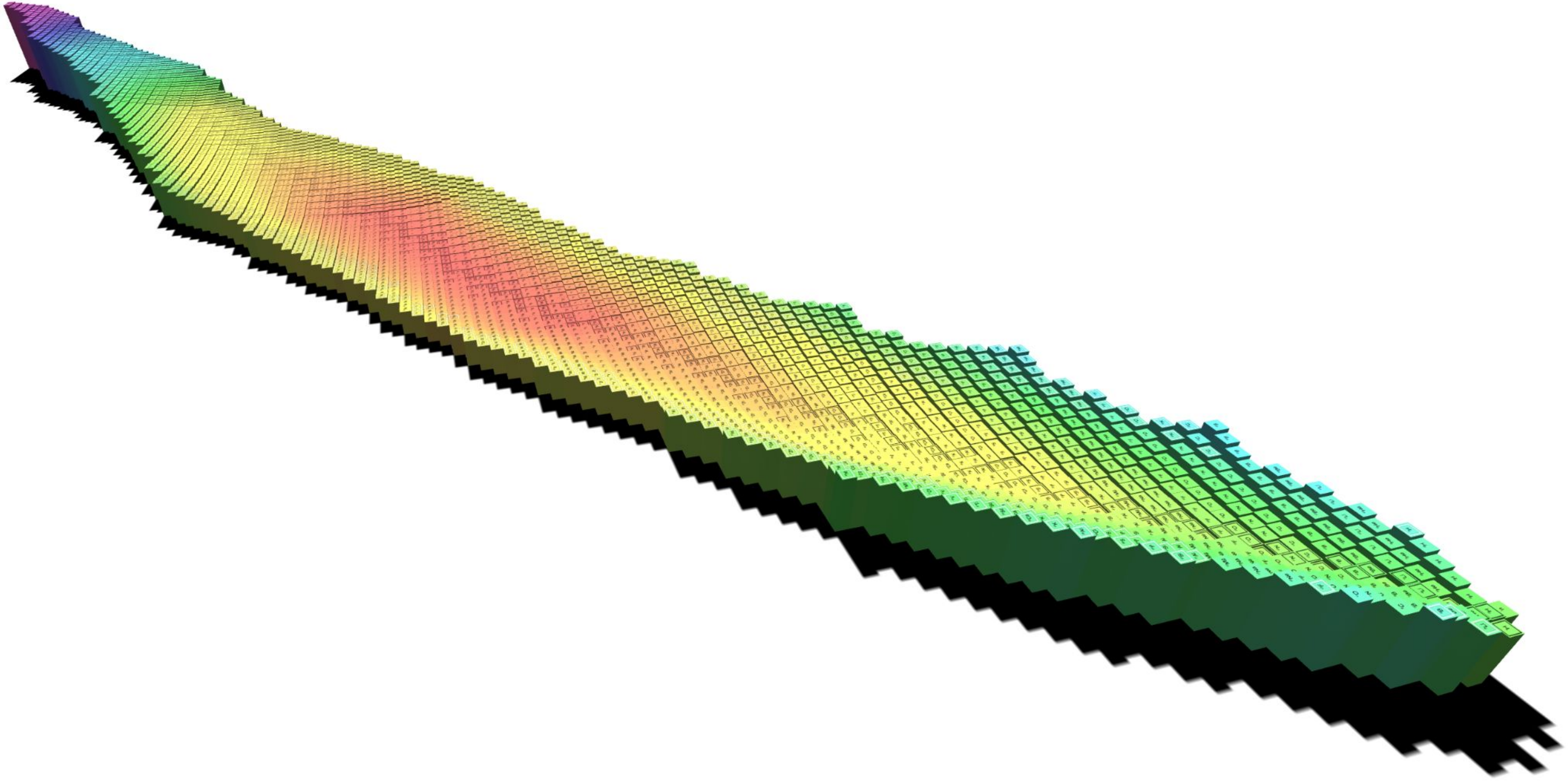
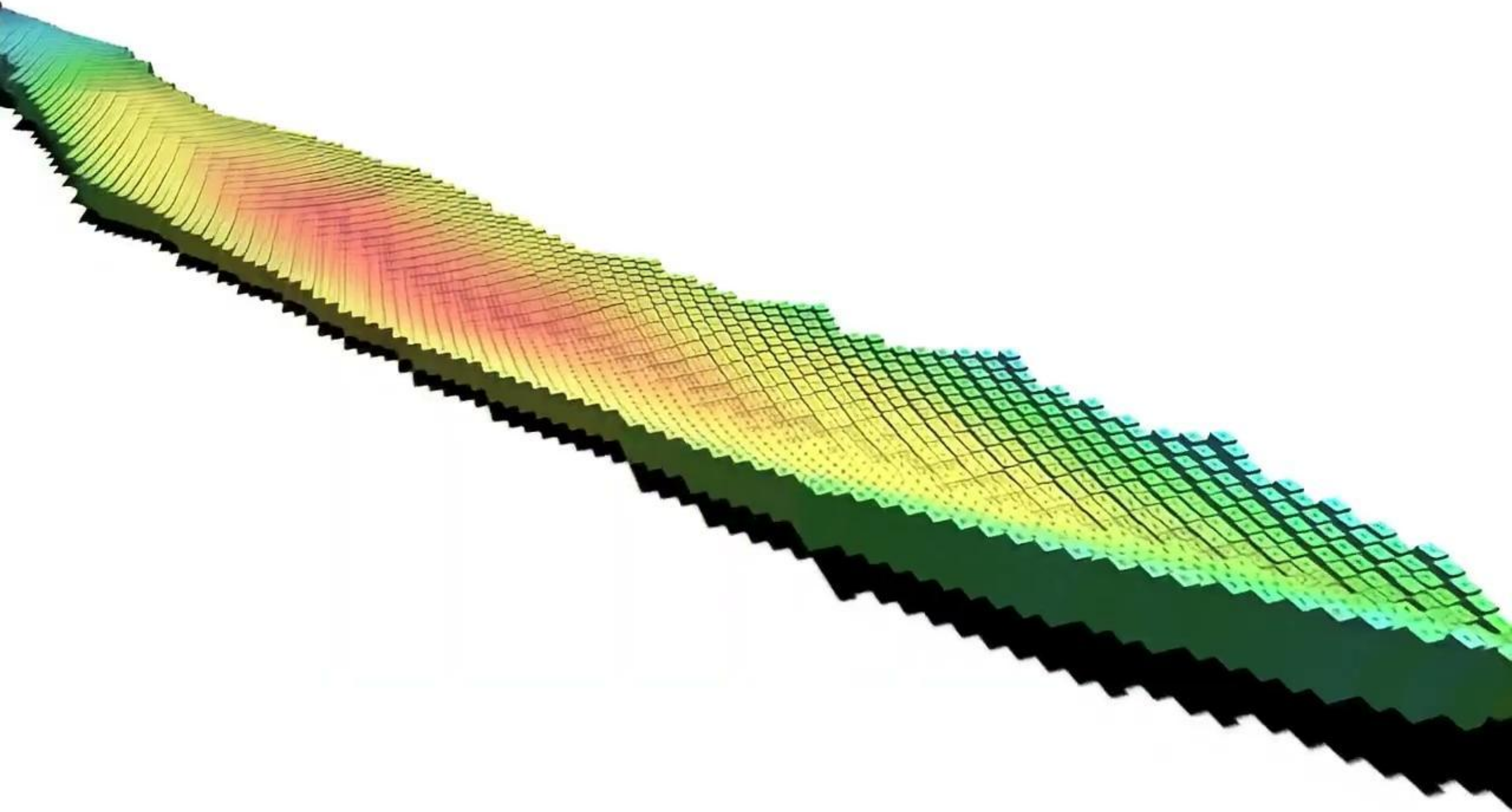
Efficient Bayesian Model Combination With pybmc

Kyle Godbey



Slides with videos: https://docs.google.com/presentation/d/1xRZo72xdc_uwMq-Jl6Gk9myKCNaJ4I2mLoLPHqvQcQo/edit?usp=sharing





What's Model Combination?

Model Combination

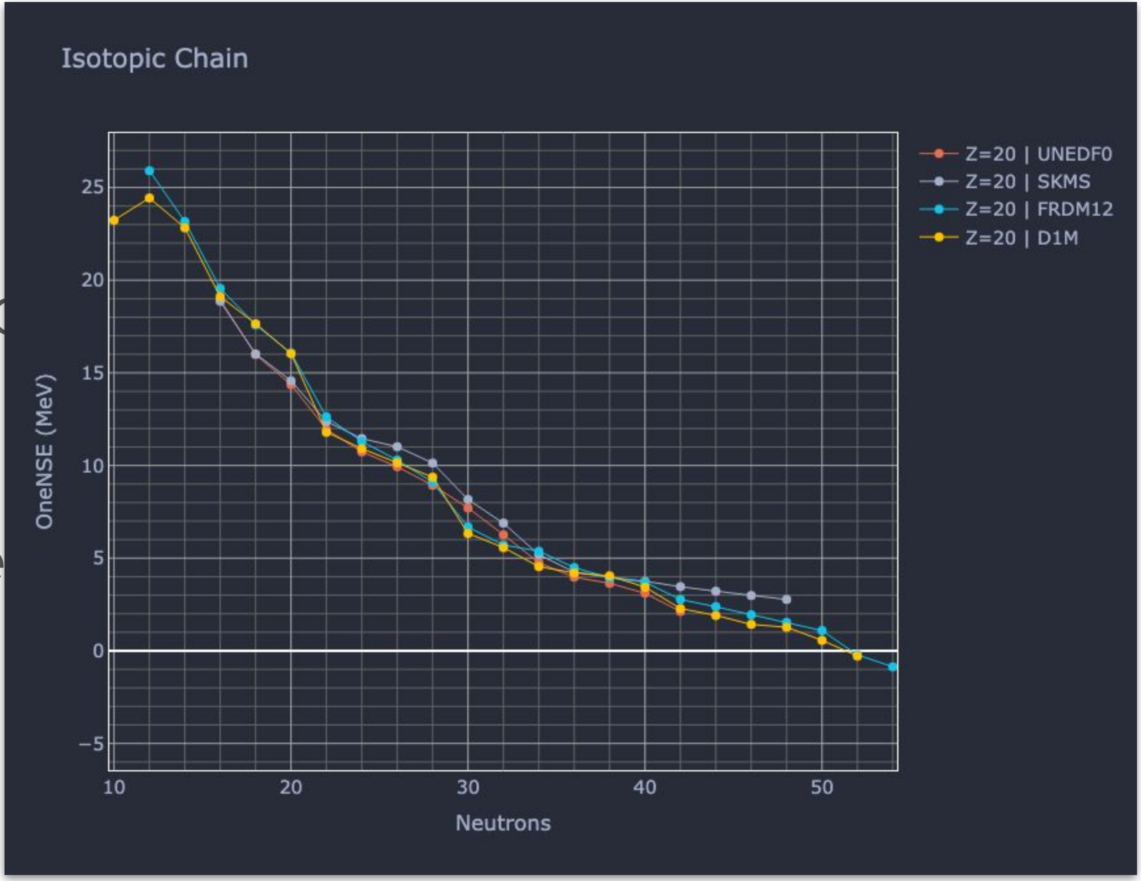
Imagine a scenario where you have many wrong (and potentially useful) models

Will they be wrong in the same ways and places? Are they equally useful?

Model Combination

Imagine a sc
wrong (and

Will they be
places? Are

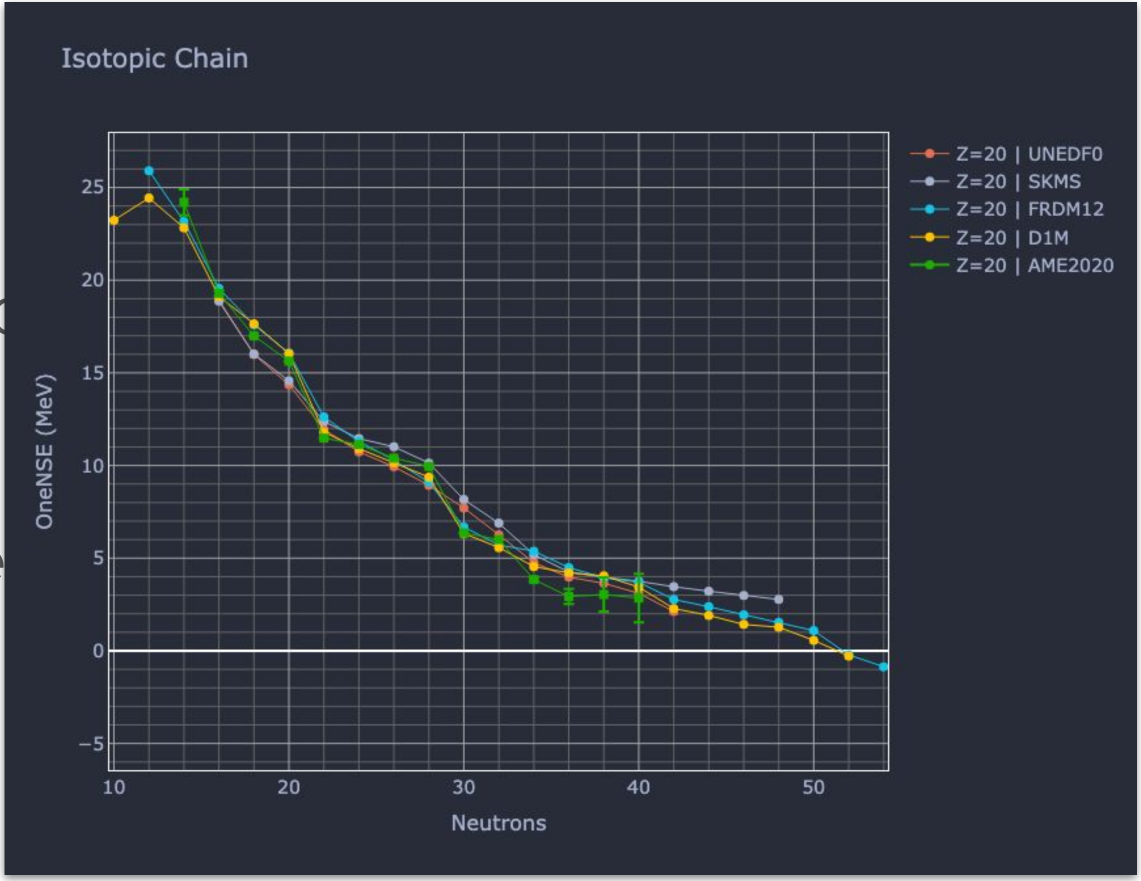


The Bayesian Mass Explorer
https://bmex.dev

Model Combination

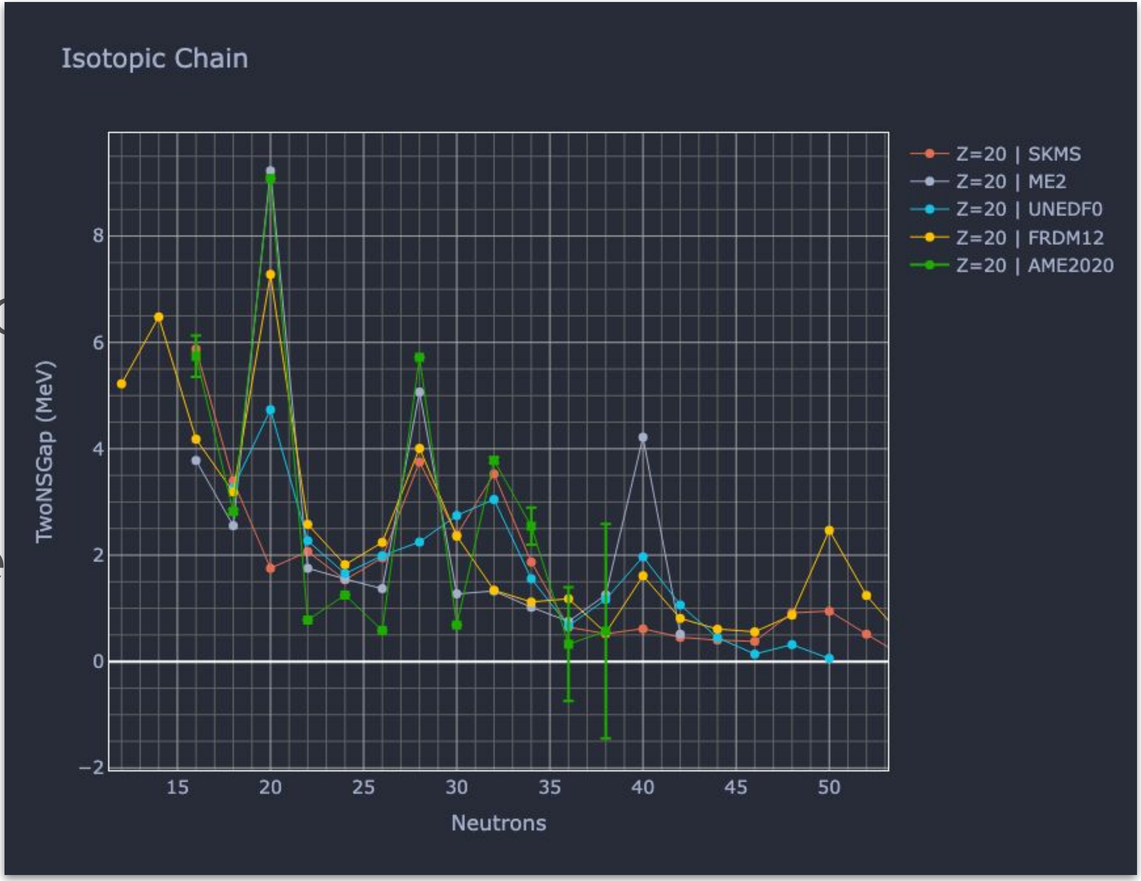
Imagine a sc
wrong (and

Will they be
places? Are



The Bayesian Mass Explorer
https://bmex.dev

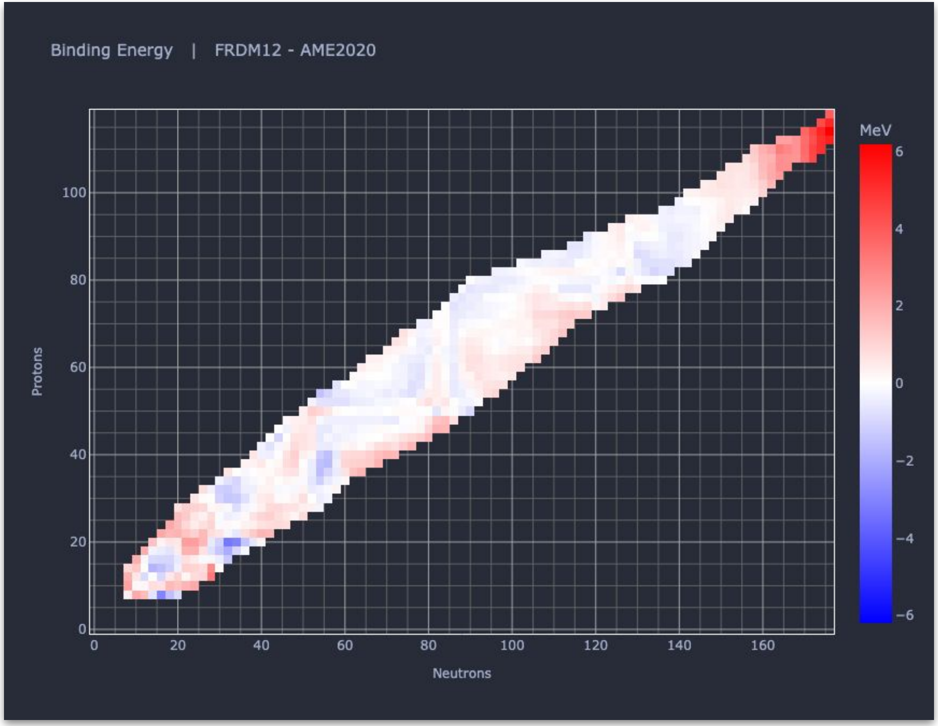
Model Combination

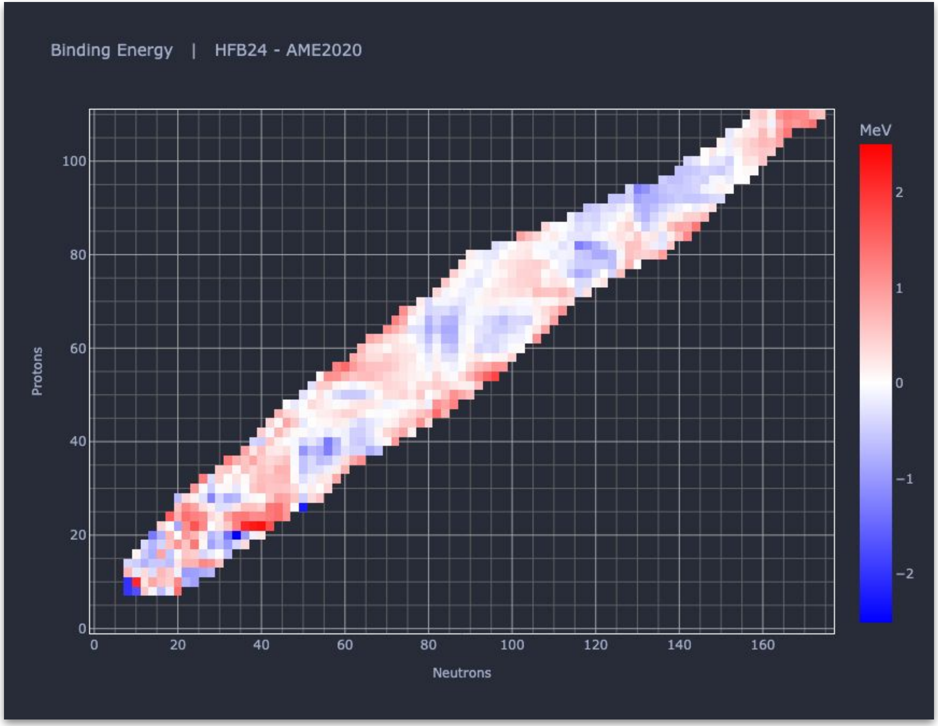


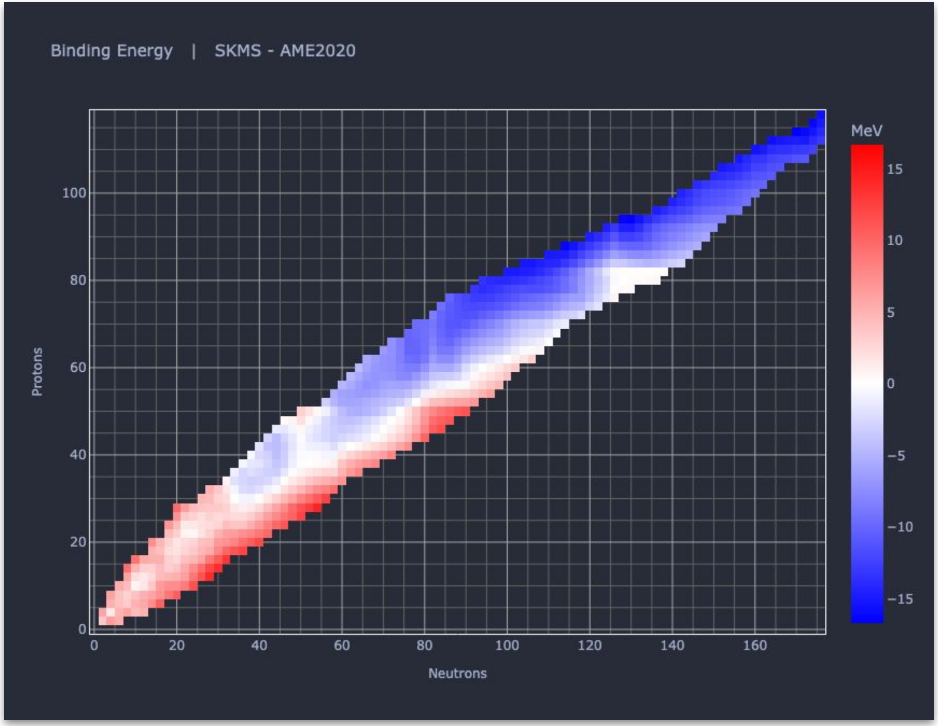
Imagine a sc
wrong (and

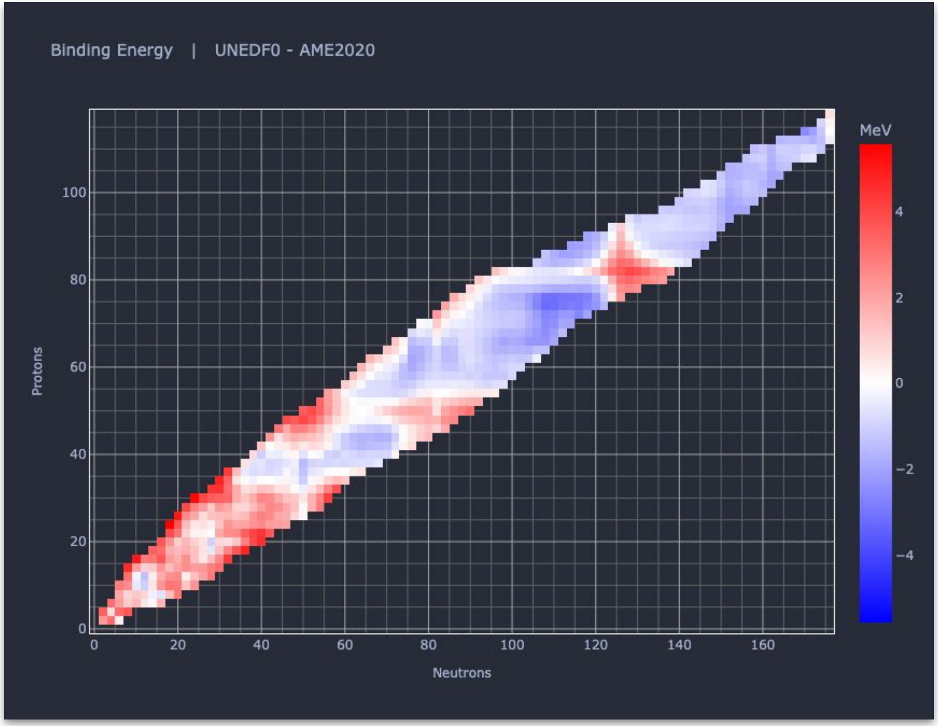
Will they be
places? Are

The Bayesian Mass Explorer
https://bmex.dev


Binding Energy | FRDM12 - AME2020
MeV
6
4
2
0
−2
−4
−6
Protons
0
20
40
60
80
100
Neutrons
0
20
40
60
80
100
120
140
160


Binding Energy | HFB24 - AME2020
MeV
2
1
0
−1
−2
Protons
0
20
40
60
80
100
Neutrons
0
20
40
60
80
100
120
140
160





The Bayesian Mass Explorer
https://bmex.dev

Model Combination

Model combination techniques provide a prescription on how to handle this!

This talk is going to focus on **forecast combination** specifically, i.e. where we already have independent model forecasts prepared

Model Combination (+ orthogonalization)

Additionally, we'll orthogonalize our forecasts to remove model redundancies

Side benefit: we can truncate the principal components to get less overfitting

Model Combination (+ orthogonalization)

First, center your data

$$X_{i,k}^{c} = X_{i,k}^{0} - \phi_0\left(x_i\right)$$

Your models

where

$$\phi_0\left(x_i\right) = \frac{1}{m}\sum_{k=1}^{m} X_{i,k}^{0} = \frac{1}{m}\sum_{k=1}^{m} f_k\left(x_i\right).$$

Model Combination (+ orthogonalization)

Then, reduce the dimensionality:

$$\boldsymbol{X^c}_{n\times m} = \boldsymbol{U}_{n\times n} \quad \boldsymbol{S}_{n\times m} \quad \boldsymbol{V}^T_{m\times m}$$

$$\approx \quad \widehat{\boldsymbol{X^c}}_{n\times p} = \hat{\boldsymbol{U}}_{n\times p} \quad \hat{\boldsymbol{S}}_{p\times p} \quad \hat{\boldsymbol{V}}^T_{p\times m}.$$

n = # of data points, **m** = # of models, **p** = # of kept components

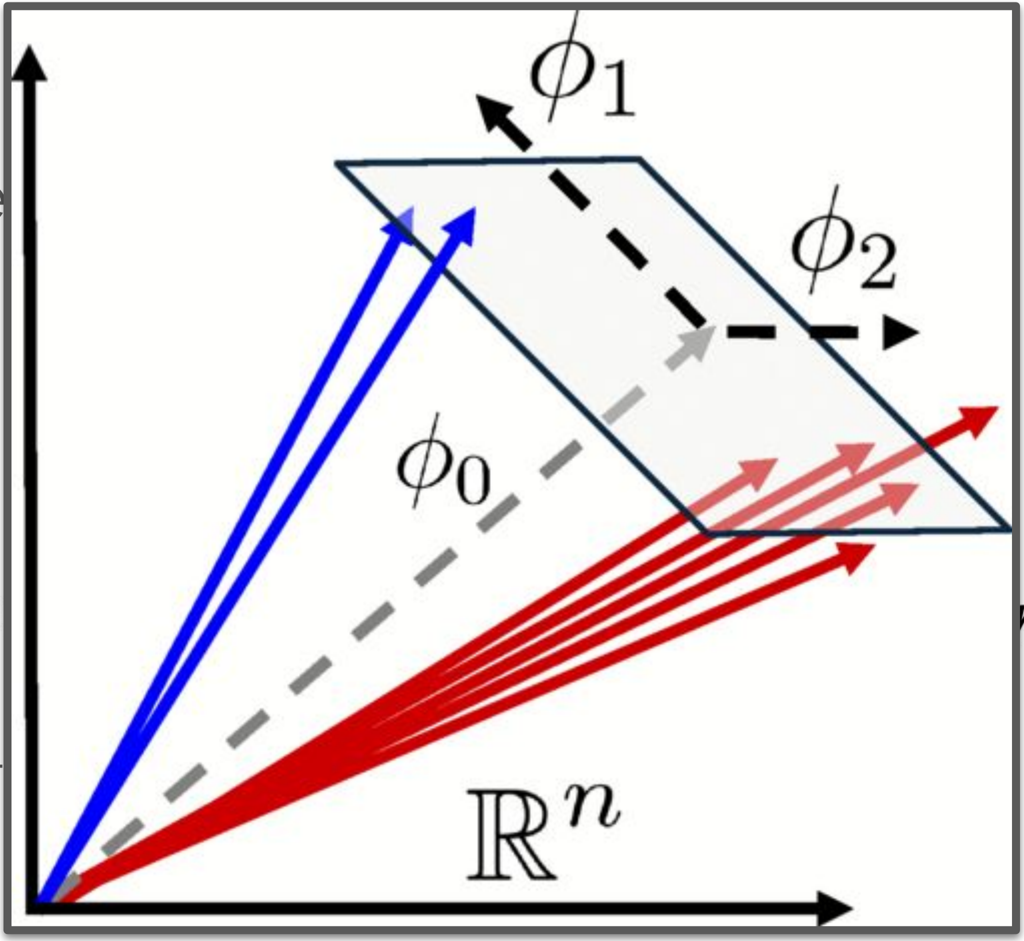
Model Combination (+ orthogonalization)

Then, reduce the dime

$$\boldsymbol{X^c}_{n \times m} = \boldsymbol{U}_{n \times n}$$

$$\approx \widehat{\boldsymbol{X^c}}_{n}$$

m.



n = # of data points, **m** = # ts

Model Combination (+ orthogonalization)

Now construct your combined forecast:

$$f^{\dagger}(\boldsymbol{x};\boldsymbol{b}) = \phi_0(\boldsymbol{x}) + \sum_{j=1}^{p} b_j \phi_j(\boldsymbol{x})$$

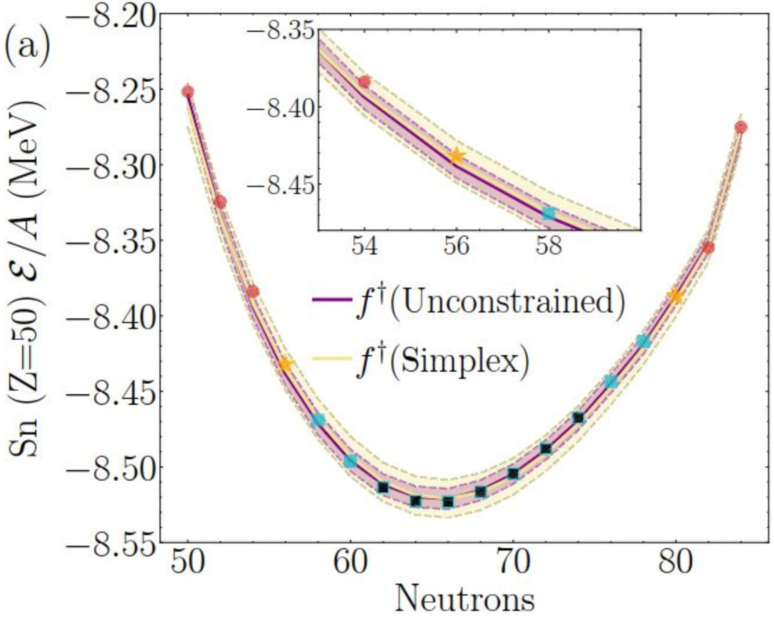
Bringing it Together

Dataset construction is important!

Once you determine the weights, you're ready to go!

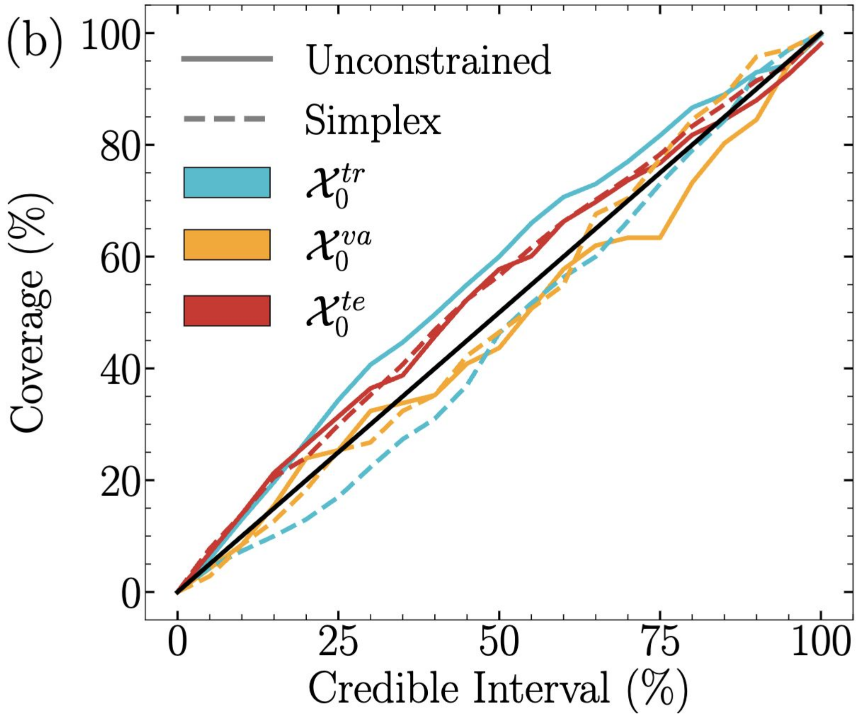
P. Giuliani, K. Godbey, V. Kejzlar, and W. Nazarewicz
Phys. Rev. Research 6, 033266

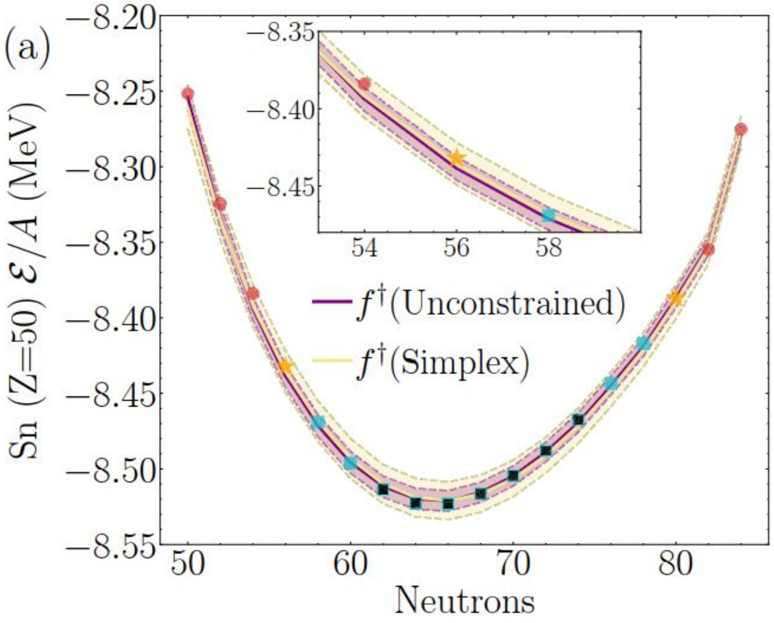
Model	RMSE $\mathcal{X}_0^{\text{tr}}$	RMSE $\mathcal{X}_0^{\text{va}}$	RMSE $\mathcal{X}_0^{\text{te}}$	RMSE $\mathcal{X}_0$
DD-ME2 [39]	2.47	2.48	2.25	2.38
DD-MEδ [40]	2.46	2.19	2.18	2.32
DD-PC1 [41]	1.94	1.77	2.19	2.03
NL3* [42]	2.18	2.15	3.59	2.84
SkM* [43]	4.91	6.34	9.75	7.42
SkP [44]	3.06	3.50	4.41	3.72
SLy4 [45]	4.53	4.84	6.27	5.34
SV-min [46]	2.97	2.99	3.98	3.42
UNEDF0 [47]	1.47	2.02	1.51	1.56
UNEDF1 [48]	1.82	1.83	2.06	1.92
UNEDF2 [49]	1.83	1.66	2.04	1.90
FRDM-12 [38]	0.62	0.62	0.65	0.63
HFB-24 [37]	0.52	0.51	0.52	0.52
BCPM [50]	2.57	2.34	2.44	2.49
D1M [51]	5.02	4.91	5.63	5.27
$f^\dagger(p=7)$	0.38	0.46	0.47	0.43
$f^\dagger(\text{simplex}, p=7)$	0.74	0.69	0.71	0.72



P. Giuliani, K. Godbey, V. Kejzlar, and W. Nazarewicz
Phys. Rev. Research 6, 033266

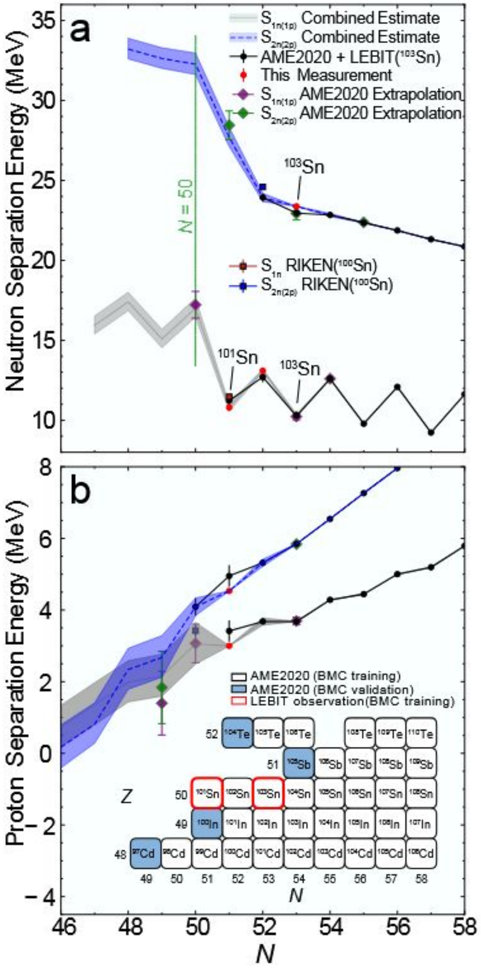
$f^{\dagger}(p=7)$	0.38	0.46	0.47	0.43
$f^{\dagger}(\text{simplex}, p=7)$	0.74	0.69	0.71	0.72





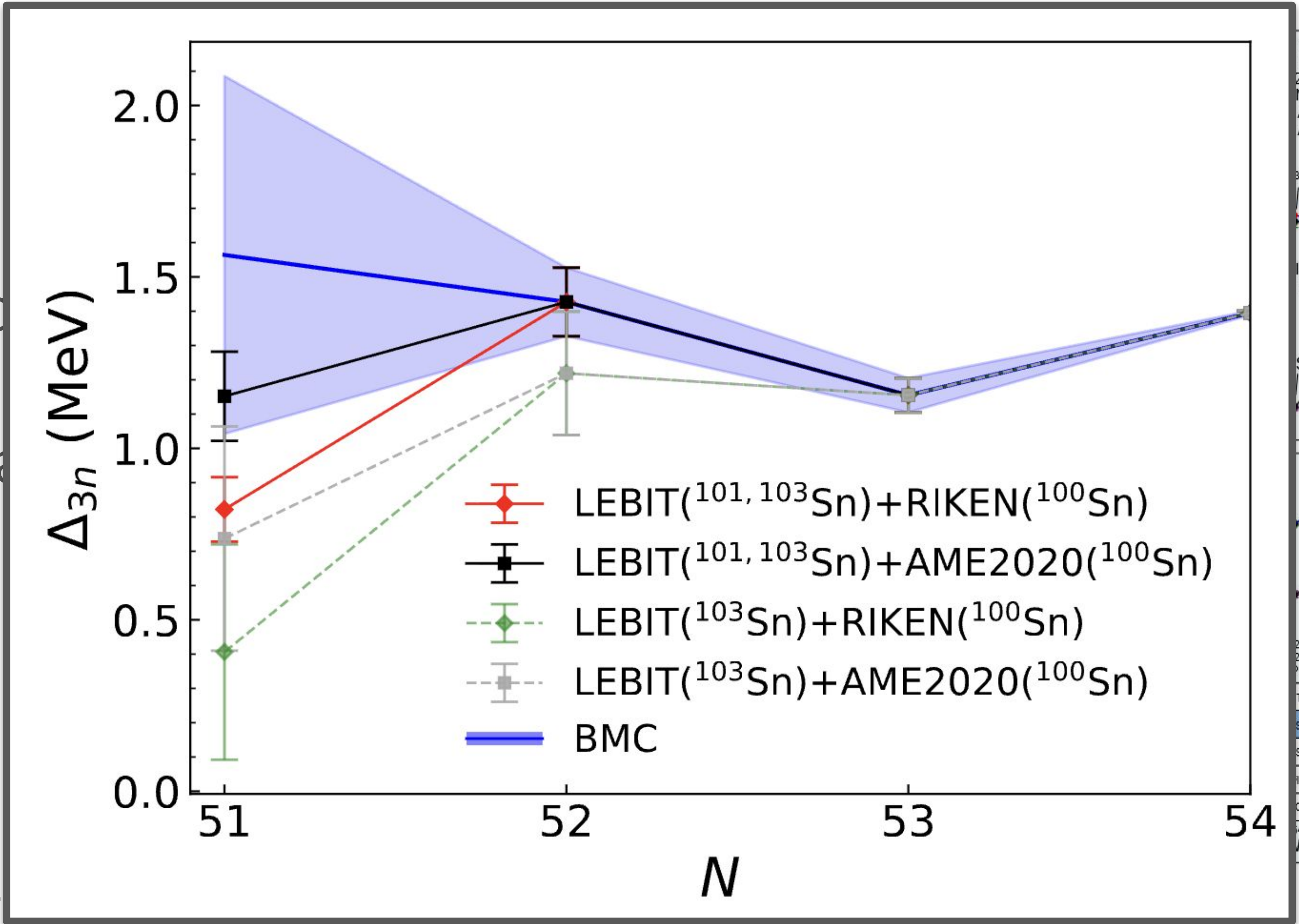
P. Giuliani, K. Godbey, V. Kejzlar, and W. Nazarewicz
Phys. Rev. Research 6, 033266

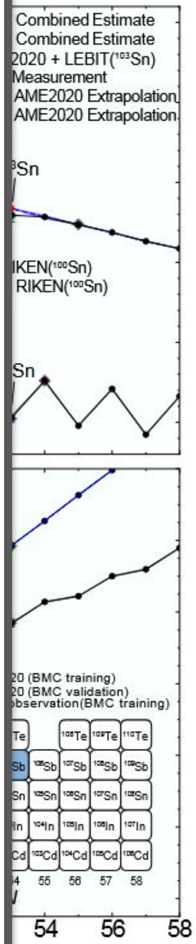
The combination scheme is *very* quick! New (or proposed) experimental data can be easily incorporated



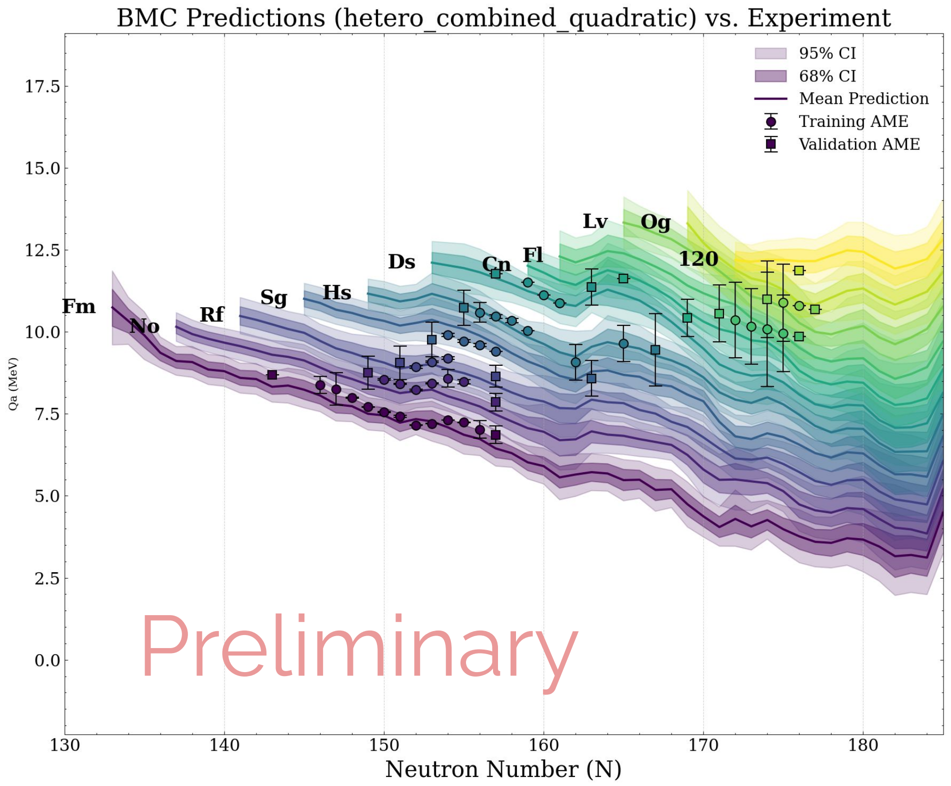
https://arxiv.org/pdf/2510.11815

The co
quick!
data ca

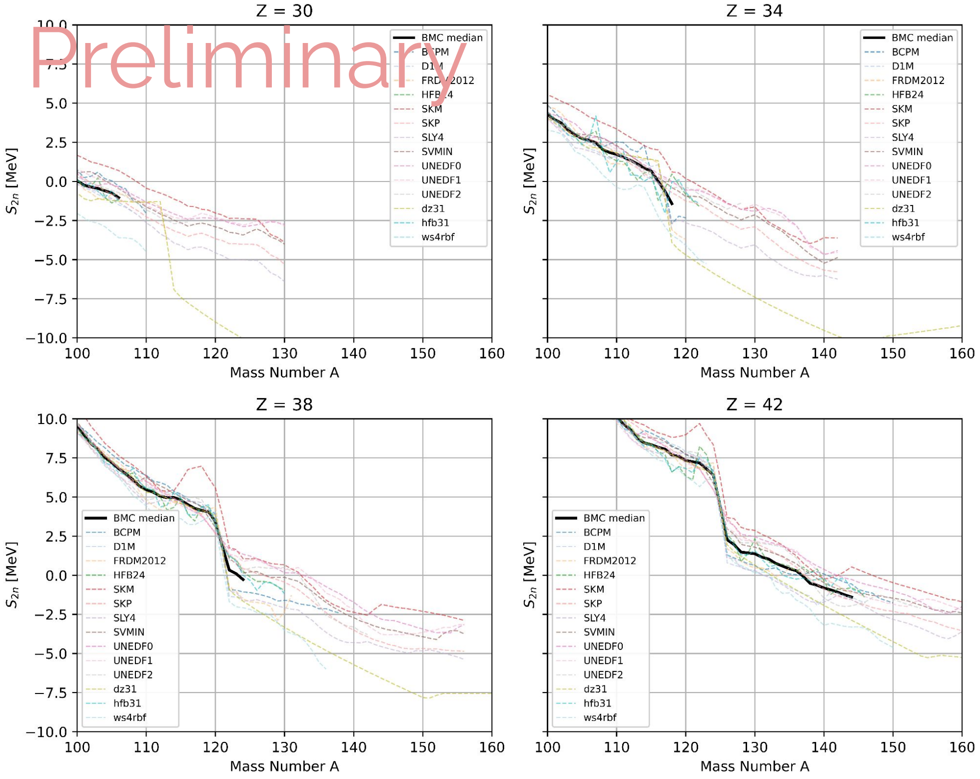


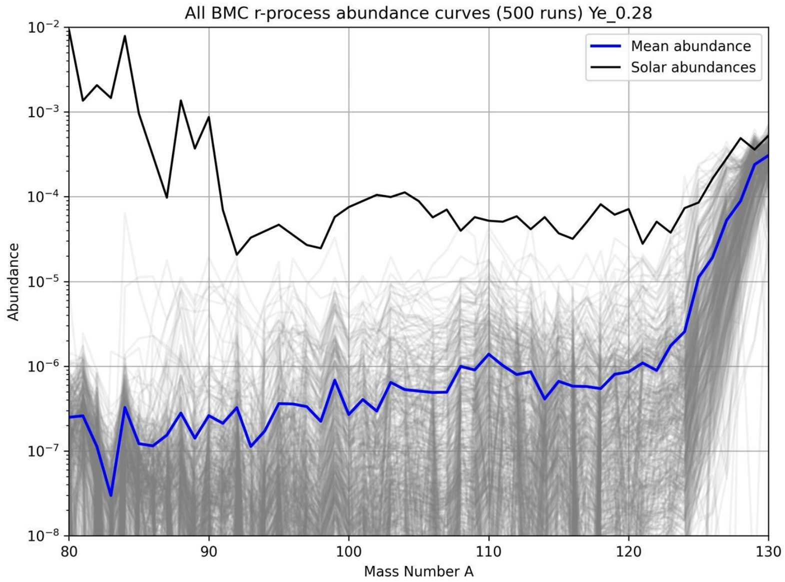


https://arxiv.org/pdf/2510.11815



Knight, et al. In Preparation





Agarwal et al. In Preparation

Enter: `pybmc`

Lightweight Python package designed to streamline model orthogonalization and combination tasks

Currently, the core algorithm has been implemented by Undergraduate students Troy Dasher, An Le, and Pranav Agarwal

http://pybmc.ascsn.net/

Sidebar

Putting the effort into making (and releasing) software can pay off immensely!

http://pybmc.ascsn.net/

Wrapping Up: Known Issues

What if your original model space doesn't contain the right behavior (missing physics)?

Model contamination?

Should the error scale be global?

What about the use of global principal components?

Wrapping Up: Known Issues

What if your original model space doesn't contain the right behavior (missing physics)?

Model contamination?

Care to help?

Should the error scale be global?

What about the use of global principal components?

Wrapping Up: Known Issues

What if your original model space doesn't contain the right behavior (missing physics)?

Model contamination?

Should the error scale be global?

What about the use of global principal components?

Care to help?

ASCSN

Advanced Scientific Computing and Statistics Network

ASCSN

Innovation Through Connection

A global community of researchers advancing the frontiers of computational science, statistics, and applied mathematics through collaboration.

Join the Community

Explore Resources

https://ascsn.net/

Our Platforms

Community Forum →

Engage in discussions, ask questions, and share knowledge.

GitHub Organization →

Explore our open-source projects and contribute to development.

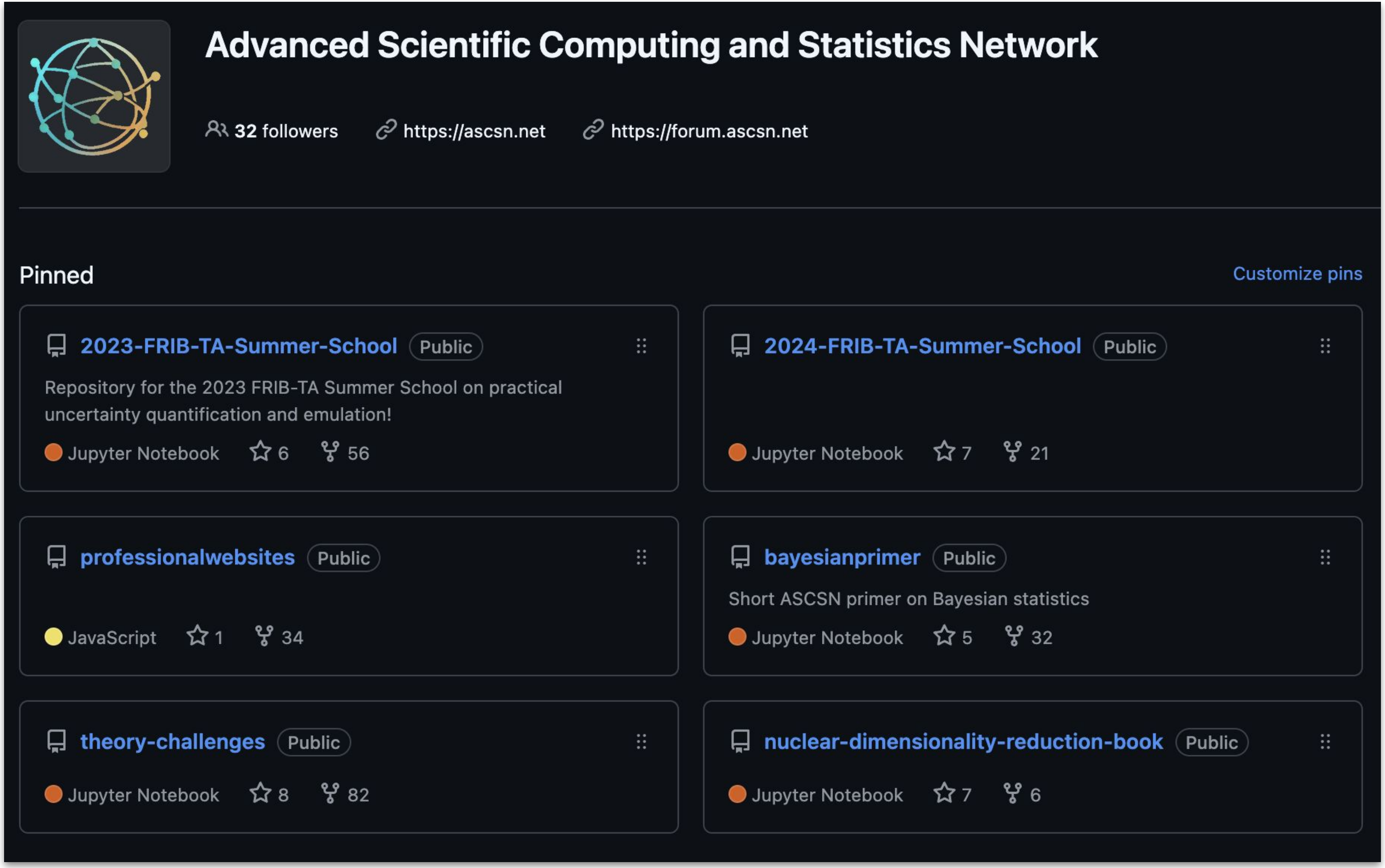
YouTube Channel →

Watch tutorials, webinars, and talks from our community.

Forum

GitHub

YouTube


Advanced Scientific Computing and Statistics Network
32 followers
https://ascsn.net
https://forum.ascsn.net
Pinned
Customize pins
2023-FRIB-TA-Summer-School
Public
Repository for the 2023 FRIB-TA Summer School on practical uncertainty quantification and emulation!
Jupyter Notebook
6
56
2024-FRIB-TA-Summer-School
Public
Jupyter Notebook
7
21
professionalwebsites
Public
JavaScript
1
34
bayesianprimer
Public
Short ASCSN primer on Bayesian statistics
Jupyter Notebook
5
32
theory-challenges
Public
Jupyter Notebook
8
82
nuclear-dimensionality-reduction-book
Public
Jupyter Notebook
7
6

Dimensionality Reduction in Nuclear Physics
Presented by ASCSN

Aplication 5: Black-Box Methods

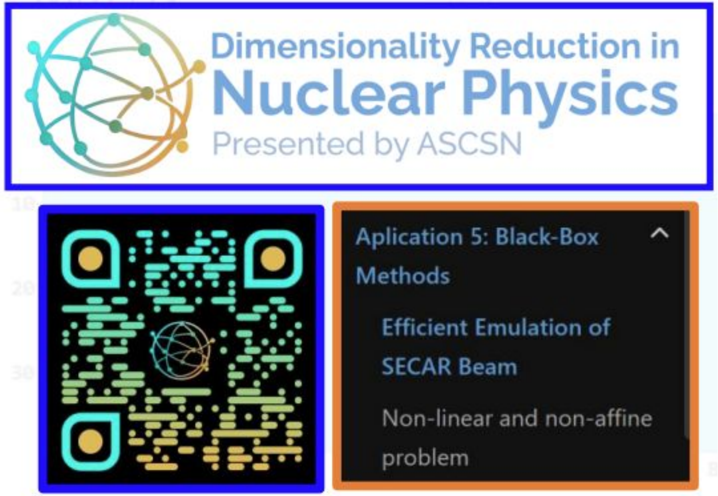
- Efficient Emulation of SECAR Beam
- Non-linear and non-affine problem

https://dr.ascsn.net

Dimensionality Reduction in Nuclear Physics
Presented by ASCSN

Introduction to Dimensionality Reduction in Nuclear Physics

- Introduction
- Application 1: The Quantum Harmonic Oscillator
- Application 2: Two body single channel nuclear scattering
- Application 3: The Empirical Interpolation Method
- Application 4: Time Dependent Systems (evolution in the reduced space)
- Aplication 5: Black-Box Methods
- Contributors



https://dr.ascsn.net

Contributors

- Eric Flynn | Content
- Pablo Giuliani | Content
- Edgard Bonilla | Content
- Daniel Odell | Content
- Ruchi Garg | Content
- Kyle Beyer | Content
- Daniel Lay | Content
- Diógenes Figueroa | Content
- Megan Campbell | Content
- Kyle Godbey | Content, Chief Editor

Immense Gratitude to All Collaborators!

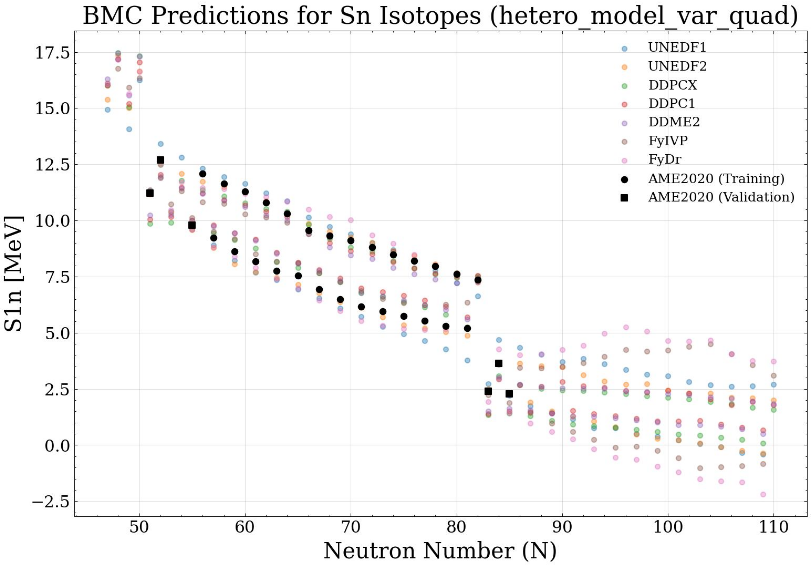
Funding

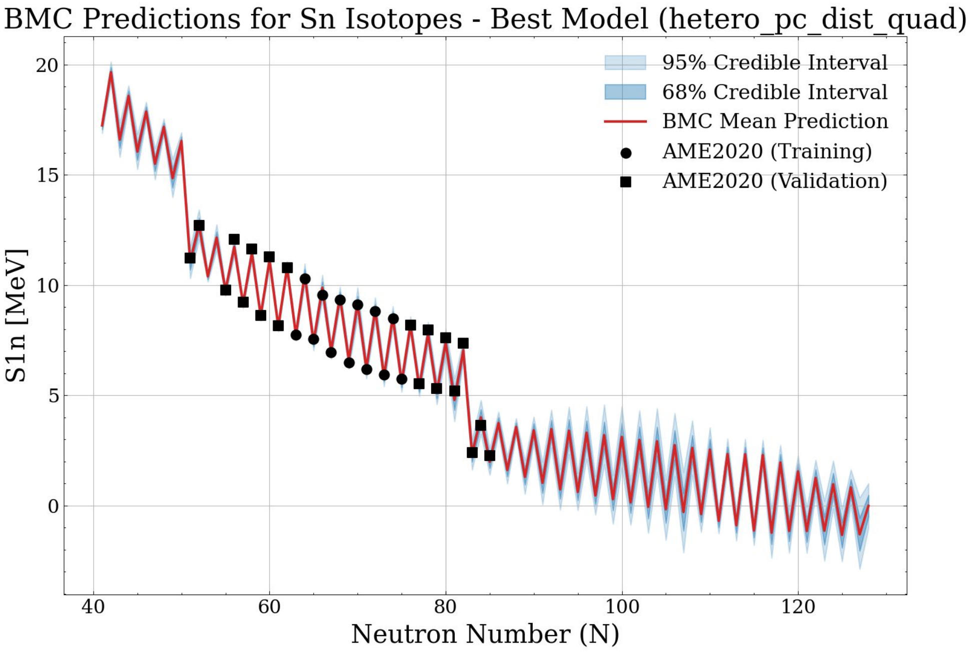
DOE NNSA Grant No. DE-NA0004074
DOE Grant Nos. DE-SC0013365, DE-SC0023175
NSF CSSI Program No. 2004601

Computing Resources

Australian National Computational Infrastructure Raijin and Gadi
Oak Ridge Leadership Computing Facility Summit and Frontier
Argonne Leadership Computing Facility Polaris
Texas A&M High Performance Research Computing Terra and Ada
Michigan State University HPCC

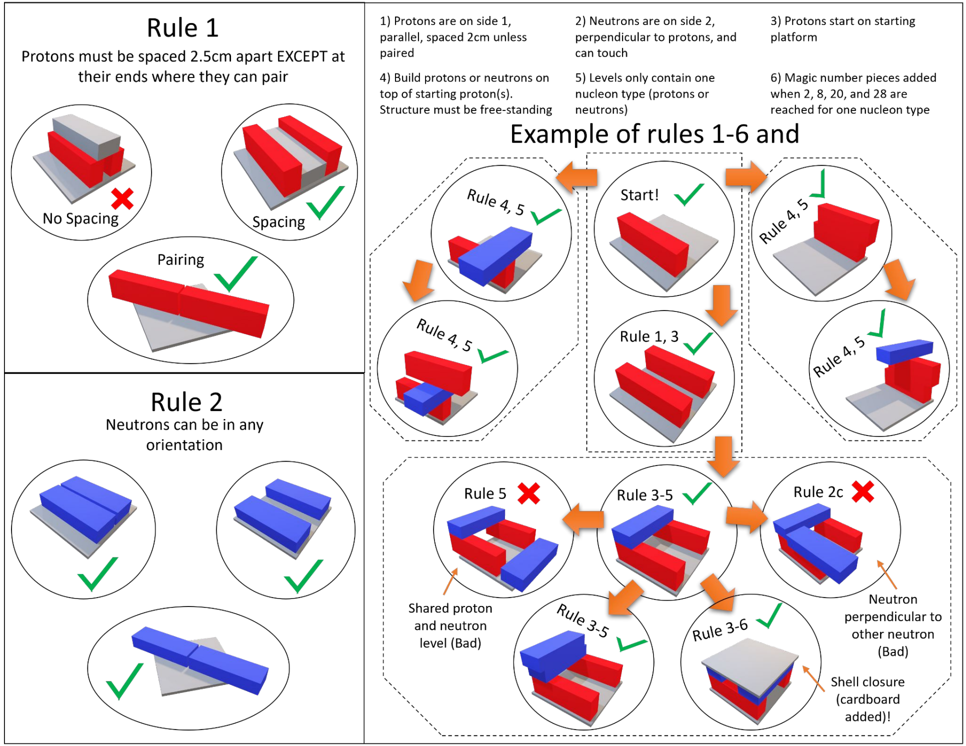
Heteroscedastic Uncertainty





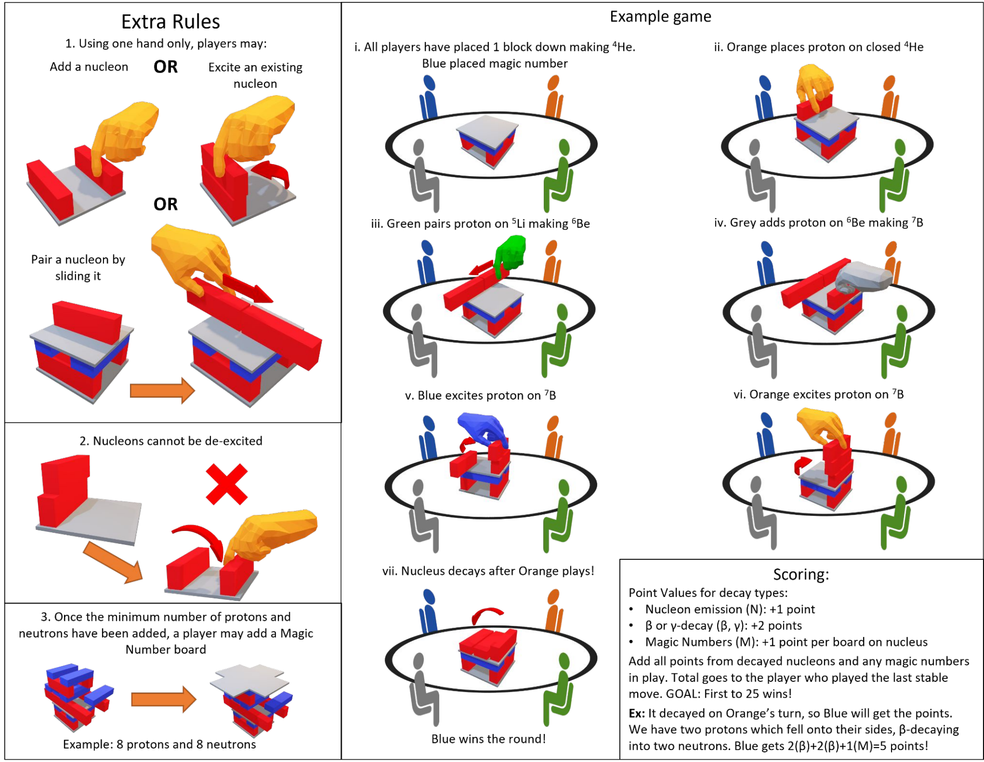
Interface of Technology and Education

- We focus on highlighting structure with dynamic decays
- Rule sets for varying student level (younger students to adults)
    - Coulomb repulsion
    - Pairing
    - Nucleon Excitations
    - Magic numbers



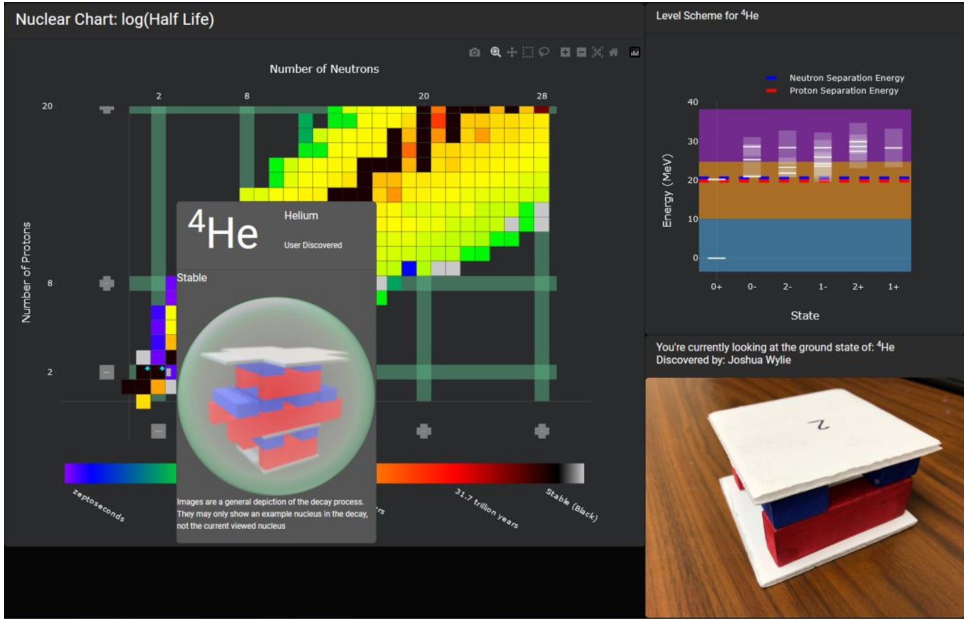
Interface of Technology and Education

- We focus on highlighting structure with dynamic decays
- Alternate multiplayer rule set:
  - Highlights competing forces and effects that may make nuclei more unstable
  - Promotes focus on excitation or more proton or neutron-rich systems



Interface of Technology and Education

- Website will be available:
  - User may click on a nucleus to see the relation between in-game view and real nuclear properties (a translation tool)
  - User-input for “discoveries” (blue diamonds) of ground state, excited state, and second excited state configurations
  - “Discovered” nuclear states are shown with user-submitted photos



Interface of Technology and Education

- Website will be available:
  - User may click on a nucleus to see the relation between in-game view and real nuclear properties (a translation tool)
  - Supplement so users can see game representation of actual physical phenomena (decay modes)