Quantified nuclear structure theory

Witold Nazarewicz Michigan State University/FRIB

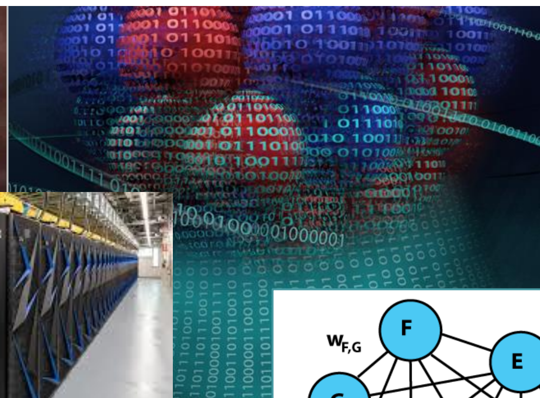
INT seminar, May 14, 2020

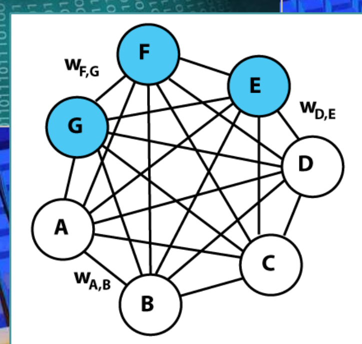
TALENT/INT Lectures on "Physics of weakly bound and unbound nuclear states: structure, decays, reactions" (INT-20-2a)

Due to the COVID-19 pandemic, the school has been cancelled.

June 22 - July 10, 2020

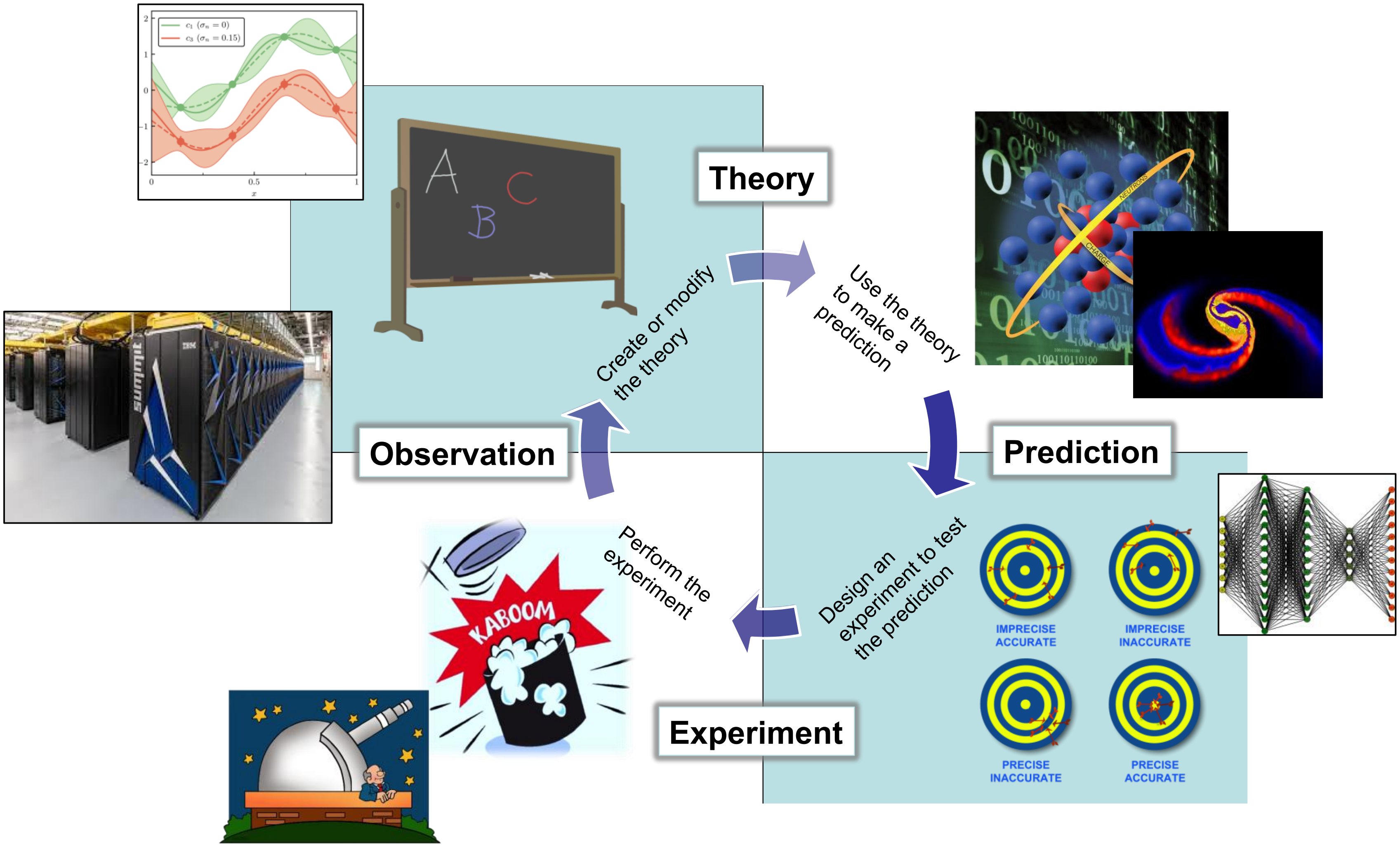
C. Forssen, K. Fossez, W. Nazarewicz, M. Ploszajczak, A. Volya

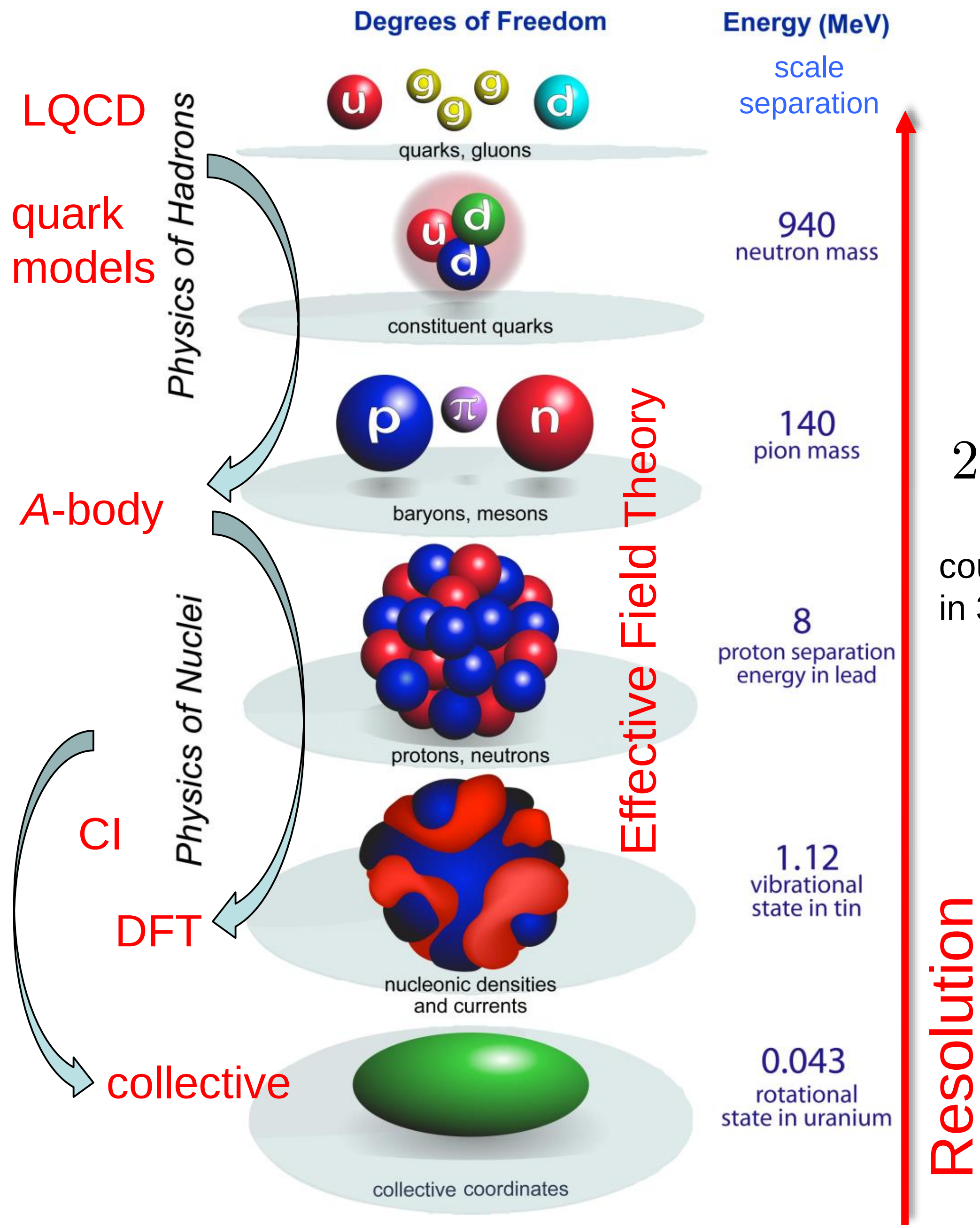




- Guiding principles
- Machine learning in nuclear structure theory
- Examples of recent work
- Bayesian extrapolations
- Bayesian model averaging
- Model calibration
- Summary

Guiding principle: the scientific method





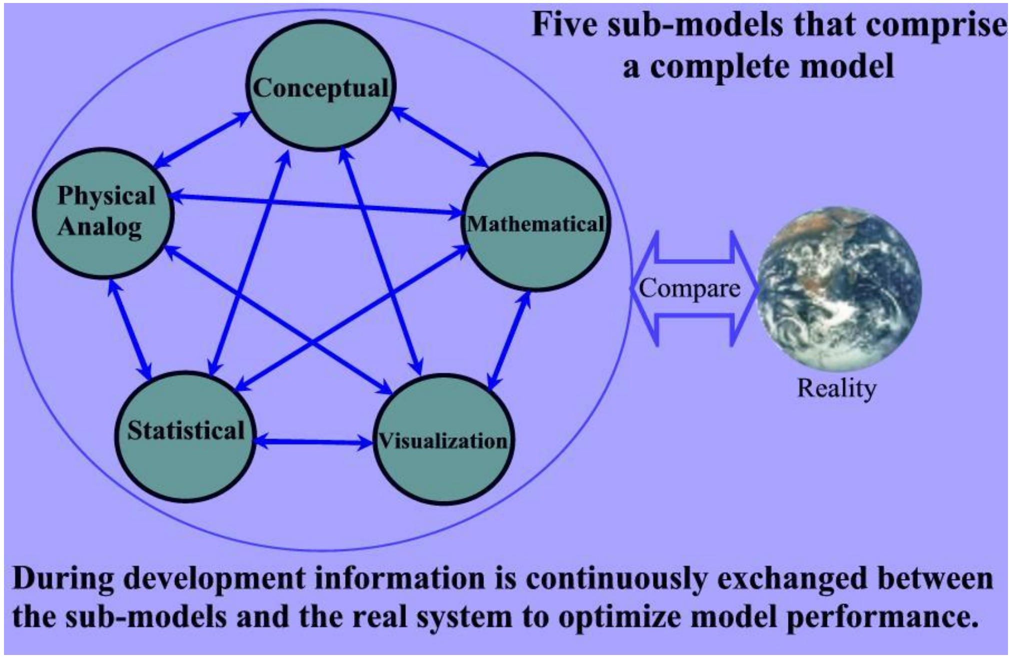
$$2^A \times \frac{A!}{N!Z!}$$

coupled equations in 3A dimensions

Hilbert-space truncations necessary

What is a Model?

In the context of the following discussion, it is useful to clarify the notion of a “model”. In this talk, by a model I will understand the combination of a raw theoretical model (i.e., mathematical framework), the calibration dataset used for its parameter determination, and a statistical model that describes the error structure.



https://serc.carleton.edu/

Machine Learning Trivia

Learning algorithms

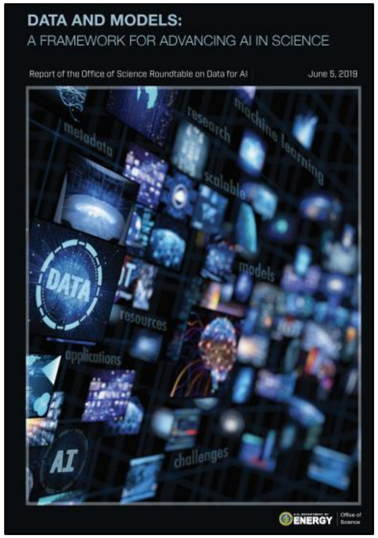
- Supervised machine learning: *Training* with known data and using this knowledge to predict the *test* data.
- Unsupervised machine learning: Finding patterns and relationship in datasets without any prior knowledge of the system (creating clusters and assigning data to these clusters).
- Reinforcement learning: Learning is achieved by trial-and-error, solely from rewards and punishment (learning from experience).
- …

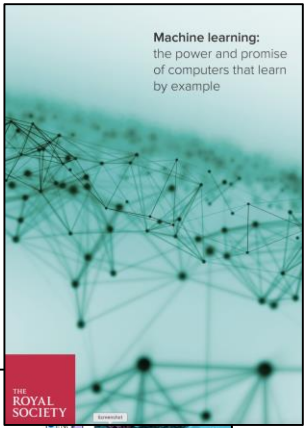
Typical applications

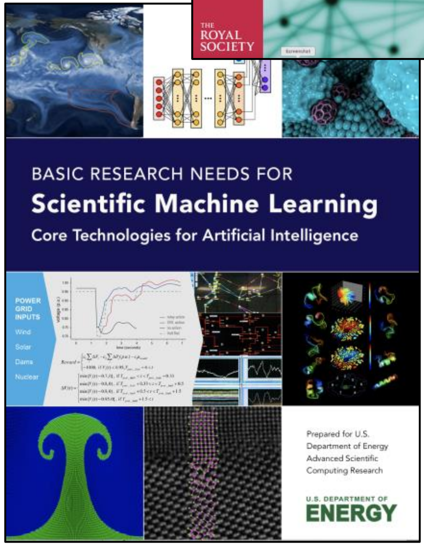
- Interpolation. Finding missing information within the known domain.
- Extrapolation. Finding missing information outside the known domain.
- Accelerating simulation and model emulation.
- Improving the interpretability.
- Estimation of bias and uncertainty.
- Model calibration and model reduction.
- Model mixing.
- …

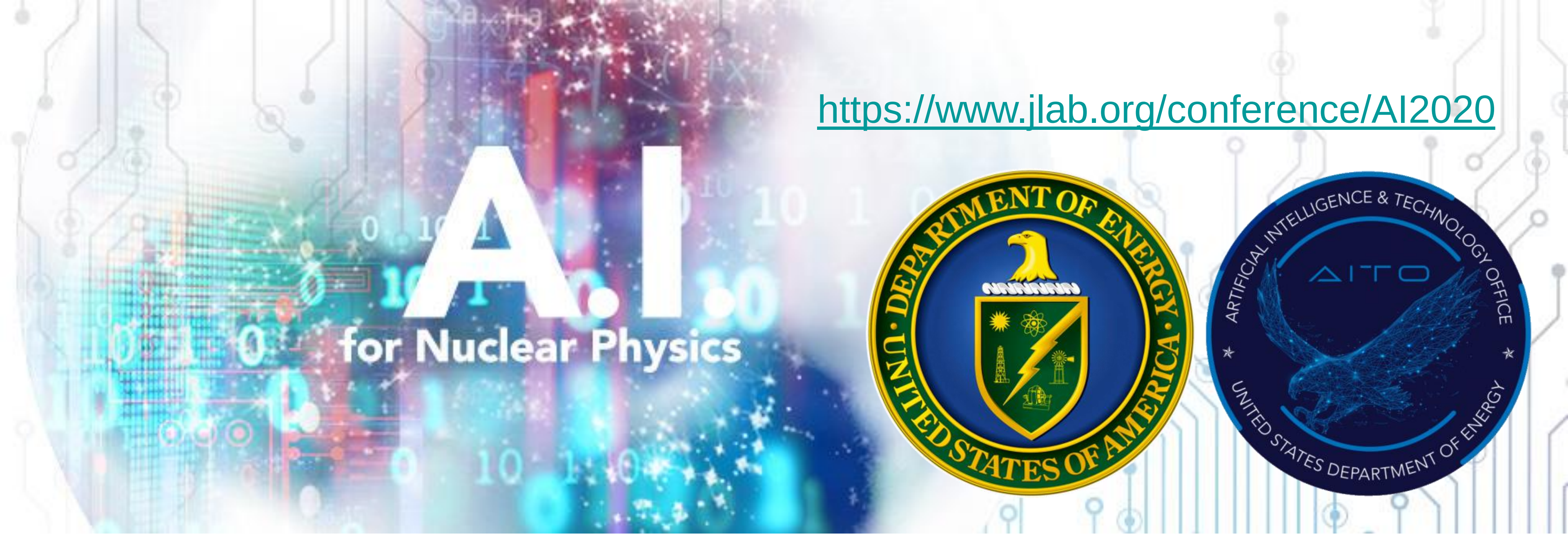
Tools

- Neural networks
- Bayesian networks
- Decision trees
- Support vector machines
- Regression analysis
- …









https://www.jlab.org/conference/AI2020

March 4-6, 2020, Thomas Jefferson National Accelerator Facility

- Explore the ways in which A.I. can be used to advance research in fundamental nuclear physics and in the design and operation of large-scale accelerator facilities.
- Explore applications and research needed on several time frames, ranging from immediate benefit.
- The results of the workshop will be summarized in a report that can serve as a roadmap for the future application of A.I. and a guide to areas for possible collaboration. (The report will be out very soon!)

Machine learning & low-energy nuclear theory: Why?

ML tools can help us to speed up the scientific process cycle and hence facilitate discoveries: Beam time and compute cycles are expensive!

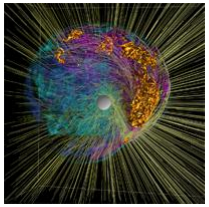
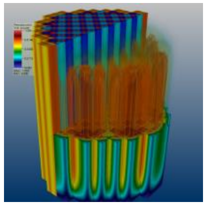
- Enabling fast emulation for big simulations
- Revealing the information content of measured observables
- Identifying crucial experimental data for better constraining theory

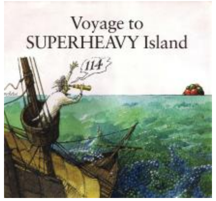
ML tools can help us to reveal the structure of our models

- Parameter estimation with heterogeneous/multi-scale data
- Model reduction
- Uncertainty quantification

ML tools can help us to provide predictive capability

- Theoretical results often involve ultraviolet and infrared extrapolations due to Hilbert-space truncations
- Providing meaningful input to applications and planned measurements
- Theoretical models are often applied to entirely new nuclear systems and conditions that are not accessible to experiment



Early applications of neural networks: 1992-2015

Nuclear Physics A540 (1992) 1–26
North-Holland

NUCLEAR
PHYSICS A

Learning and prediction of nuclear stability by neural networks

S. Gazula and J.W. Clark

McDonnell Center for the Space Sciences and Department of Physics, Washington University, St. Louis, MO 63130, USA

H. Bohr

School of Chemical Sciences, University of Illinois, Urbana, IL 61801, USA

Received 17 June 1991
(Revised 21 November 1991)

The backpropagation learning algorithm is used to teach layered feedforward networks of model neurons the existing data on nuclear stability and atomic masses. Specific applications include (i) the construction of networks that decide stability, (ii) learning and prediction of nuclear mass excesses and (iii) analysis of the systematics of neutron separation energies. With suitable architecture and representation of input and output data, learning can be accomplished with high accuracy. Evidence is presented that these new adaptive computational systems can grasp essential regularities of nuclear physics including the valley of β-stability, the pairing effect and the existence of shell structure. Significant predictive ability is demonstrated, opening the prospect that neural networks may provide a valuable new tool for computing nuclear properties and, more broadly, for phenomenological description of complex many-body systems.

Several applications of NN to various nuclear structure problems by Clark and his collaborators, including: nuclear systematics[1,2]; backpropagation algorithm for training NN[3]; Nuclear mass systematics[4]; Global properties: Support Vector Machines[5]

[1]Phys. Lett. B 300, 1 (1993); [2]Neural Networks 8, 291 (1995); [3]CPC 88, 1 (1995);
[4]Nucl. Phys. A 743, 222 (2004); [5]IJMP B 20, 5015 (2006)

The current universe of AI applications to nuclear structure theory

BC=Bayesian calibration; BGP=Bayesian Gaussian processes; BMA=Bayesian model averaging; BNN=Bayesian neural networks; NN=Neural Networks; RBF=Radial basis function

Recent applications (collected in March 2020)

Nuclear structure

- **Masses**: NN[1]; RBF, BNN[2]; Multilayer Perceptron NN[3]; BNN[4,5,6]
  [1]Ann. Nucl. Energy 63, 172 (2014); [2]Phys. Rev. C 100, 054311 (2019); [3]1912.11365; [4]Phys. Rev. C 93, 014311 (2016); [5]Phys. Rev. C 97, 014306 (2018); [6]Phys. Lett. B 778, 48 (2018)
- **Mass extrapolations:** BGP, BNN[1]; BGP, BMA[2]; BGP, BMA[3]; BGP, BMA[4]
  [1]Phys. Rev. C 98, 034318 (2018); [2]Phys. Rev. Lett. 122, 062502 (2019); [3]Phys. Rev. C 101, 014319 (2020); [4]2001.05924
- **Masses for r-process:** BC[1]
  [1]1901.10337
- **Charge radii:** naïve BNN[1]; NN[2]; BNN[3]
  [1]Phys. Rev. C 101, 014304 (2020); [2]J. Phys. G 40, 055106 (2012); J. Phys. G 43, 114002 (20016); [3]Phys. Rev. C 96, 044308 (2017)
- **Excited states with CI:** shell model with BC[1]; NN[2]
  [1]Phys. Rev. C 98, 061301(R) (2018), 1907.04974; [2]2001.08561
- **Excited 2^+ states:** NN[1]
  [1]2002.08218

Recent applications of AI (cont.)

BC=Bayesian calibration; BGP=Bayesian Gaussian processes; BMA=Bayesian model averaging; BNN=Bayesian neural networks; NN=Neural Networks; RBF=Radial basis function

Nuclear structure: Dealing with Hilbert-space truncations

- **Nucleon-nucleon phase shifts in EFT:** BC[1]
  [1]J. Phys. G 46, 045102 (2019)
- **Collective states:** emulating beyond-DFT with committee of deep NN[1]
  [1]1910.04132
- **Potential Energy Surfaces:** feedforward NN[1]
  [1]Phys. Part. Nucl. Lett. 10, 528 (2013)
- **Truncation errors in EFT:** BGP[1,2]
  [1]Phys. Rev. C 100, 044001 (2019); [2]Phys. Rev. C 96, 024003 (2017)
- **Finite-size corrections to A-body models:** NN[1]; feedforward NN[2]
  [1]Phys. Rev. C 100, 054326 (2019); [2]Phys. Rev. C 99, 054308 (2019)
- **A-body models:** Subspace projected A-body technique[1]; BC[2;] feedforward NN[3]
  [1]Phys. Rev. Lett. 123, 252501 (2019); [2]1912.02227; J. Phys. G 46 095101 (2019);
  [3]arXiv:1911.13092

Nuclear structure: model reduction

- **Mass model structure:** BC, BGP, BMA[1]
  [1]2002.04151

Recent applications of AI (cont.)

BC=Bayesian calibration; BNN=Bayesian neural networks; NN=Neural Networks ; RBF=Radial basis function; SVM=Support Vector Machines

Nuclear decays

- **Beta-decay for r-process:** BNN[1]
  [1]Phys. Rev. C 99, 064307 (2019)
- **Beta decay half-lives:** SVM, feedforward NN[1]
  [1]0809.0383 [nucl-th]
- **Alpha decays:** BNN[1,2]; NN[3]
  [1]J. Phys. G 46, 115109 (2019); [2]EPL 127, 42001 (2019); [3]1910.12345
- **Evaluation of incomplete fission yields:** BNN[1] and mixture-density NN[2]
  [1]Phys. Rev. Lett. 123, 122501 (2019); [2]EPJ Web Conf. 211, 04006 (2019)

Nuclear reactions

- **UQ for direct nuclear reactions:** BC[1,2,3]
  [1]Phys. Rev. Lett. 122, 232502 (2019); [2]Phys. Rev. C 97 064612 (2018);
  [3]Phys. Rev. C 100, 064615 (2019)
- **Cross sections in proton induced spallation reactions:** BNN[1]
  [1]Chinese Phys. C 44, 014104 (2020)
- **Fusion reaction cross-sections:** NN[1]
  [1]NIM B 462, 51 (2020)

EXAMPLES OF RECENT WORK

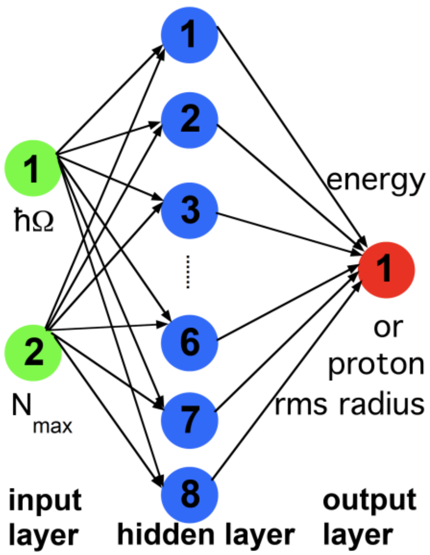
ML for Nuclear Structure Theory

Extrapolations of A-body results with ANN

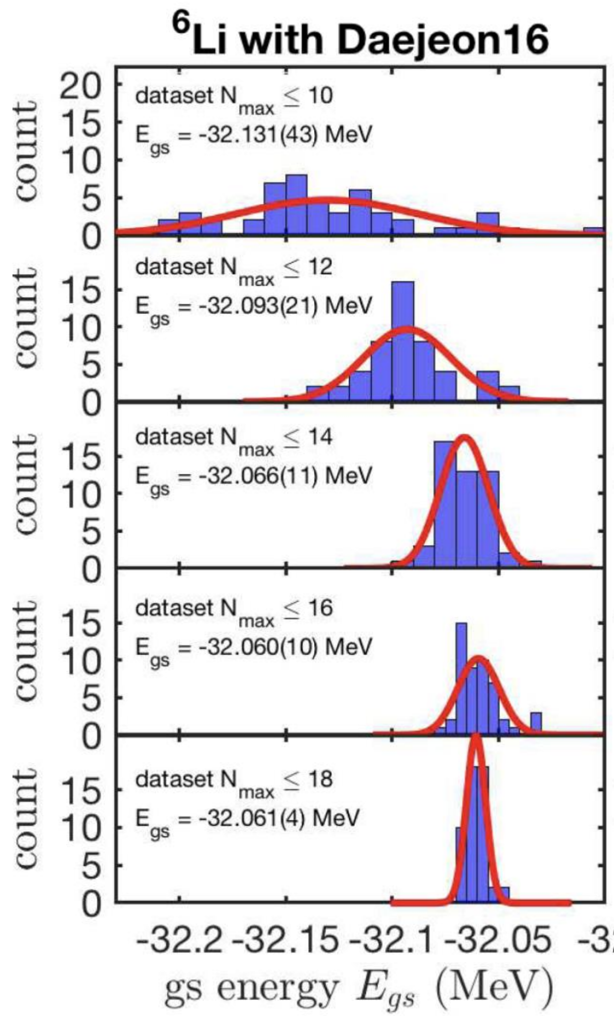
Jiang et al. Phys. Rev. 100, 054326 (2019)

Negoita et al. Phys. Rev. C 99, 054308 (2019)

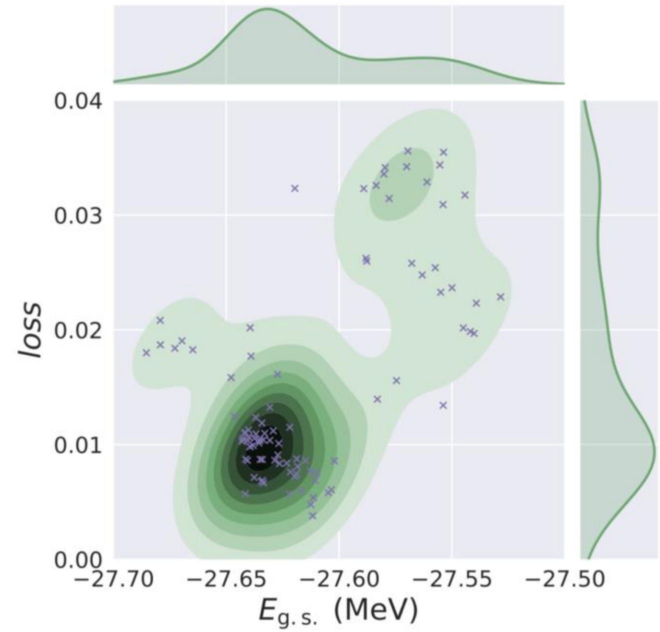
NCSM



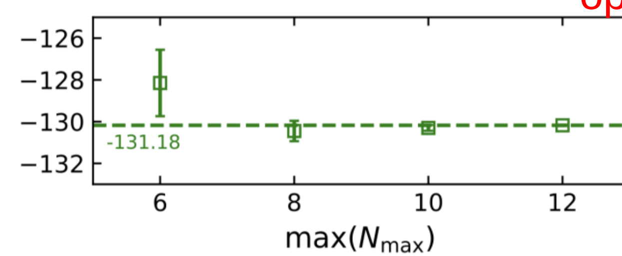
Topological structure of the feed-forward artificial neural network used



^{4}He with NCSM+NNLO$_{opt}$



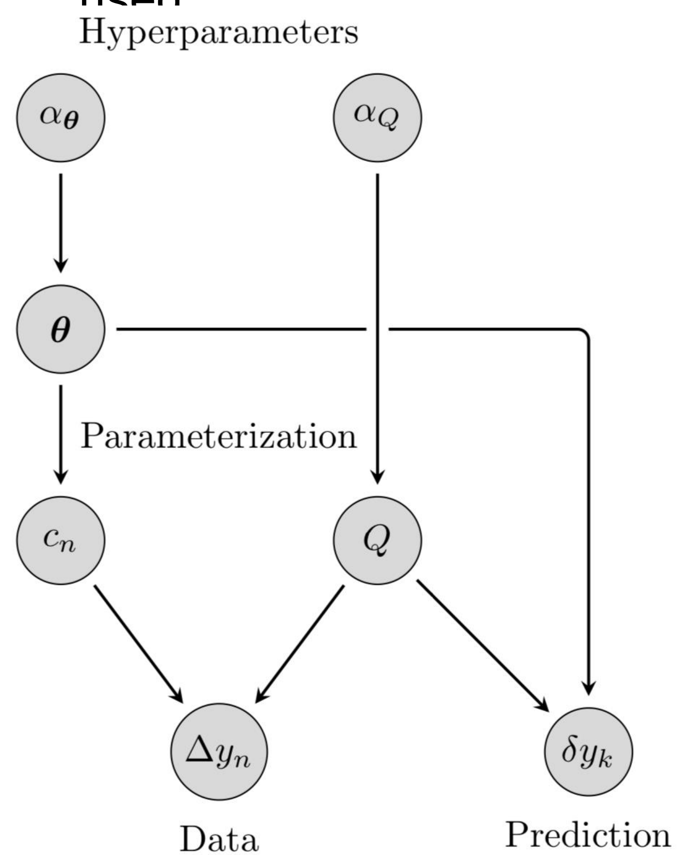
^{16}O with CC+NNLO$_{opt}$



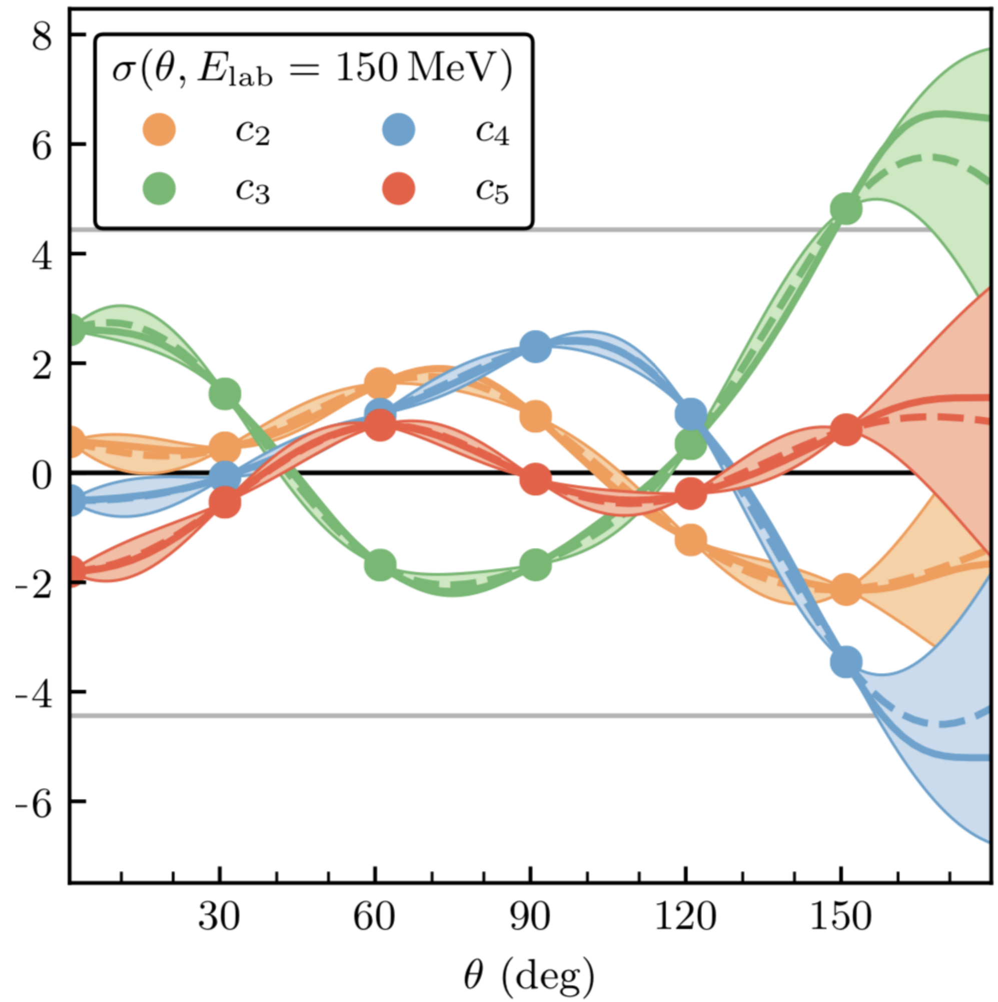
Bayesian GP model for Quantifying Correlated Truncation Errors in EFT

Melendez et al., Phys. Rev. C 100, 044001 (2019)

Bayesian network used

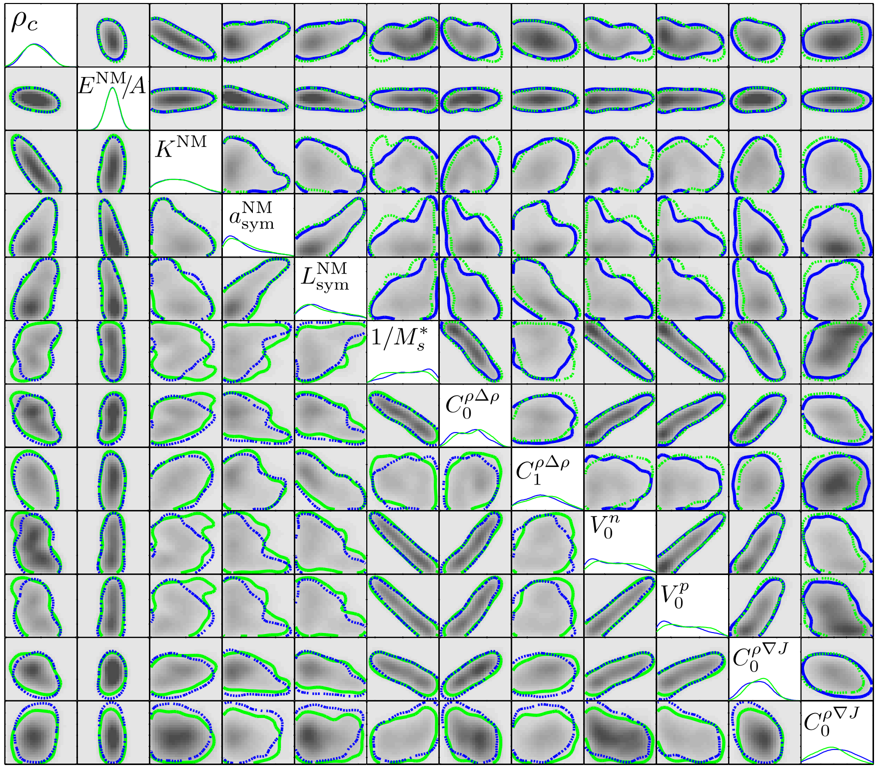


Differential NN cross section coefficients from the semilocal EKM potential



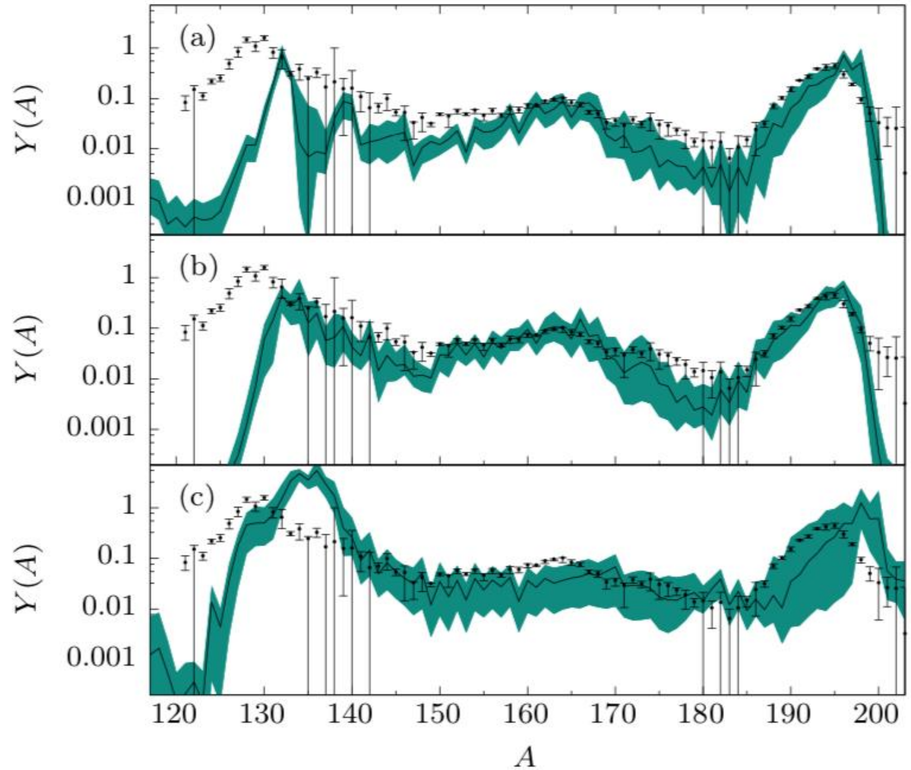
Parameter estimation: Quantified DFT

McDonnell et al.
Phys. Rev. Lett. 114, 122501 (2015)



Bivariate marginal estimates of the posterior distribution for the 12-dimensional DFT UNEDF$_1$ parameterization.

Sprouse et al.
arXiv 1901.10337



Simulated r-process abundances with astrophysical conditions corresponding to high-entropy (a), low-entropy (b), and fission-recycling (c).

Now: Extrapolations using Bayesian learning

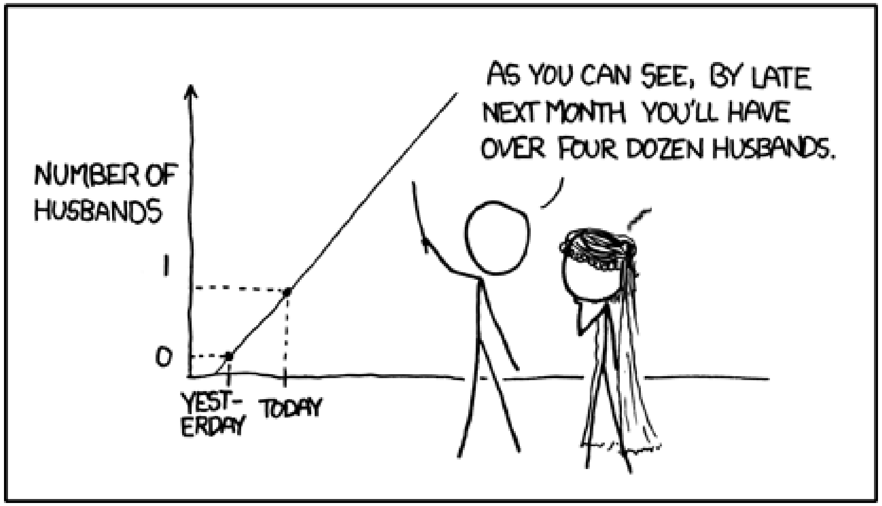
Bayesian inference

$$\underbrace{P(A|B)}_{\text{posterior}} = \frac{\overbrace{P(B|A)}^{\text{likelihood}}\,\overbrace{P(A)}^{\text{prior}}}{\underbrace{P(B)}_{\text{model evidence}}}$$

- Posterior: the degree of belief after incorporating news that B is true. Posterior probability is obtained from a prior probability, given evidence
- Likelihood: measures the goodness of fit of a statistical model to a sample of data for given values of the parameters.
- Prior: initial degree of belief in A
- Model evidence: this factor is the same for all possible hypotheses being considered.

In many cases, nuclear input MUST involve massive extrapolations based on predicted quantities. And extrapolations are ~~impossible~~ tough.

MY HOBBY: EXTRAPOLATING



BML and quantified extrapolations

Residual of an observable O:

$$\delta_{\mathcal{O}}(Z,N) = \mathcal{O}^{\mathrm{exp}}(Z,N) - \mathcal{O}^{\mathrm{th}}(Z,N)$$ small number!

$|\delta_{\mathcal{O}}| \ll |\mathcal{O}|$ Smooth part of the residual represents missing physics

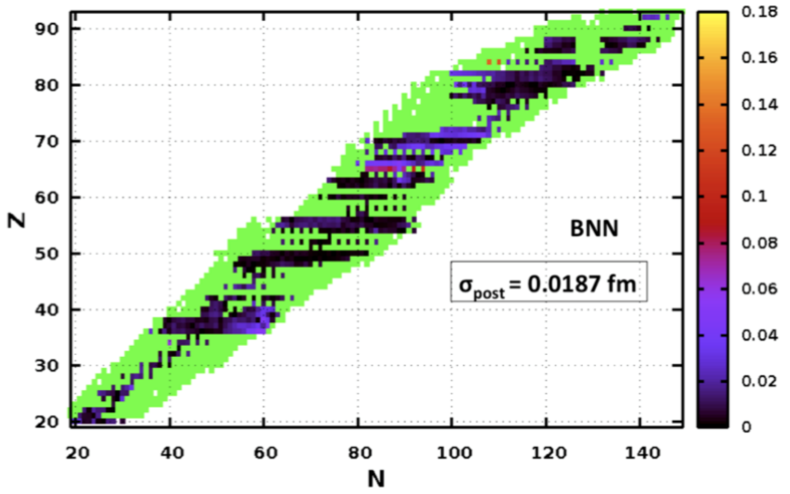
Estimate of an observable O:

$$\mathcal{O}^{\mathrm{est}}(Z,N) = \mathcal{O}^{\mathrm{th}}(Z,N) + \delta^{\mathrm{em}}_{\mathcal{O}}(Z,N)$$

emulator of the residual

Supervised learning: the nuclear modeling and the choice of priors represent two aspects of the supervision

Charge radii with BNN
R.Utama et al., J. Phys. G 43, 114002 (2016)



Bayesian approach

residual $$y_i = f(x_i, \theta) + \sigma\epsilon_i$$

(Z,N)$_i$

$$p(y^*|y) = \int p(y^*|y, \underbrace{\theta, \sigma) p(\theta, \sigma}_{\text{model parameters}}|y)\, d\theta d\sigma$$

Prediction of unknown observable *y** given known data *y*

Two statistical models used:

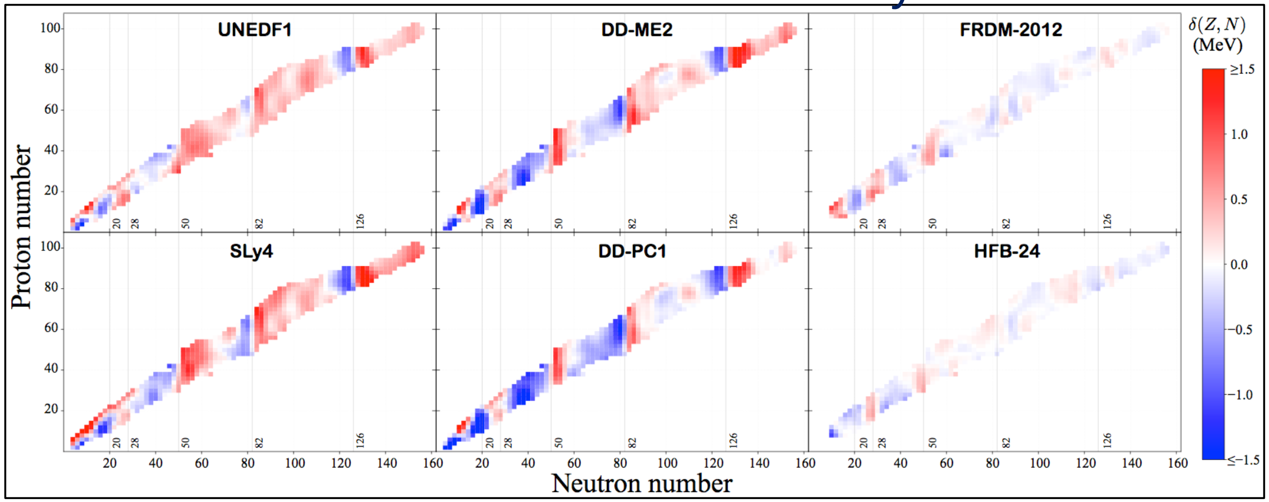
- Gaussian process (**3** parameters)
- Bayesian neural network with sigmoid function (30 neurons, 1 layer; **181** parameters)

100,000+ iterations of an ergodic Markov chain produced by the Metropolis-Hastings algorithm

Some refinements added based on our knowledge of trends

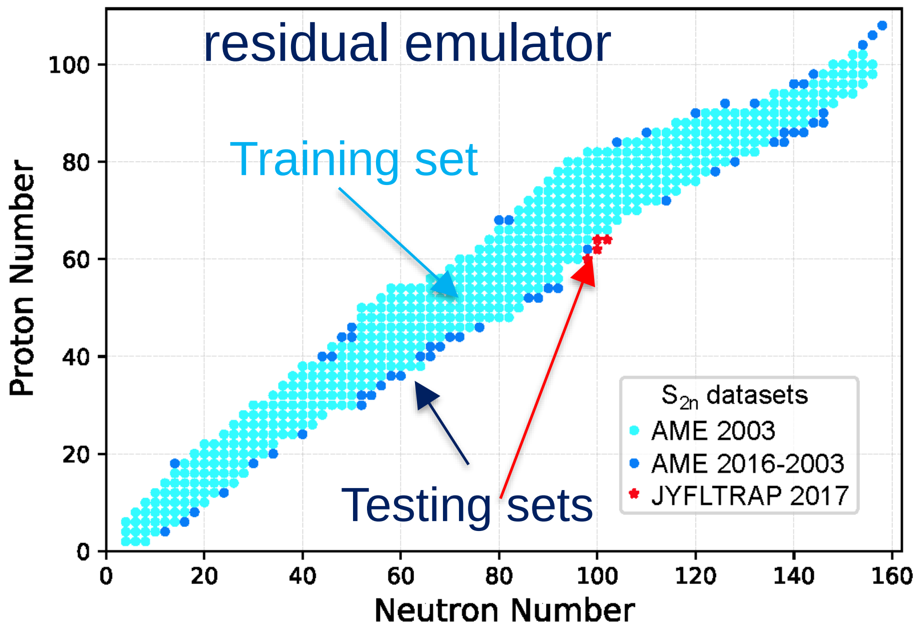
Mass extrapolations with BNN and GP
Neufcourt et al. Phys. Rev. C 98, 034318 (2018)

Models of different fidelity

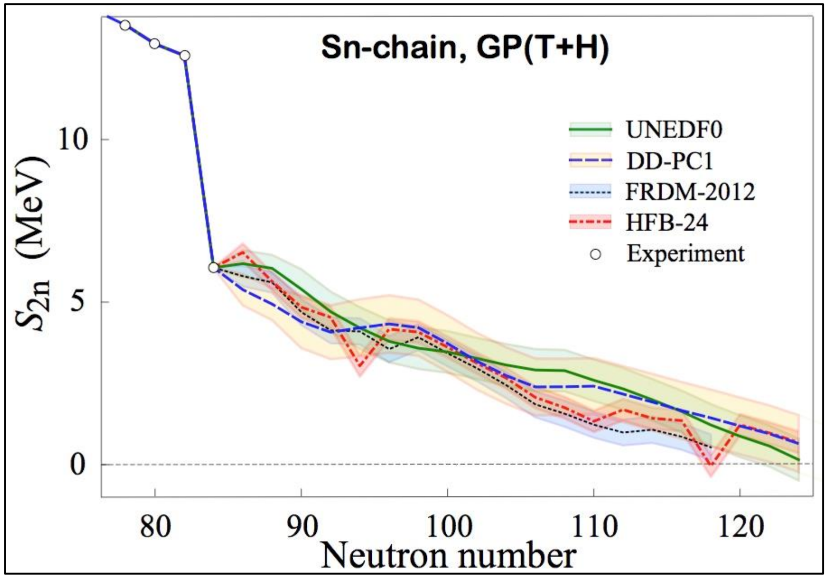


Residuals (based on data and theory) exhibit patterns

- This information can be used to our advantage to improve model-based predictions!
- It can also be used to improve models themselves



outcome

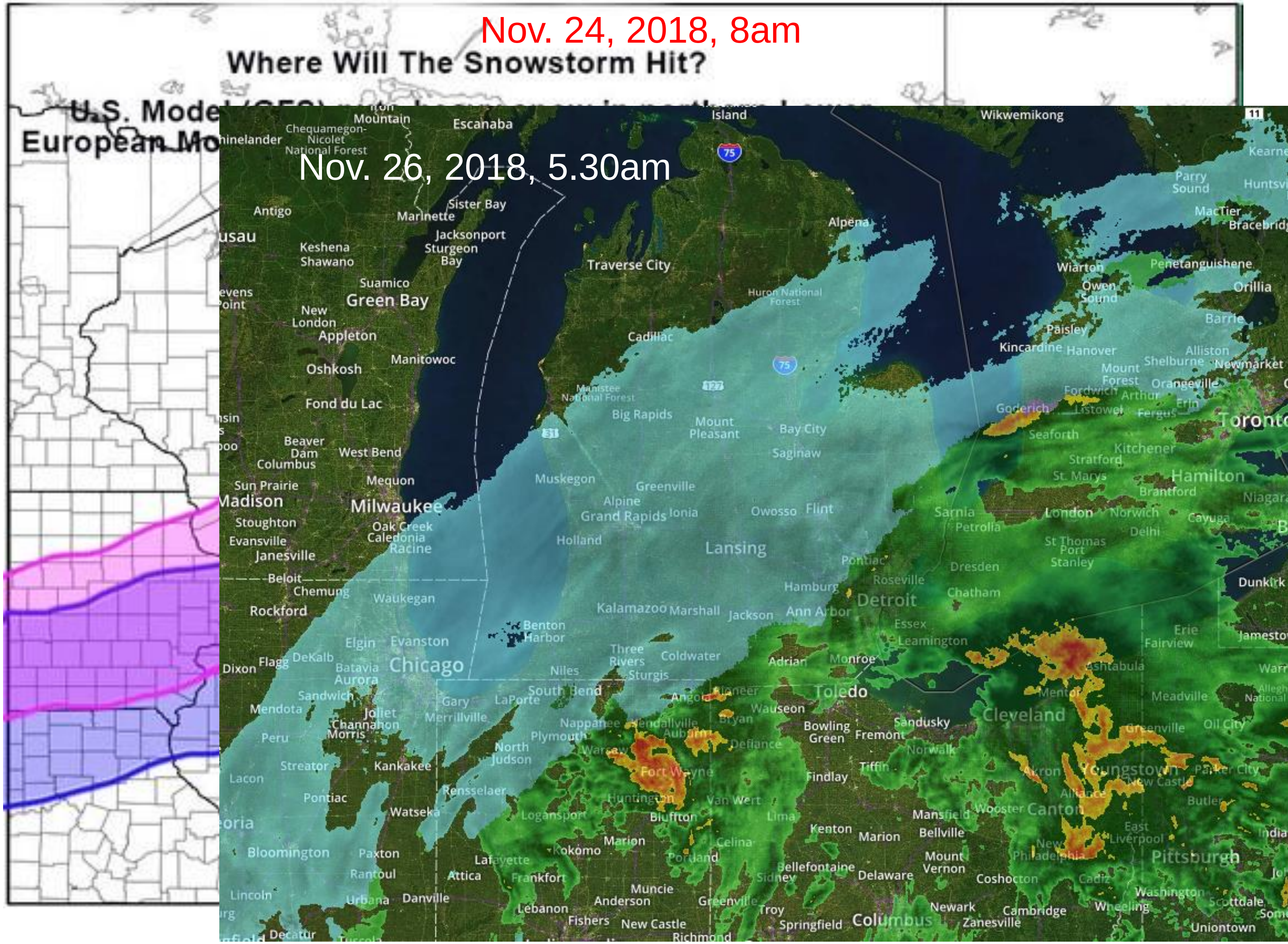


Naïve nuclear theorist's approach to a systematic (model) error estimate:

- Take a set of *reasonable* global models M_i, hopefully based on different assumptions/formalism, that satisfy basic theoretical requirements (here comes the expert belief thing).
- Make predictions.
- Compute average and variation within this set
- Compute rms deviation from existing experimental data.

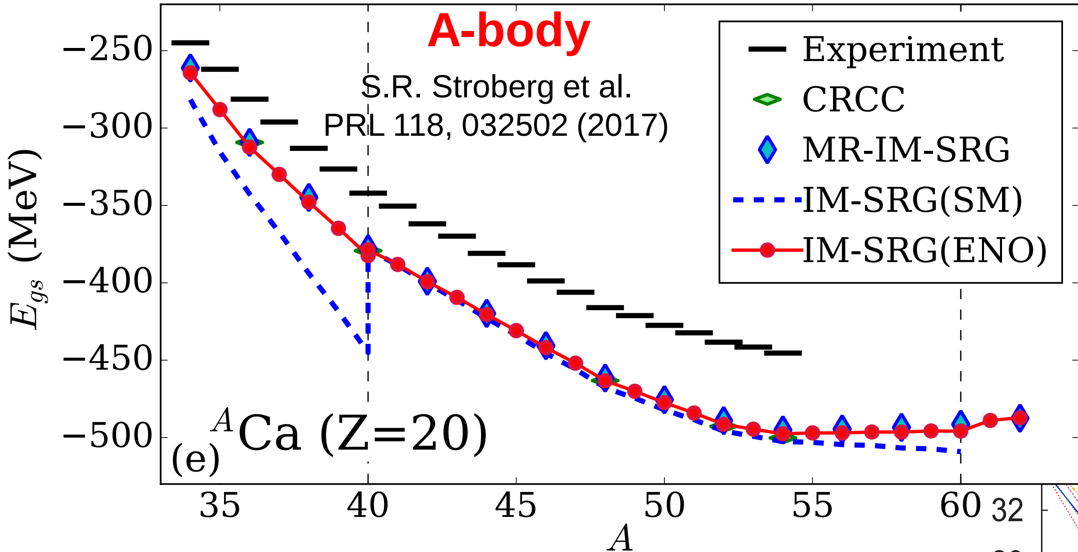
Can we do better? Yes!

Model mixing: $$p(\mathcal{M}_k|y) = \frac{p(y|\mathcal{M}_k)\pi(\mathcal{M}_k)}{\sum_{\ell=1}^{K} p(y|\mathcal{M}_\ell)\pi(\mathcal{M}_\ell)}$$


Nov. 24, 2018, 8am
Where Will The Snowstorm Hit?
Nov. 26, 2018, 5.30am
Green Bay
Milwaukee
Chicago
Traverse City
Lansing
Grand Rapids
Detroit
Toledo
Cleveland
Toronto
Hamilton
London
Pittsburgh
Columbus

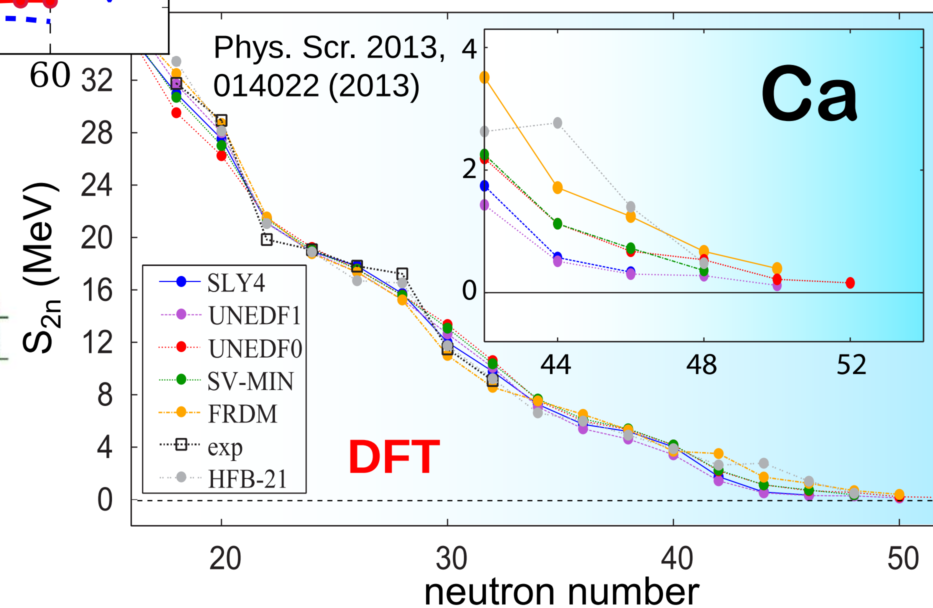
How many Ca nuclei exist?

Neufcourt et al.,Phys. Rev. Lett. 122, 062502 (2019)

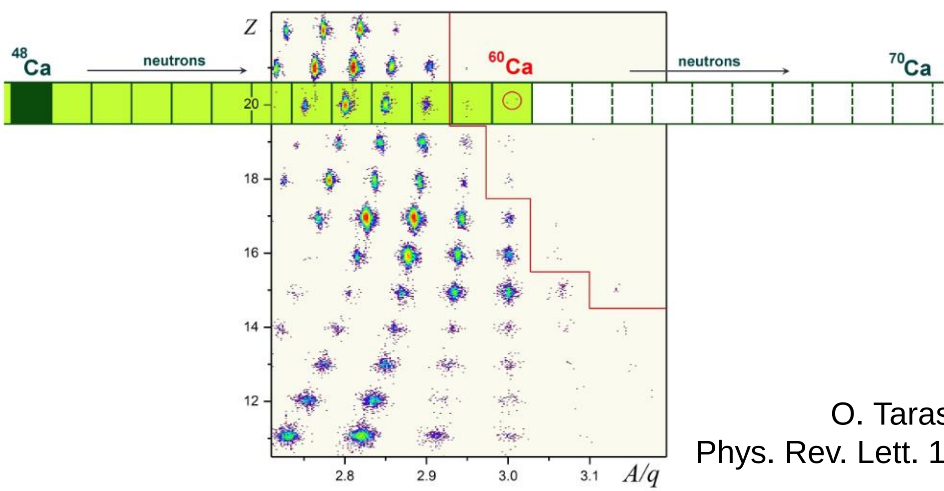


^{60}Ca weakly bound/unbound

Calcium isotopes bound out to about ^{70}Ca



Discovery of ^{60}Ca



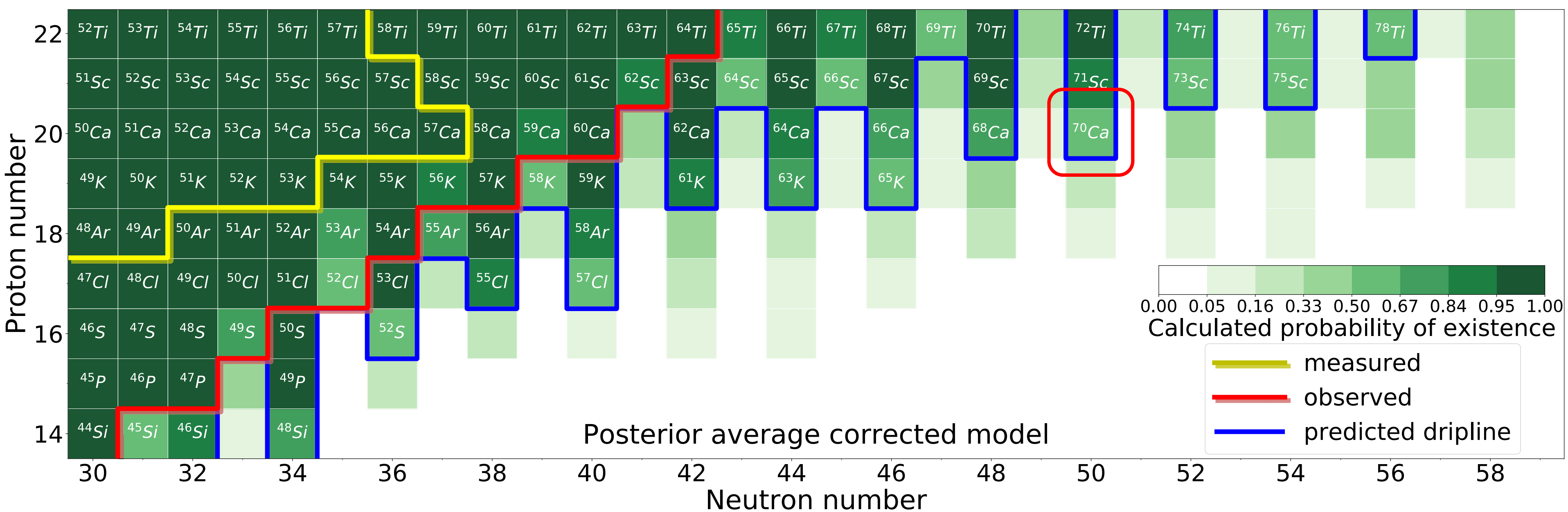
O. Tarasov et al.
Phys. Rev. Lett. 121, 022501 (2018)

Future: Quantified predictions with machine learning

Probability of existence

$$p_{ex}(Z,N) := p(S^*_{1n/2n}(Z,N) > 0 | S_{1n/2n})$$

Bayesian model mixing, see L. Neufcourt et al., Phys. Rev. Lett. 122, 062502 (2019)

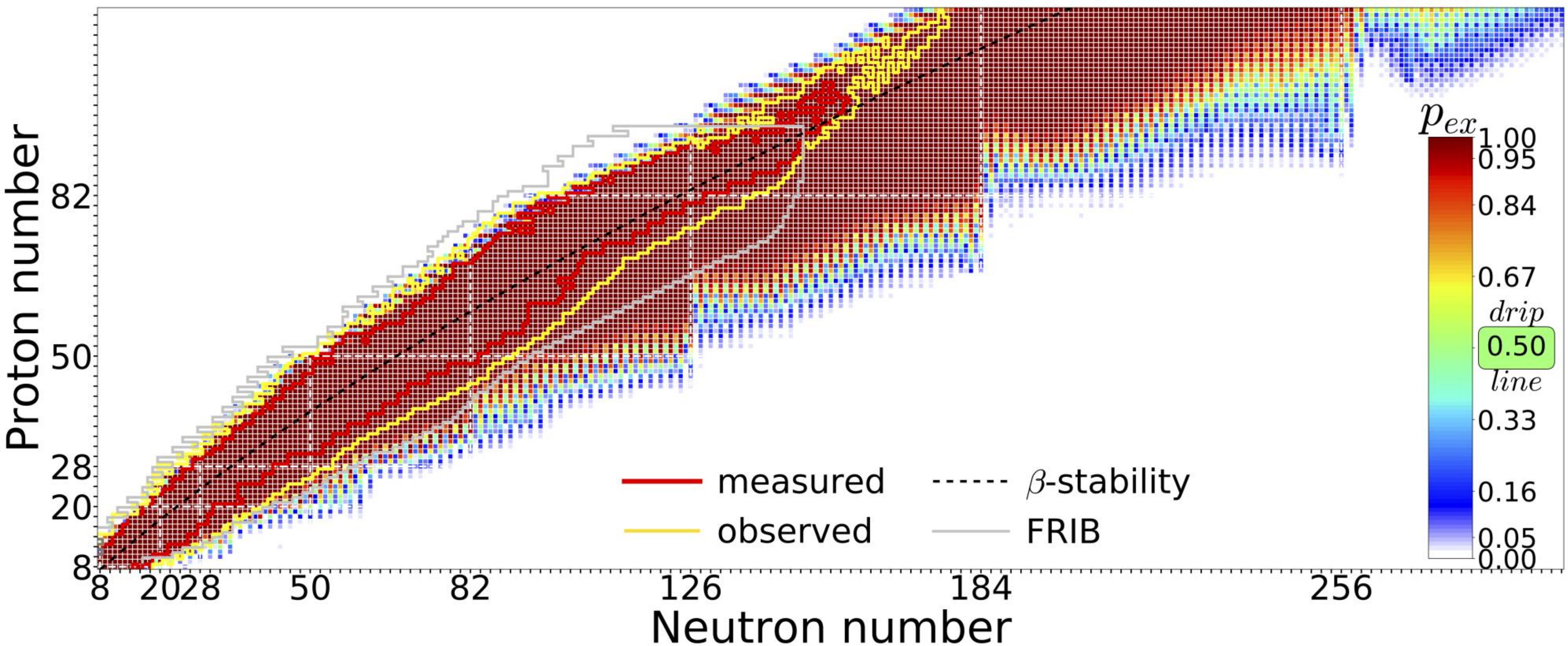


Quantified limits of the nuclear landscape

Neufcourt et al., Phys. Rev. C 101, 044307 (2020)

Predictions made with 11 global mass model and Bayesian model averaging

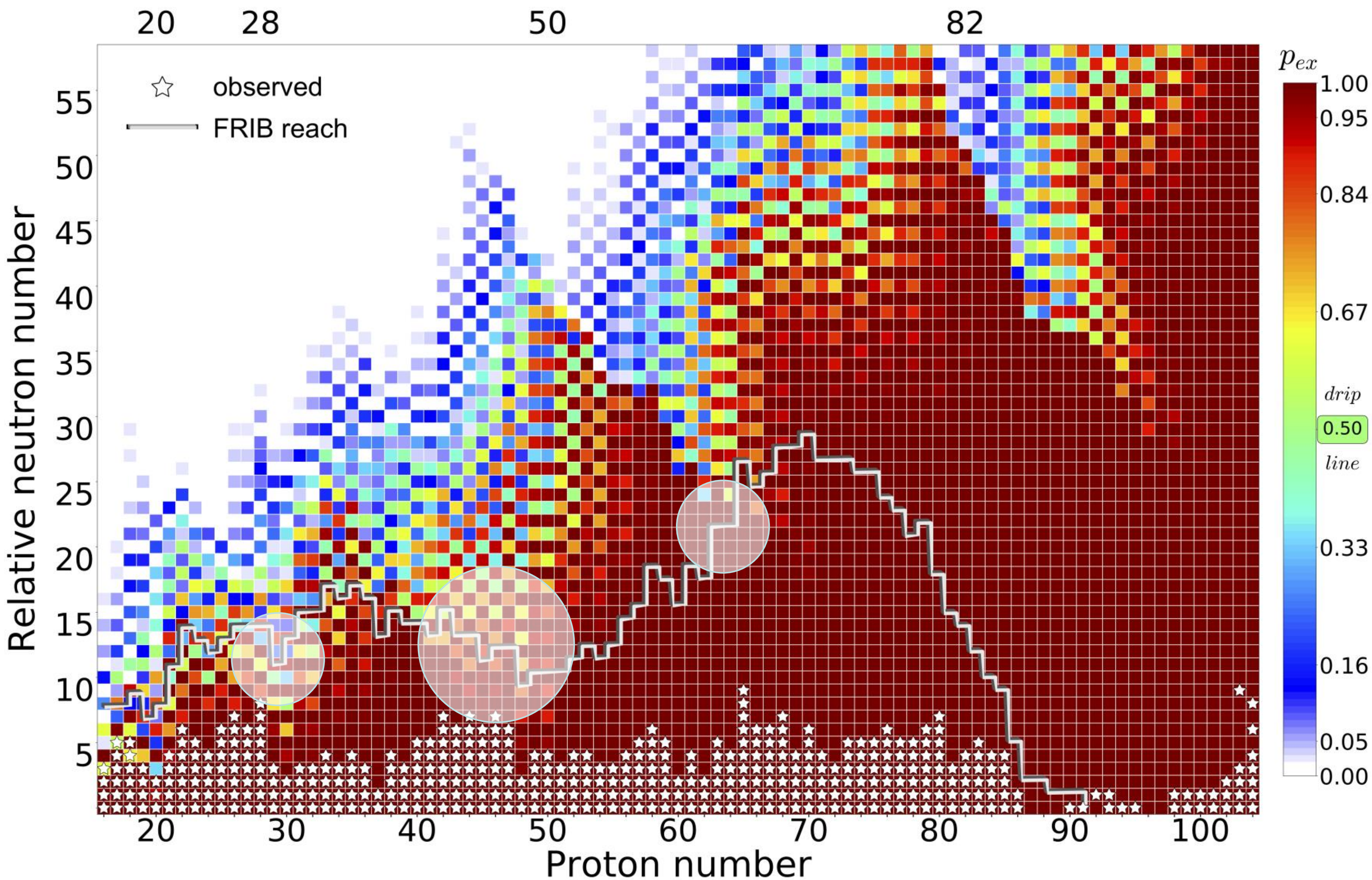
$$p_{\rm ex} := p(S^*_{1p/2p/1n/2n} > 0 | S_{1p/2p/1n/2n})$$



The FRIB production rates estimated with the LISE++ . We assumed the experimental limit for the confirmation of existence of an isotope to be 1 event/2.5 days.

INFORMATION NEEDED

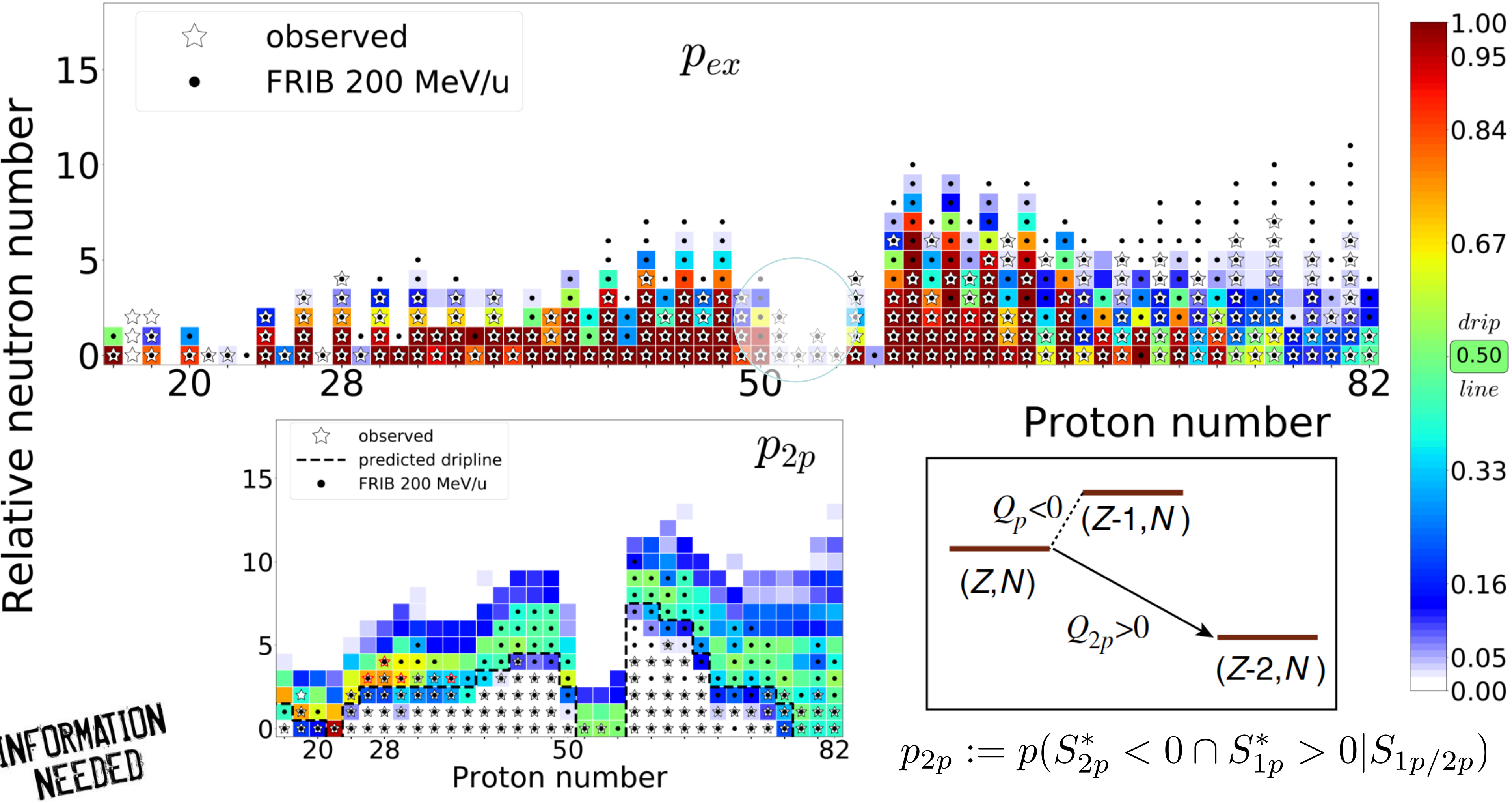
Of particular importance for constraining theory are the existence data for Z=28-30, Z=42-48, and Z=64-66



“0” corresponds is the neutron number of the heaviest isotope for which an experimental separation energy value is available

Proton drip line and beyond: Bayesian analysis of proton-emitting nuclei

Neufcourt et al., Phys. Rev. C 101, 014319 (2020)



$$p_{2p} := p(S^*_{2p} < 0 \cap S^*_{1p} > 0 | S_{1p/2p})$$

The most promising new candidates for the true 2p radioactivity are: ^{30}Ar, ^{34}Ca, ^{39}Ti, ^{42}Cr, ^{58}Ge, ^{62}Se, ^{66}Kr, ^{70}Sr, ^{74}Zr, ^{78}Mo, ^{82}Ru, ^{86}Pd, ^{90}Cd, and ^{103}Te.

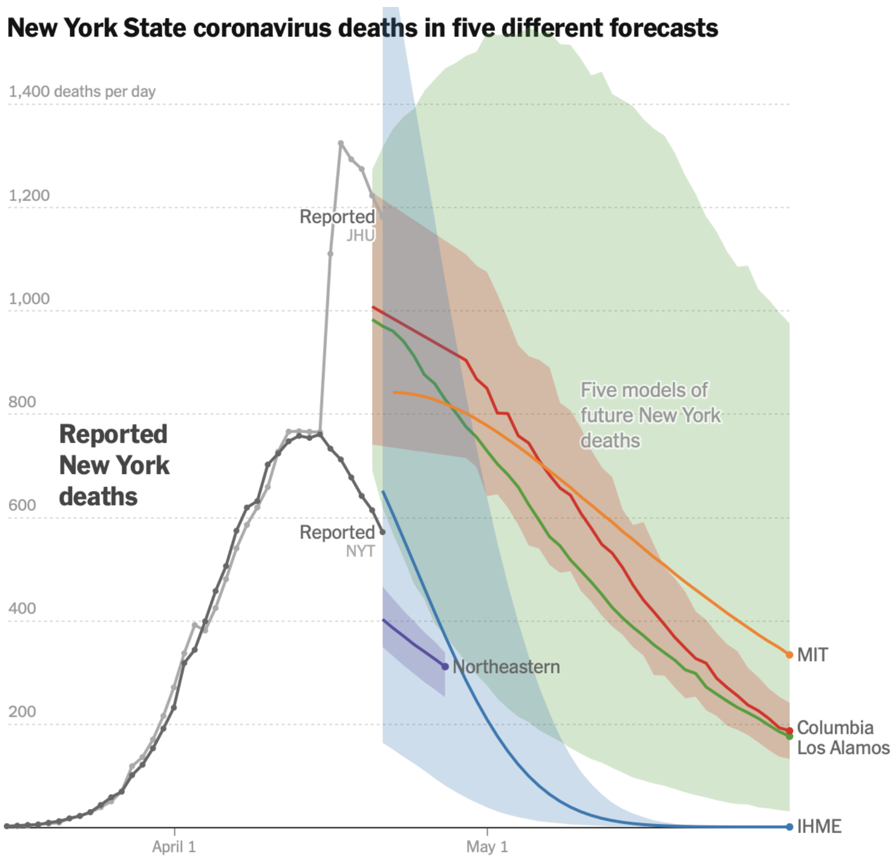
The New York Times

TheUpshot

What 5 Coronavirus Models Say the Next Month Will Look Like

By Quoctrung Bui, Josh Katz, Alicia Parlapiano and Margot Sanger-Katz April 22, 2020

New York State coronavirus deaths in five different forecasts



“They’re basing their model on very few data, and because of that you have very large uncertainty”

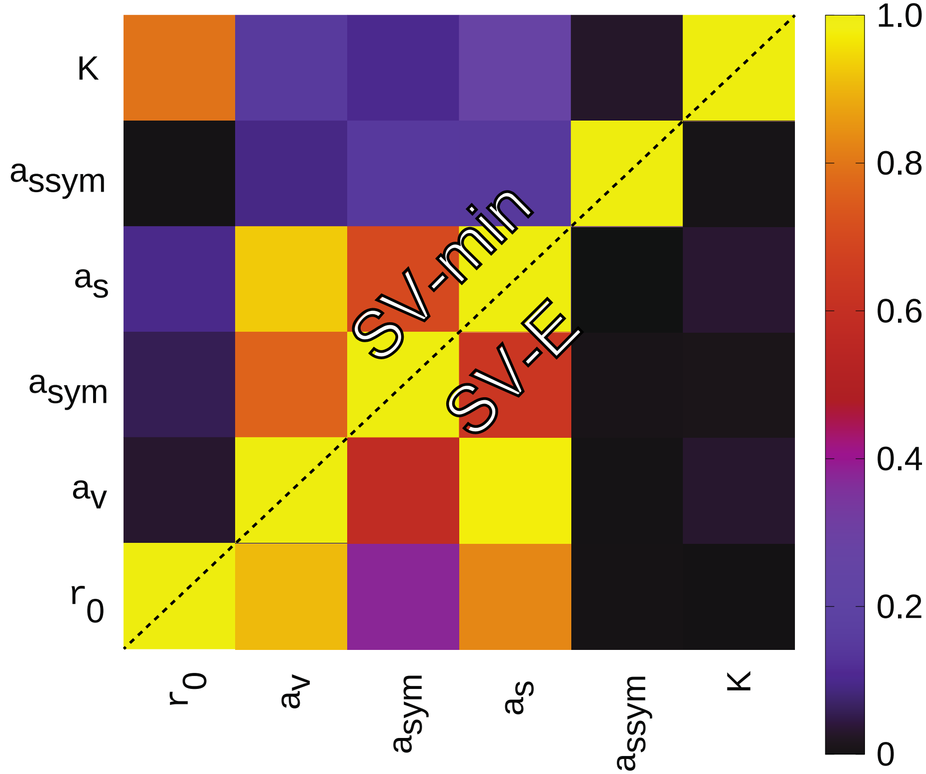
“If you’re just looking at one model, you’re not seeing the full diversity of what could happen”

Statistical aspects of nuclear models

Kejzlar et al., J. Phys. G (2000) arXiv:2002.04151

SV-min: informed by masses, sizes, pairing gaps
SV-E: informed by masses only

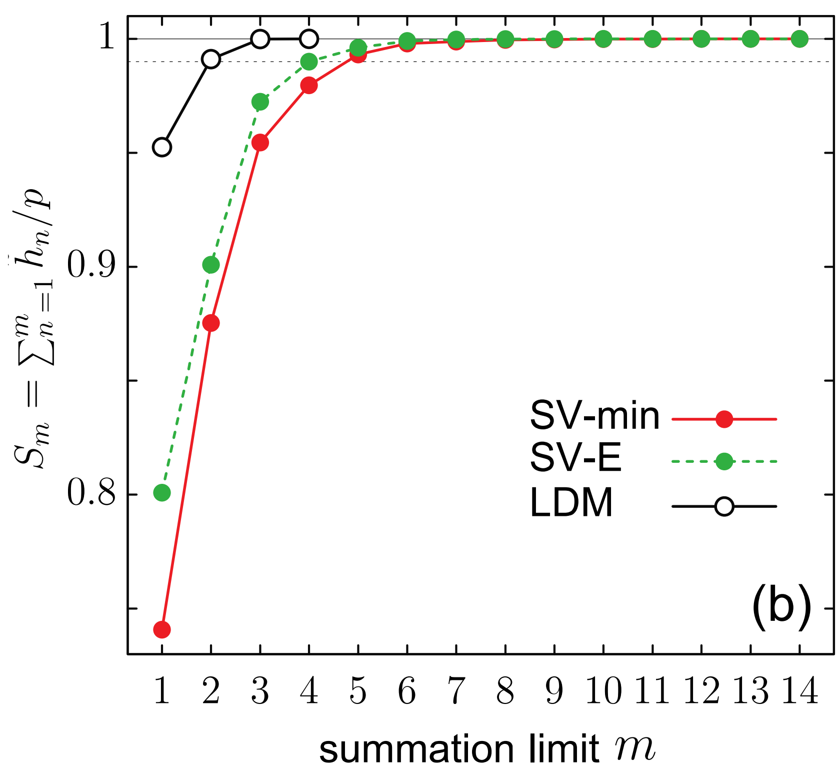
correlation matrix



Conclusion: correlations between parameters and observables strongly depend on dataset of fit-observables.

ρ_0 determined by charge radii data

principal component analysis



Conclusion: effective number of degrees of freedom is 4-6 for the 14-parameter Skyrme functional. Long way to go!

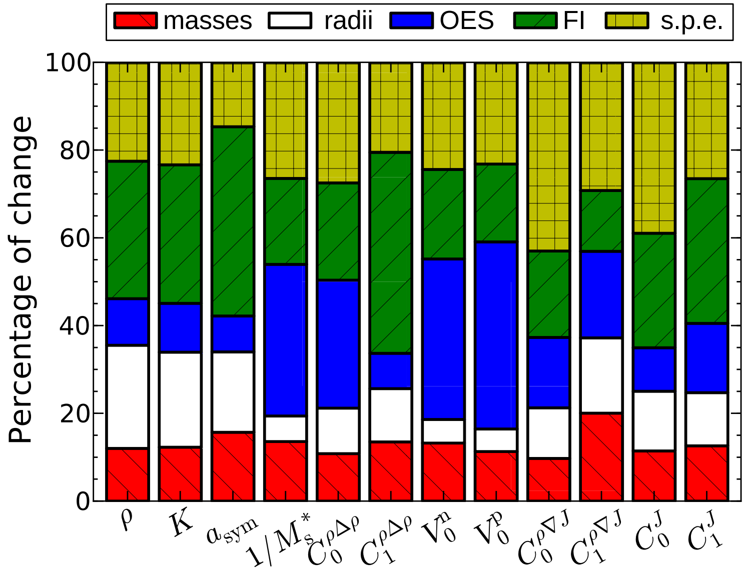
See also G.F. Bertsch et al, Phys. Rev. C 71, 054311 (2005); Phys. Rev. Lett. 119, 252501 (2017)

Sensitivity studies, model calibration

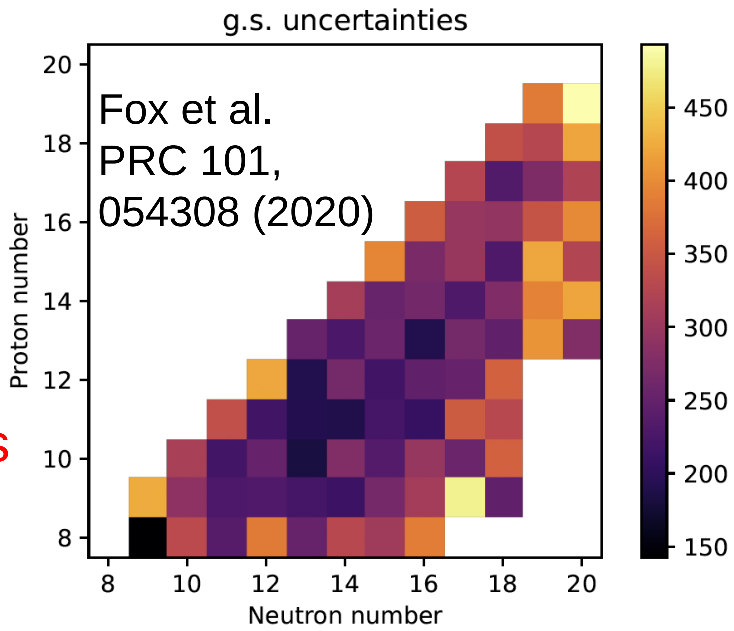
Identify key data to constraint models; Understanding model structure

Kortelainen et al. PRC 89, 054314 (2014)

Energy density functional

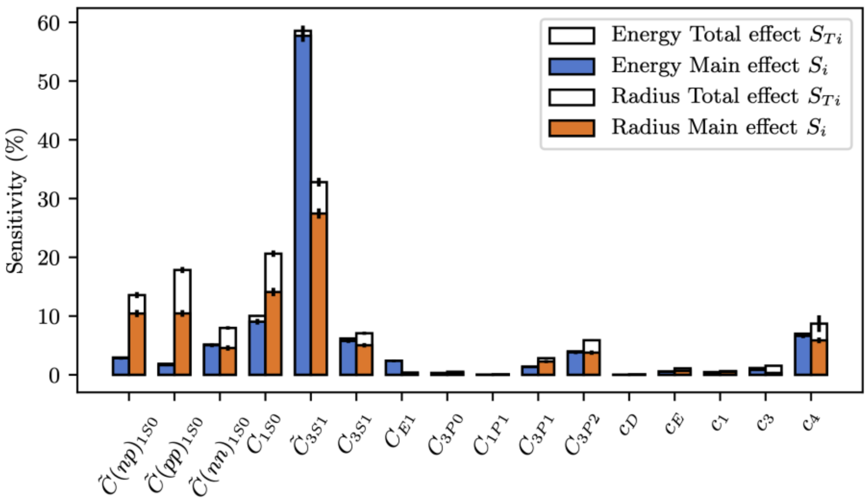


Shell-model matrix elements



Ekström and Hagen
PRL 123, 252501 (2019)

NNLO chiral-EFT Hamiltonian



Collaborators

Physics

Y. Cao

J. Dobaczewski

G. Hagen

A. Ekström

D. Furnstahl

S. Gandolfi

M. Hjorth-Jensen

Y. Jaganathen

M. Kortelainen

D. Lee

J. McDonnell

Th. Papenbrock

D. Phillips

P.-G. Reinhard

N. Schunck

…

Statistics

D. Higdon

V. Kejzlar

L. Neufcourt

F. Viens

…

Applied math/CS

J. O'Neal

J. Sarich

S. Wild

…

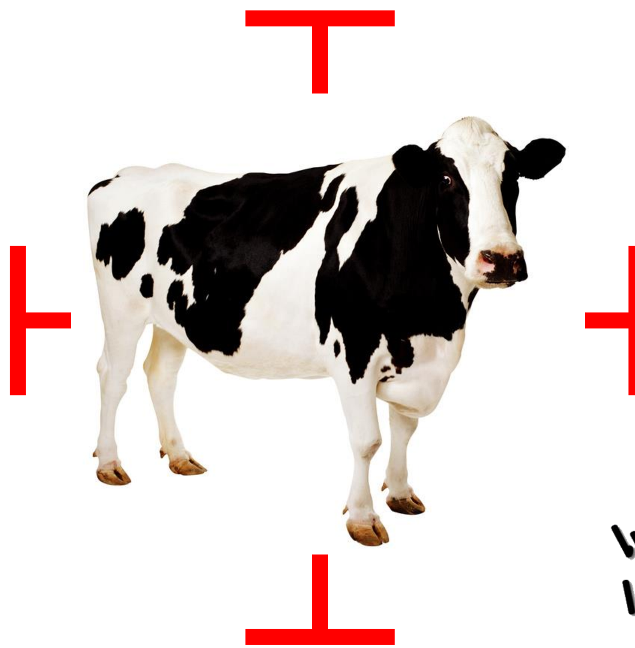
A series of annual meetings on *Enhancing the interaction between nuclear experiment and theory through information and statistics (ISNET):* Kraków, Glasgow, ECT*, **INT**, York, Darmstadt, and Göteborg.

ISNET-8 will be held (?) at FRIB Dec. 14-17th, 2020:

https://indico.frib.msu.edu/event/21/

Summary

- Many exciting results achieved but the best is yet to come!
- To solve many complex problems in the field and facilitate discoveries, multidisciplinary efforts efforts are required involving scientists in nuclear physics, computational science, applied math, and statistics.
- We need to invest in relevant educational efforts.



Thank You

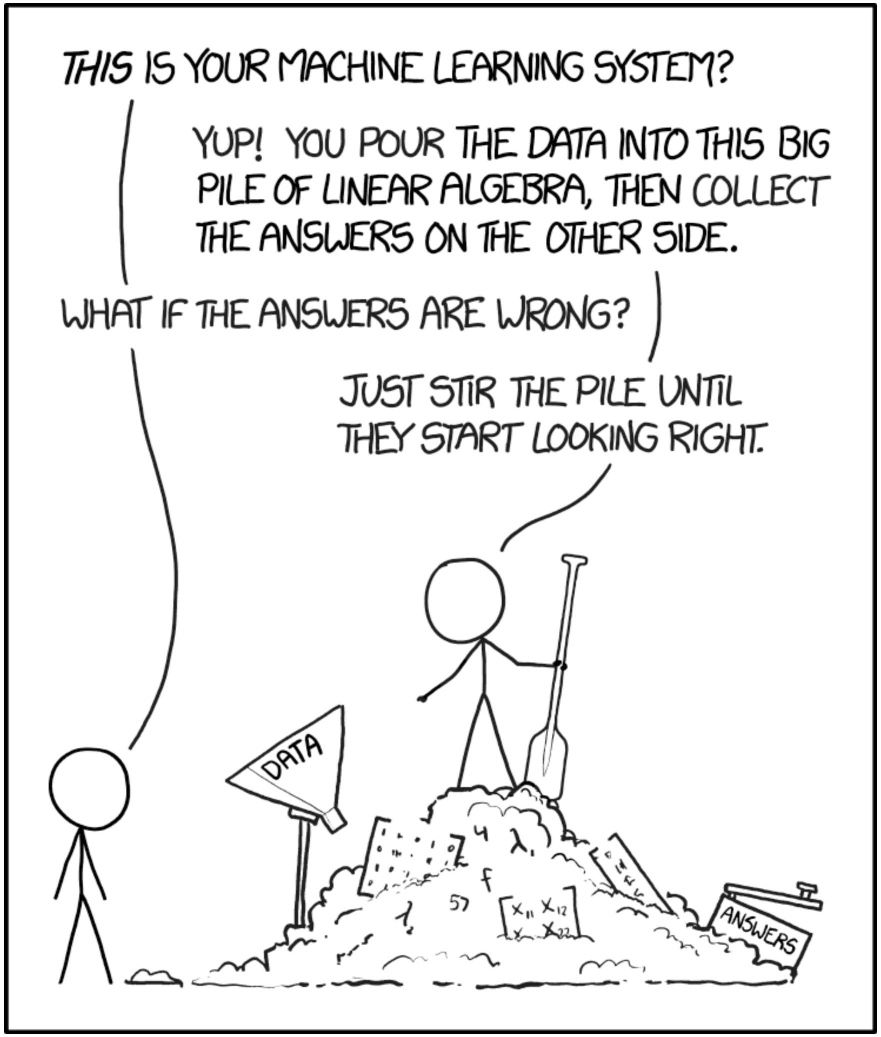
WORK IN PROGRESS

Backup

A disclaimer…

What is “microscopic”? In our field, theory is useful at a quantitative level only if the parameters, or coupling constants, of models are optimized to experiment. For that reason, *all quantitative nuclear models are phenomenological at some level*. Superlatives such as ‘fully microscopic’ or ‘from first principles’, sometimes used to characterize particular methods, may be viewed more as aspirational than the present reality. However, it is useful to distinguish the degrees of phenomenology in different theoretical approaches. The term “microscopic theory” is often used for theoretical approaches in which nucleonic degrees of freedom are explicitly present together with inter-nucleon forces. In this respect, *A*-body, CI, and DFT models all belong to this group; they of course differ in their resolution scales.

(from *Future of Nuclear Fission* 2020)



https://xkcd.com/1838/

From Hoeting and Wasserman:

When faced with several candidate models, the analyst can either choose one model or average over the models. Bayesian methods provide a set of tools for these problems. Bayesian methods also give us a numerical measure of the relative evidence in favor of competing theories.

- *Model selection* refers to the problem of using the data to select one model from the list of candidate models. *Model averaging* refers to the process of estimating some quantity under each model and then averaging the estimates according to how likely each model is.
- Bayesian model selection and model averaging is a conceptually simple, unified approach. An intrinsic Bayes factor might also be a useful approach.
- *There is no need to choose one model*. It is possible to average the predictions from several models.
- Simulation methods make it feasible to compute posterior probabilities in many problems.

It should be emphasized that BMA should not be used as an excuse for poor science... BMA is useful after careful scientific analysis of the problem at hand. Indeed, BMA offers one more tool in the toolbox of applied statisticians for improved data analysis and interpretation.