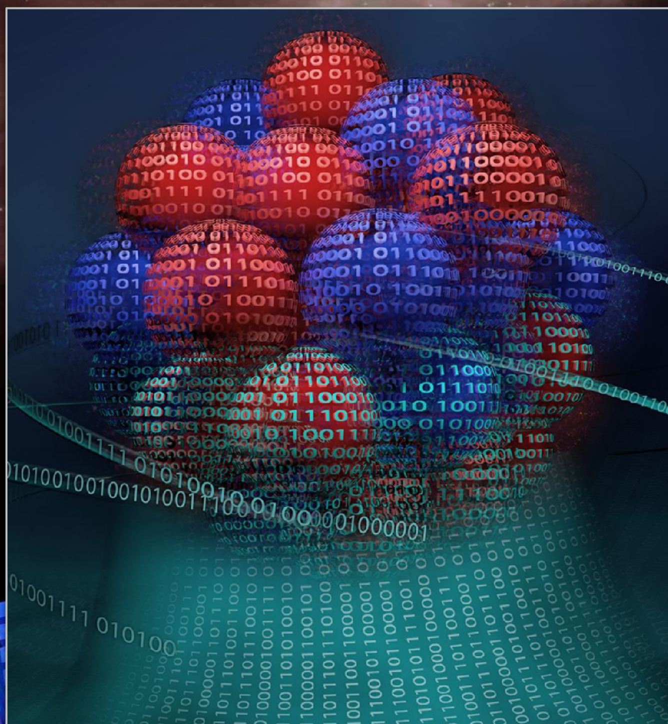
Bayesian Framework for Mining of Evaluated Nuclear Mass Data

Presented by: Kyle Godbey



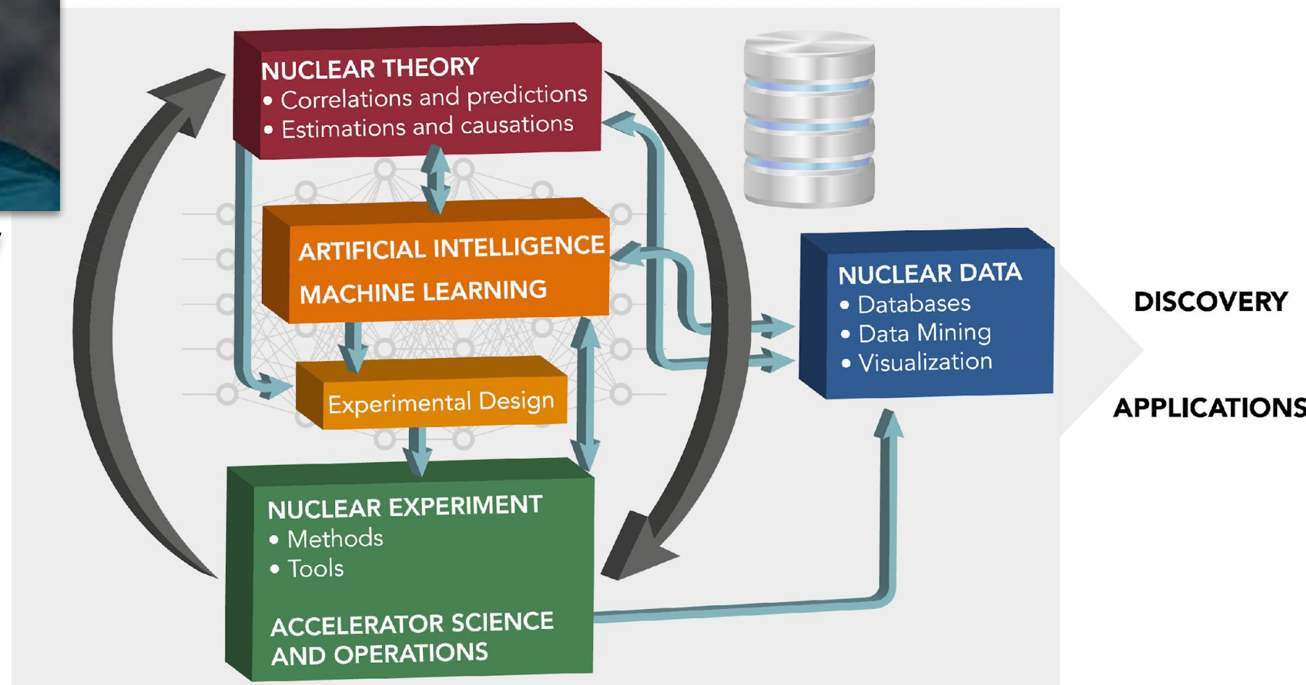
This material is based upon work supported by the US Department of Energy, Office of Science, Office of Nuclear Physics under Grant No. DE-SC0023688

Our team: nuclear theory, nuclear experiment (nuclear data, nuclear astrophysics), statistics

Witek Nazarewicz (PI)

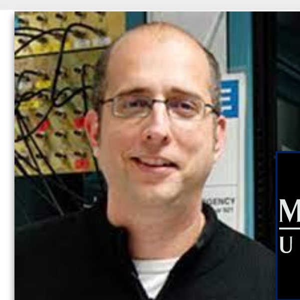
Kyle Godbey

MICHIGAN STATE UNIVERSITY



Filip Kondev

Argonne NATIONAL LABORATORY



Hendrik Schatz

MICHIGAN STATE UNIVERSITY

Our collaborators

BAND
Bayesian Analysis of Nuclear Dynamics

NUCLEI
Nuclear Computational Low-Energy Initiative

IJCLab
Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

IMP
中国科学院近代物理研究所
Institute of Modern Physics, Chinese Academy of Sciences

International Atomic Energy Agency
Nuclear Data Services
Provided by the Nuclear Data Section

U.S. Nuclear Data Program

What, Why, How?

The main objective: to quantify the nuclear binding by employing state-of-the-art global nuclear models, the most current nuclear mass data, and Bayesian machine learning that enables us to make quantified predictions (⇒ extrapolations) by building on the most relevant theoretical and experimental information.

Why? Atomic Mass Evaluation recommends mass values and their uncertainties for all nuclei. AME also provides masses for nuclei that are not-yet observed and for nuclei where the experimental data are poorly known by extrapolating data from neighboring known nuclei. Improvements in the value of estimated masses and their uncertainties provided by us will benefit not only AME, but also other related nuclear data, such as nuclear cross sections, half-lives and delayed particle emission probabilities, which are essential input ingredients in nucleosynthesis simulations. Our extrapolated masses will guide experimental efforts in various areas of rare-isotope science and provide benchmarks for nuclear models.

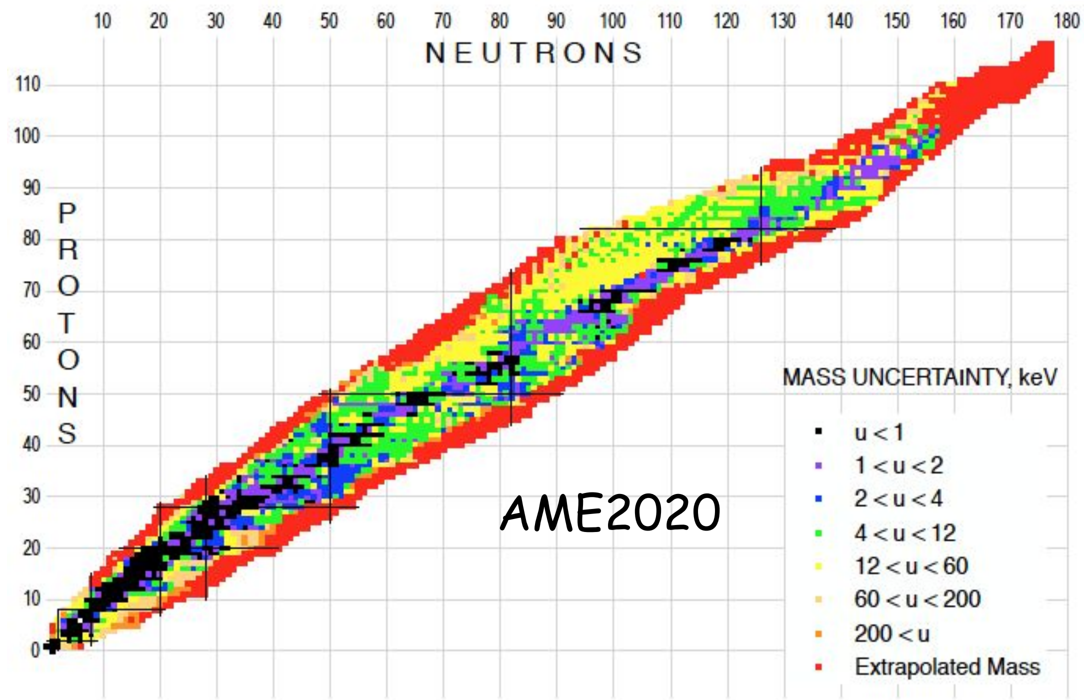
Tools: use nuclear density functional theory (DFT) and statistical Bayesian machine learning (Bayesian Model Mixing) to provide quantified theoretical predictions.

Member of the US Nuclear Data Program - DOE/SC/NP

ANL component: provide Nuclear Data input Atomic Mass Evaluation & NuBASE

- provide updated values for the atomic masses and other basic nuclear properties for all nuclei where experimental data exist

- provide extrapolated mass values for a limited number of nuclei where experimental data are limited or do not exist

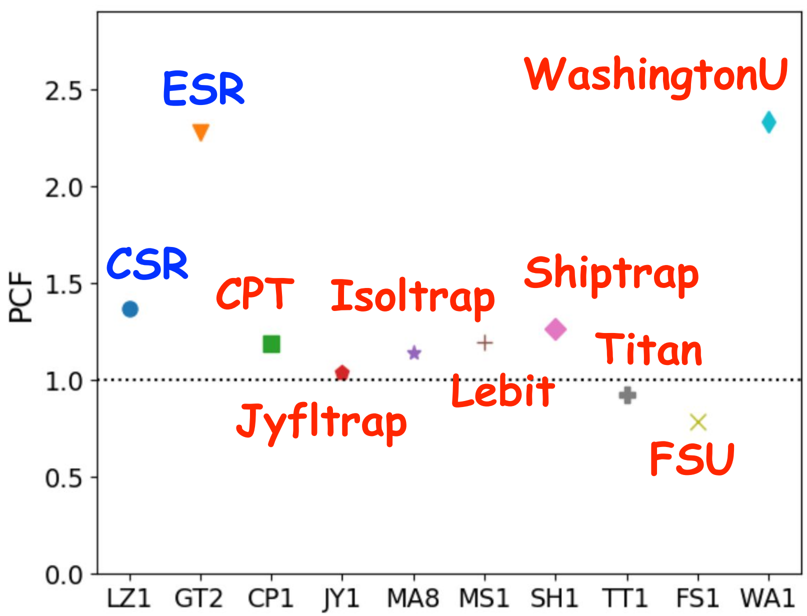


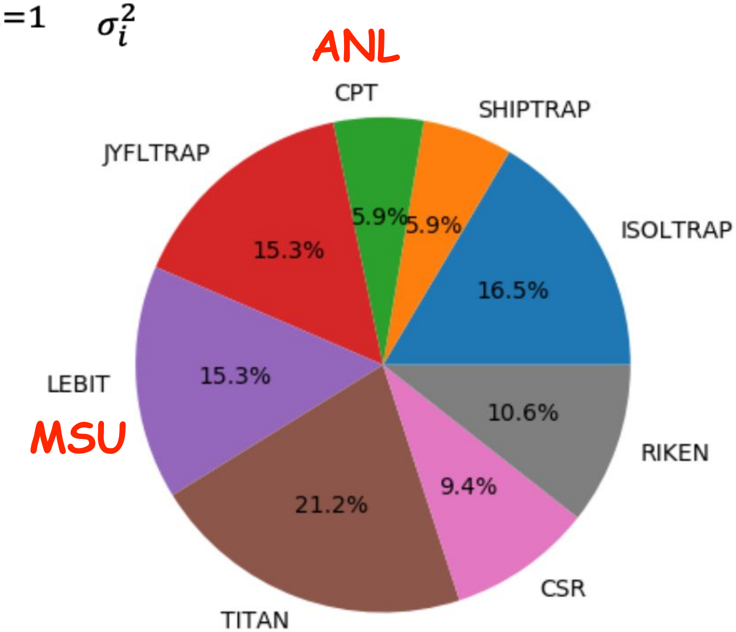
Which data are considered

- **Direct (mass spectrometry) and Indirect (reactions and decays) Data produced worldwide**
  - TOF & MR-TOF, Storage Rings & Penning Traps
  - Decay Energies in β^-, β^+, α and p decays – far from stability

⇒ critically evaluate the experimental data & combine the accepted ones using the least-squares fit approach -> mass values & covariances for all known nuclei

Partial Consistency Factor (PCF): $\chi^p = \sqrt{\frac{Q}{Q-M}\frac{1}{p}\sum_{i=1}^{p}\frac{(q_i-\bar{q})^2}{\sigma_i^2}}$





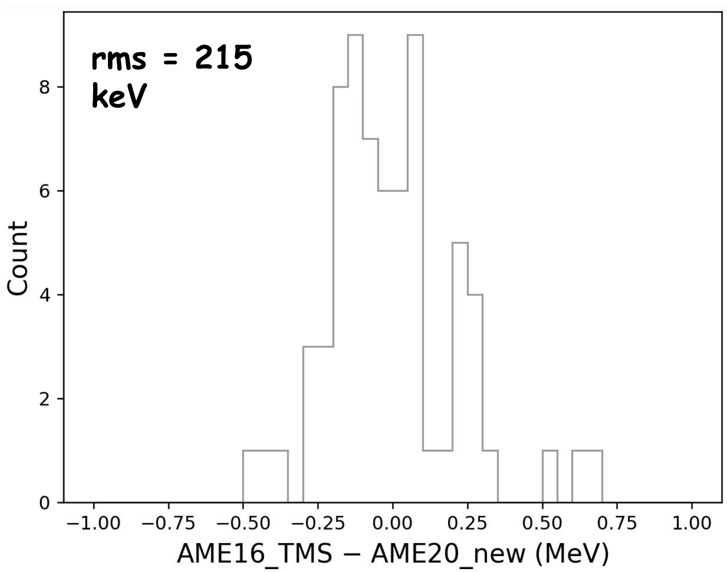
published mass data since AME2016

AME extrapolations

⇒ using an empirical approach by assuming that the Trend of the Mass Surface (TMS) is smooth

- TMS extrapolated mass values for a limited number of unknown nuclei
- replace "irregular" experimental masses by TMS extrapolated values – 77 cases in AME2020

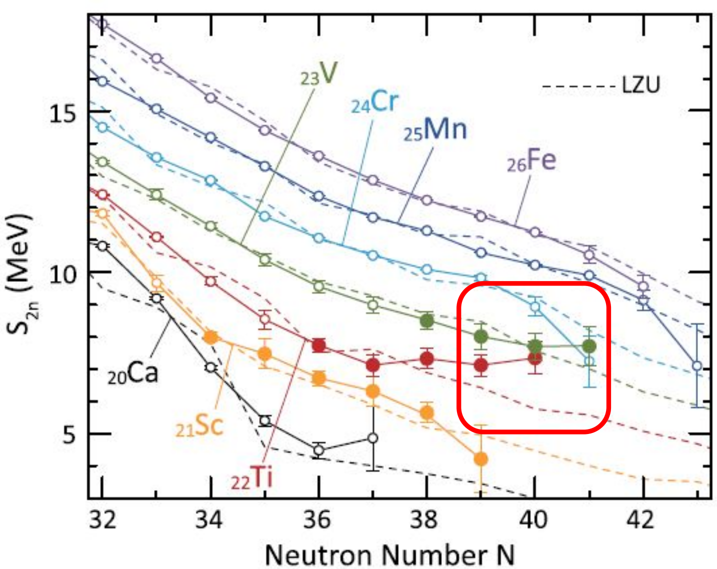
accuracy of the AME extrapolation



TMS in AME2016, BUT exp in AME2020

not always justified … new physics?

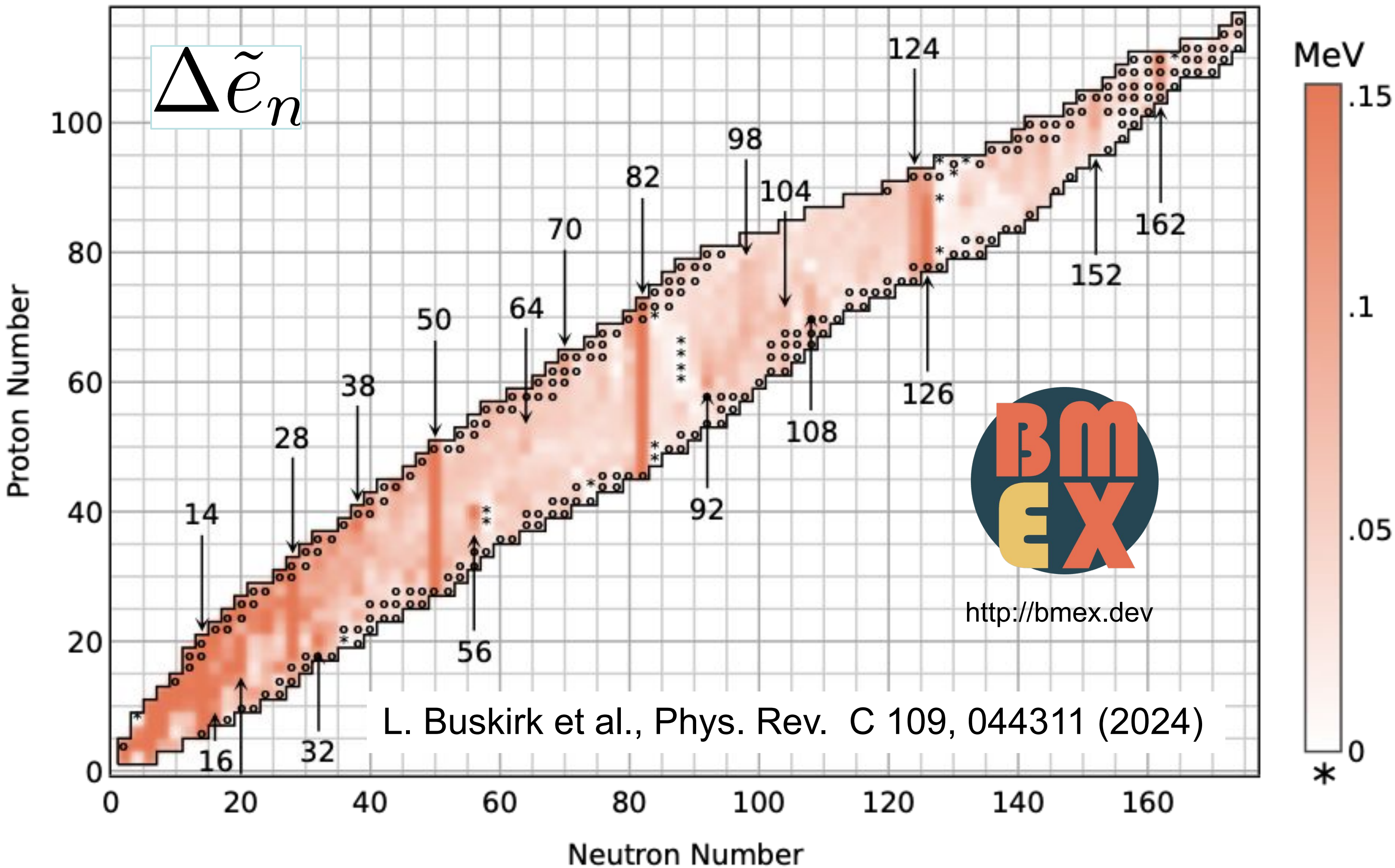
S. Michimasa et al., PRL125 (2020) 122501



build up of deformation around N=40

Nucleonic shells from nuclear masses

$$\Delta e_n(N=2k,Z)=S_n(N,Z)-S_n(N+2,Z)$$
$$\Delta e_n(N=2k,Z)\approx e^n_{k+1}-e^n_k$$



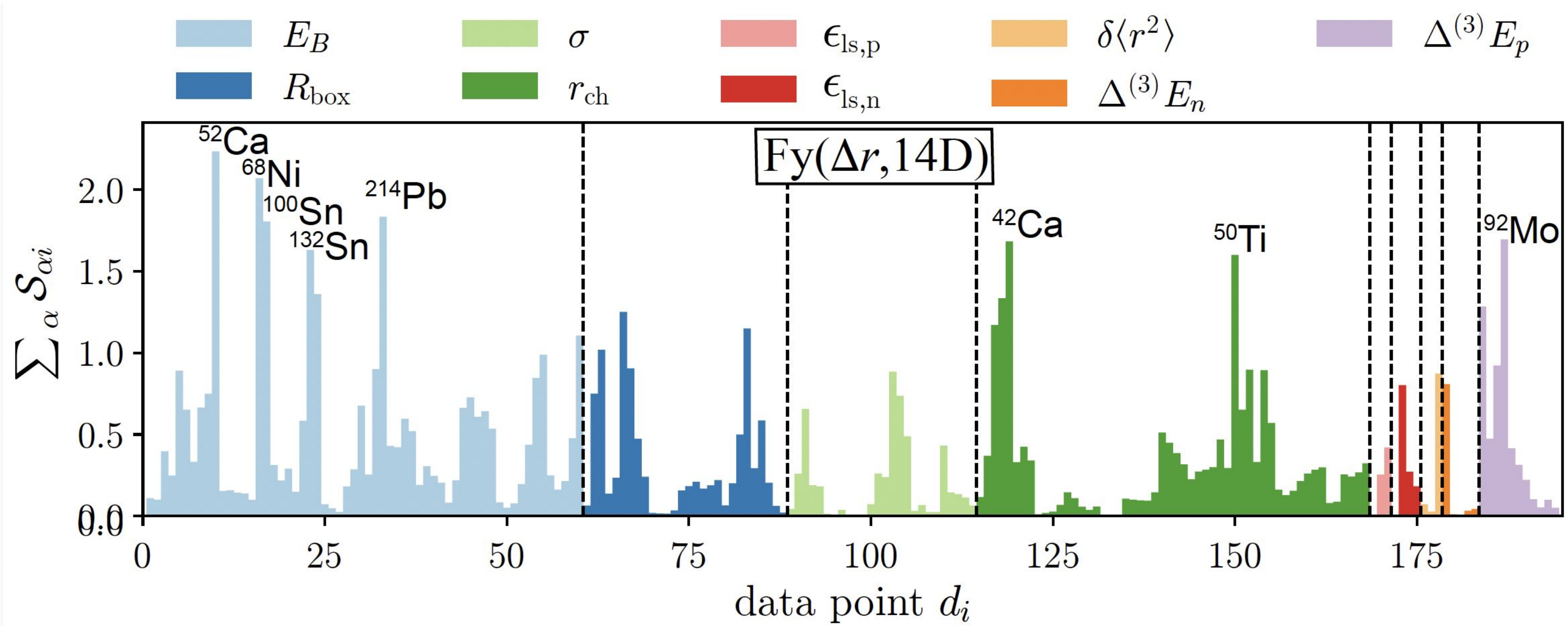
L. Buskirk et al., Phys. Rev. C 109, 044311 (2024)

Identifying key experimental data for model calibration

Extended Fayans energy density functional: optimization and analysis

Paul-Gerhard Reinhard, Jared O'Neal, Stefan M Wild and Witold Nazarewicz

Published 21 August 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Journal of Physics G: Nuclear and Particle Physics, Volume 51, Number 10
Citation Paul-Gerhard Reinhard *et al* 2024 *J. Phys. G: Nucl. Part. Phys.* **51** 105101
DOI 10.1088/1361-6471/ad633a



The total impact of a data point on the parameters of the functional

Many mass models based on machine learning!

See examples in Rev. Mod. Phys. Rev. Mod. Phys. 94, 031003 (2022)

- Physically interpretable machine learning for nuclear masses, M. R. Mumpower et al., C Phys. Rev. C 106, L021301 (2022)
- Controlling extrapolations of nuclear properties with feature selection, R. Navarro Perez and N. Schunck, Phys. Lett. B833, 137336 (2022)
- Predicting nuclear masses with product-unit networks, B. Dellen et al., Phys. Lett. B, 852, 138608 (2024)
- Uncertainty Quantification of Mass Models using Ensemble Bayesian Model Averaging, Y. Saito et al., Phys. Rev. C 109, 054301 (2024)
- …

There's still room to improve!

Local Bayesian Dirichlet mixing of imperfect models

Vojtech Kejzlar, Léo Neufcourt & Witold Nazarewicz

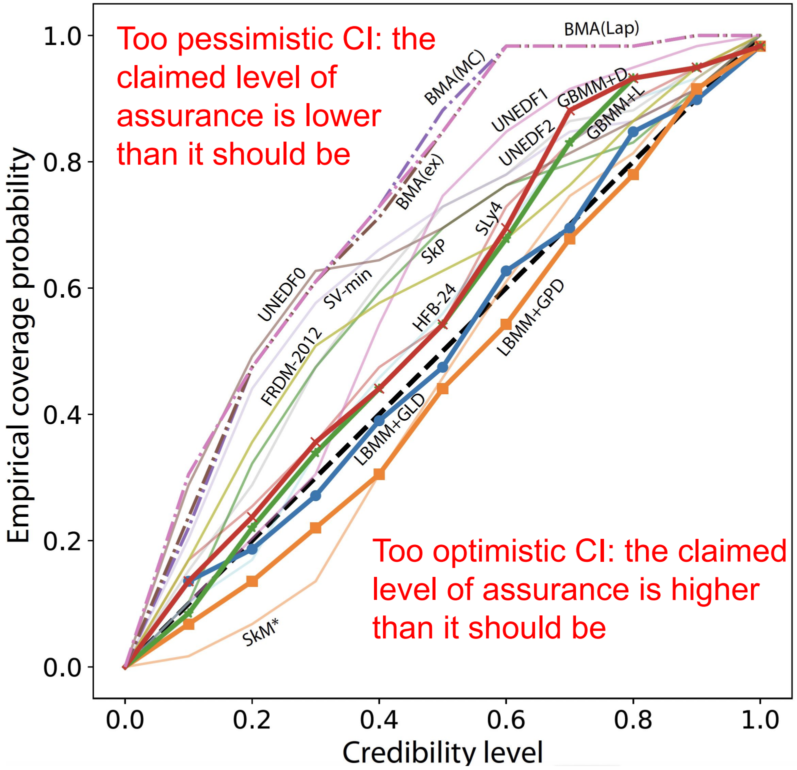
Scientific Reports **13**, Article number: 19600 (2023) | Cite this article

SKIDMORE COLLEGE

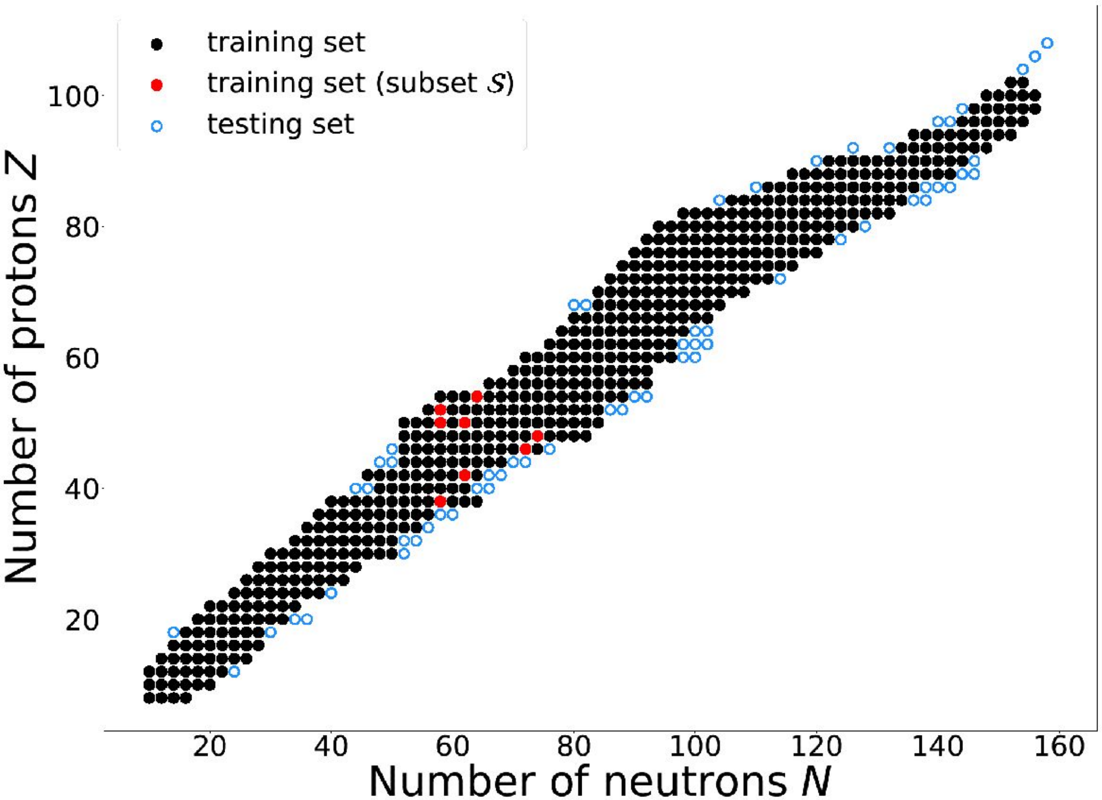
We proposed a Bayesian statistical machine learning framework utilizing the (hierarchical) Dirichlet distribution that combines the results of several imperfect models. To illustrate the method, we studied the ability of Bayesian model averaging and mixing techniques to mine nuclear masses.

The global and local mixtures of models reach excellent performance on both prediction accuracy and uncertainty quantification and are preferable to classical Bayesian model averaging. We also show that improving model predictions through mixing rather than mixing of corrected models leads to more robust extrapolations

ECP: the fraction of predictions which should theoretically fall in a confidence interval (CI) centered around the posterior mean prediction, for any given interval width.



Local Bayesian Model Mixing: Short-range extrapolations



Two-neutron separation energies

	Test RMSE (MeV)
Best Model	0.40
Best Model (with a GP model for residuals)	0.38
BMA (global)	0.37
LBMM (local)	**0.33**

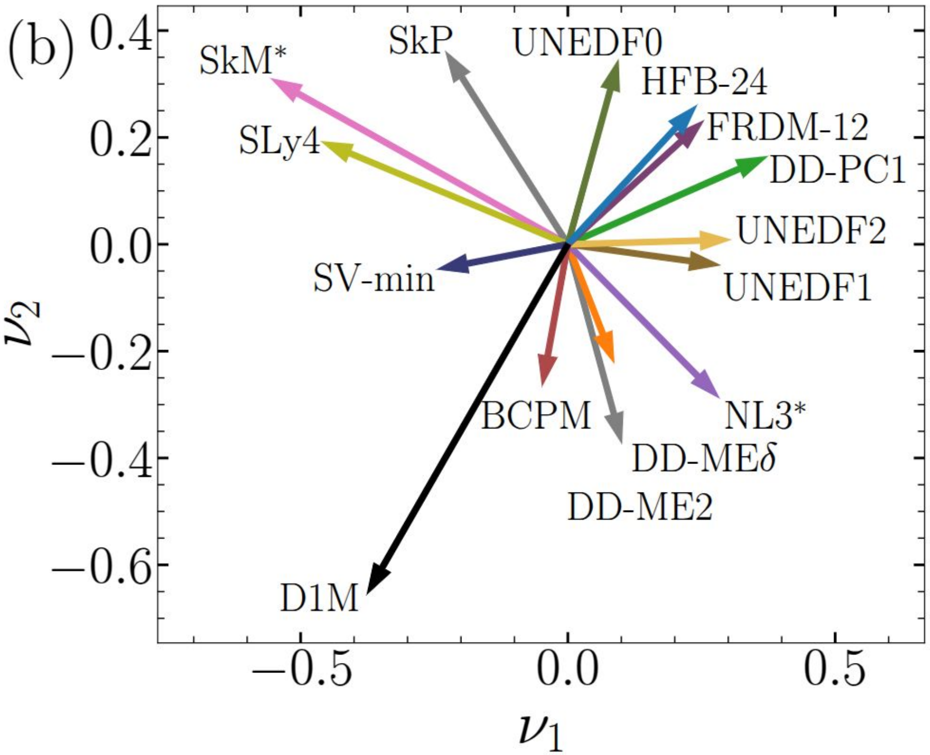
- **21%** improvement to the best theoretical model, **15%** improvement to the best model with GP residuals, **12%** improvement to BMA (global)
- **Future LBMM refinements:** Dirichlet process for $\omega(x)$, model pre-selection

In many cases physics models may have a similar mathematical foundation but their parameters are calibrated using different methodologies. It is also possible that models are identical despite their different formulations. How to eliminate model redundancy?

Model orthogonalization and Bayesian mixing of nuclear mass models via Principal Component Analysis

P. Giuliani, K. Godbey, V. Kejzlar, W. Nazarewicz, Phys. Rev. Research 6, 033266 (2024)

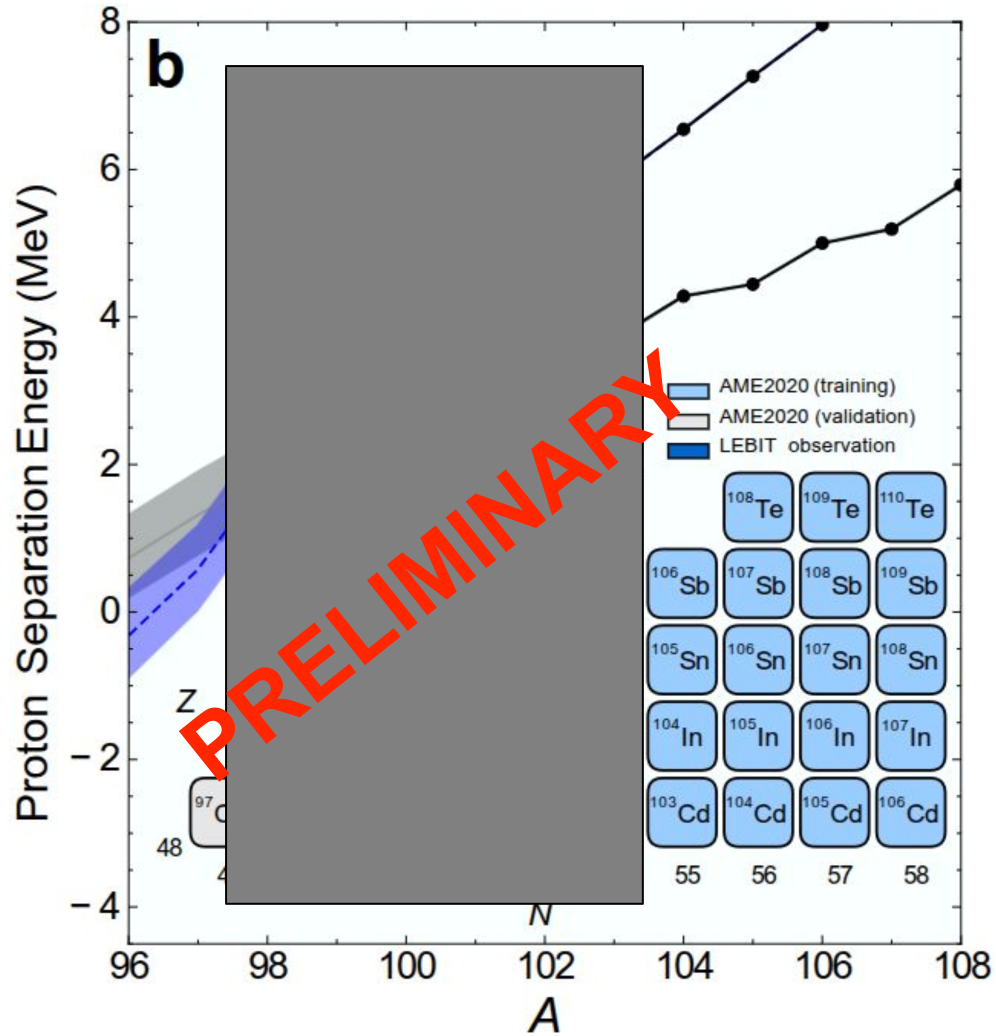
The resulting scheme: the Bayesian Model Combination (BMC)



Projections of 15 realistic models of the nuclear binding energy into the first two principal components. This representation allows us to visualize inter-model relationships.

^{X}Sn mass measurement at FRIB

Currently we are applying this framework in a joint project with experimentalists at FRIB to examine the impact of new mass measurements on short range extrapolations towards the proton drip line. Expect the data soon!



Year-2 summary

- The paper on "Model orthogonalization and Bayesian forecast mixing via principal component analysis" has been published in Phys. Rev. Research 6, 033266 (2024). We show that by adding model orthogonalization to the proposed Bayesian model combination (BMC) framework, one can arrive at better prediction accuracy and reach excellent uncertainty quantification performance. We are generalizing the method into non-heterogenous datasets. This ongoing works involves An Le, an undergraduate student, hired on another grant.
- In the paper on "Extended Fayans energy density functional: optimization and analysis", we carried out advanced calibration of the Fayans energy density functional (EDF) that will be used in our BMC studies.
- Lalit emulated the Tolmann-Oppenheimer-Volkoff (TOV) equations by building an emulator that has sub-percent error in making predictions and is able to understand the dynamics of the underlying nonlinear system. This emulator is novel as it requires much less data than an artificial neural network, and has high prediction capability and speed up (~3×10^4). The goal of building such an emulator is to speed up the process of creating mass tables from given EDF parameters.

- We developed the Bayesian Mass Explorer (BMEX; https://bmex.dev) that aims to provide an open-source suite of user-friendly web applications for on-the-fly data retrieval, visualization, and Bayesian uncertainty quantification. In the future BMEX will continue to provide new user-focused, accessible tools to the nuclear physics community in the fields of model emulation, online model calibration, and experimental design. BMEX is being developed by Godbey and Troy Dasher, an undergraduate student.
- We are working towards further integration into r-process frameworks. This work is led by Lalit.
- Work at ANL was focused on updating the recommended AME values for atomic masses and their uncertainties for all nuclei. During the period covered by this report, more than 300 journal articles were compiled and evaluated into the AME and NUBASE data files. A suite of computer codes was applied, and updated values for the masses were obtained. The ANL team is preparing an interim AME and NUBASE distribution to the MSU collaborators—a meeting is planned for March 2025 to implement this distribution.
- We helped organizing the BAND-Data 2024 workshop, Dec. 16-17, in Dublin OH, which aim was to apply Bayesian tools & techniques to aid the compilation & analysis of nuclear data. A number of promising collaborative projects, involving nuclear theorists and members of the nuclear data community, have been identified.

Next Steps?

Top priority is the local extension of the model orthogonalization BMC method and application to heterogeneous datasets (BE, charge radii, etc.)

Continued development of quantified nuclear models and applications to large scale calculations and incorporation into our theoretical database

Development and deployment of BMC package for nuclear masses in our web portal. A more general package is envisioned as well, for use by the community

Incorporation of ‘fresh’ evaluation into our frameworks to explore the impact of recently measured masses and the updated uncertainties on nuclear astrophysics applications

Thank you!