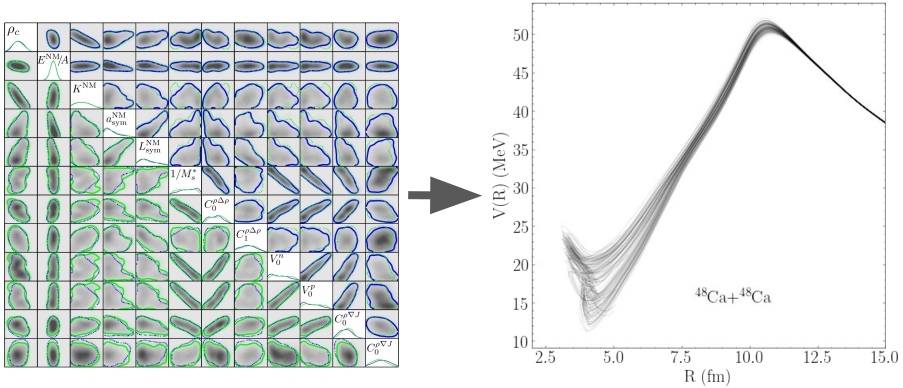
Perspectives for Accessible and Reproducible Bayesian Workflows

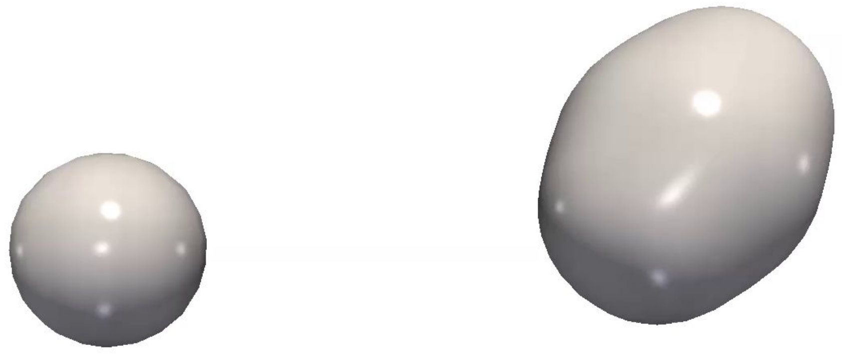
Kyle Godbey

https://docs.google.com/presentation/d/1phyqtNiJb8sEwlwbjZNaVGrCc5Oz3XRIJvSSIWRblak/edit?usp=sharing

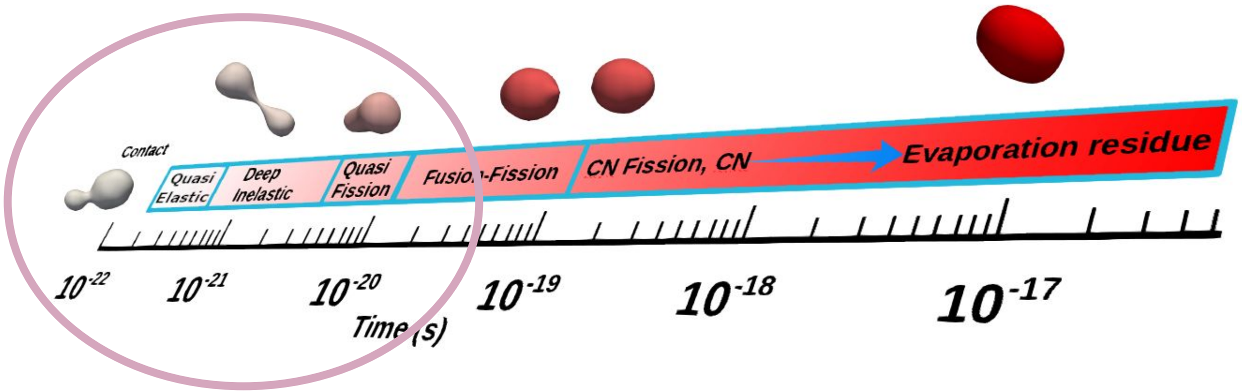


INT 24-88W | 2024 INT Workshop 24-88W
Seattle, WA

Some Physics Context

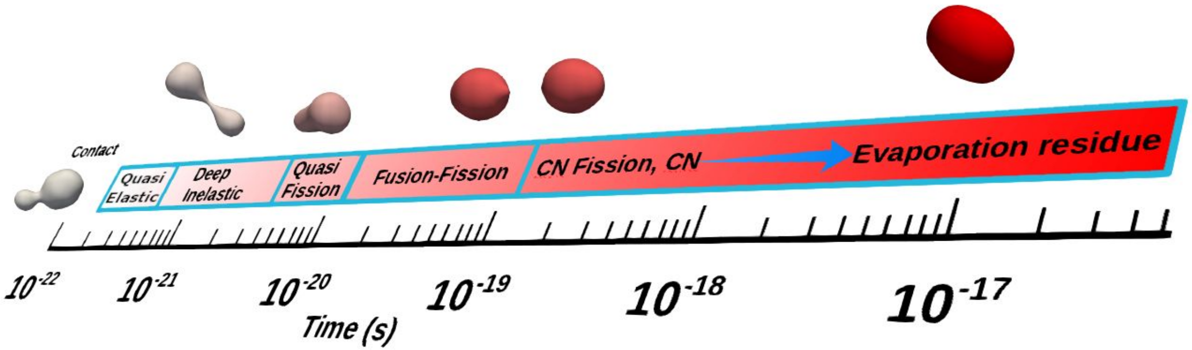


Some Physics Context



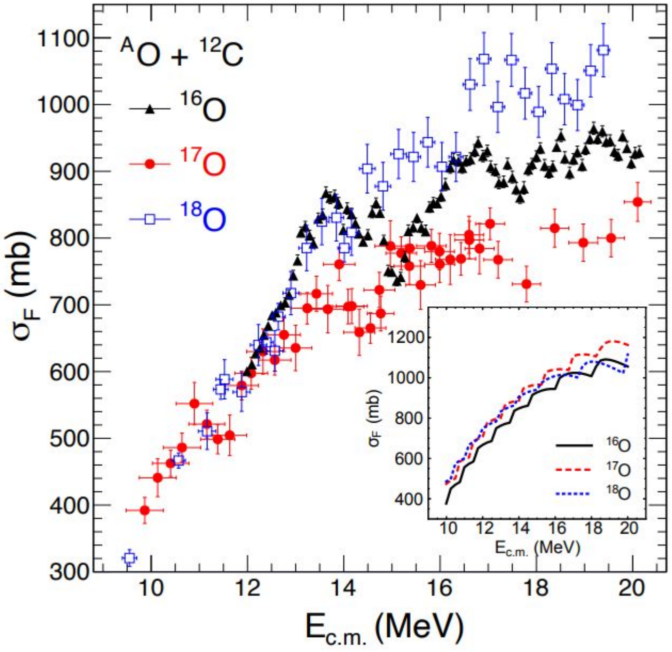
Time-dependent, microscopic theories offer a rich depiction of the many complicated things nuclei might do within the characteristic nuclear timescale

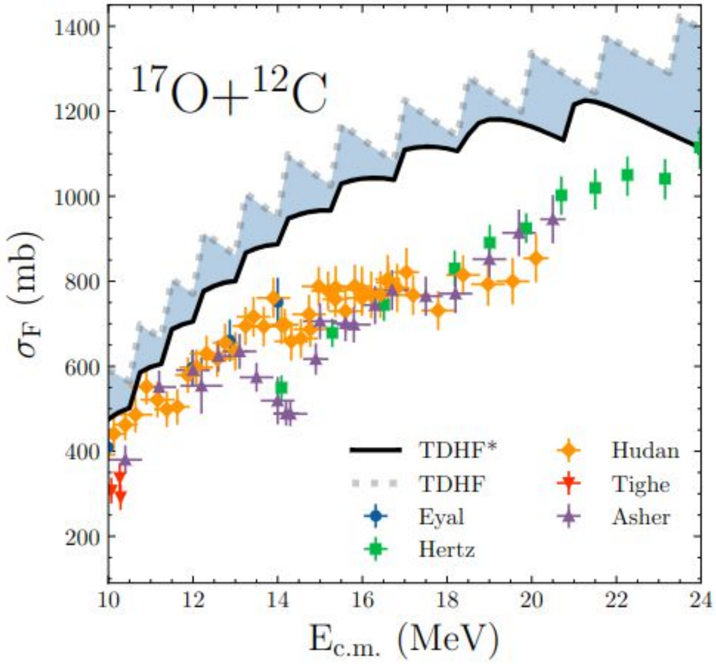
Some Physics Context



Despite my focus on reactions, **dynamics** encompass a whole lot more! Decays, collective excitations, fission etc.

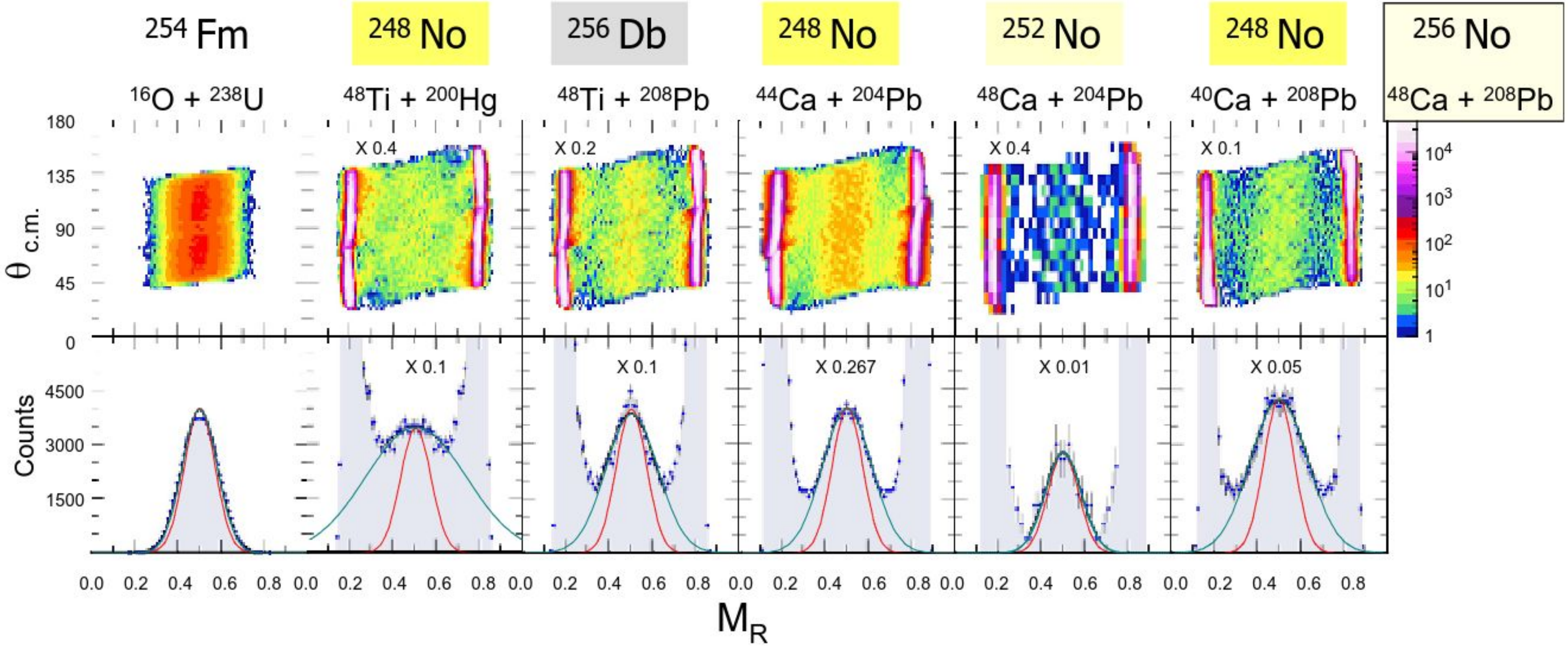
Data Showcase - Reactions





R. T. deSouza, K. Godbey, S. Hudan, W. Nazarewicz, In search of beyond mean-field signatures in heavy-ion fusion reactions. PRC(L) (2024)

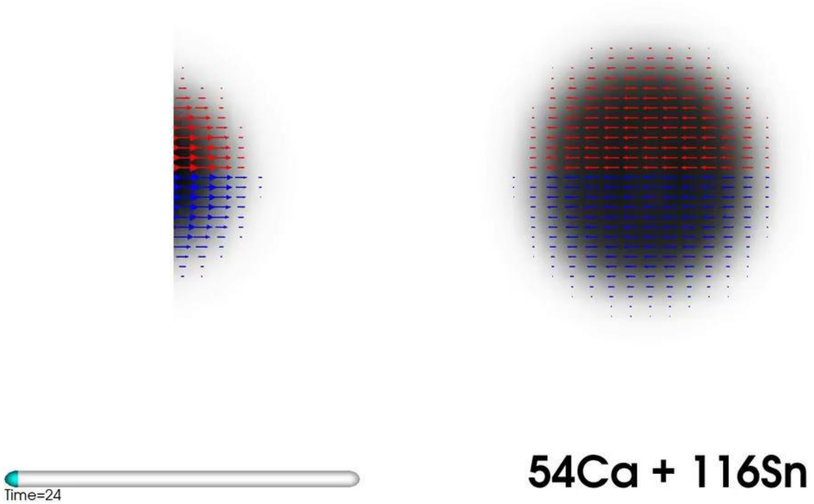
Data Showcase - Reactions



D. Hinde et al, Mass-angle distributions - Insights into the dynamics of heavy element formation EPJ conf. 66 03037 (2014)

Quasifission Example

Transfer Example



One Issue: Simulations are Expensive!

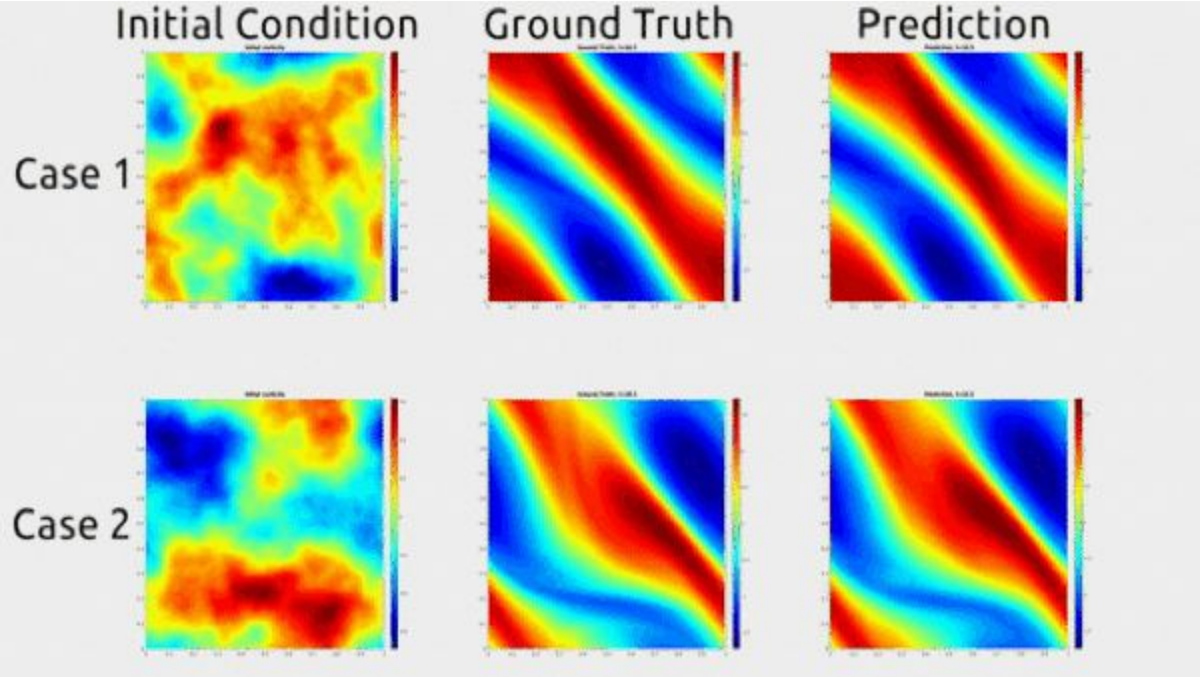
We have a wealth of observables on offer if we can afford to incorporate time-dependent dynamics into our Bayesian analyses

Costs range, however, from a few node hours to full system runs on Summit/Frontier

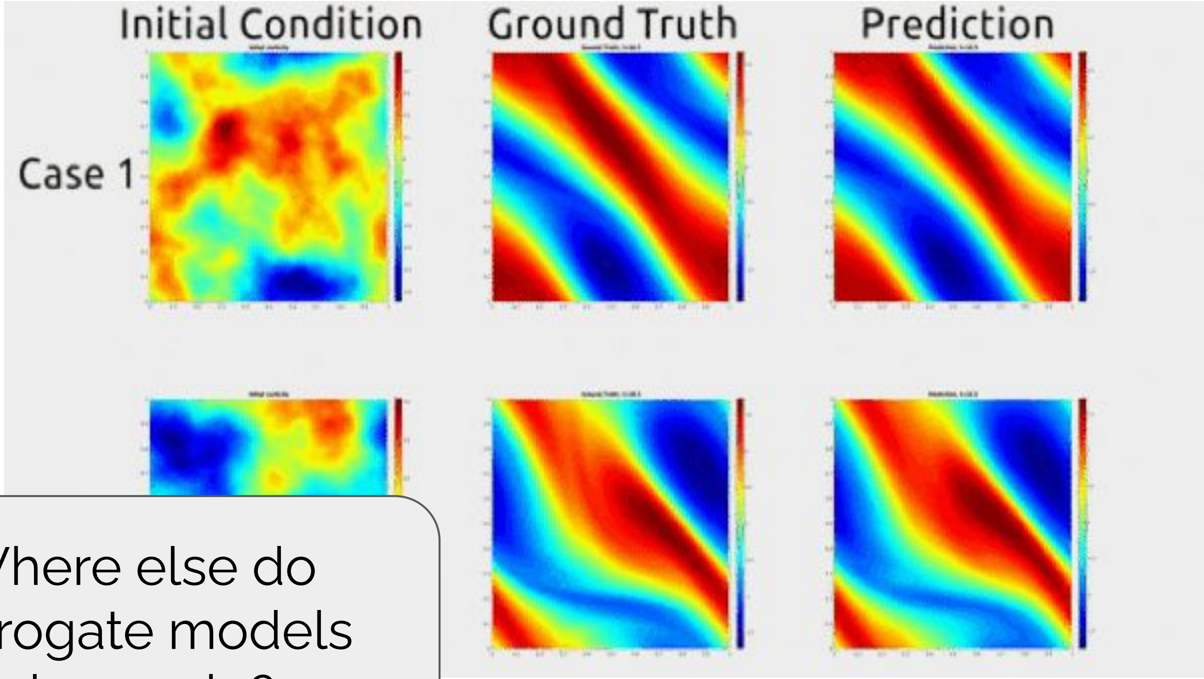
One Issue: Simulations are Expensive!

Emulators and surrogate models are one key direction for us – directly replacing the simulation output gives us access to correlated observables

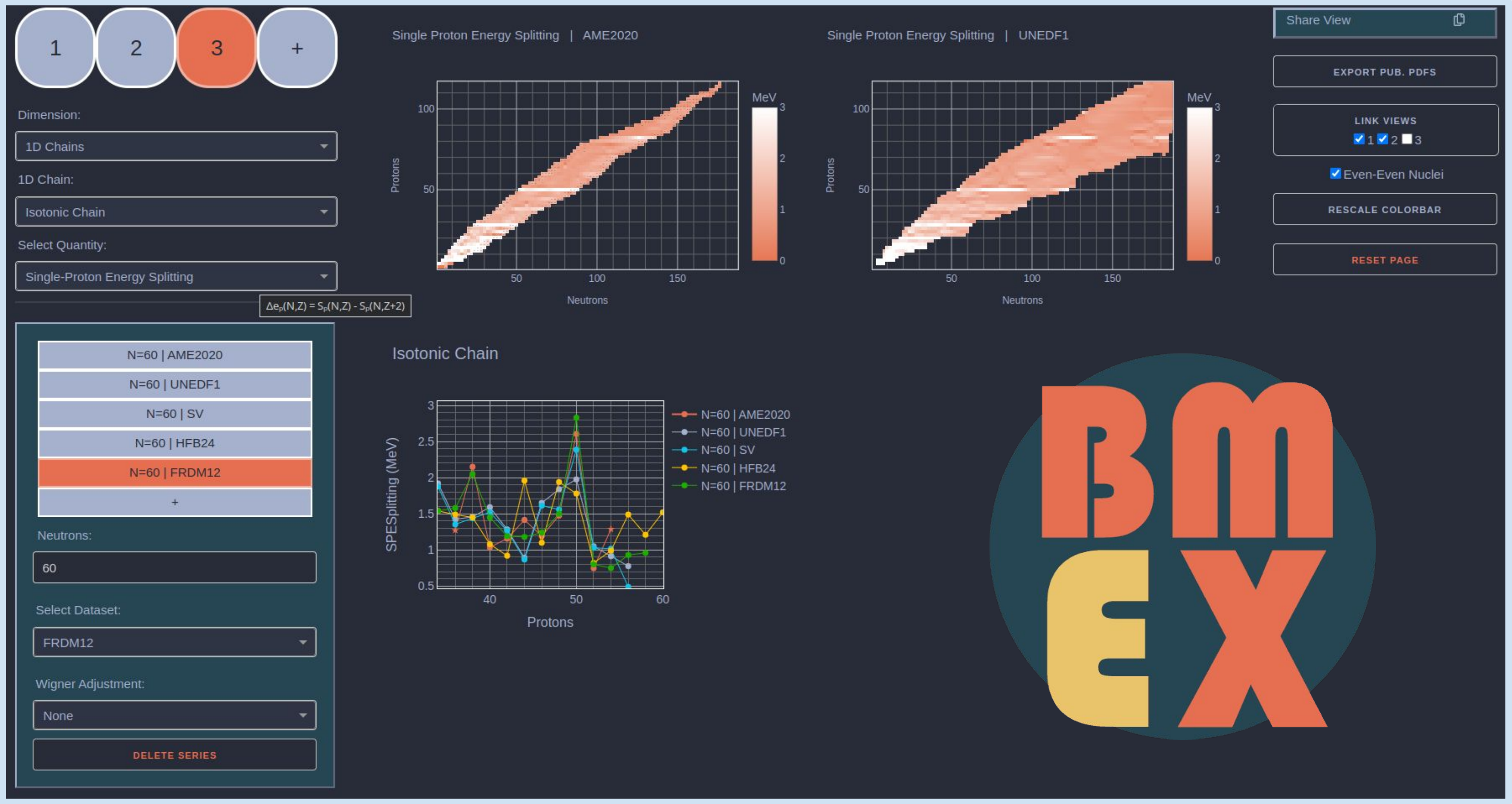
Current Ideas for Dynamics



Current Ideas for Dynamics

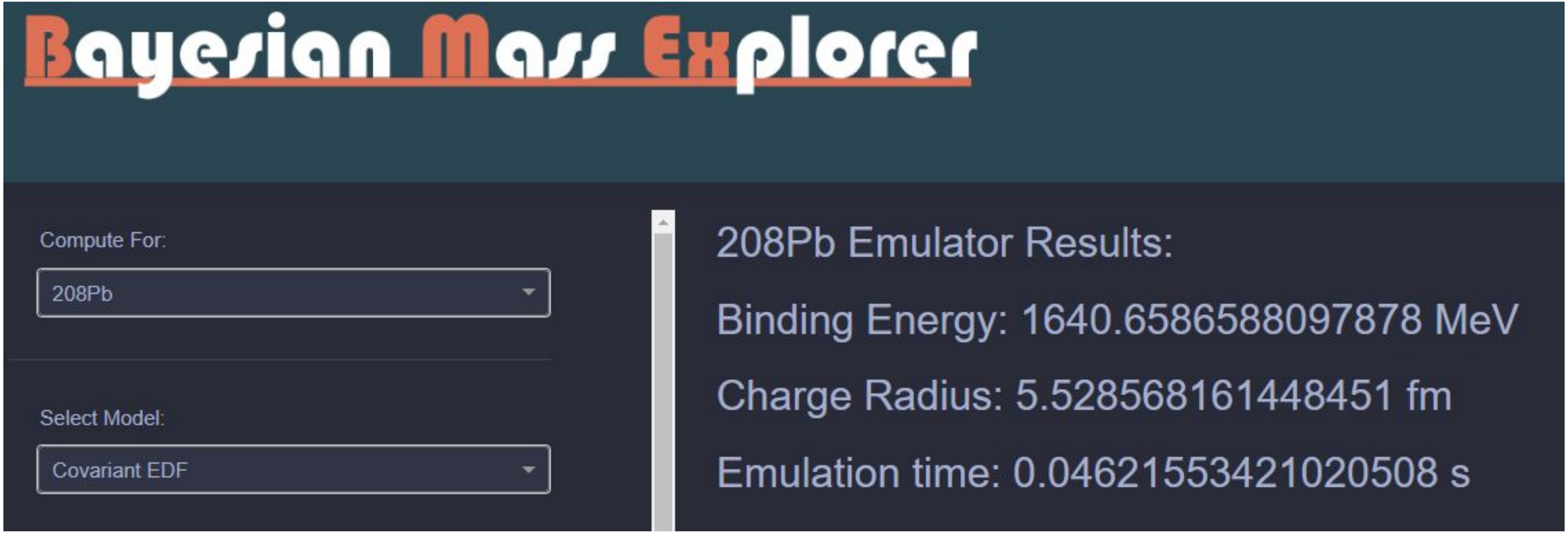


Where else do surrogate models play a role?



INT 24-88W | 2024 INT Workshop 24-88W
Seattle, WA

https://bmex.dev


Bayesian Mass Explorer
Compute For:
208Pb
Select Model:
Covariant EDF
208Pb Emulator Results:
Binding Energy: 1640.6586588097878 MeV
Charge Radius: 5.528568161448451 fm
Emulation time: 0.04621553421020508 s

Compute For:

Single Nucleus

Select Quantity:

Potential Energy Surface

Select Dataset:

UNEDF1

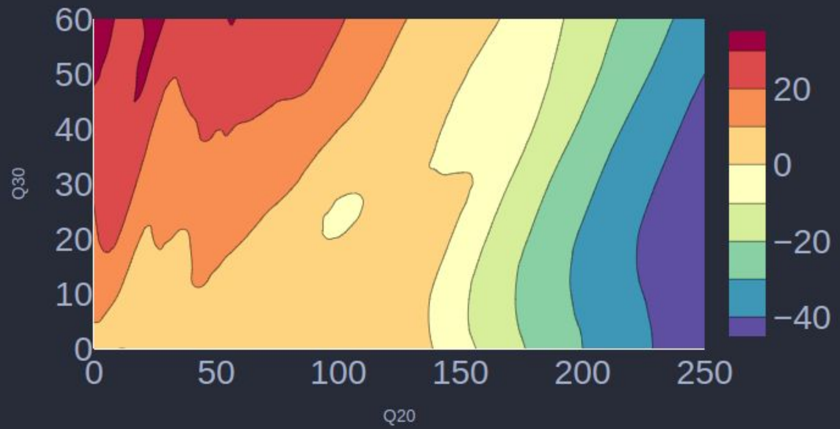
Protons:

92

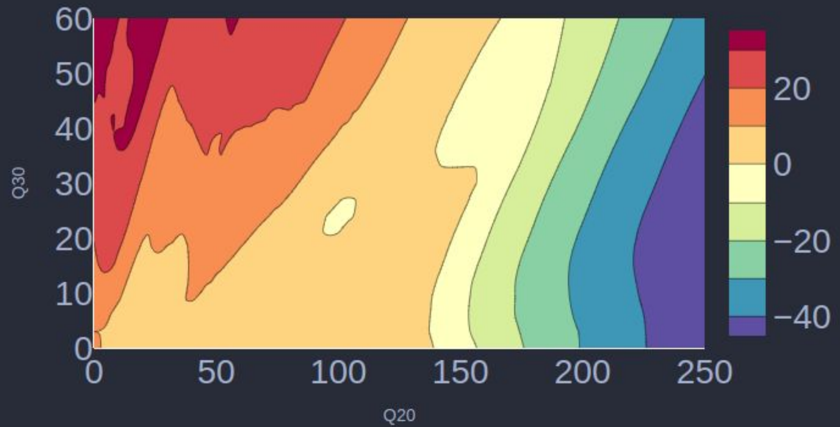
Neutrons:

158

PESnet Prediction



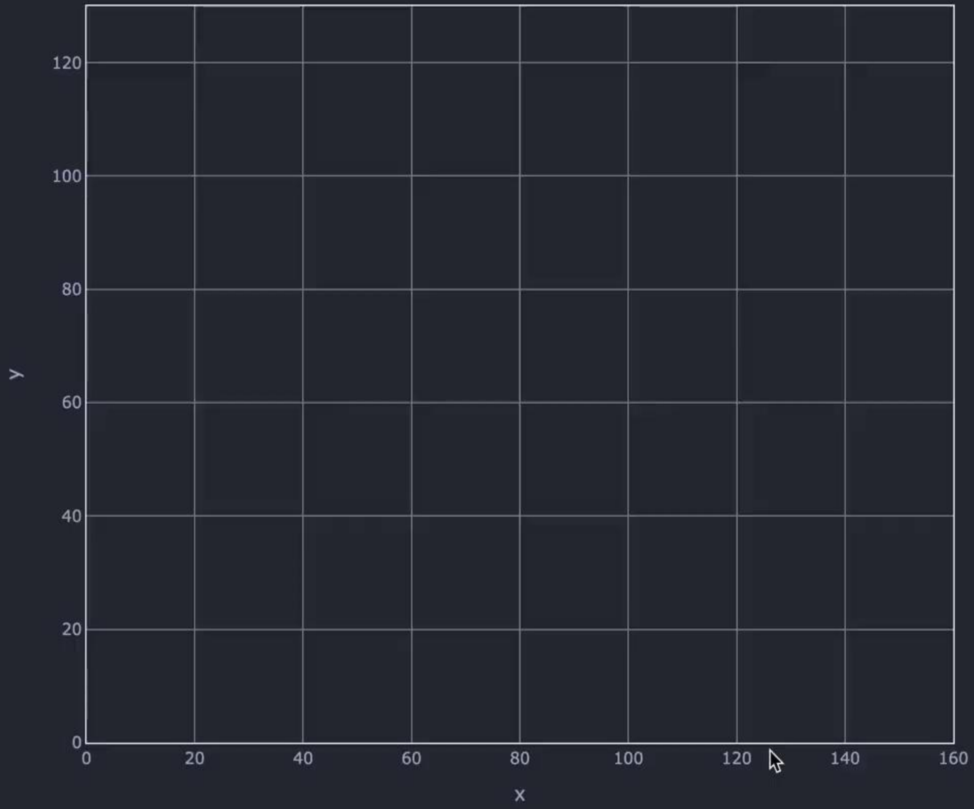
True PES

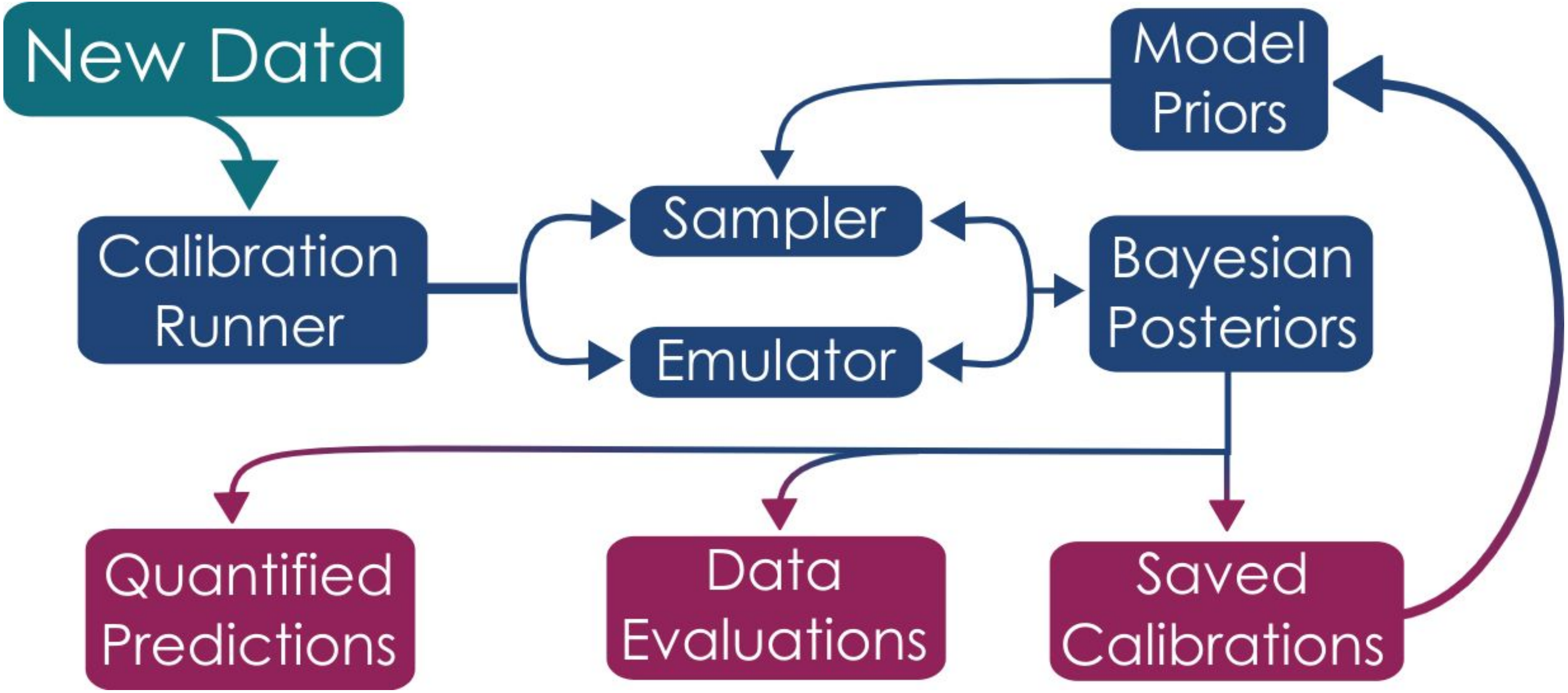


FIT

CLEAR

	x	y	δy
×			




New Data
Model Priors
Sampler
Calibration Runner
Bayesian Posteriors
Emulator
Quantified Predictions
Data Evaluations
Saved Calibrations

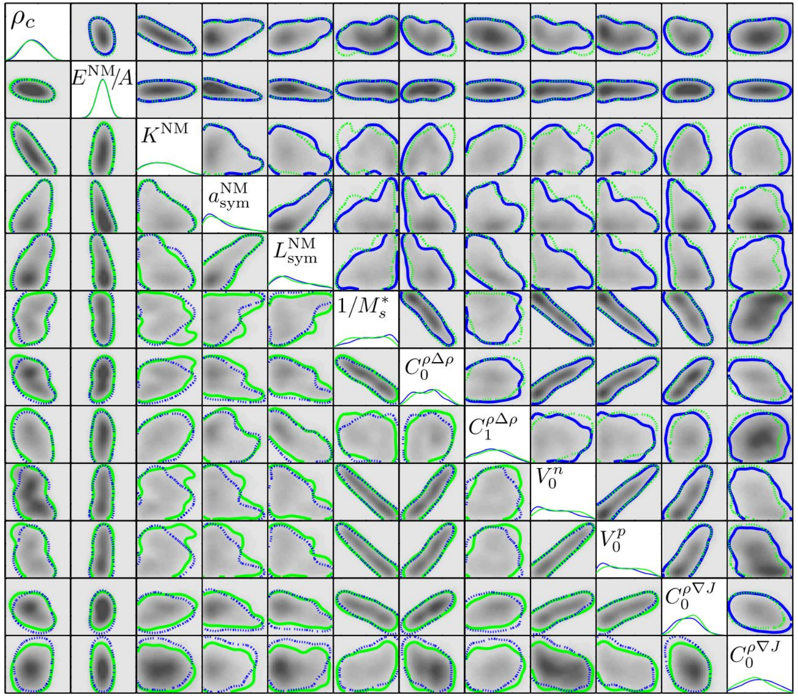
Reproducibility and Accessibility

A few persistent challenges include:

- Agility in the face of new data
- Efficiency of calibration
- Distribution of Bayesian posteriors (not just samples!)
- Traceability and reproducibility of results

Y. Yamauchi, L. Buskirk, P. Giuliani, K. Godbey, Normalizing Flows for Bayesian Posteriors: Reproducibility and Deployment, (submitted) (2023).

Reproducibility and Accessibility



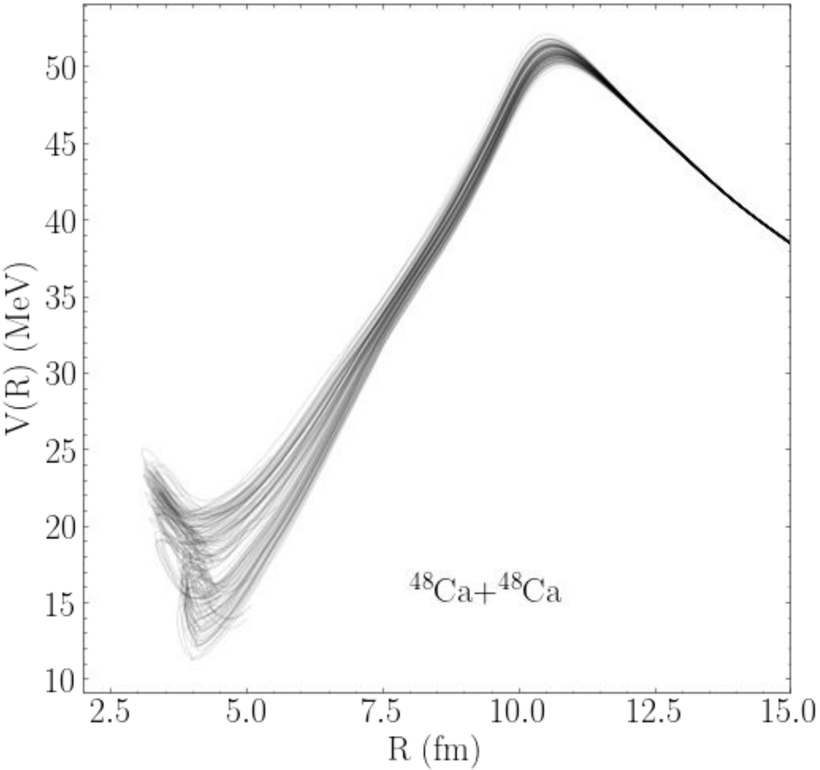
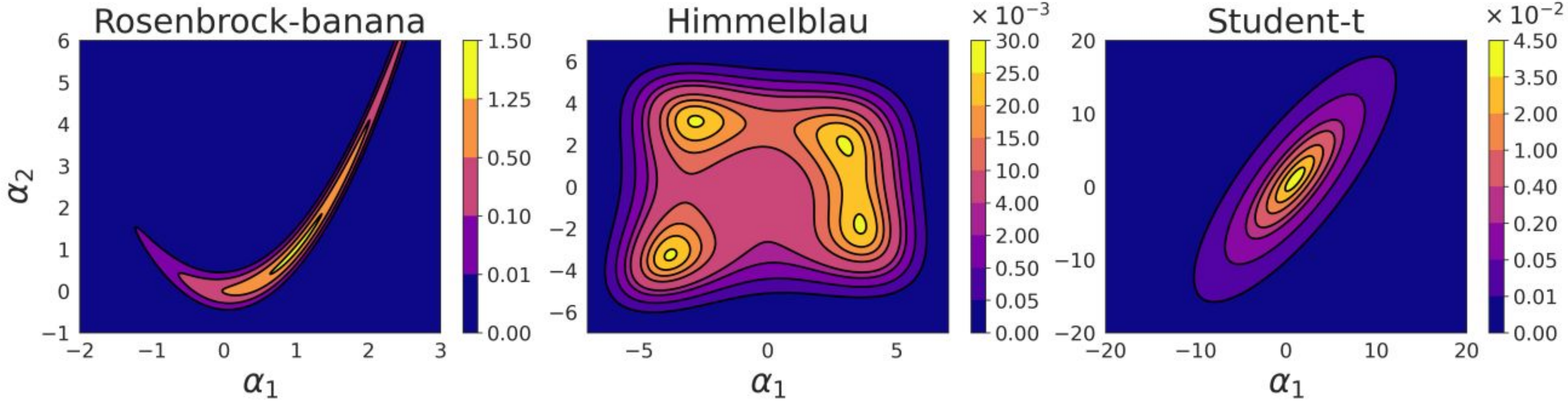


Image Credit:
J. D. McDonnell, N. Schunck, D. Higdon, J. Sarich, S. M. Wild, and W. Nazarewicz, Uncertainty Quantification for Nuclear Density Functional Theory and Information Content of New Measurements, Phys. Rev. Lett.114, 122501 (2015).

Reproducibility and Accessibility

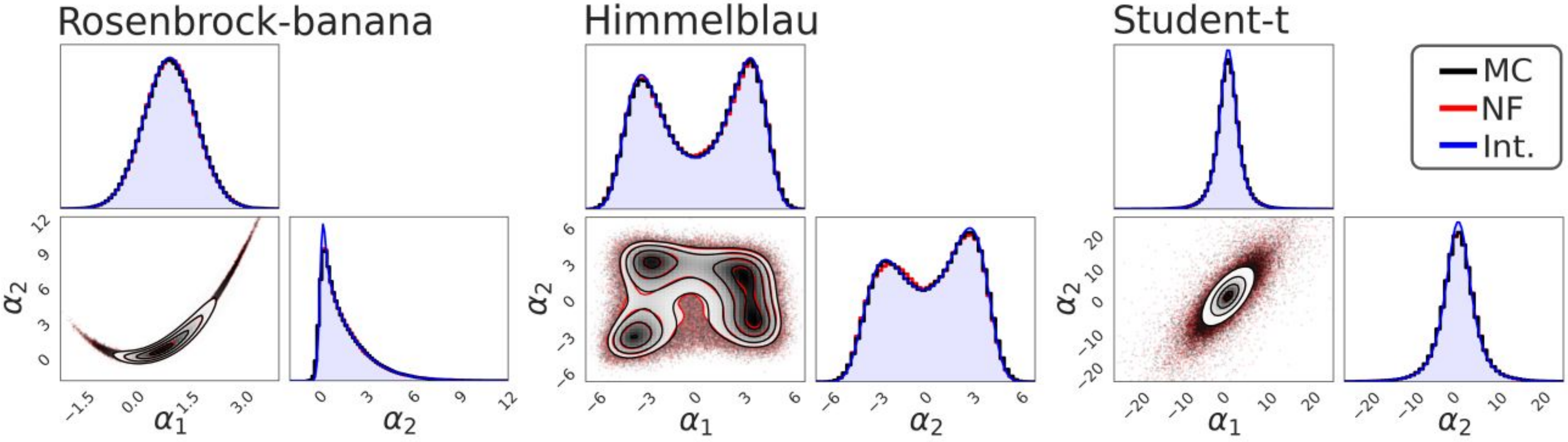
Our approach: use an ML approach to learn normalizing flows for the high-dimensional posterior distributions



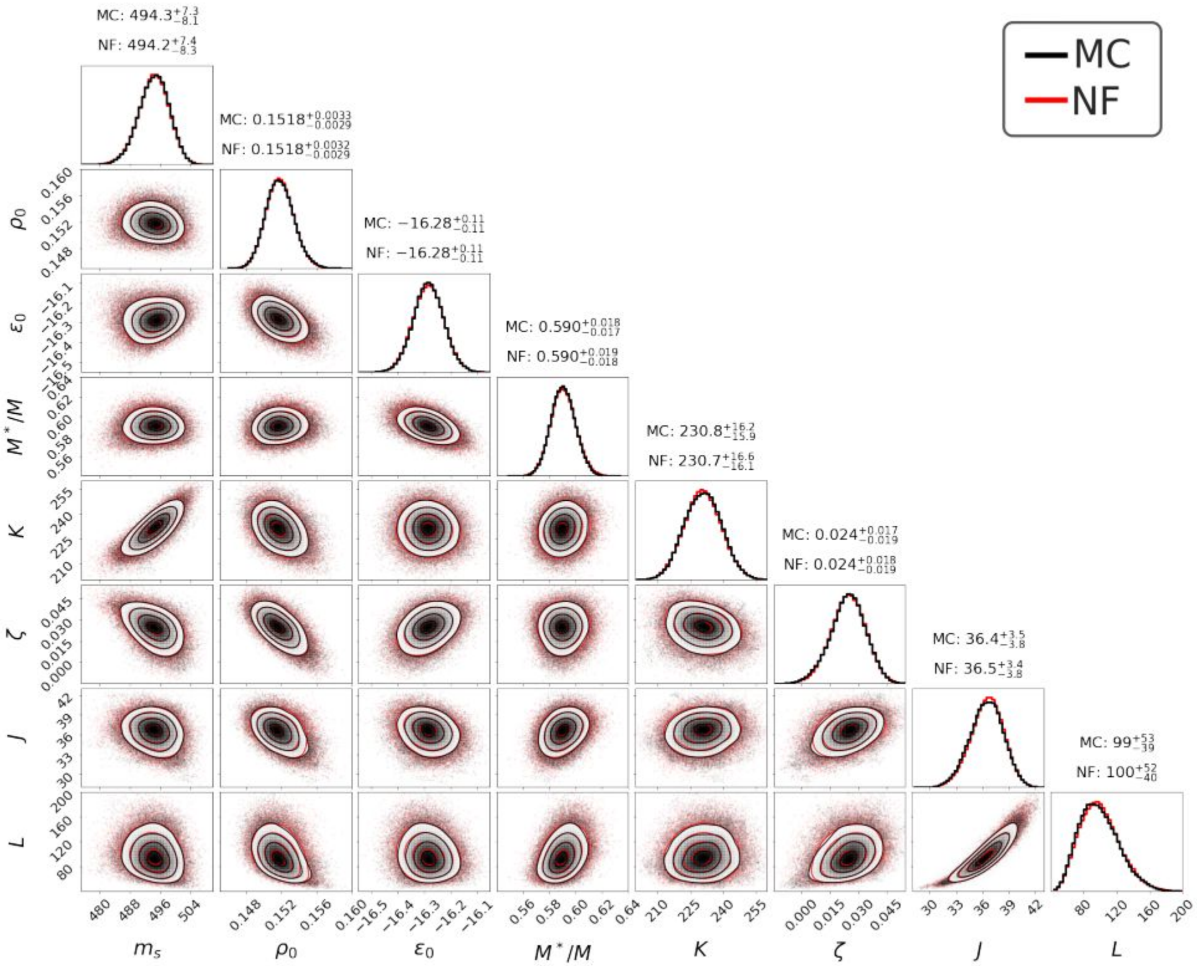
Y. Yamauchi, L. Buskirk, P. Giuliani, K. Godbey, Normalizing Flows for Bayesian Posteriors: Reproducibility and Deployment, (submitted) (2023).

Reproducibility and Accessibility

Our approach: use an ML approach to learn normalizing flows for the high-dimensional posterior distributions



Y. Yamauchi, L. Buskirk, P. Giuliani, K. Godbey, Normalizing Flows for Bayesian Posteriors: Reproducibility and Deployment, (submitted) (2023).



Y. Yamauchi, L. Buskirk, P. Giuliani, K. Godbey, Normalizing Flows for Bayesian Posteriors: Reproducibility and Deployment, (submitted) (2023).

Dimensionality Reduction in Nuclear Physics
Presented by ASCSN

Aplication 5: Black-Box Methods

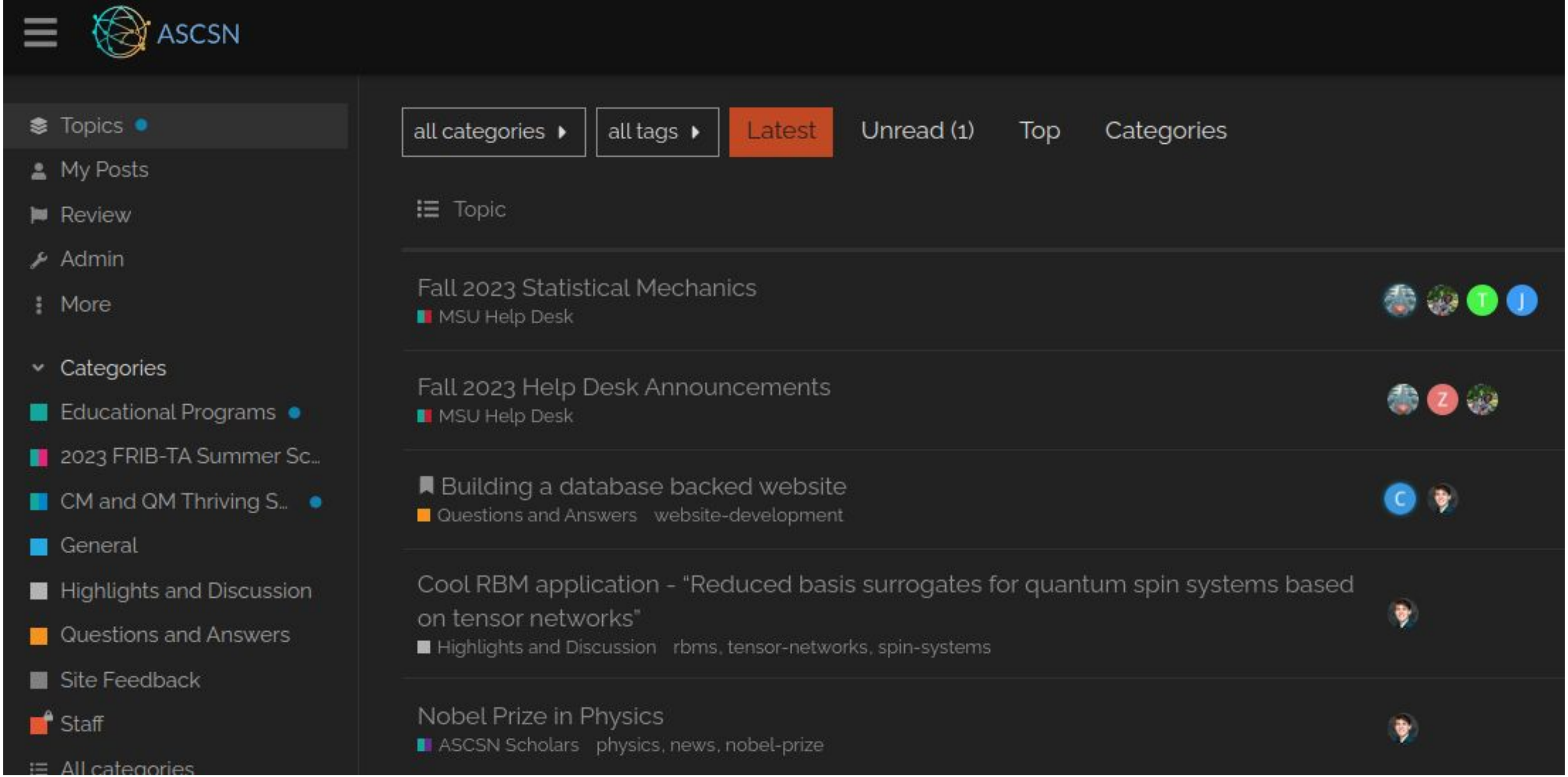
Efficient Emulation of SECAR Beam

Non-linear and non-affine problem

Always accepting new examples!

https://dr.ascsn.net

Dimensionality Reduction in Nuclear Physics
Presented by ASCSN



https://forum.ascsn.net

Immense Gratitude to All Collaborators!

Funding

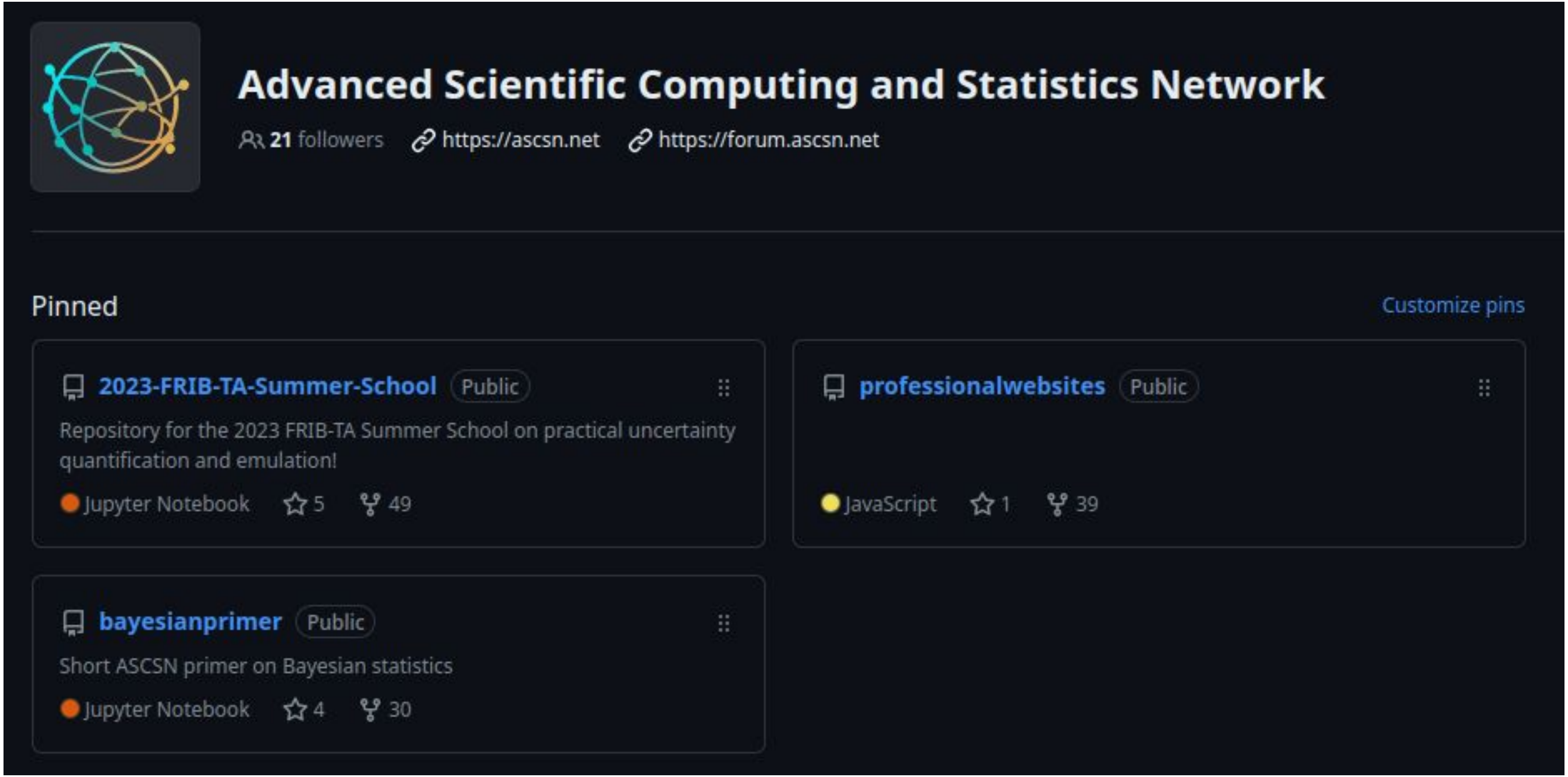
DOE NNSA Grant Nos. DE-NA0004074, DE-NA0003885
DOE Grant No. DE-SC0013365

Computing Resources

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

Coming Soon: NLDBench

To make all of this easier, we're currently working on a benchmark suite for nonlinear dynamics – if you've got a use case, reach out!



https://github.com/ascsn

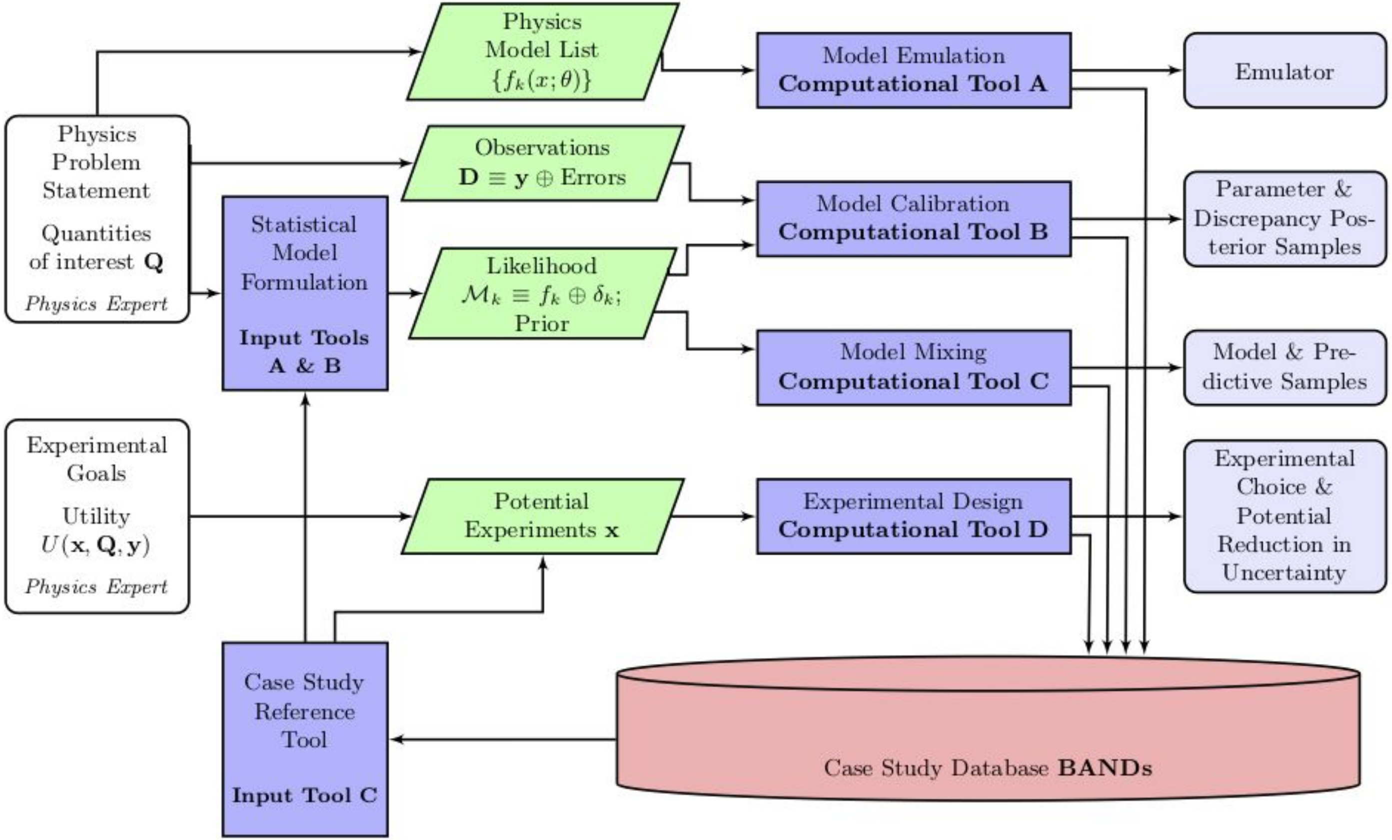
Introducing: BAND

BAND

Bayesian Analysis of Nuclear Dynamics

Supported by the NSF CSSI program under grant OAC-2004601

The Framework



The Framework

Online:
https://bandframework.github.io/

On GitHub:
https://github.com/bandframework/bandframework

Software

External code delivery will be from the bandframework github repository

surmise

A Python package designed to provide a surrogate model interface for calibration, uncertainty quantification, and other tools.

O. Surer, M. Plumlee, S.M. Wild, M. Y-H. Chan
surmise Read the Docs

Taweret

A versatile Python package containing multiple model mixing techniques for a variety of use cases.

K. Ingles, D. Liyanage, A. C. Semposki, J. C. Yannotty
Taweret documentation

SAMBA

The SAndbox for Mixing using Bayesian Analysis, developed as a testing ground for multivariate model mixing on a toy model setup.

A. C. Semposki, R. J. Furnstahl, D. R. Phillips
SAMBA repository

ParMOO

ParMOO is a parallel multiobjective optimization solver that seeks to exploit simulation-based structure in objective and constraint functions.

T.H. Chang, S.M. Wild, H. Dickinson
parmoo Read the Docs

BMEX

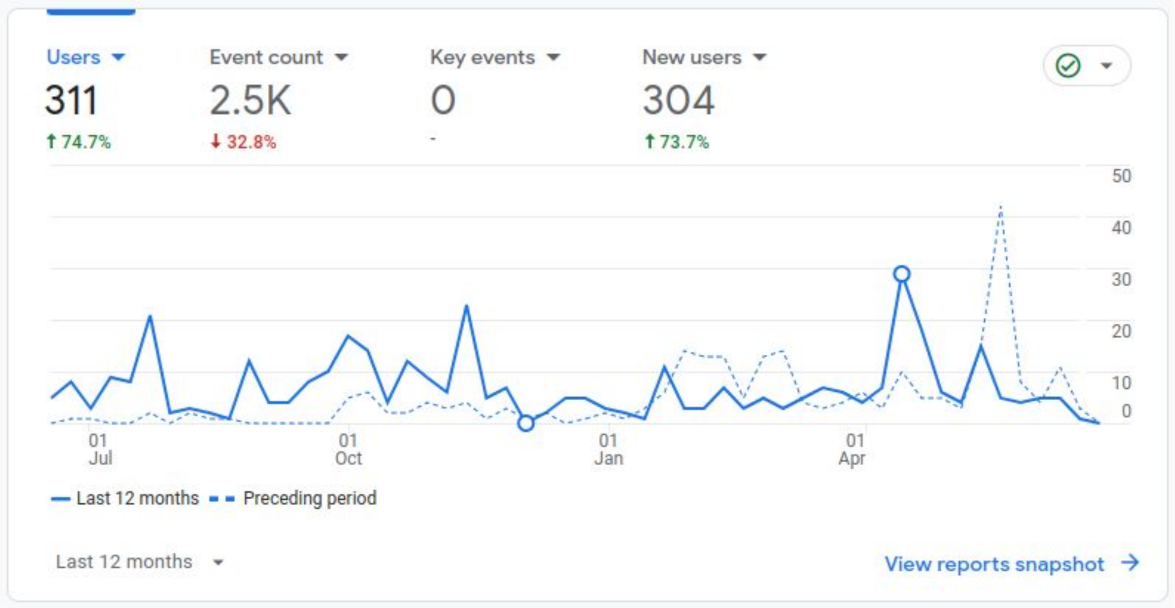
The Bayesian Mass Explorer (BMEX) is a user-focused web application that provides a one-stop-shop for quantified theoretical model predictions of nuclear masses and related quantities.

K. Godbey, L. Buskirk, P. Giuliani
BMEX Web Application

ROSE

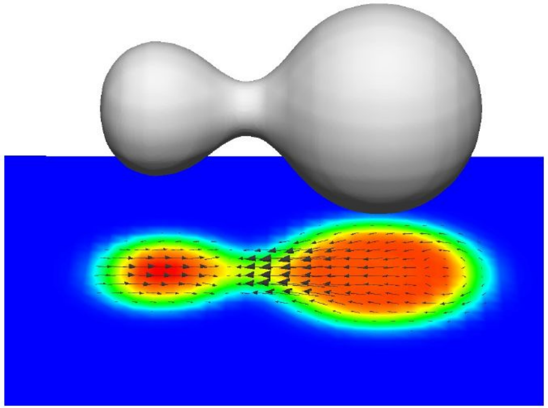
The Reduced Order Scattering Emulator (ROSE) is a Python package for building emulators using reduced basis methods for calculating nuclear scattering observables for user-defined interactions, including optical potentials.

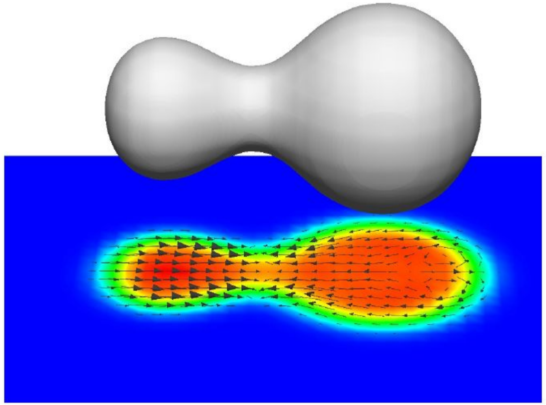
D. Odell, P. Giuliani, K. Godbey, K. Beyer, M. Y. Chan
ROSE Github



https://bmex.dev

Features of DFT: Dynamics





What Drives the Dynamics (and Structure)?

The energy density functional! A functional of various densities and currents that defines the system

e.g.

$$\mathcal{H}_{\mathrm{I}}(\mathbf{r}) = C_{\mathrm{I}}^{\rho}\rho_{\mathrm{I}}^{2} + C_{\mathrm{I}}^{s}\mathbf{s}_{\mathrm{I}}^{2} + C_{\mathrm{I}}^{\Delta\rho}\rho_{\mathrm{I}}\Delta\rho_{\mathrm{I}} + C_{\mathrm{I}}^{\Delta s}\mathbf{s}_{\mathrm{I}}\cdot\Delta\mathbf{s}_{\mathrm{I}} + C_{\mathrm{I}}^{\tau}\left(\rho_{\mathrm{I}}\tau_{\mathrm{I}} - \mathbf{j}_{\mathrm{I}}^{2}\right) + C_{\mathrm{I}}^{T}\left(\mathbf{s}_{\mathrm{I}}\cdot\mathbf{T}_{\mathrm{I}} - \overleftrightarrow{J}_{\mathrm{I}}^{2}\right) + C_{\mathrm{I}}^{\nabla J}\left(\rho_{\mathrm{I}}\nabla\cdot\mathbf{J}_{\mathrm{I}} + \mathbf{s}_{\mathrm{I}}\cdot(\nabla\times\mathbf{j}_{\mathrm{I}})\right)$$

Parameter Determination

One method is, given an EDF, fit the constants to some experimental data

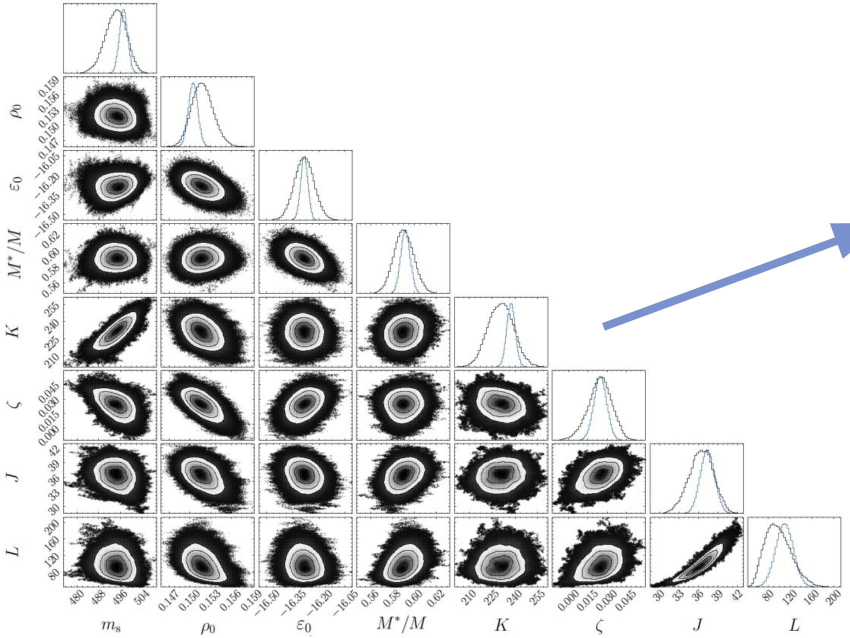
The result? Many, many, many 'forces' for a given functional. Some fave Skyrme-types include SkM*, SLy4d, SLy5t, SV-Min, UNEDF1, etc.

Parameter Determination

But.. nobody is perfect, sorry SLy4d. Given a certain set of data there is some uncertainty on what the optimal parameters are

So why do we even need optimal parameters? Instead we can define our physical model as a distribution of reasonable parameters

“Doing UQ”



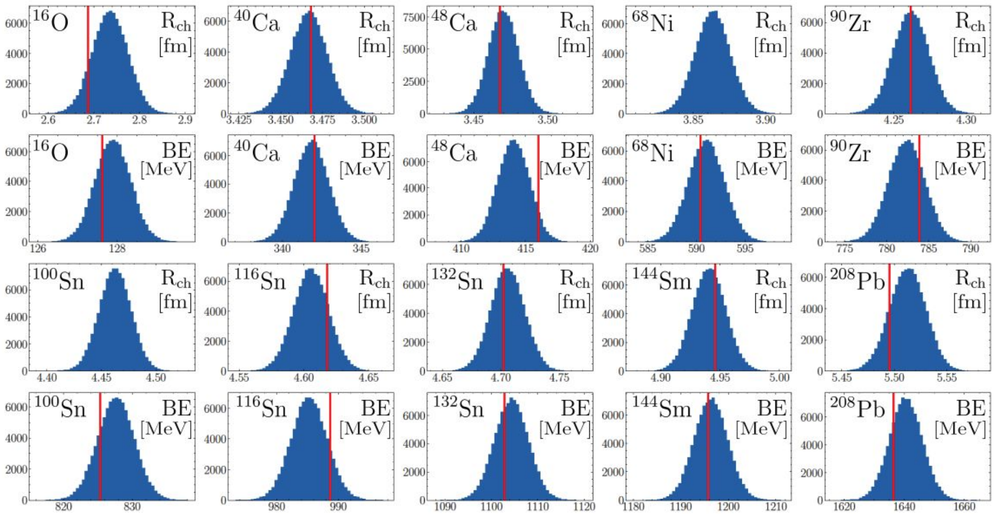
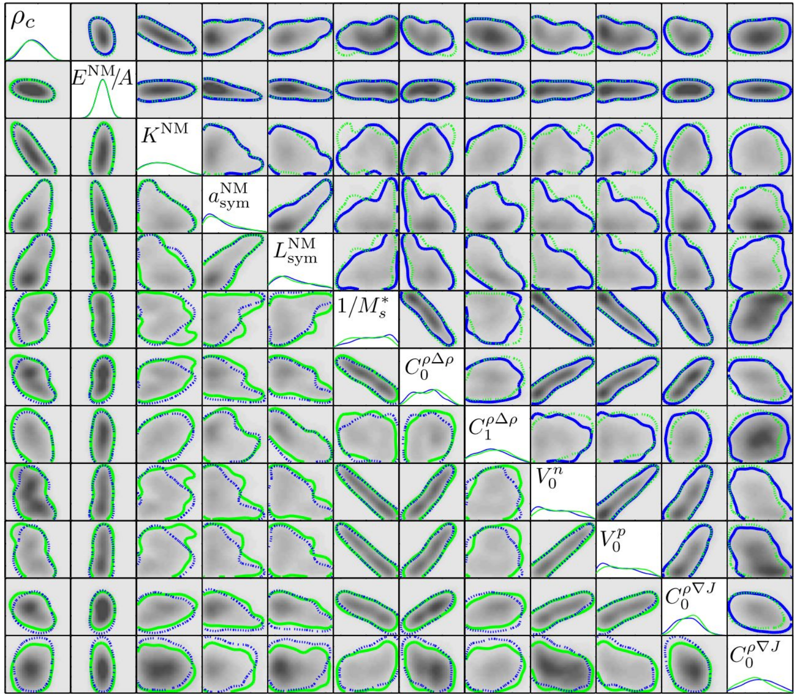


Image Credit:
P Giuliani, **K Godbey,** E Bonilla, F Viens, J Piekarewicz, Bayes goes fast: Uncertainty Quantification for a Covariant Energy Density Functional emulated by the Reduced Basis Method

“Doing UQ”



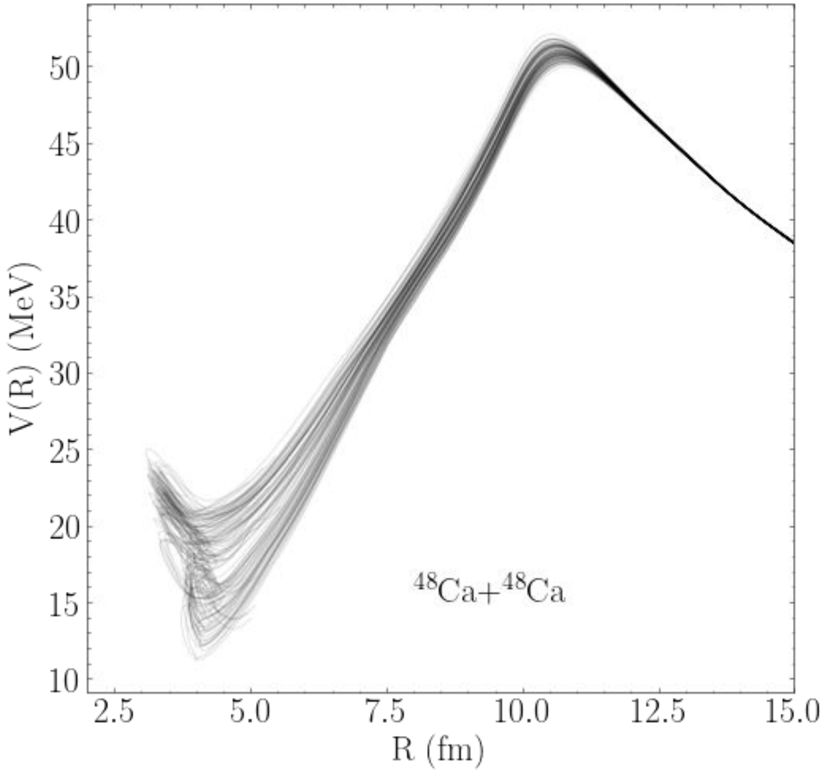
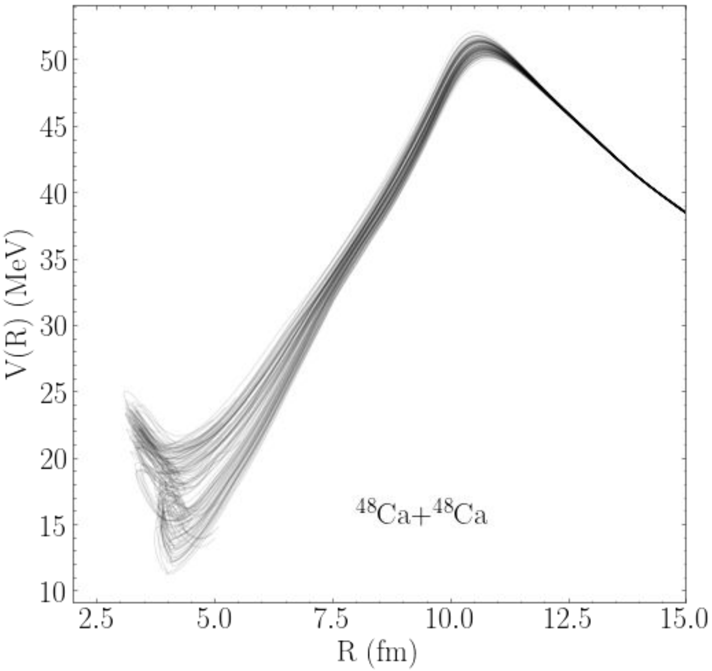
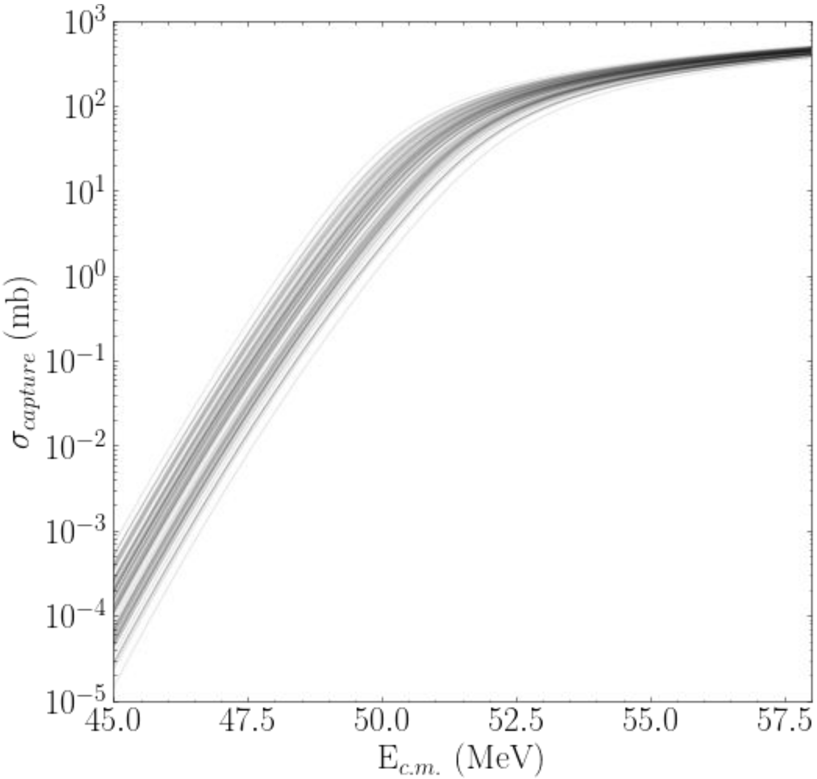


Image Credit:
J. D. McDonnell, N. Schunck, D. Higdon, J. Sarich, S. M. Wild, and W. Nazarewicz, Uncertainty Quantification for Nuclear Density Functional Theory and Information Content of New Measurements, Phys. Rev. Lett.114, 122501 (2015).

"Doing UQ"



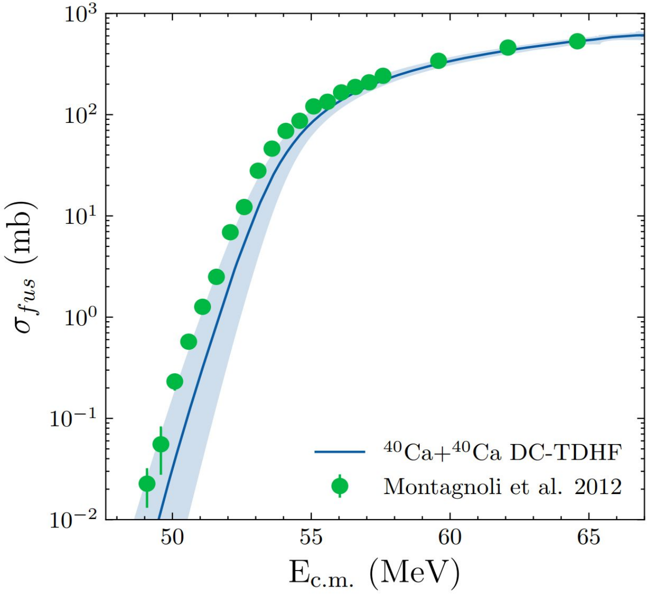
V(R) (MeV)
50
45
40
35
30
25
20
15
10
2.5
5.0
7.5
10.0
12.5
15.0
R (fm)
^{48}Ca+^{48}Ca



$\sigma_{capture}$ (mb)
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
45.0
47.5
50.0
52.5
55.0
57.5
$E_{c.m.}$ (MeV)

“Doing UQ”

$^{40}Ca + ^{40}Ca$



$^{48}Ca + ^{48}Ca$

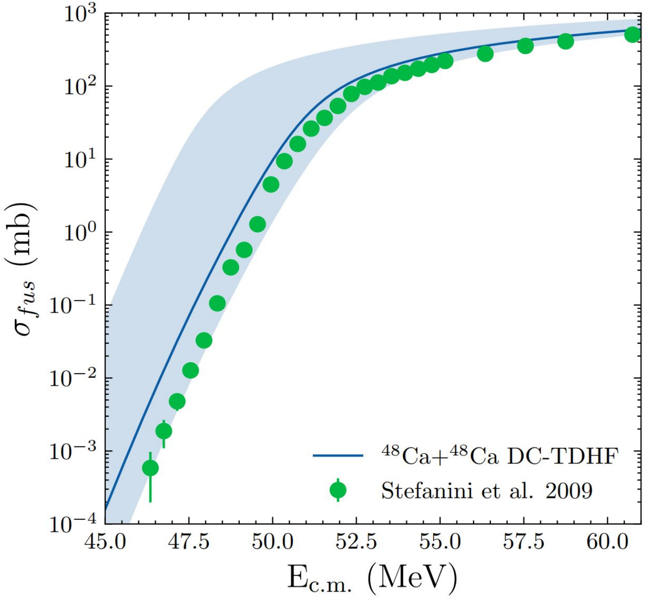


Image Credit:
K. Godbey, A. S. Umar, and C. Simenel, Uncertainty quantification in microscopic heavy-ion fusion simulations Phys. Rev. C 106, L051602 (2022) (Editors Suggestion),

Extracting Barrier Height

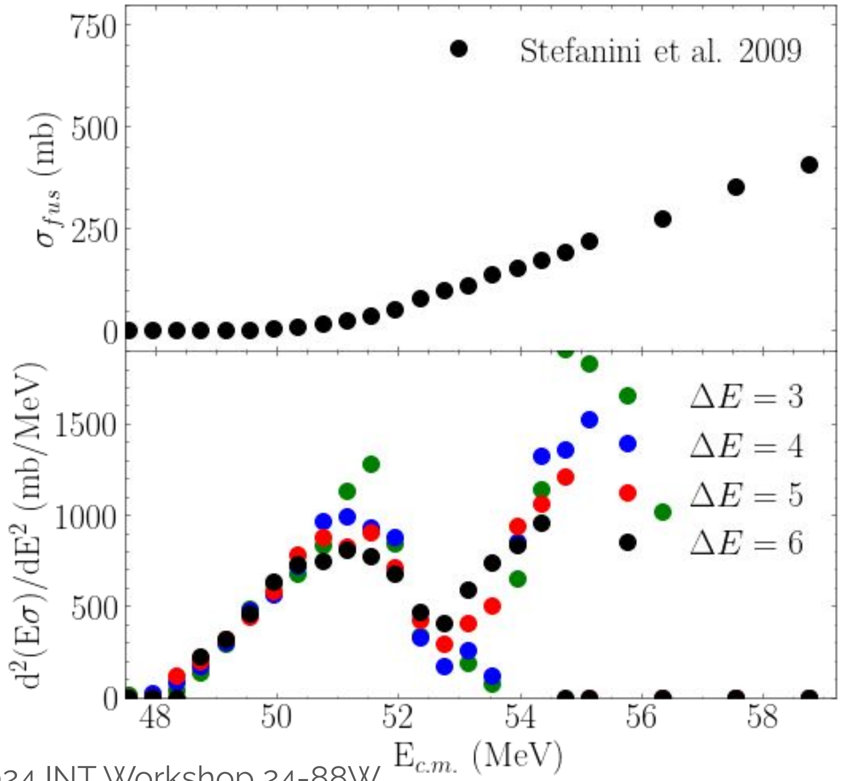
To get the barrier height from experimental data without any model dependence, we should deal only with the data

To this end, let's consider the experimental barrier distribution:

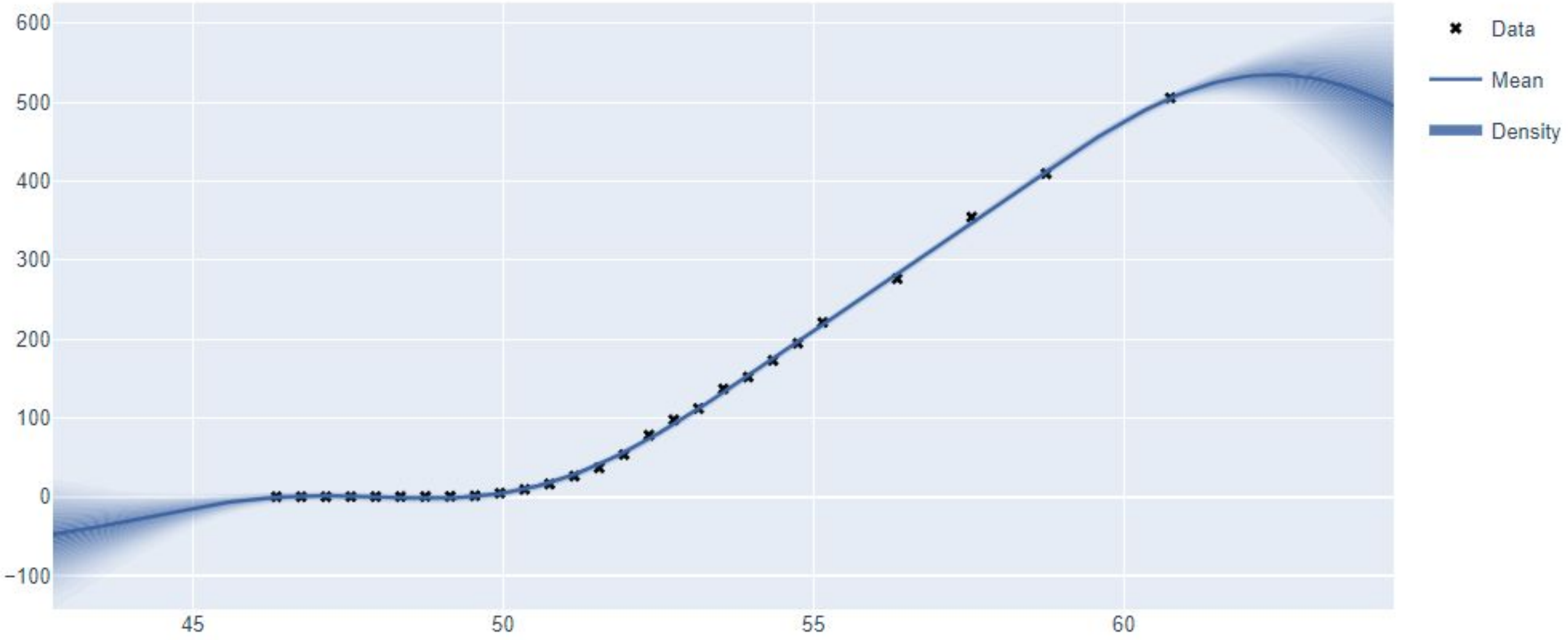
$$D_{\mathrm{exp}}(E) = \frac{d^2(E\sigma)}{dE^2}$$

M. Dasgupta, D. J. Hinde, N. Rowley, A. M. Stefanini, "MEASURING BARRIERS TO FUSION", Annual Review of Nuclear and Particle Science, Vol. 48:401-461 (1998)

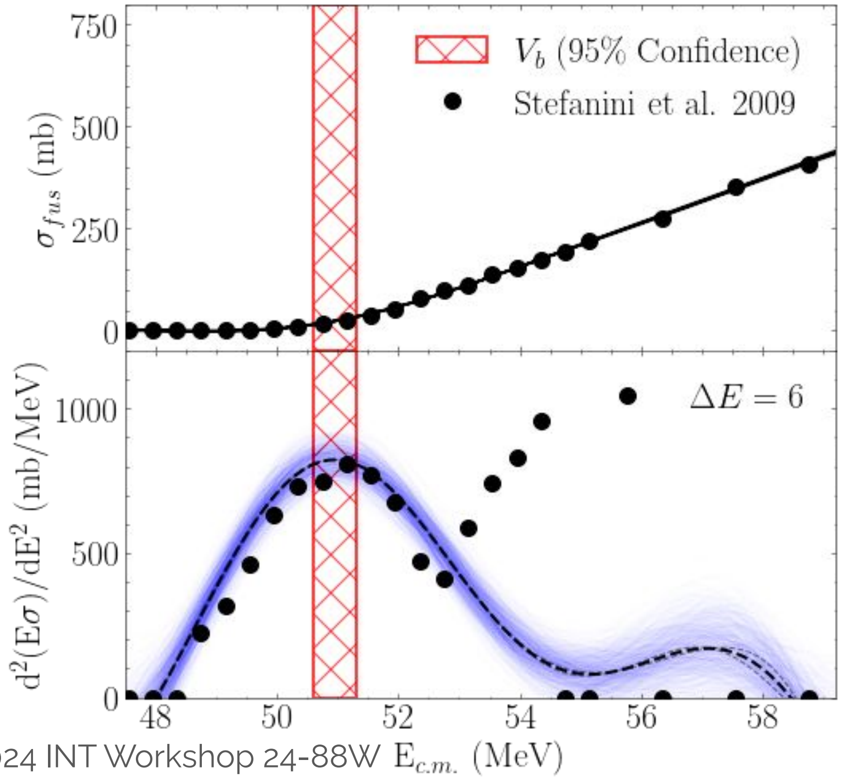
Extracting Barrier Height



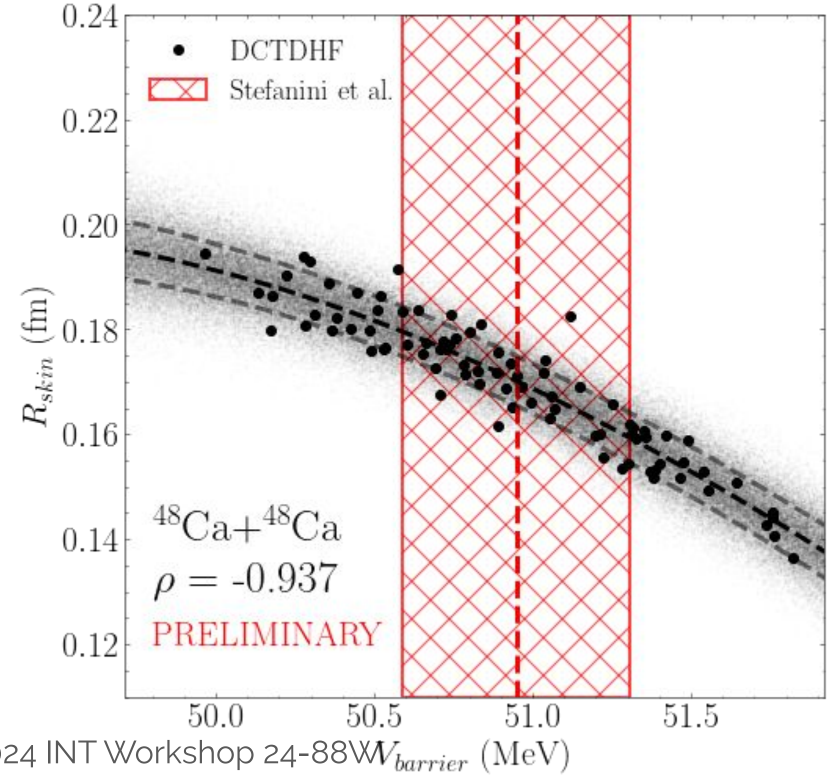
Extracting Barrier Height



Extracting Barrier Height



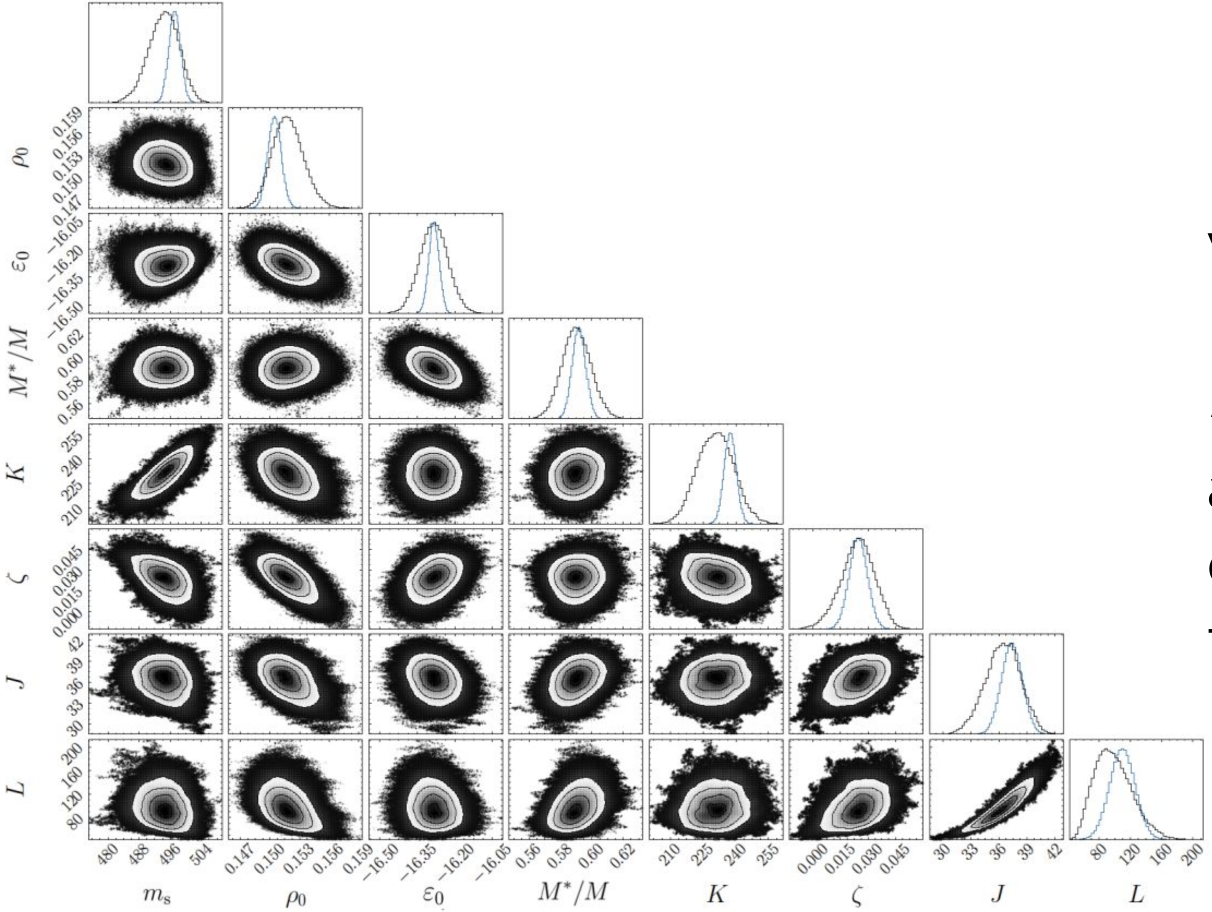
Checking Skin vs. Barrier Correlations



What Next?

A strong correlation is a strong indicator of opportunity, and this is just one system

Constraints are great, but including fusion cross sections from TDDFT in a direct Bayesian calibration is unlikely without advances in emulation



What about RBMs?

~5,000,000 samples in about a day on commodity hardware for covariant DFT

Image Credit:
P Giuliani, **K Godbey,** E Bonilla, F Viens, J Piekarewicz, Bayes goes fast: Uncertainty Quantification for a Covariant Energy Density Functional emulated by the Reduced Basis Method

Challenges

Robust calibration requires great emulators, but we're behind on time-dependent emulation

RBMs have proven great for DFT and scattering, but generally perform worse for time evolution

Thus current direction is on data-driven approaches like neural implicit flow or Fourier neural operator

Challenges

Even with powerful emulation, direct Bayes may still be out of reach

Techniques such as the polynomial chaos expansion (like an RBM for your random distribution) might be crucial to lower the total number of samples required

Conclusions

Linking structure to reactions must consider the uncertainties lurking at every step, no matter the energy scale

Nuclear structure often contains lots of intricacies and possible refinements. Each of these intricacies are affected by the model uncertainties that generated them

Background - Density Constraint + TDDFT

Perform standard time evolution to time *t* and internuclear separation *R* and save the density

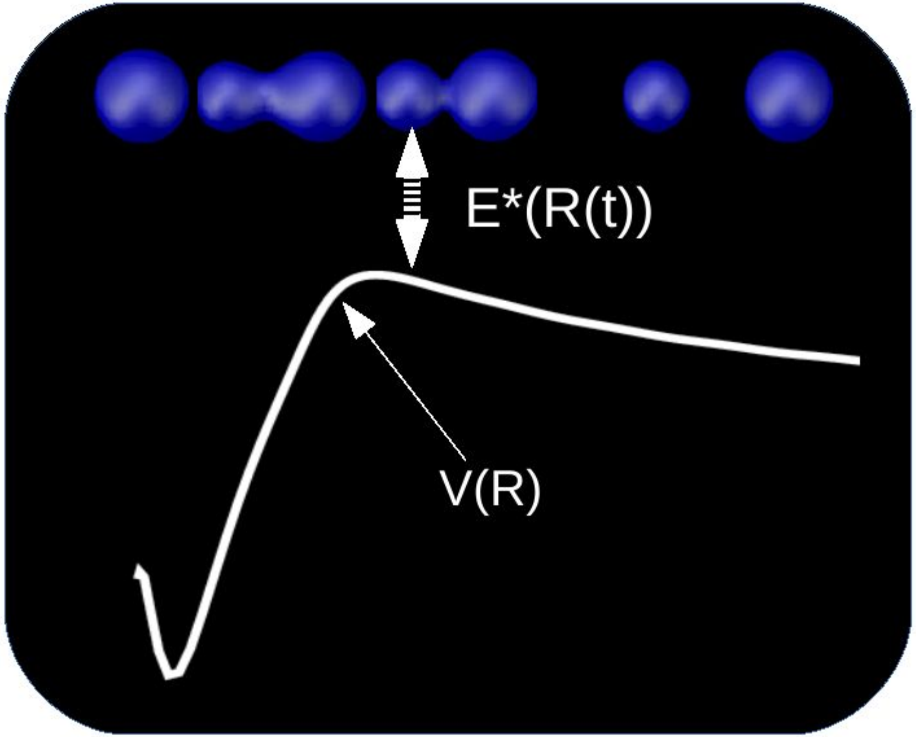
Start a static iteration to minimize the energy using the density from the time-dependent calculation at *t* and *R*

The converged energy is to then be interpreted as the static collective energy, E_{DC}

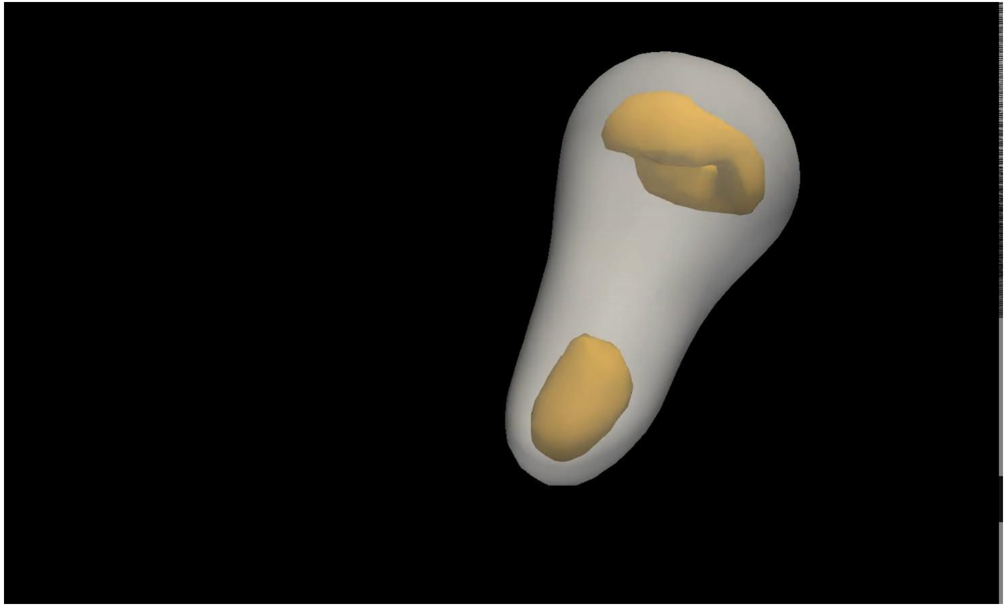
Background - Density Constraint + TDDFT

The potential is obtained by subtracting the static binding energies of the incoming fragments from E_{DC}

$$V(R) = E_{DC} - E_{A1} - E_{A2}$$



Features of DFT: Dynamics



Features of DFT: Dynamics