Constraining Fission Yields Using Machine Learning

Amy E. Lovell

In collaboration with:
Arvind Mohan, Patrick Talou, and Michael Chertkov

WONDER 2018, Aix-en-Provence
October 11, 2018

LA-UR-18-29884

Outline

- Introduction and motivation
- Brosa yield model
  – Parameter optimization for ^{252}Cf Y(A,TKE)
- Machine Learning methods
  – Probabilistic approach for Neural Networks
  – Learning mass yields for spontaneous fission
- Conclusion
- Future work

Consistency and Correlations

A recent study showed that there are inconsistencies within evaluations

P. Jaffke, NSE **190**, 258 (2018)

With tools like CGMF, we can form a consistent picture of fission – post-scission to prompt fragment emissions

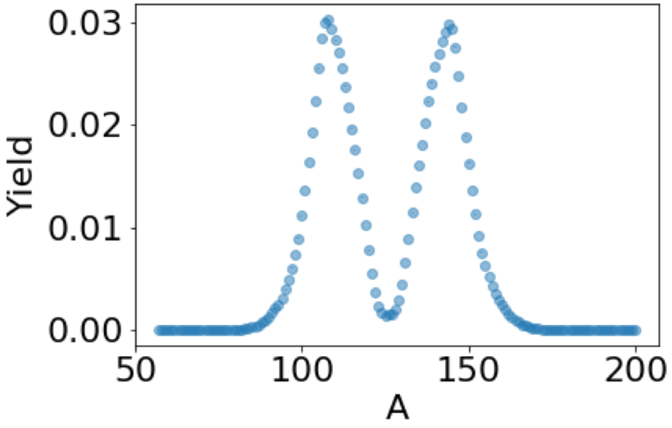
	JEFF-3.1.1[a]	JENDL-4.0u2[b]	ENDF/B-VII.1[c]	Evaluation or Experiment
^{237}Np(n_{th}, f)	2.74 ± 0.06	1.93 ± 0.07	1.94 ± 0.07	2.52 ± 0.016 (Ref. 53)
^{238}Np(n_i, f)	2.87 ± 0.09	2.37 ± 0.09	2.37 ± 0.09	2.77 ± 0.14 (Ref. 55)
^{238}Pu(n_i, f)	3.09 ± 0.10	1.93 ± 0.07	1.95 ± 0.07	2.88 ± 0.14 (Ref. 55)
^{241}Am(n_{th}, f)	3.35 ± 0.08	2.69 ± 0.10	2.71 ± 0.10	3.21 ± 0.032 (Ref. 53)
^{243}Cm(n_{th}, f)	3.76 ± 0.09	2.75 ± 0.10	2.78 ± 0.10	3.43 ± 0.047 (Ref. 53)
^{244}Cm(n_i, f)	4.04 ± 0.12	3.61 ± 0.13	3.64 ± 0.13	3.33 ± 0.17 (Ref. 55)
^{245}Cm(n_{th}, f)	4.31 ± 0.09	3.12 ± 0.11	3.14 ± 0.10	3.72 ± 0.004 (Ref. 53)

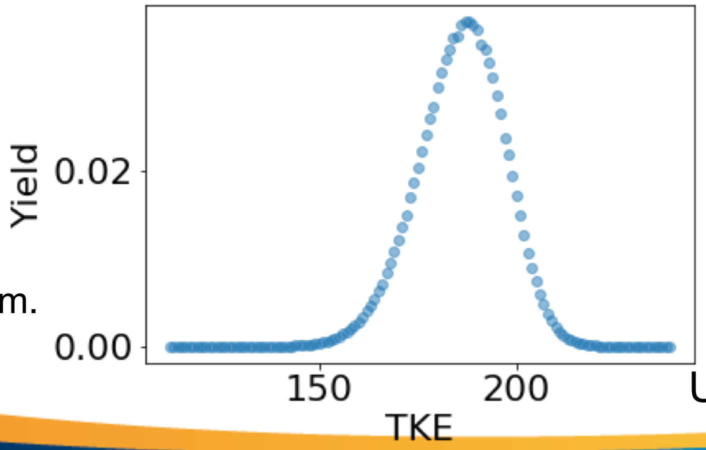
[53] S.F. Mughabghab, *Atlas of Nuclear Resonances: Resonance Parameters and Thermal Cross section Z=1-100* (2006)

[55] R. Howerton, NSE **62**, 438 (1977)

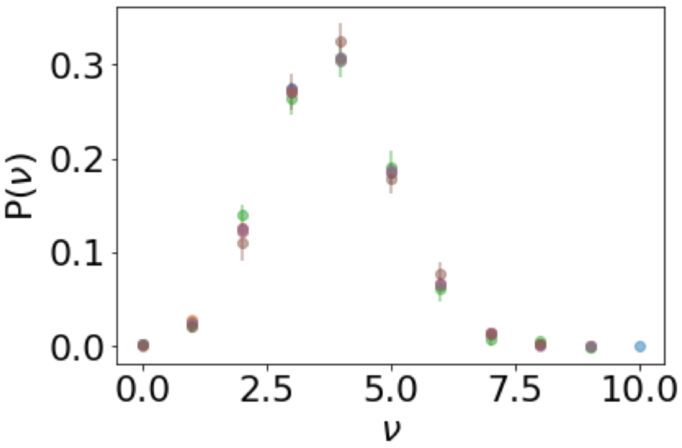
CGMF

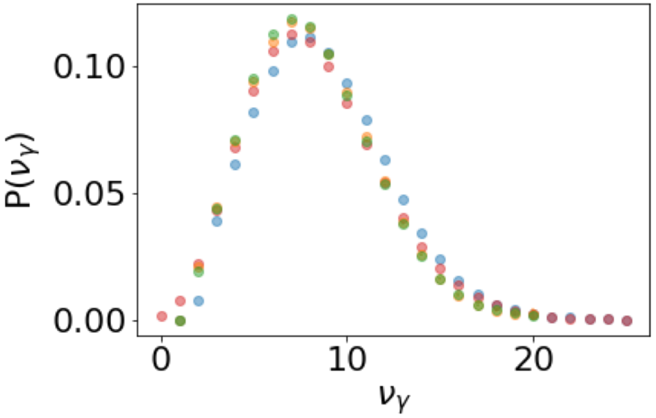
(Ac, Zc, En)





Hauser-Feschbach statistical decay

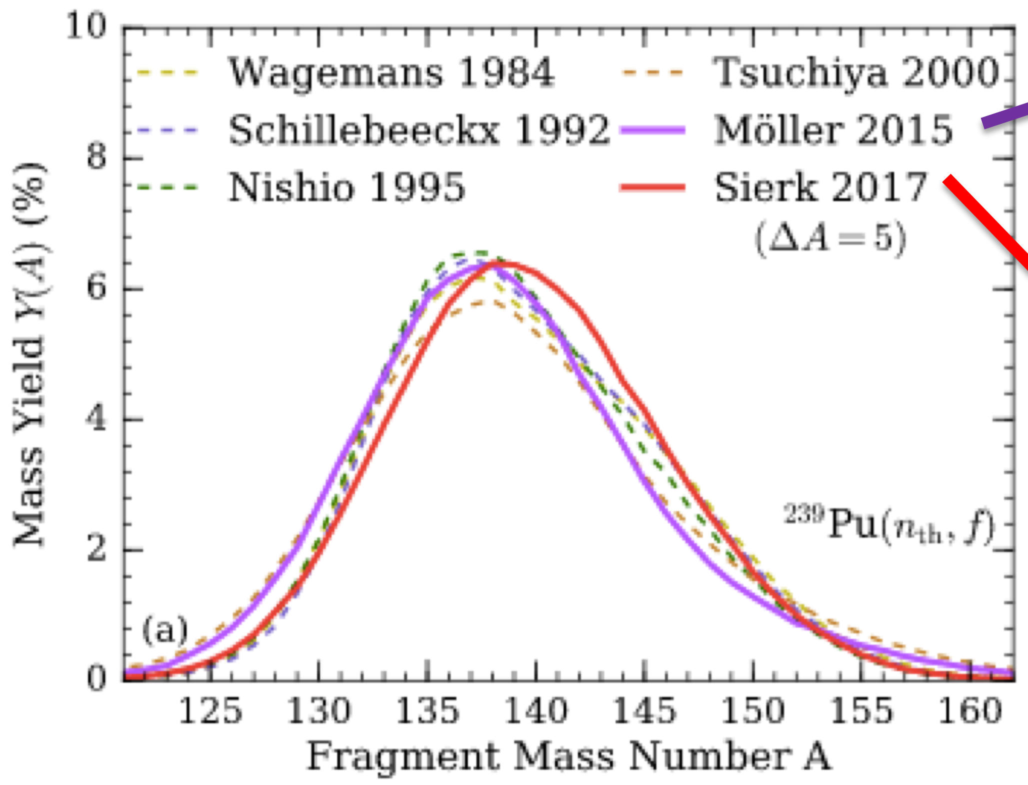




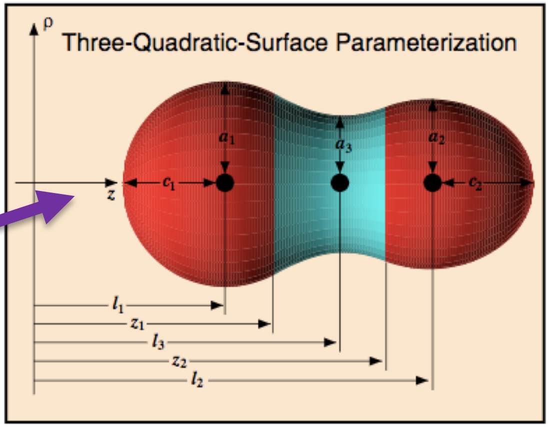
Compound nucleus ➤ yields ➤ prompt fission observables

P. Talou, et. al., Comp. Phys. Comm. in preparation

Fundamental Fission Information can be Included

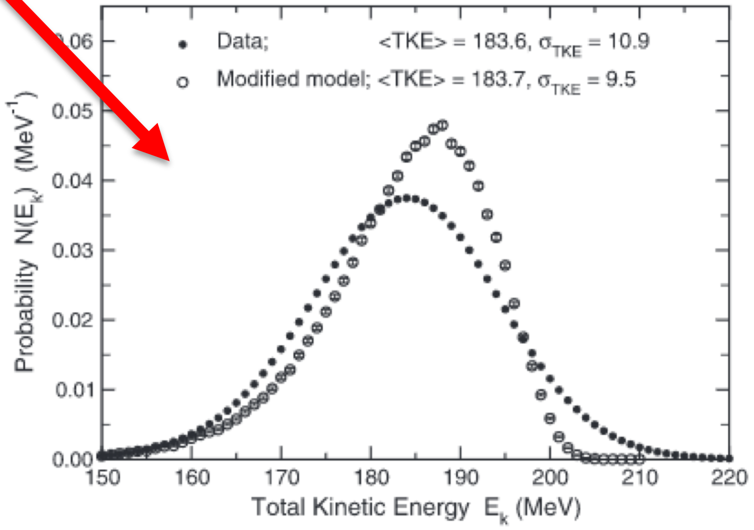


P. Jaffke, et. al., PRC **97**, 034608 (2018)



P. Moller and T. Ichikawa, Eur. Phys. J. A **51**, 173 (2015)

Brownian Motion: Mass and charge



A. Sierk, Phys. Rev. C **96**, 034603 (2017)

Langevin: Also give TKE

Brosa Modes for Y(A,TKE)

$$Y(A,TKE) = \sum_m Y_m(A) Y_m(TKE|A)$$

$$Y_m(A) = \frac{w_m}{\sqrt{8\pi\sigma_m^2}} \left[\exp\left(-\frac{(A-\bar{A}_m)^2}{2\sigma_m^2} \right) + \exp\left(-\frac{(A-A_{cn}+\bar{A}_m)^2}{2\sigma_m^2} \right) \right]$$

$$Y_m(TKE|A) = \left(\frac{200}{TKE} \right)^2 \exp\left(2\frac{d_m^{max}-d_m^{min}}{d_m^{dec}} - \frac{T_m(A)}{d_m^{dec}} - \frac{(d_m^{max}-d_m^{min})^2}{T_m(A) d_m^{dec}} \right)$$

$$T_m(A) = \frac{(Z_{cn}/A_{cn})^2 (A_{cn}-A) A e^2}{TKE} - d_m^{min}$$

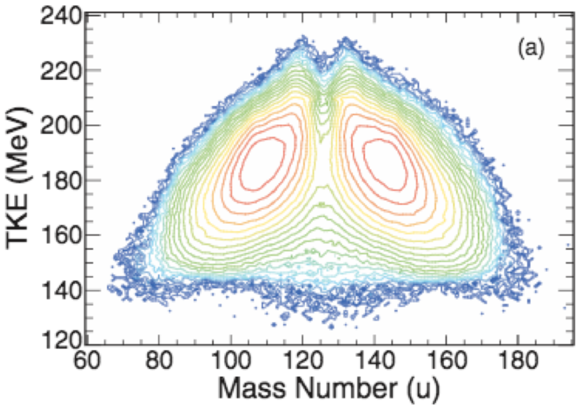
6 free parameters per mode:

w – weight of mode
$\bar{A}$ – mean heavy mass
σ – width of mass distribution

d^{max} – most probable fission semilength
d^{min} – semilength below which fission will not occur
d^{dec} – length scale for exponential decay

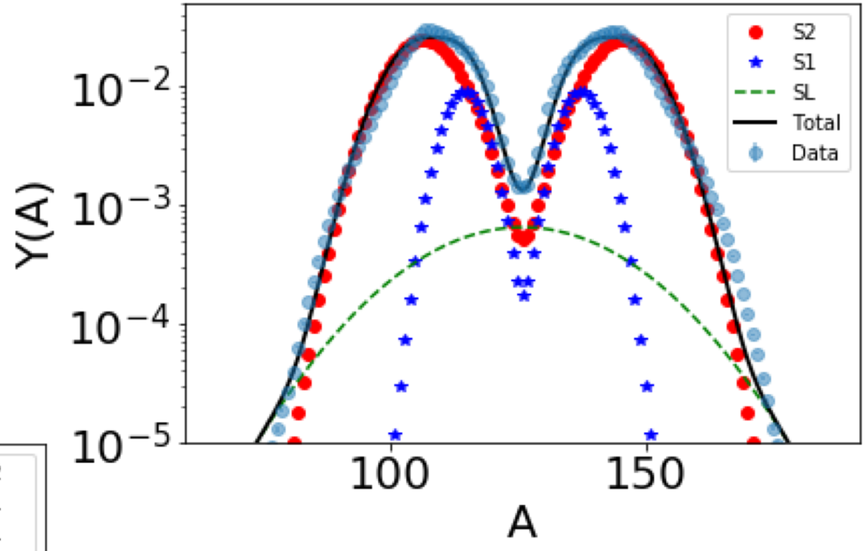
U. Brosa, et. al., Phys. Rep. **197**, 167 (1990)

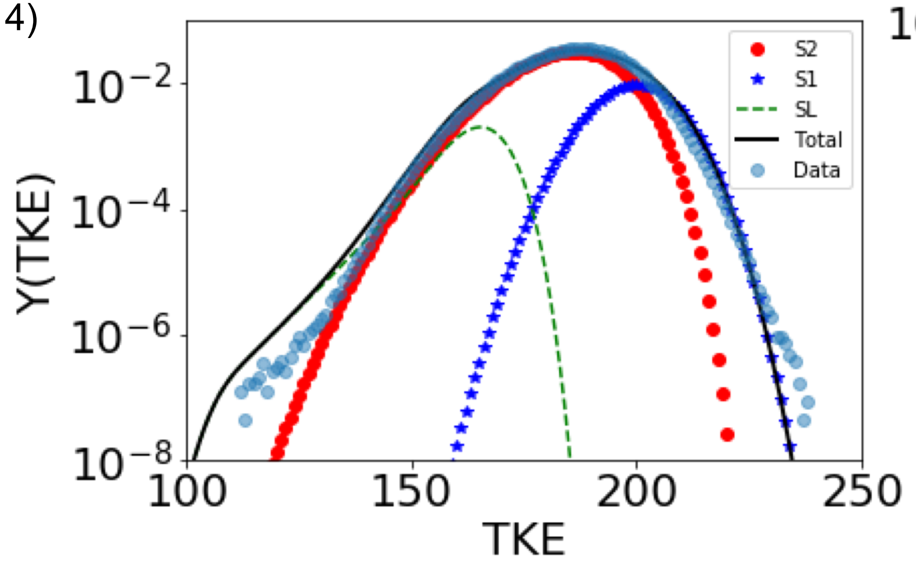
Previous work related to Brosa modes



A. Gook, et. al., PRC **90**, 064611 (2014)

^{252}Cf(sf)





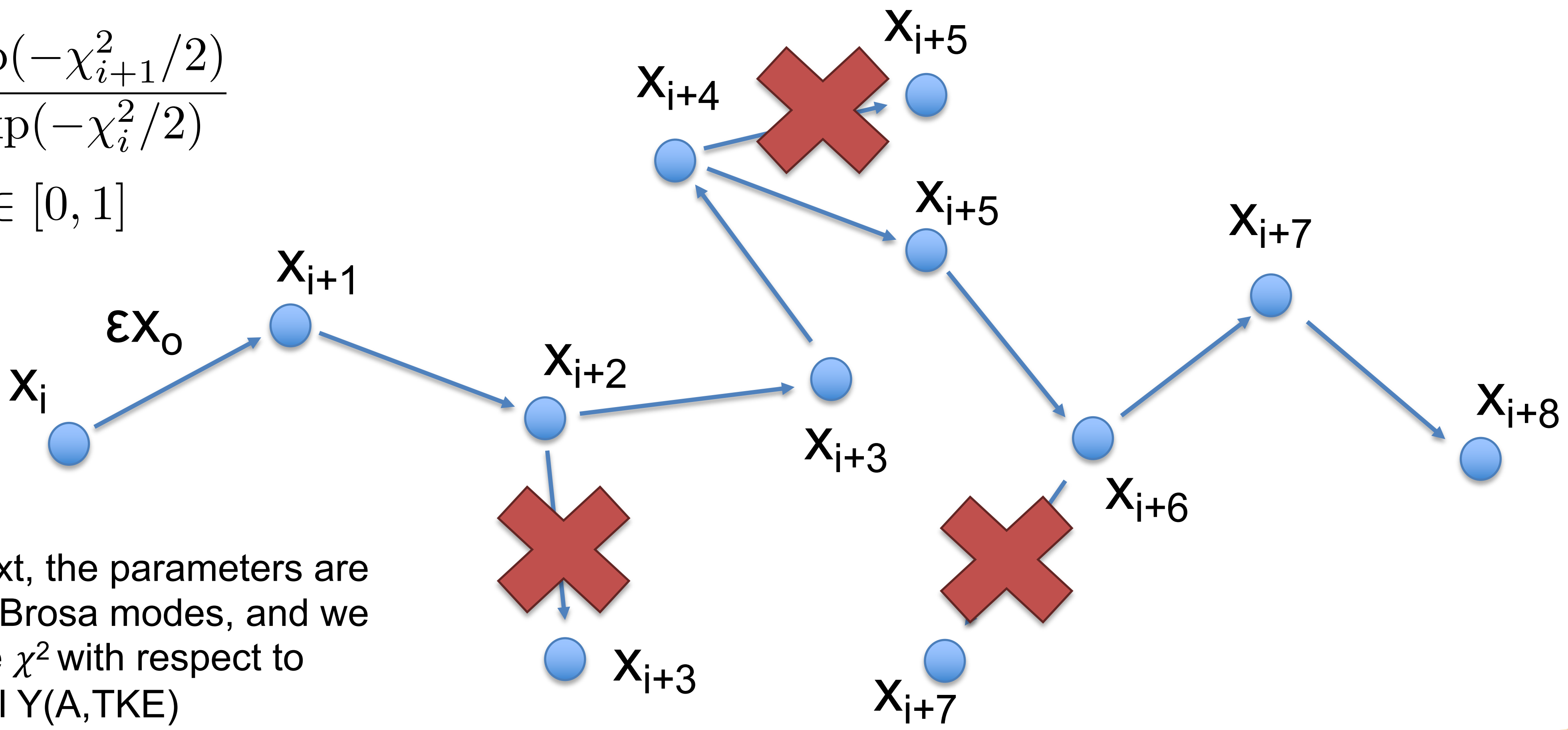
Parameterization from the Master's thesis of A. Carter, University of Michigan, *Sensitivity Analysis and Optimization of Cf-252(sf) Observables in CGMF* LA-UR-17-31022

$$\bar{\upsilon} = 3.82$$
$$\overline{TKE} = 184.91\ MeV$$

Markov Chain Monte Carlo

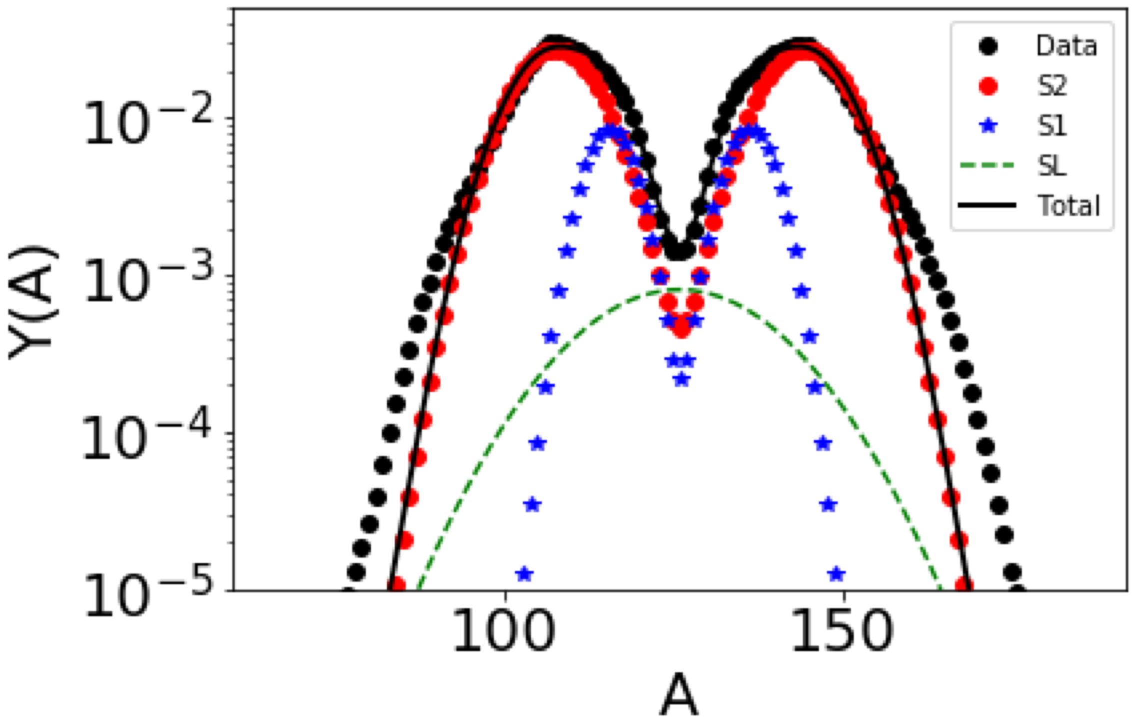
$$R < \frac{\exp(-\chi^2_{i+1}/2)}{\exp(-\chi^2_i/2)}$$

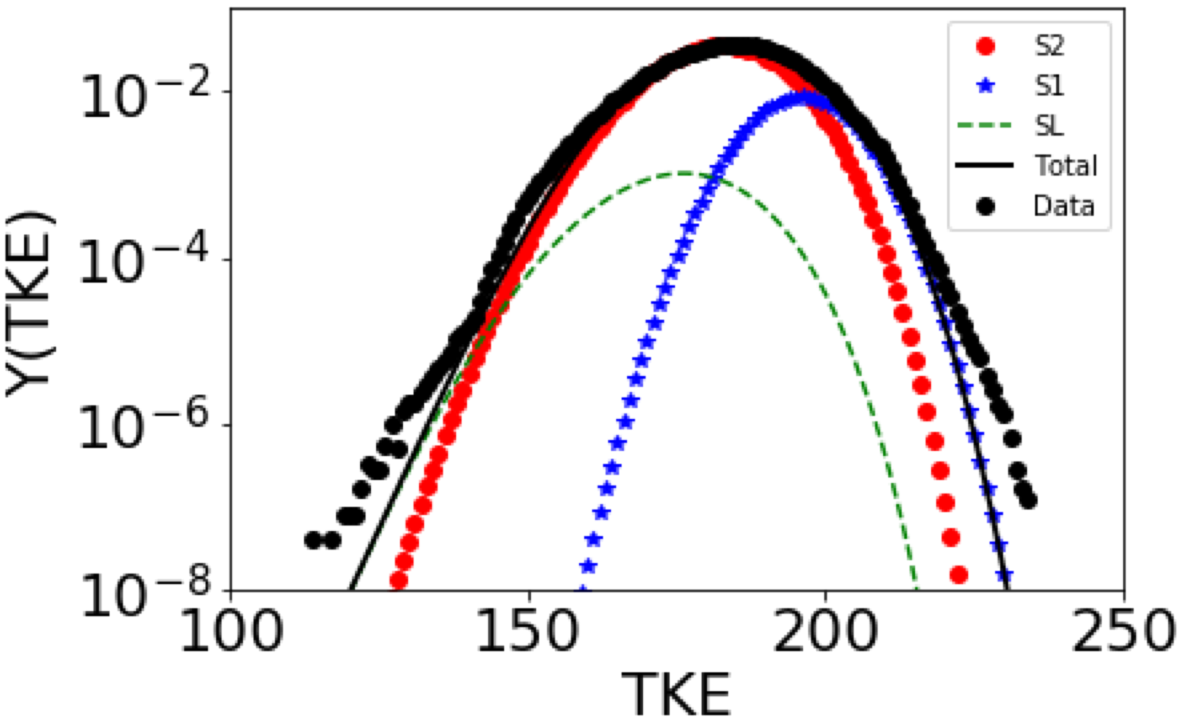
$$R \in [0,1]$$



In this context, the parameters are those of the Brosa modes, and we calculate the χ^2 with respect to experimental Y(A,TKE)

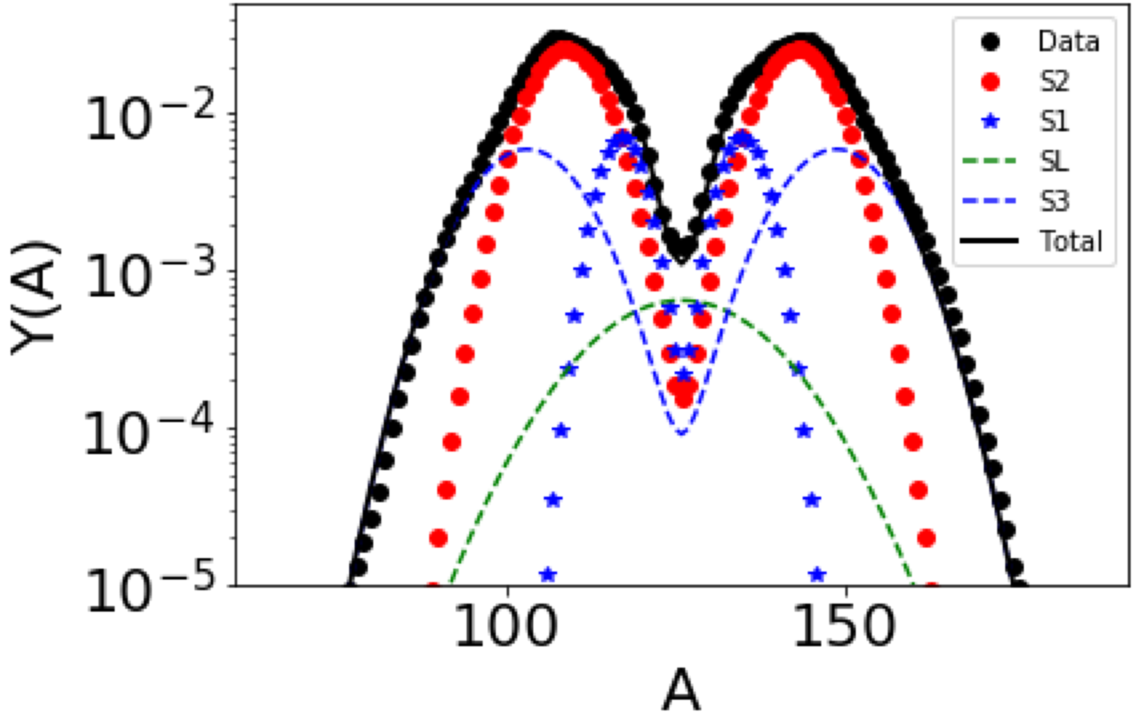
Three Brosa Modes: S2, S1, SL

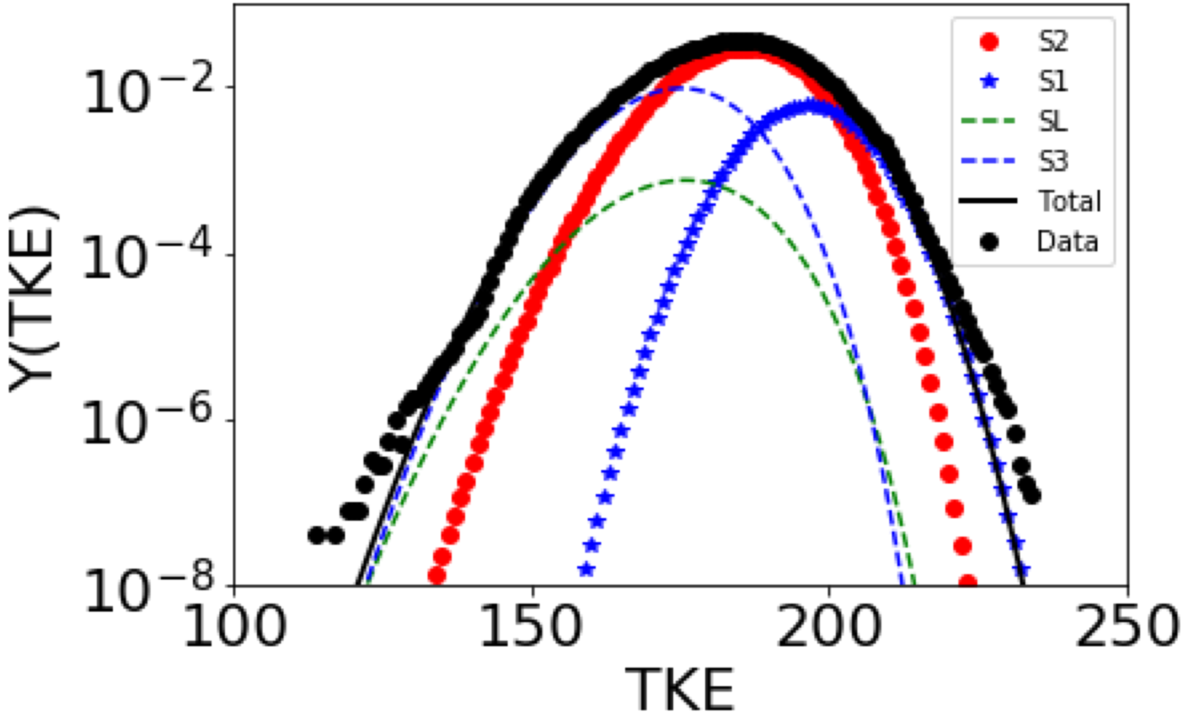




Data: A. Gook, et. al., PRC **90**, 064611 (2014)
Austin, et. al., Master's Thesis (2017)

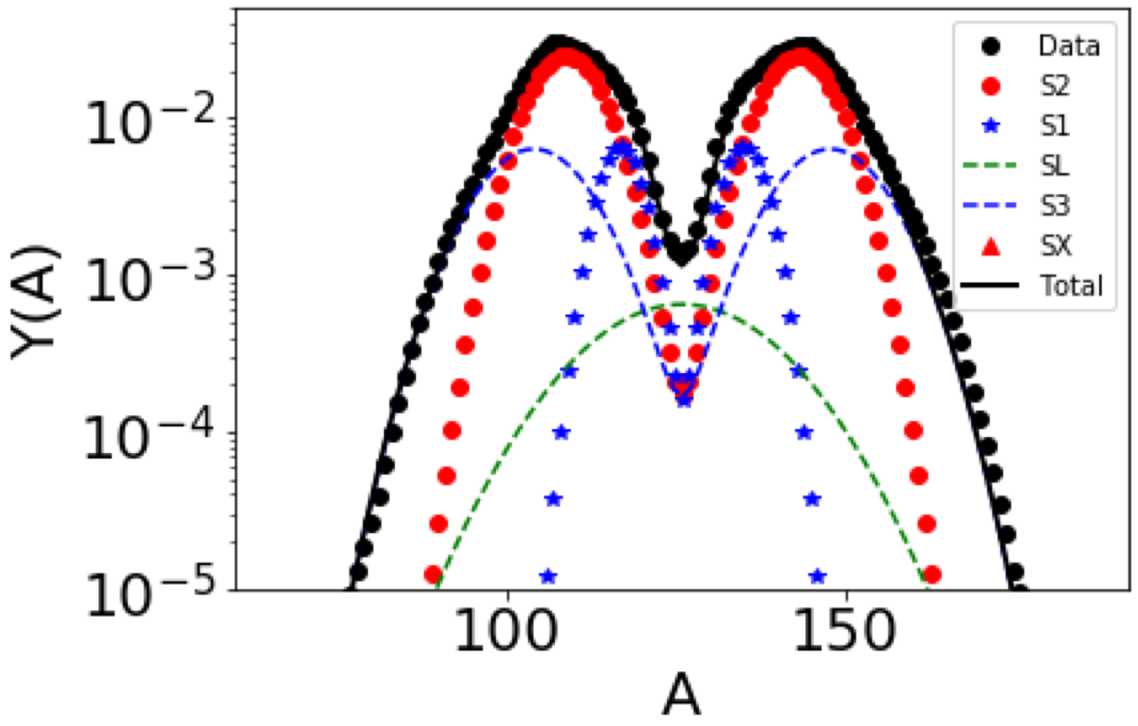
Including the S3 Mode



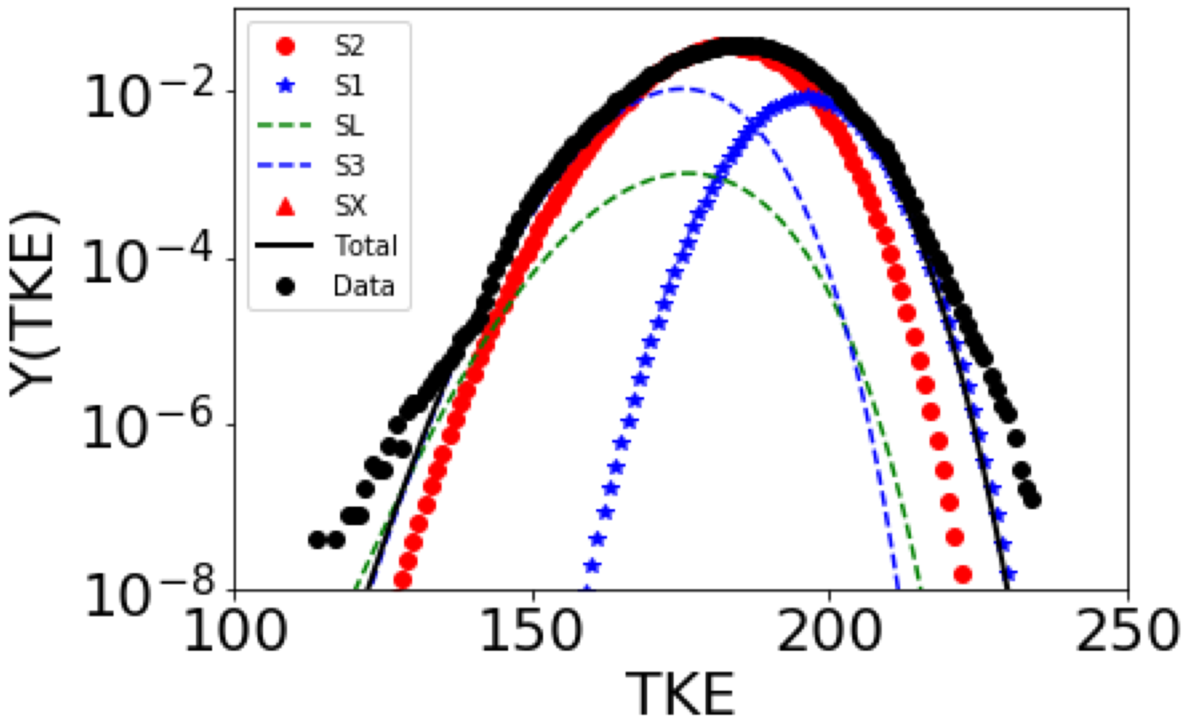


Data: A. Gook, et. al., PRC **90**, 064611 (2014)

Including the SX mode



The SX mode is orders of magnitude lower than everything else, it has no noticeable effect



There is still a tail of the TKE distribution that we are not reproducing

Data: A. Gook, et. al., PRC **90**, 064611 (2014)

Comparison of Fission Observables

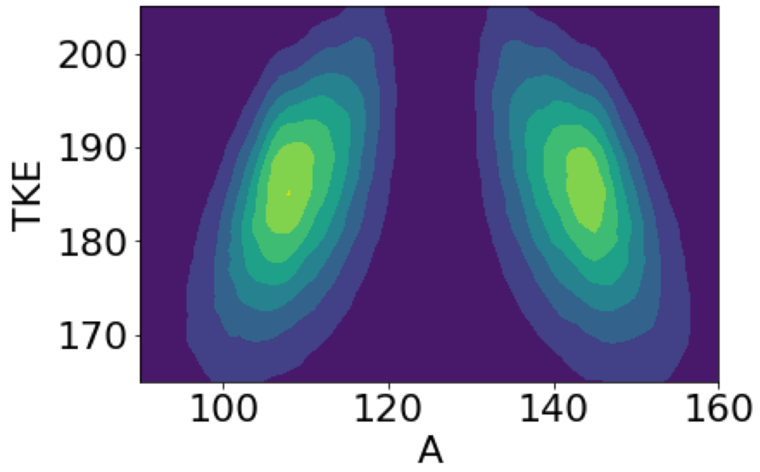
Calculation	<TKE>	σ_{TKE}	<v>
Default CGMF	185.77	8.966	3.758
S2, S1, and SL	184.46	10.20	3.956
Including S3	184.16	10.74	3.946
Including SX	184.27	10.71	3.948
Experiment	184.1	10.85	3.76

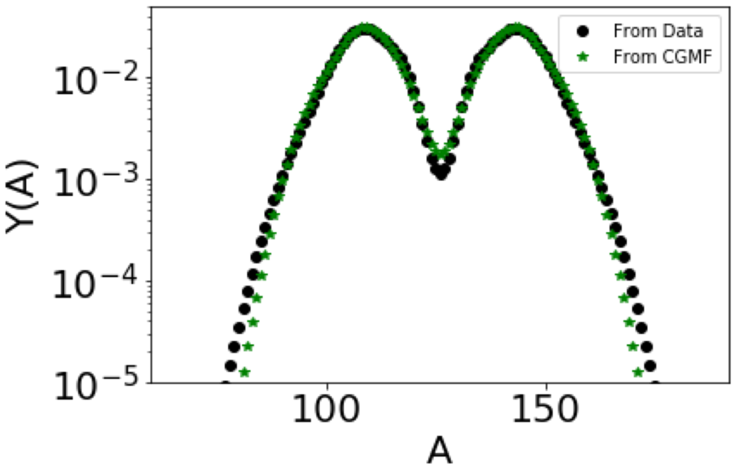
It is important to be able to reproduce <v> since this value is very well known; the discrepancy is most likely due to the value of <TKE>

Experimental values from:
A. Carlson, et. al., Nucl. Data Sheets **148**, 143 (2018)
C. Wagemans, *The Nuclear Fission Process*, CRC-Press (09 1991)

Constraining with Y(A,TKE) in CGMF

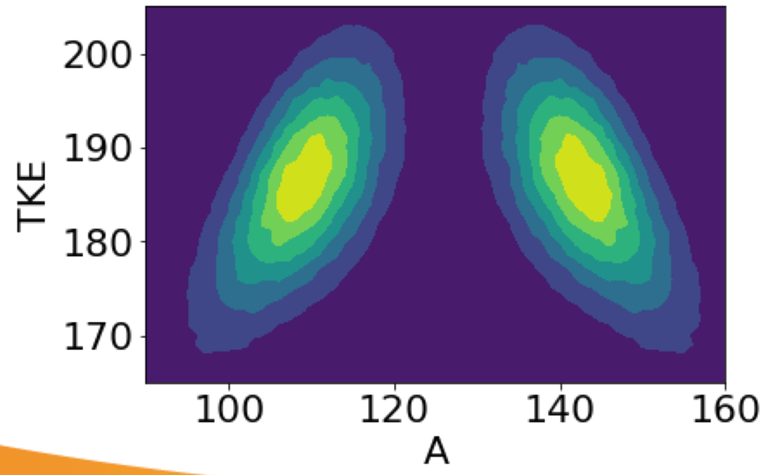
Gook

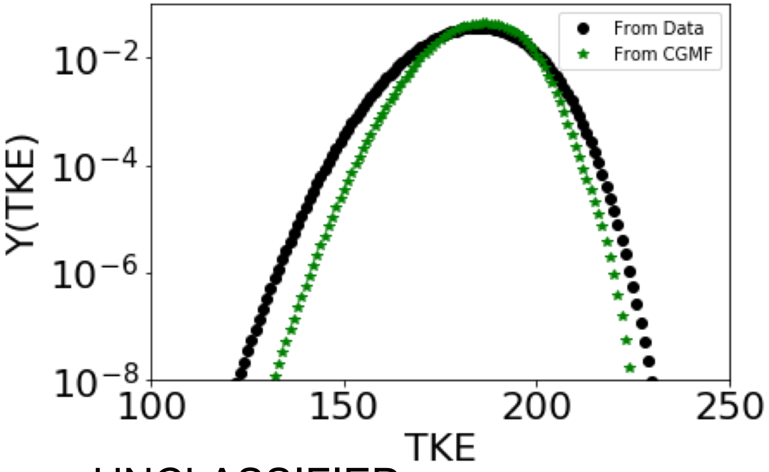




$$\bar{\nu}(Data) = 3.948$$
$$\overline{TKE}(Data) = 184.27\ MeV$$
$$\sigma_{TKE}(Data) = 10.71\ MeV$$

CGMF

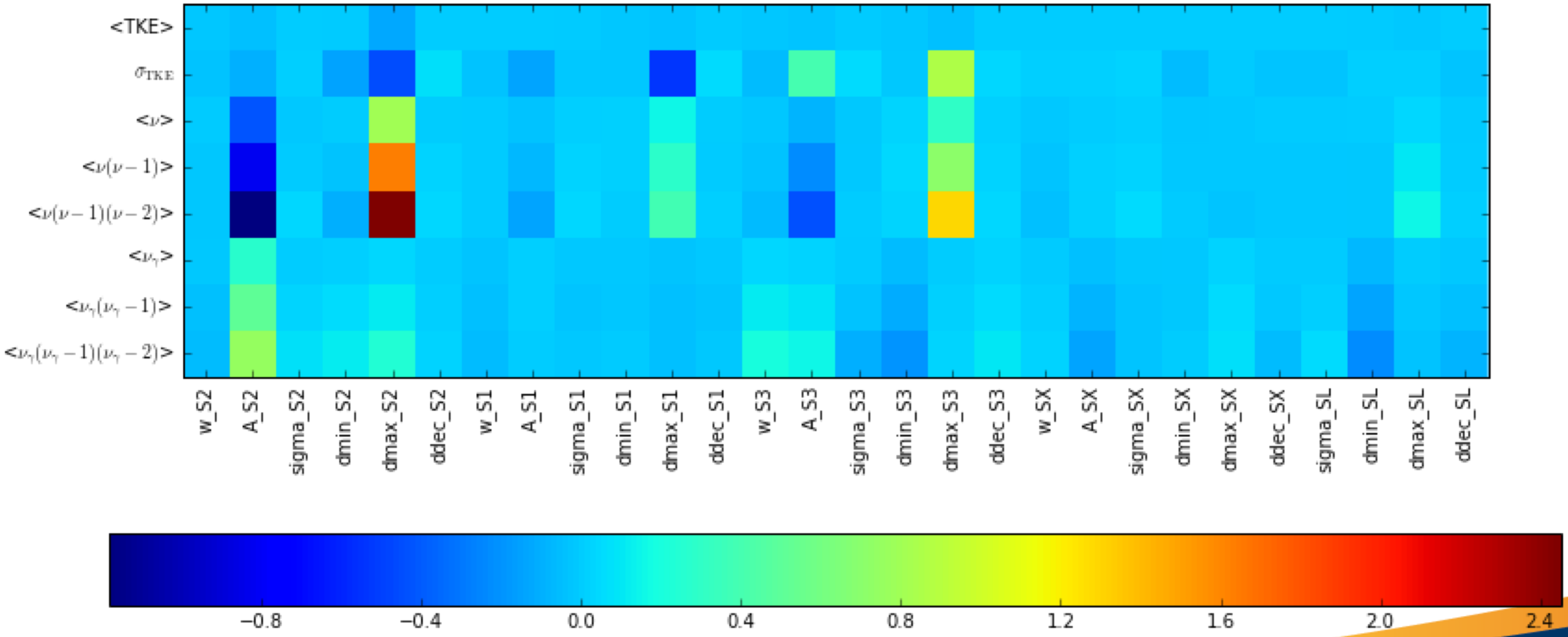




$$\bar{\nu}(CGMF) = 3.767$$
$$\overline{TKE}(CGMF) = 185.85\ MeV$$
$$\sigma_{TKE}(CGMF) = 8.85\ MeV$$

Sensitivity Analysis

$$S = \frac{x}{R}\frac{\partial R}{\partial x}$$



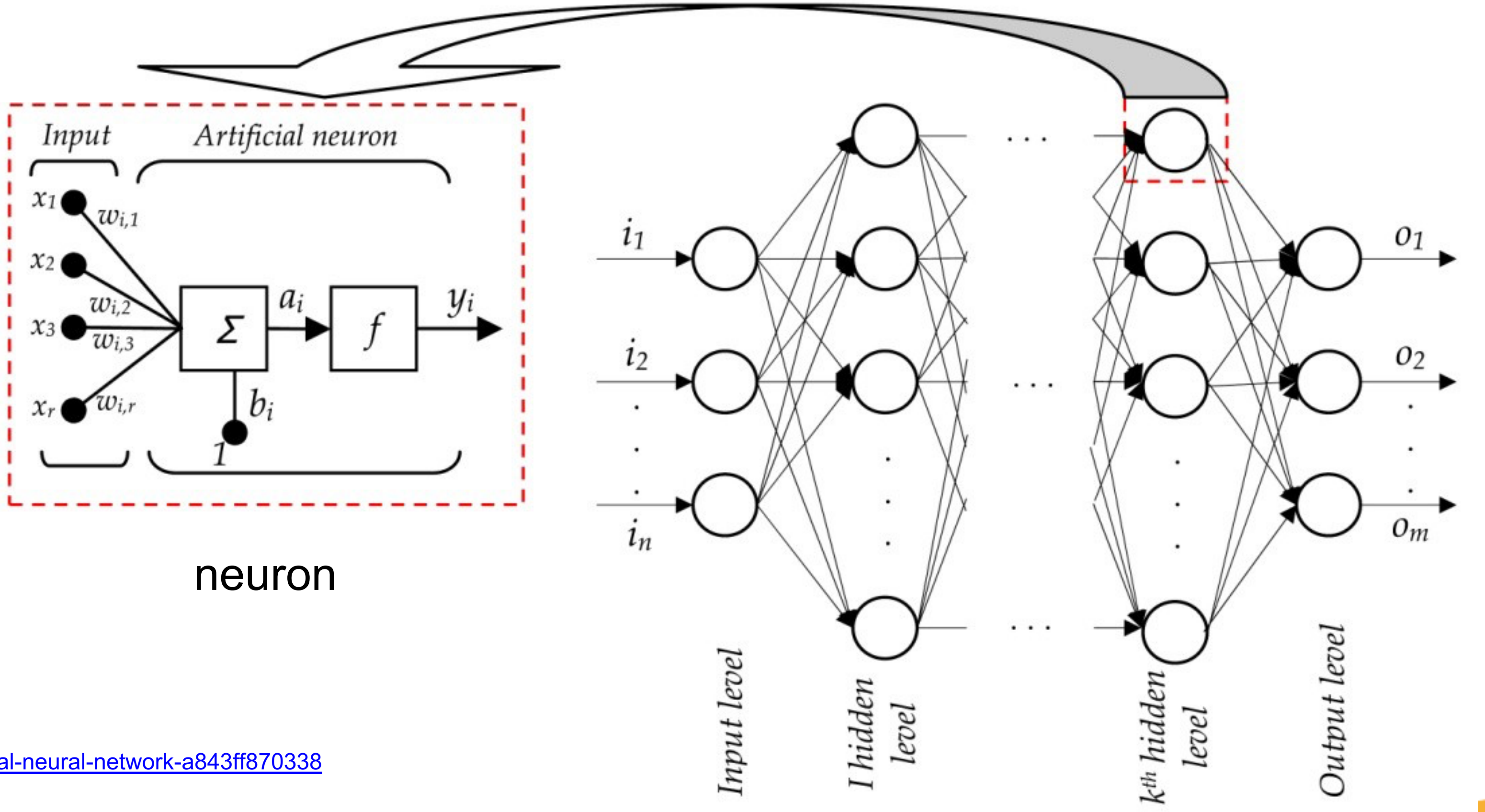
UNCLASSIFIED

Machine Learning Goals

- The problem has been broken down into two parts:
  - constructing the complete yields in A, Z, and KE
  - calculating the prompt neutron and gamma observables
- For the first, we want to construct yields based on experimental data
  - If this is done systematically, we have a way to fill in missing information from the table and make predictions for any fissioning system
- For the second, we want to calculate prompt neutron and gamma observables given a fissioning system
  - This will require a significant amount of data to capture the complex correlations between these observables

Neural Networks

- A neural network attempts to approximate a nonlinear mapping of $y = f(x)$ using large scale, data-driven optimization over hundreds/thousands/millions of parameters



neuron

Credit: https://hackernoon.com/artificial-neural-network-a843ff870338

Probabilistic Approach to Neural Networks

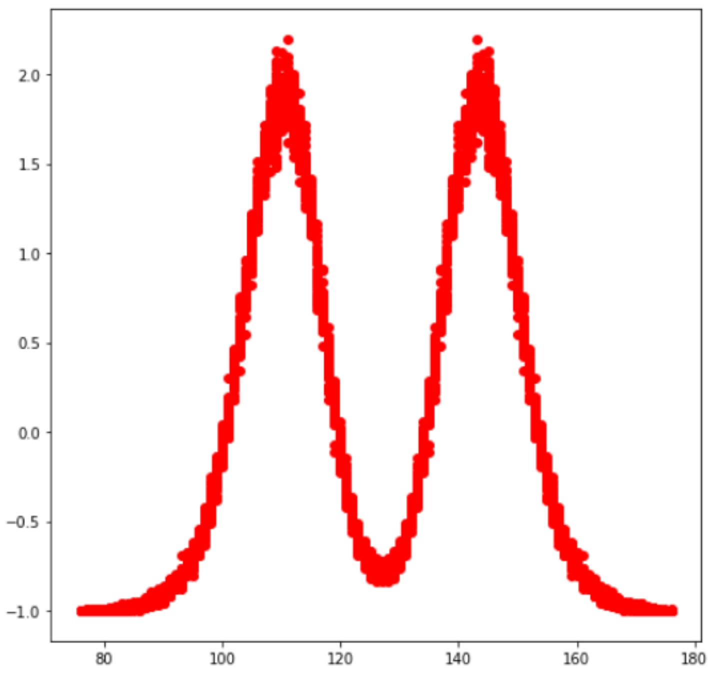
- Weights w_i and b_i for a standard NN are optimized based on Maximum Likelihood Estimation (MLE) → does not properly account for uncertainties in data (assumes all output are weighted equally).
- Our data – fission yields – and nuclear data in general consists of probability distributions and uncertainties (from experiment or model)
- **Standard NN approach:** Given A, predict a single value for Y(A). However, data contains true Y(A) mixed with an error ϵ → direct prediction can be erroneous/misleading.
- If the data has uncertainty, our predictions should also have uncertainty (and confidence bounds) → probabilistic predictions are necessary.
- **Our approach:** Mixture Density Networks (MDN) → predicts Y(A) as a mixture of Gaussians.

$$Y(A) = \alpha_1 G(\mu_1, \sigma_1) + \alpha_2 G(\mu_2, \sigma_2) + \ldots + \alpha_i G(\mu_i, \sigma_i)$$

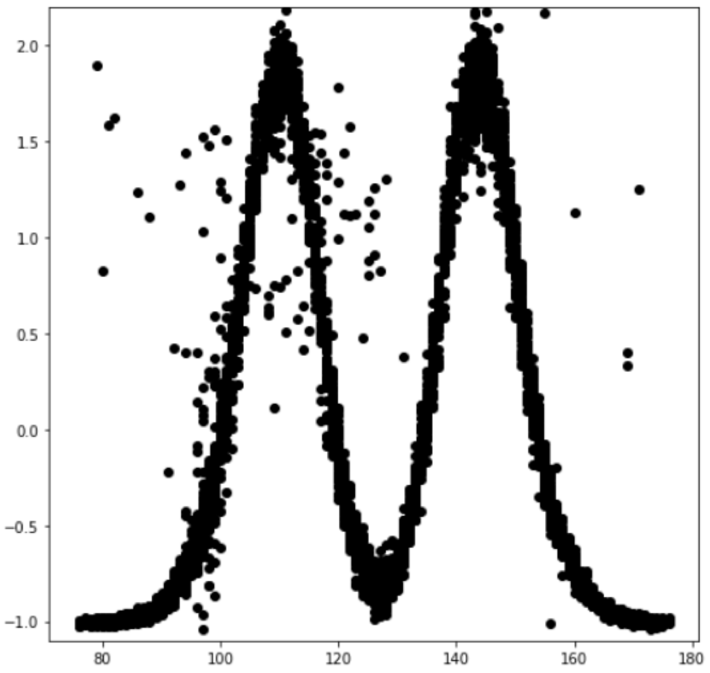
Parameters μ_i, σ_i of the Gaussians G and their combination coefficients α_i are predicted by the MDN → not the absolute values of Y(A). User has control over number of Gaussians in the mixture.

Task 1: MDN to reproduce ^{252}Cf yields

20 training data sets were simulated from CGMF



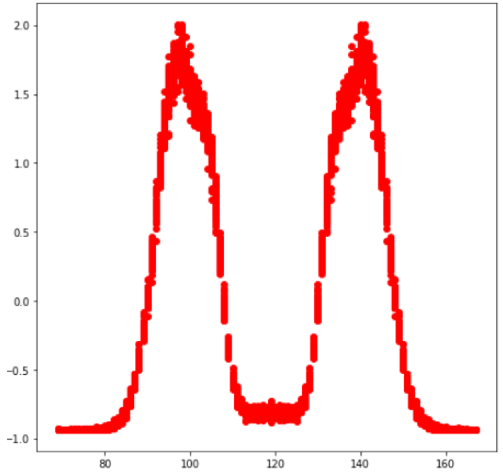
Training dataset ^{252}Cf



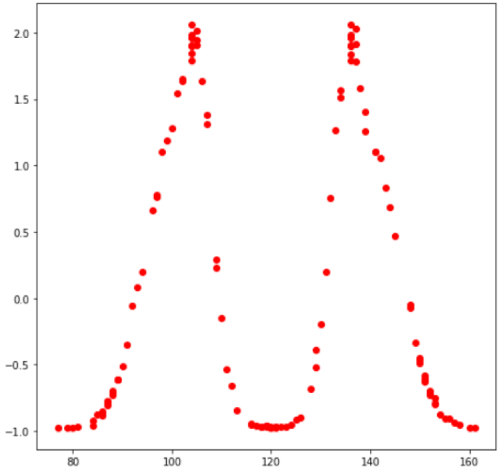
MDN Prediction

MDN can capture features of the data and samples multiple values for a single data point = can be used to understand confidence intervals in predictions and outliers.

Task 2: MDN Transfer learning from Cf → Pu

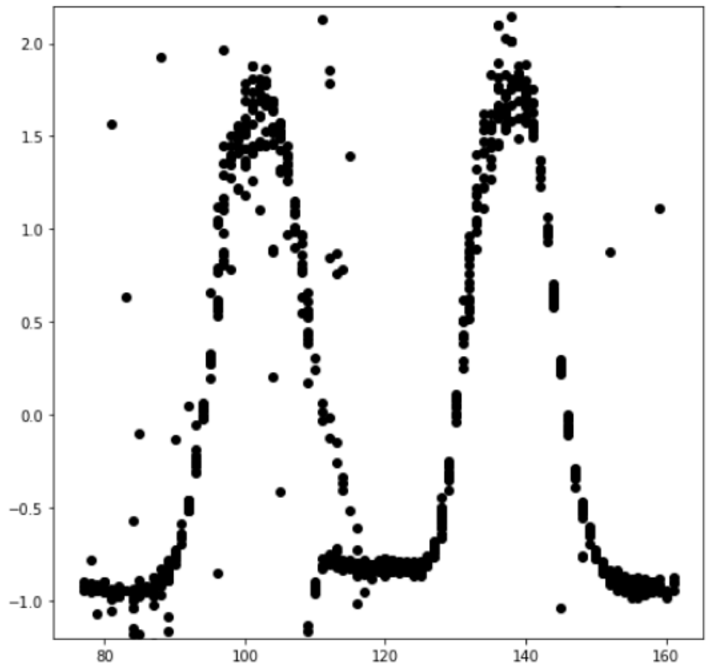


Full dataset

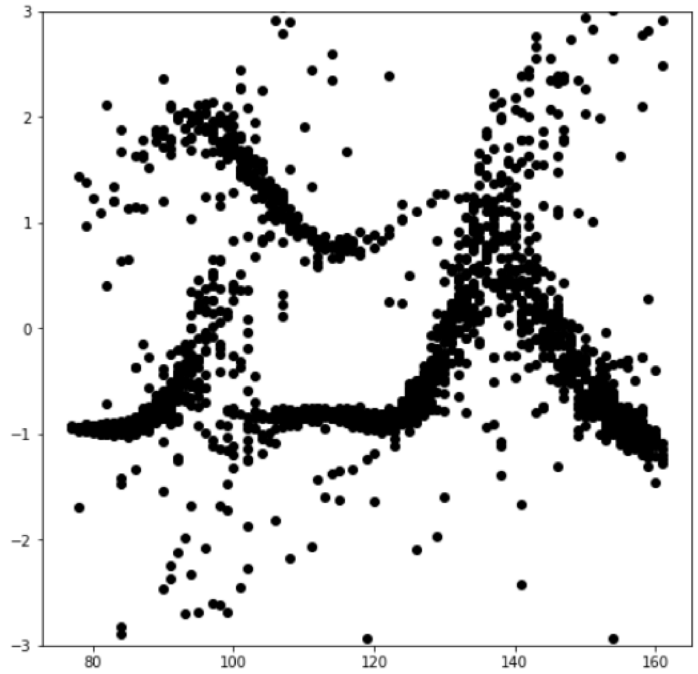


3% sparse dataset

- We use just 3 % of the data to train **MDN with Transfer learning.**
- *Methodology:* In 3 layer network, we use weights from the Cf-252 trained. Model in the 1st layer and train only the other 2 layers with the sparse data.



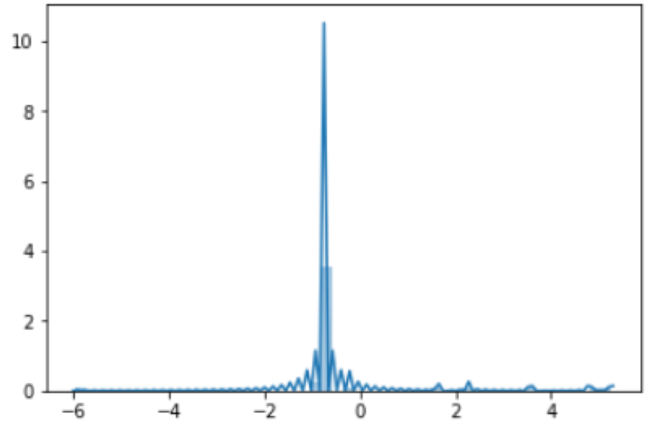
Transfer Learning - Predicted dense dataset
Structural similarities in *Cf* makes learning faster + interpolate gaps in sparse data!



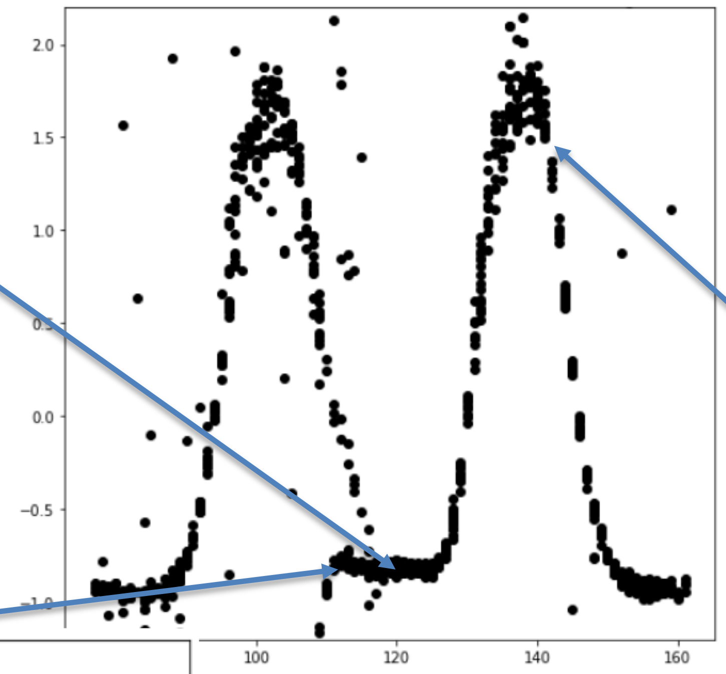
Direct Learning - Predicted dense dataset
Network struggles to learn full *Pu* distribution without information from *Cf*

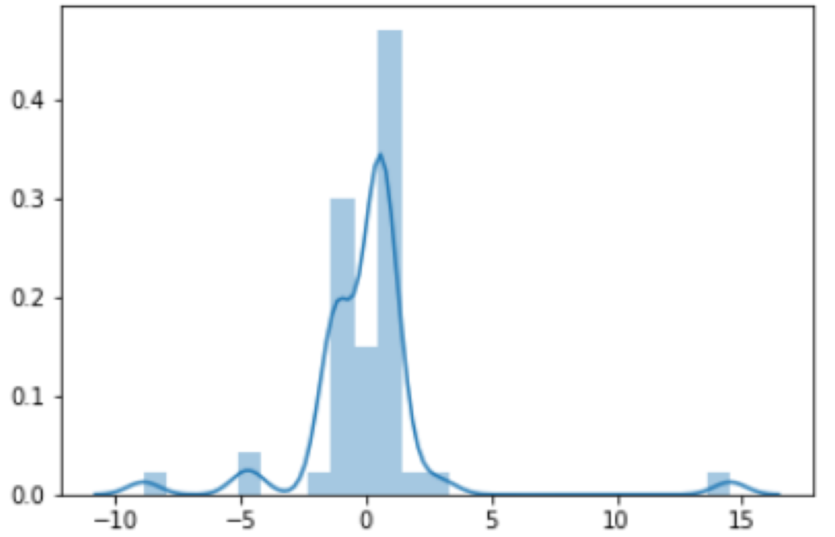
UNCLASSIFIED

Variance in MDN Predictions

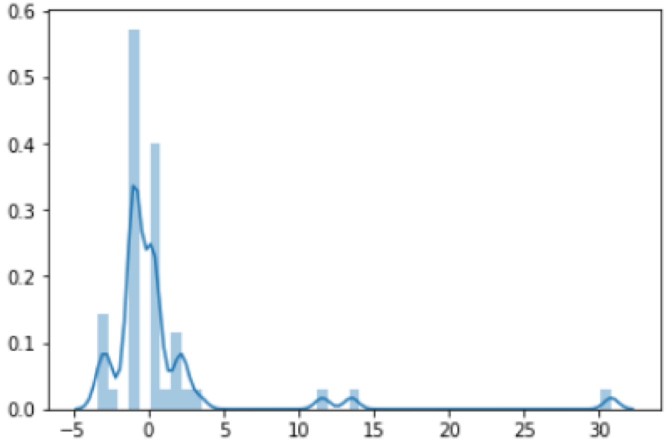


Low Variance region





High Variance region



High Variance region

- Each point in the predicted domain can be analyzed for MDN's confidence in prediction
- Further processing of these variances can estimate overall uncertainty in predictions.
- Acknowledges there are uncertainties in input data.

UNCLASSIFIED

Summary and Conclusions

- We have started using the Brosa parameterization of Y(A,TKE) for spontaneous fission of ^{252}Cf.
- Markov Chain Monte Carlo has been successfully used to optimize the parameters of this model.
- The response of several observables to changes in these parameters have been calculated in preparation for a global optimization.
- Probabilistic approach to machine learning along with transfer learning show considerable promise in building efficient and fast emulators for fission yields.

Future Work

- The Brosa modes are being implemented into CGMF in order to optimize the parameters with respect to the prompt neutron and gamma observables.
- We are working with others at LANL to use more sophisticated optimization techniques to constrain these parameters.
- We are working on stability issues in MDN training, allowing us to study its characteristics and add physics based constraints (symmetry, normalization, etc.).
- Use MDN with transfer learning for to fill in missing nuclear data across isotopes and energies.

Acknowledgements

- Ionel Stetcu, Samuel Jones, Harsha Nagarajan, LANL

- NNSA

- Center for Non-Linear Studies