Uncertainty Propagation with Mixture Density Networks

A.E. Lovell
In collaboration with A. Mohan, P. Talou, and M. Chertkov
September 19, 2019
Bayesian Inference in Subatomic Physics
Highlights from MSU/NSCL few-body group: Uncertainty quantification for optical potentials

\[ P(\mathcal{H}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{H})P(\mathcal{H})}{P(\mathcal{D})} \]

\[ f(r; V_o, R_o, a_o) = -\frac{V_o}{1 + e^{(r-R_o)/a_o}} \]


Highlights from MSU/NSCL few-body group: Comparison between Bayesian and frequentist

G.B. King, A.E. Lovell, L. Neufcourt, and F.M. Nunes

*PRL 122*, 232502 (2019)

\[ ^{48}\text{Ca}(n,n) @ 12 \text{ MeV} \]
Highlights from MSU/NSCL few-body group: Experimental constraints to reduce uncertainties

Information content of the angular distributions

Including other reaction data

Including scattering at nearby energies

Size of the experimental error bars

M. Catacora-Rios, G.B. King, A.E. Lovell, and F.M. Nunes

In prep.

\[ \begin{array}{|c|c|c|}
\hline
\text{Reaction} & \Delta \varepsilon_{20/10} & \Delta \varepsilon_{10/5} \\
\hline
^{48}\text{Ca}(n,n) \text{ at } 12 \text{ MeV} & 1.53 & 1.94 \\
^{48}\text{Ca}(p,p) \text{ at } 12 \text{ MeV} & 1.68 & 1.71 \\
^{48}\text{Ca}(p,p) \text{ at } 21 \text{ MeV} & 1.55 & 1.74 \\
^{48}\text{Ca}(d,p) \text{ at } 21 \text{ MeV} & 1.68 & 1.52 \\
^{208}\text{Pb}(n,n) \text{ at } 30 \text{ MeV} & 1.62 & 1.79 \\
^{208}\text{Pb}(p,p) \text{ at } 30 \text{ MeV} & 1.39 & 1.61 \\
^{208}\text{Pb}(p,p) \text{ at } 61 \text{ MeV} & 1.99 & 1.74 \\
^{208}\text{Pb}(d,p) \text{ at } 61 \text{ MeV} & 1.41 & 1.58 \\
\hline
\end{array} \]
Uncertainty Propagation with Mixture Density Networks

A.E. Lovell
In collaboration with A. Mohan, P. Talou, and M. Chertkov
September 19, 2019
Bayesian Inference in Subatomic Physics
The fission process

Diagram showing the fission process with labels for different stages and time scales.

- \( Y_{\text{pre}}(A,Z) \)
- \( Y_{\text{post}}(A,Z) \)
- \( Y_{\text{post, } \beta\text{-decay}}(A,Z) \)

Time scale in seconds:
- \( 10^{-14} \)
- \( 10^{-14} \)
- \( 10^{-21} \)

Stages:
- Ground State
- Saddle
- Scission

Processes:
- Prompt Neutron Emission
- Prompt Gamma Emission
- Delayed Neutron and Gamma Emission

\( n \rightarrow p + \epsilon + \nu_e \)
Describing correlated fission observables

Initial conditions are fragment yields: \( Y(A,Z,TKE,J,\pi) \)

These initial conditions are strongly correlated to the resulting neutron and \( \gamma \)-ray observables (e.g. energies and numbers). They often cannot be calculated or measured directly.

CGMF: P. Talou, et al., *in prep*
Our current model is limited to fissile targets where there is measured data for many observables – predictive power is limited.


More fundamental models exists to calculate mass yields, but are still not at the accuracy needed for many applications.
Nuclear data is often sparse
Mixture Density Network (MDN)

Standard neural network

Input → output

\[ y = f(x) \]

Neural network learns the Gaussian variables instead of the mapping between x and y directly

\[ f(x) = \alpha_1 \mathcal{N}(\mu_1, \sigma_1) + \alpha_2 \mathcal{N}(\mu_2, \sigma_2) + \ldots + \alpha_n \mathcal{N}(\mu_n, \sigma_n) \]

Sources of uncertainty:
- Inconsistent experimental data
- Experimental errors
- Detector response

Figure: [https://hackernoon.com/artificial-neural-network-a843ff870338](https://hackernoon.com/artificial-neural-network-a843ff870338)

Uncertainty quantification with MDN

\[ f(x) = \alpha_1 \mathcal{N}(\mu_1, \sigma_1) + \alpha_2 \mathcal{N}(\mu_2, \sigma_2) + \ldots + \alpha_n \mathcal{N}(\mu_n, \sigma_n) \]

When making predictions, the full posterior distribution can be sampled – instead of just calculating a mean and standard deviation.

Correlations between data can be included in the training set.
Numeric details for the MDN

• Mixture Density Network
  – 1 Gaussian mixture*
  – 4 layers
  – 10 nodes per layer

• Written in pytorch
• Runs on CPU or GPU

* For training sets with smaller uncertainties (~1%), more Gaussians were required in the network for numerical stability (even if they ended up with no weight)

• Training Sets
  – 3000 samples of A in \([A_c/2-50,A_c/2+50]\)
  – Each Y(A) sampled independently from a Gaussian centered at the accepted value
  – Y(A) mean values taken from fits to experimental data or from data directly

• Testing Sets
  – A in \([A_c/2-50,A_c/2+50]\)
  – 10,000 values of A sampled
Convergence study: Number of training data

Los Alamos National Laboratory
Convergence study: Reflection of the experimental uncertainty

1% uncertainty

5% uncertainty

10% uncertainty

252Cf(sf)

15% uncertainty

20% uncertainty
Convergence study: Reflection of the experimental uncertainty

$^{252}\text{Cf(sf)}$

2σ

1σ

3σ

Percent within Interval

Percent Error on Training Set

MDN calculations

Training

Testing
Training on experimental data

100 samples drawn from each experimental data point – all assumed to be uncorrelated (poor assumption, we can do better)
Comparison to the MDN prediction

(b) $^{252}\text{Cf}(sf)$

(c) $^{252}\text{Cf}(sf)$
Verifying physical constraints

Yields are normalized to 2

Yields should be symmetric about $A_c/2$

(a) $^{252}\text{Cf}(sf)$

(b) Relative Peak difference vs. $A$
Interpolating between training sets

Only train on the energy grid $E = [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$ MeV

Exact $Y(A)$ taken from fit to data as a function of neutron incident energy – 5% uncertainty included
Interpolating between training sets

Predictions made from the MDN between training points

\[ E_{\text{inc}} = 1.56 \text{ MeV} \]

\[ ^{235}\text{U}(n,f) \]

\[ E_{\text{inc}} = 2.98 \text{ MeV} \]

\[ ^{235}\text{U}(n,f) \]
Summary

• Fission observables, in particular those needed to constrain inputs for decay models, are sparse across isotope and incident energy

• To fill in this sparse landscape, we propose using a novel machine learning technique, the Mixture Density Network, which enables us to make predictions with full posterior distributions without having a parametric model

• We are able to reproduce fission mass yields for $^{252}$Cf(sf) and $^{235}$U(n,f) with a similar level of uncertainty as is included on the training data, as well as interpolate the mass yields as a function of incident energy
Outlook and future work

• How well does the method extrapolate beyond the range of the fitted data?

• Can we easily incorporate partial information (e.g. only knowing what the most likely heavy/light masses are)?

• The next study will be on multi-dimensional data – making predictions across an isotopic chain or in a specific region of the nuclear chart (can we also do this globally?)

• In addition, we ultimately would predict the full fission yields in mass, charge, total kinetic energy, spin, and parity.