Simulating cosmic ray muon spallation in Hyper-Kamiokande for a DSNB analysis

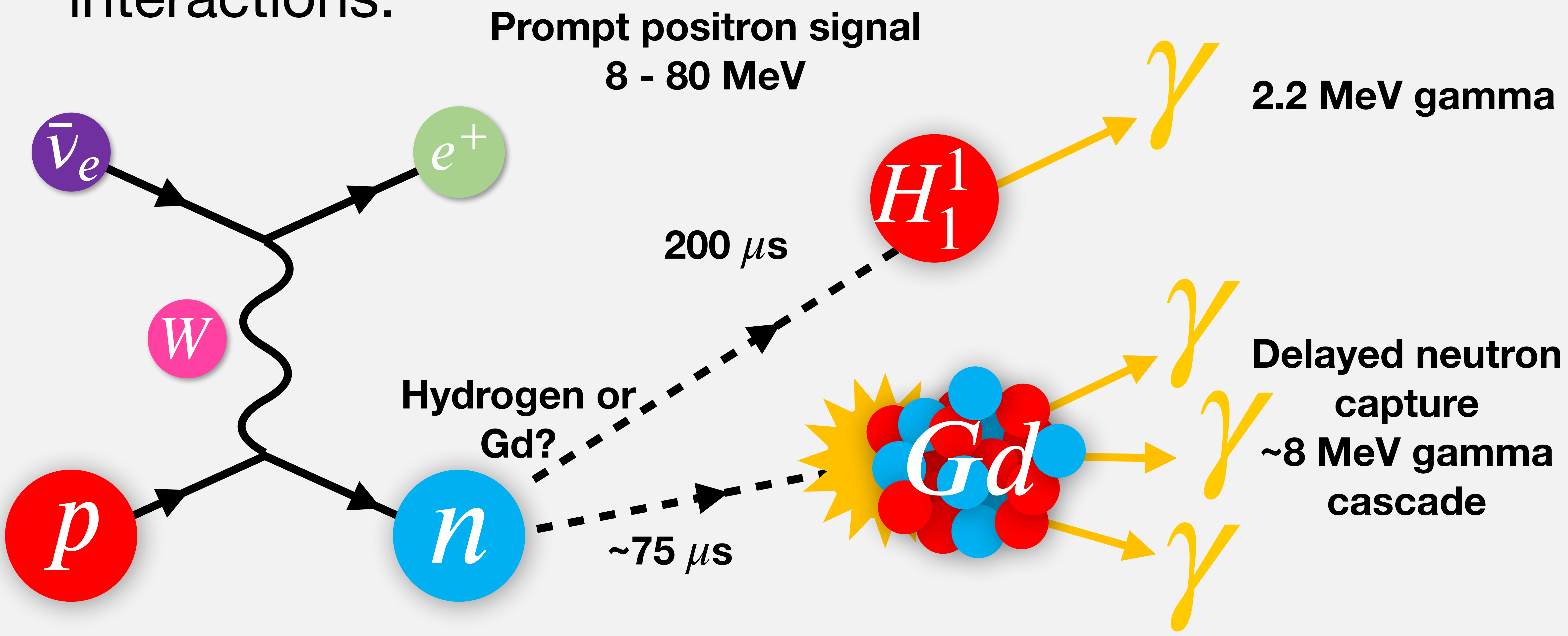
Jack Fannon

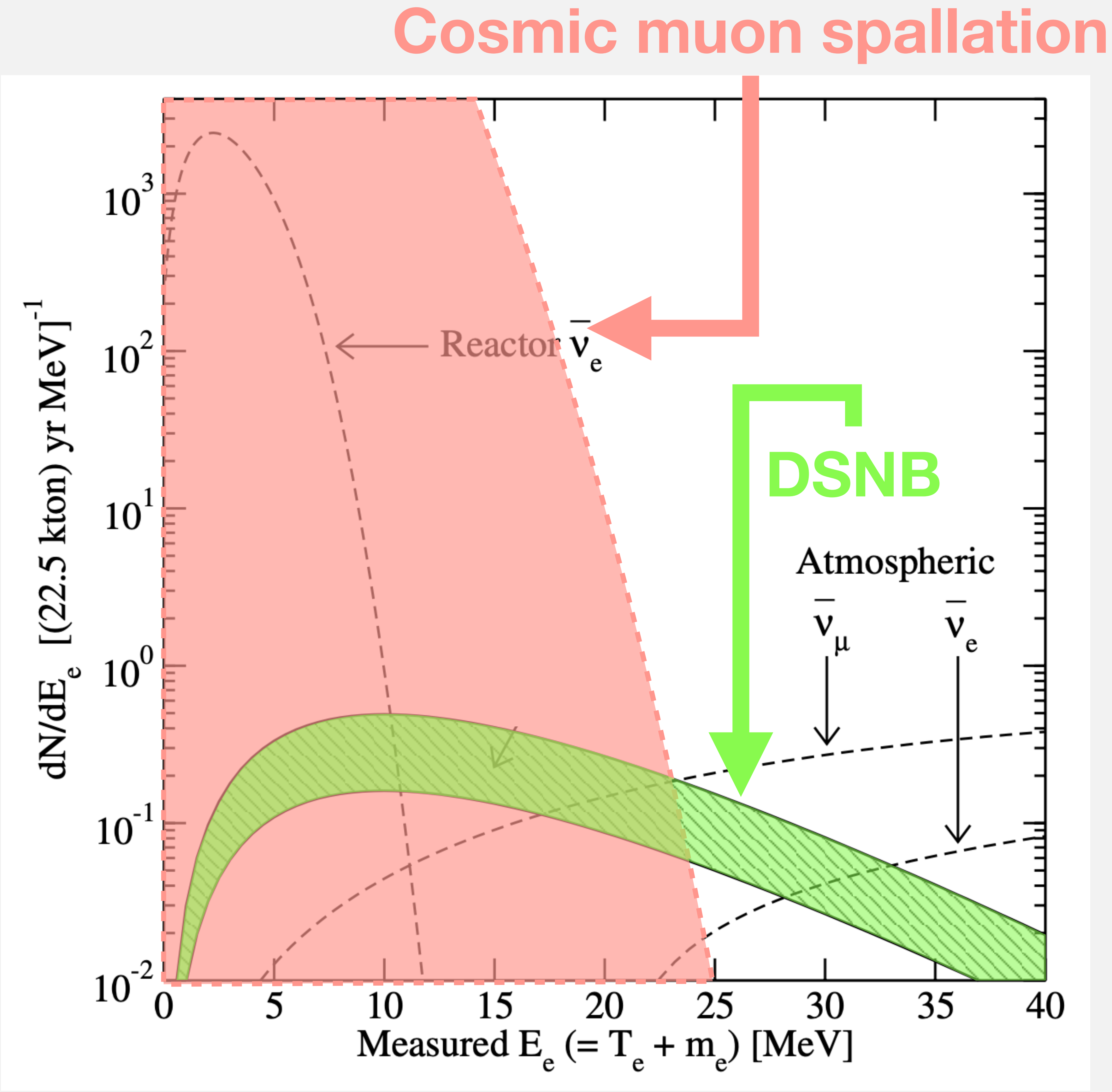
University of Sheffield

What are we trying to detect?

University of Sheffield | Jack E.P. Fannon

- **Diffuse supernova neutrino background** (DSNB) is the cumulative neutrino flux created by all past core collapse supernovae in the universe.
- Estimate (1 - 4) $\times 10^{-4}$ events/kton year from 8-80 MeV.
- Can be identified through inverse beta decay interactions:



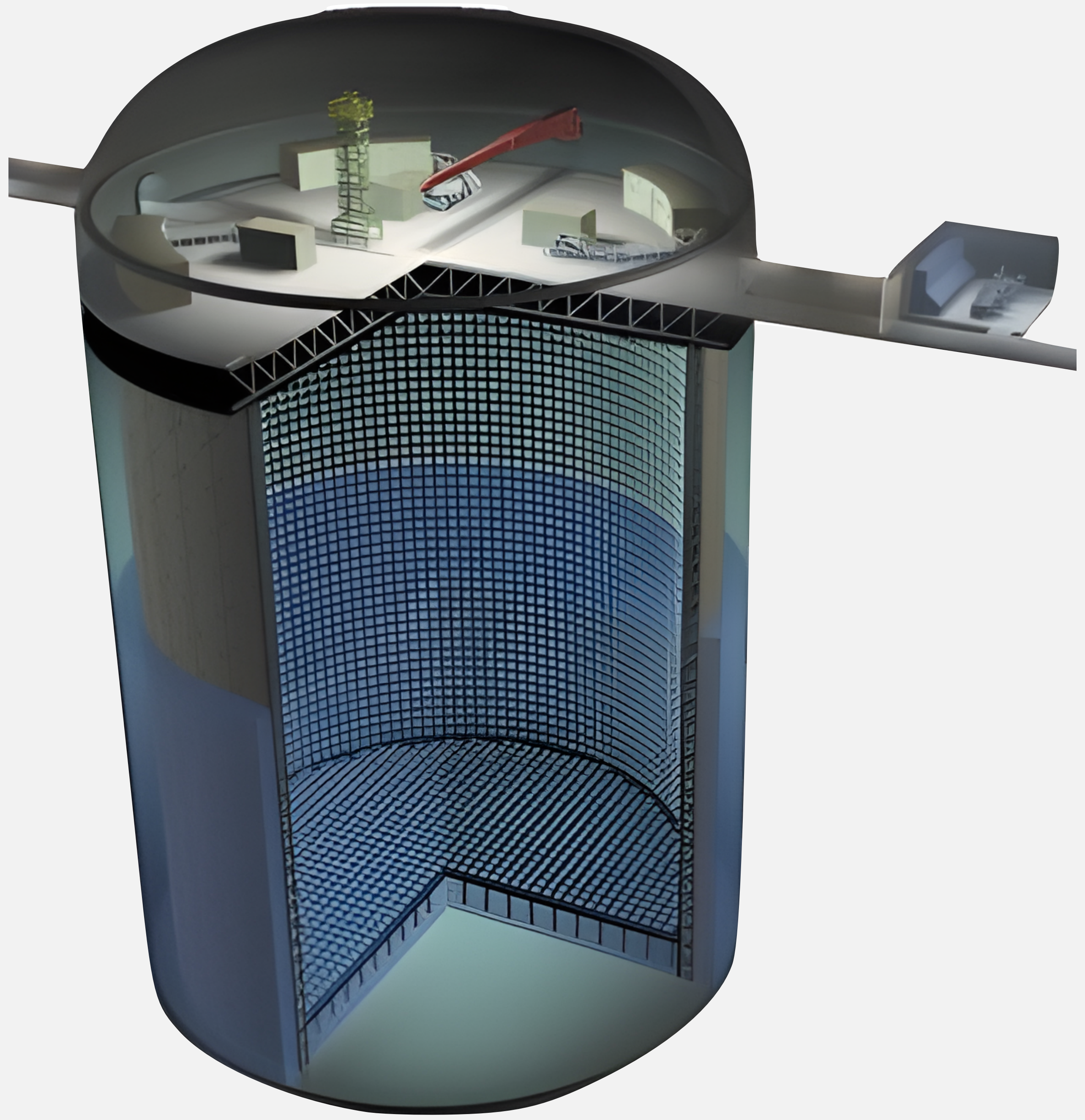


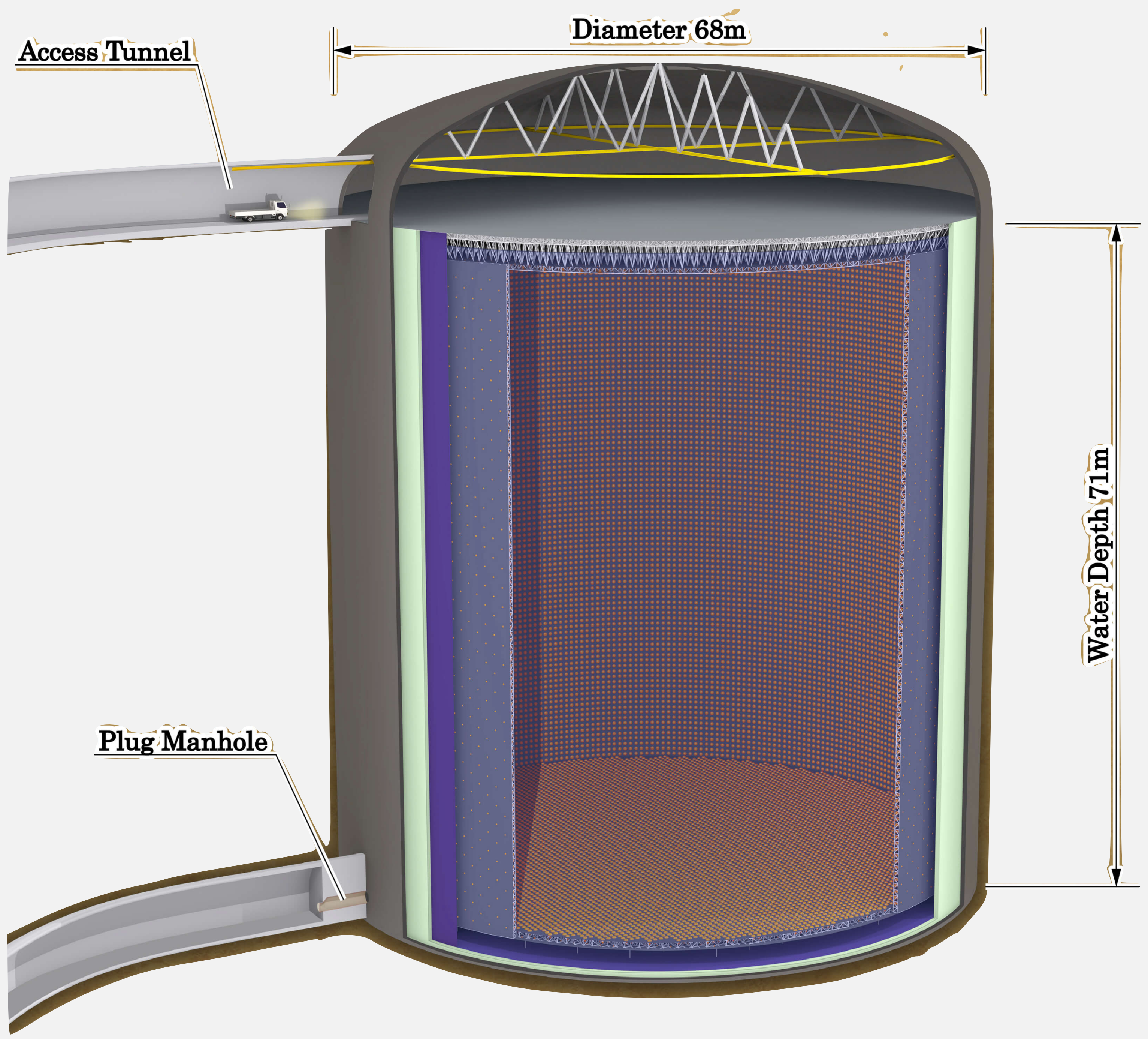
Spallation background overlayed onto the positron kinetic energy spectra for low energy antineutrino events.

PhysRevLett.93.171101

How are we going to detect it?

University of Sheffield | Jack E.P. Fannon





Super-Kamiokande

- ~27 years of data taking!
- Almost 4 years with Gd. Loaded to 0.03 % as of July 2022, $75\ \mu s$ average time for neutrons to capture.
- Has a 1000 m (2700 m.w.e) rock overburden.
- Observes muons at a rate of 2.5 Hz

Hyper-Kamiokande

- 8 x fiducial volume of Super-K
- Currently under construction and planned to begin data taking in 2027.
- Has a 650 m (1750 m.w.e) overburden.
- Will see muons at a rate pf ~45 Hz
- No plan to dope with Gd.

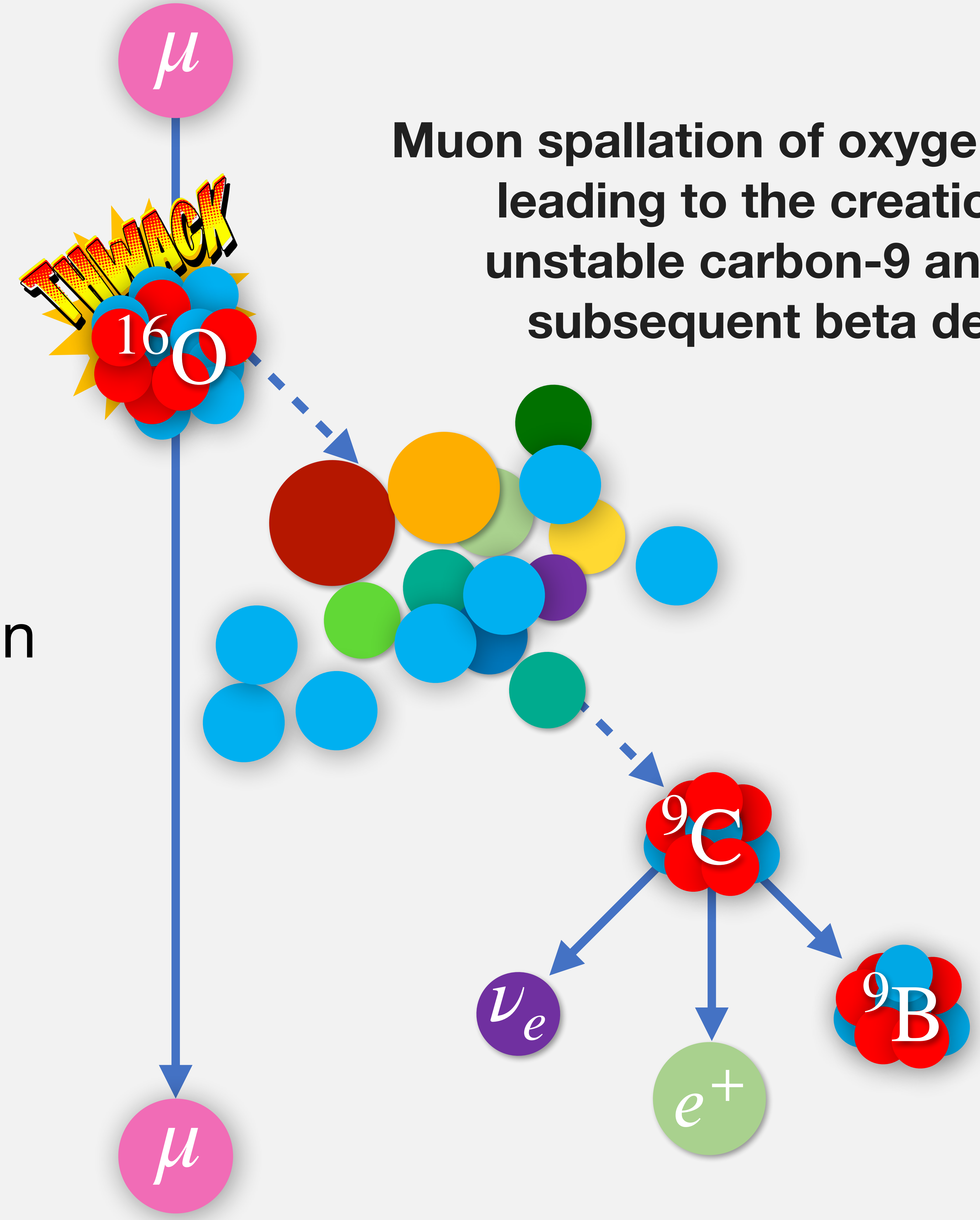
Spallation process

- Muons produce daughter particles that initiate **electromagnetic and hadronic showers through spallation processes**.

- **Hadronic showers** are the dominant production route for **unstable isotopes**.
  - 89% vs 11% for direct interactions between muon and oxygen nuclei.

- Decay products $e^{\pm}$, n, γ, are **produced at low energy** — $< 20\ \mathrm{MeV}$

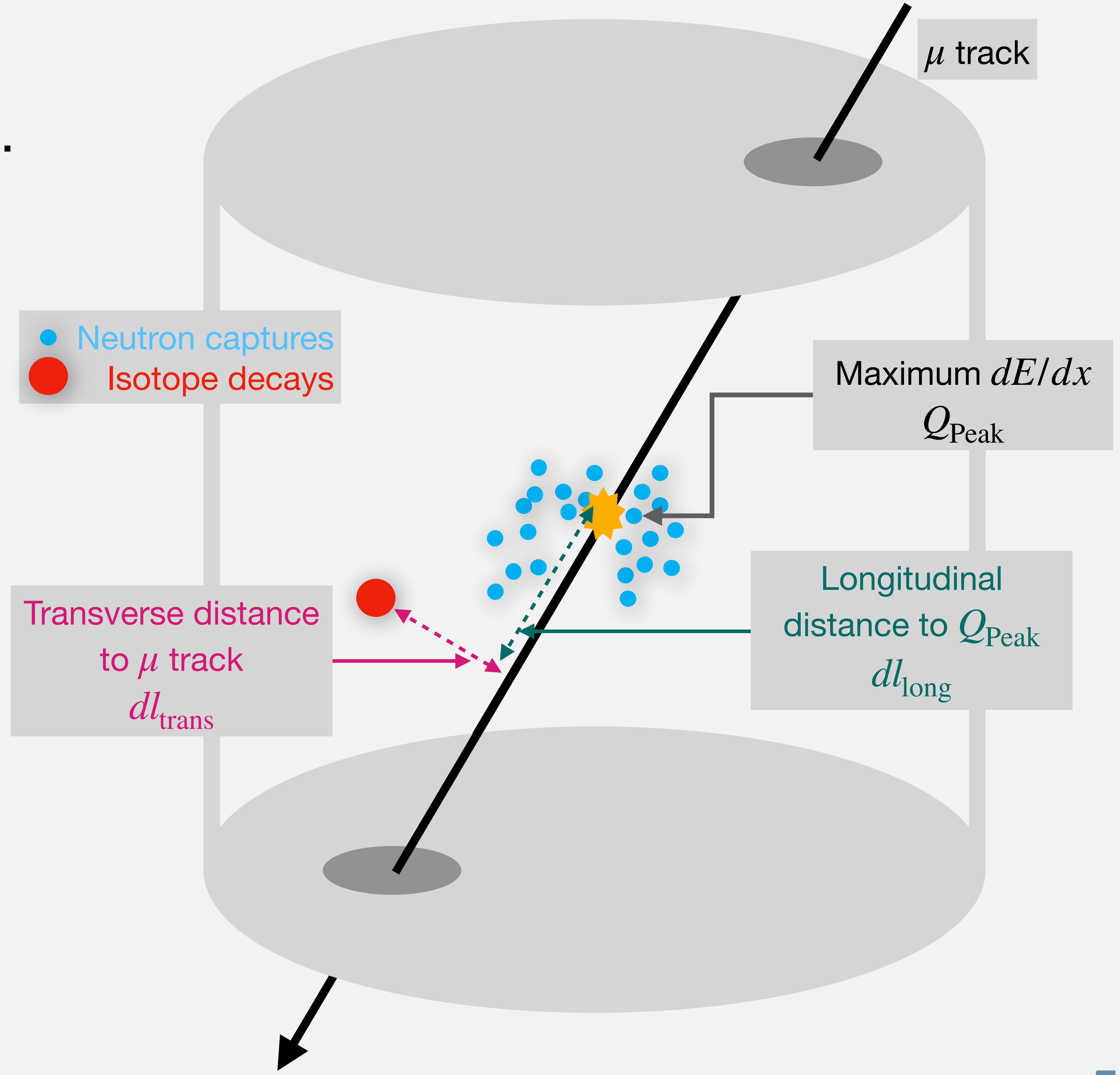
- Showers can be **located by neutron captures**…



Muon spallation of oxygen-16 leading to the creation of unstable carbon-9 and its subsequent beta decay.

The current SK reduction

- The current SK reduction is **based on data**.
  - Spallation **likelihood** calculated using 5 variables.
    - $dl_{\mathrm{trans}}, dl_{\mathrm{long}}, dt, Q_{\mu}, Q_{\mathrm{res}}$
- Spallation background **removal efficiency of ~90%**.
- Estimated **50-90% SRN signal acceptance**
- Introduces a dead-time for SRN events of ~10%.



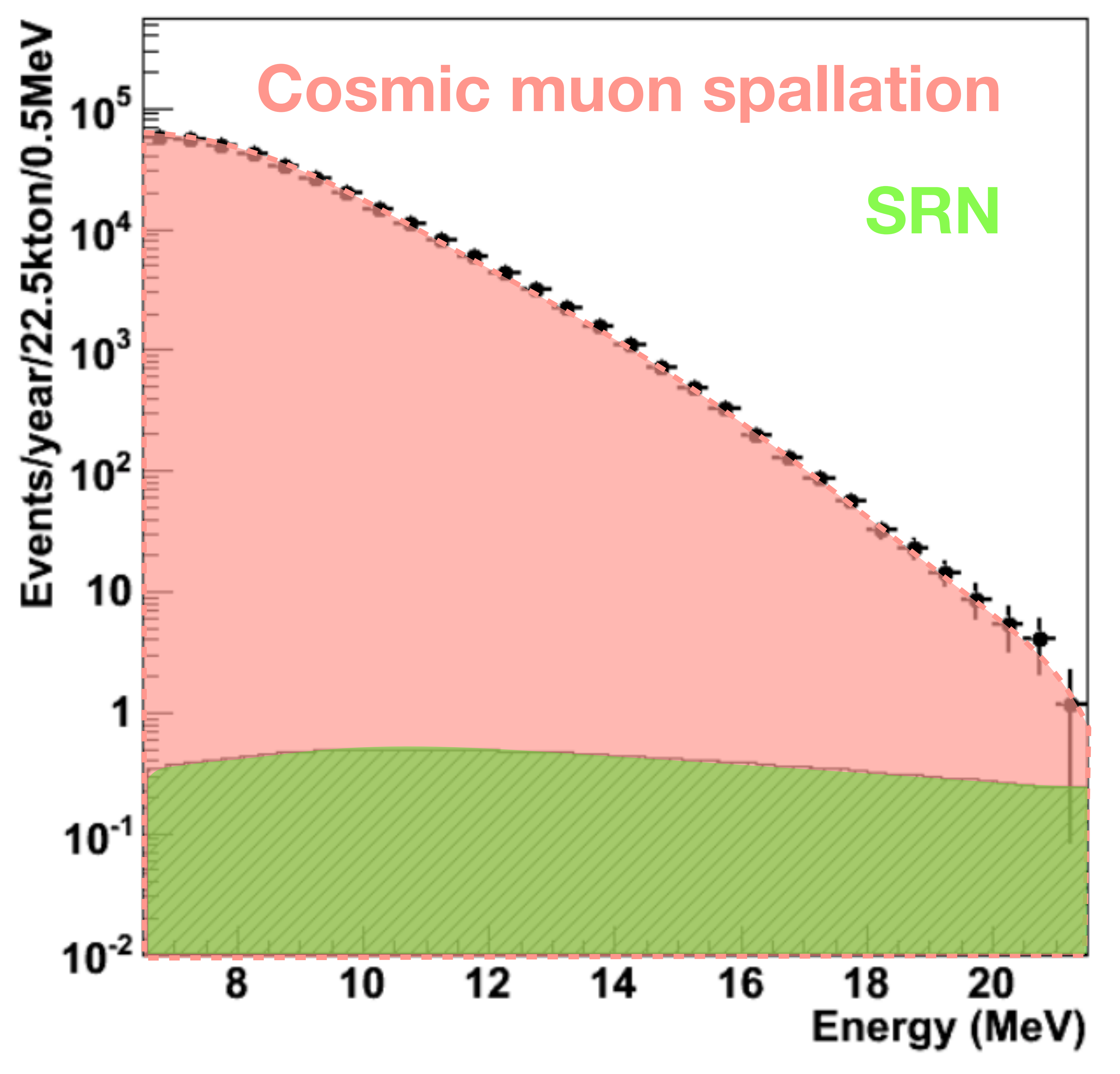
Why change anything?

- **Two previous studies** simulating cosmic ray muon spallation at SK have been done (Li and Beacom (2014) and A. Coffani (2021)).
    - A. Coffani study done up to 0.01% Gd (SK-VI), no SK-VII simulations.

- Spallation events remain a **major background to the DSNB search**.
    - Background is $10 - 10^5$ times larger than the SRN flux in the 8-20 MeV region of interest.

- Can the simulations instead be used to create the cut?
    - Use a **machine learning classification** in place of the likelihood cut.

- Data based cut may not be enough for Hyper-K!



Expected SRN flux and spallation background plotted against electron/positron energy.

ZHANG Yang PhD thesis 2015

The plan

University of Sheffield | Jack E.P. Fannon

Create a **simulation-based** cut using **machine learning** to classify spallation-caused low-E events against DSNB candidates.

Transport muons with MUSIC

Astropart.Phys.7:357-368,1997

MUSIC is an **MC muon propagation code** used to transport muons through Ikenoyama to SK.

Generates the energy and angle distributions of cosmic ray muons intersecting with the HK tank. **Used as the input for FLUKA**.

The plan

Create a **simulation-based** cut using **machine learning** to classify spallation-caused low-E events against DSNB candidates.

Transport muons with MUSIC → FLUKA

Astropart.Phys.7:357-368,1997

FLUKA — General purpose MC with a wide range of applications.

Used here to **simulate the muon interactions with water**: hadronic showers, spallation, radioactive decays etc. Difficult to do this in Geant4 with optical photon tracking enabled.

Custom FLUKA user-routines select spallation events and generate vectors for use with GHOST.

The plan

University of Sheffield | Jack E.P. Fannon

Create a **simulation-based** cut using **machine learning** to classify spallation-caused low-E events against DSNB candidates.

Transport muons with MUSIC → FLUKA → GHOST

Astropart.Phys.7:357-368,1997

FLUKA output is piped into WCSim / the **G**eant-4 **H20** **S**imulation **T**oolkit (**GHOST**) with a Hyper-K geometry.

Used here to **simulate the detector response** to decay particles and hadronic showers simulated by FLUKA.

Output from this point should **emulate data**.

The plan

Create a **simulation-based** cut using **machine learning** to classify spallation-caused low-E events against DSNB candidates.

Transport muons with MUSIC → FLUKA → GHOST → **Machine learning classification**

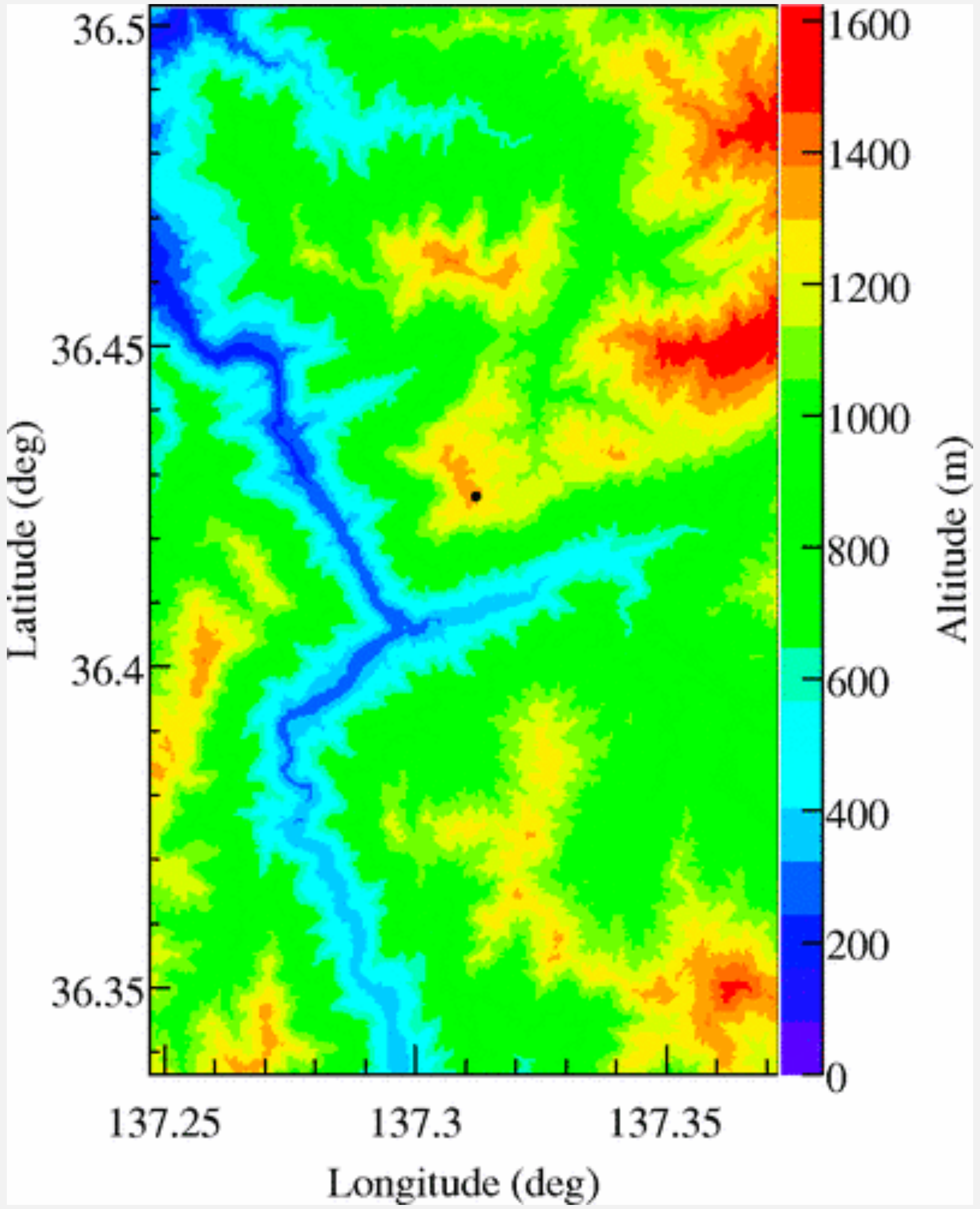
Astropart.Phys.7:357-368,1997

Reconstructed spallation-produced low-E events from GHOST then used alongside a sample of DSNB IBD events to train and test a machine learning classification.

Training **features are based on the variables currently used to calculate the data-based likelihoods** (distance to muon track etc.).

Muon transport with MUSIC

- MUSIC transports muons from surface of Mt. Ikeno and takes into account:
  - Topology
  - Bremsstrahlung
  - Pair production
  - Multiple and inelastic scattering
- Mean energy of muons reaching SK and HK — 258 GeV
- Energy range reaches several TeV.

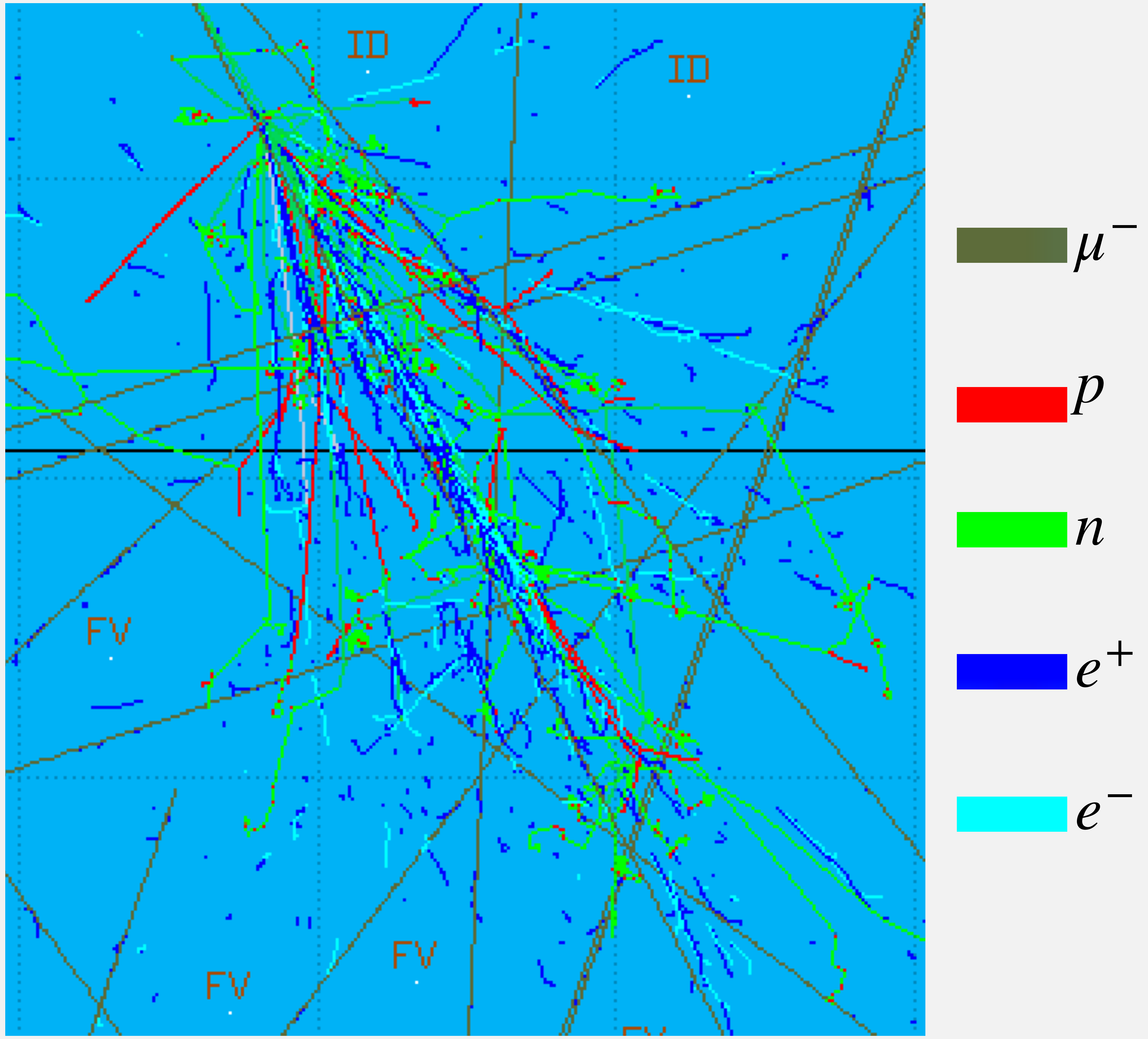


Ikenoyama topological profile
Phys.Rev.C81:025807,2010

FLUKA Spallation Simulation

University of Sheffield | Jack E.P. Fannon

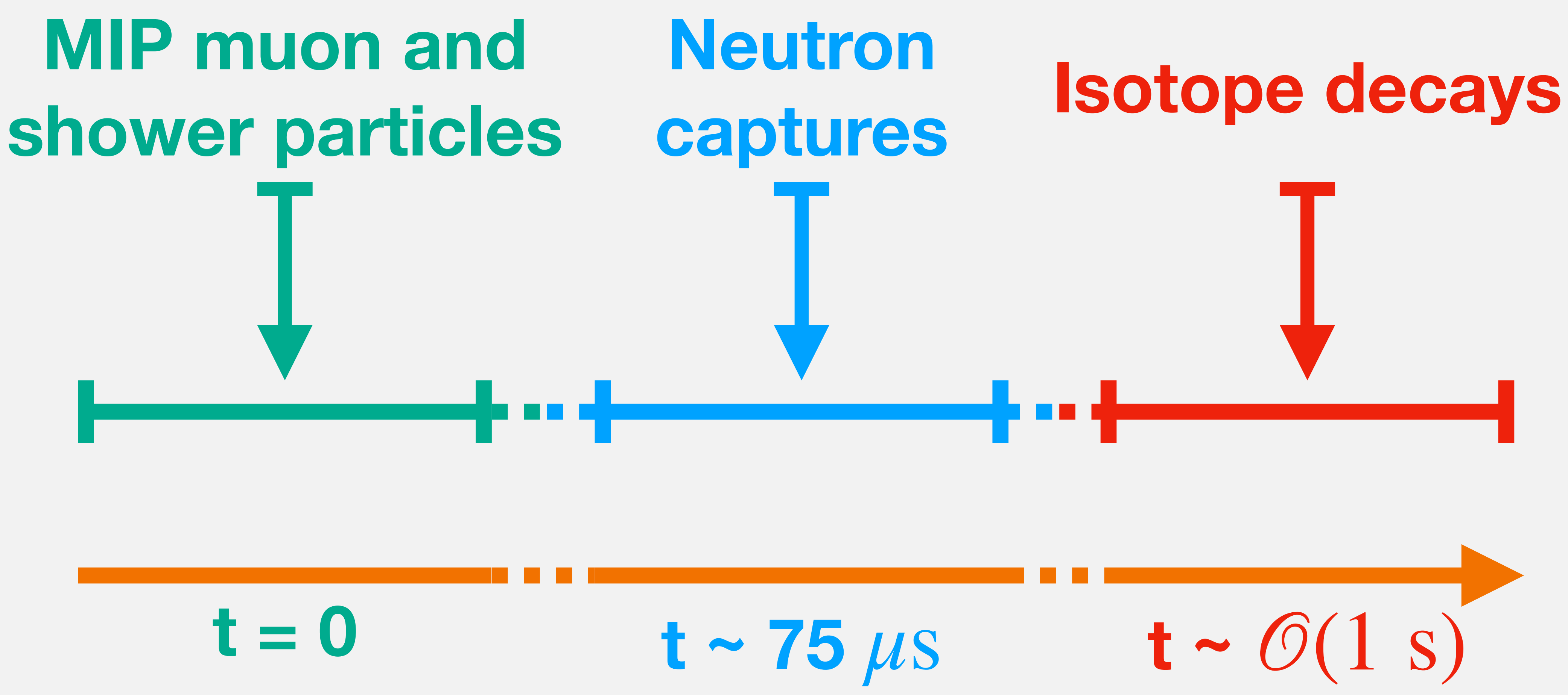
- Reproducing simulation by A. Coffani with increased Gd concentration.
- SK (HK) geometry within FLUKA:
  - 39.3 m (68 m) diameter
  - 41.4 m (71 m) tall
  - Acts as **pure-water or gadolinium-doped water target**
- Pipe transported muon primaries from MUSIC into FLUKA.
- Simulate muon interactions in Gd-water.
- Interface with GHOST



Magnified event display

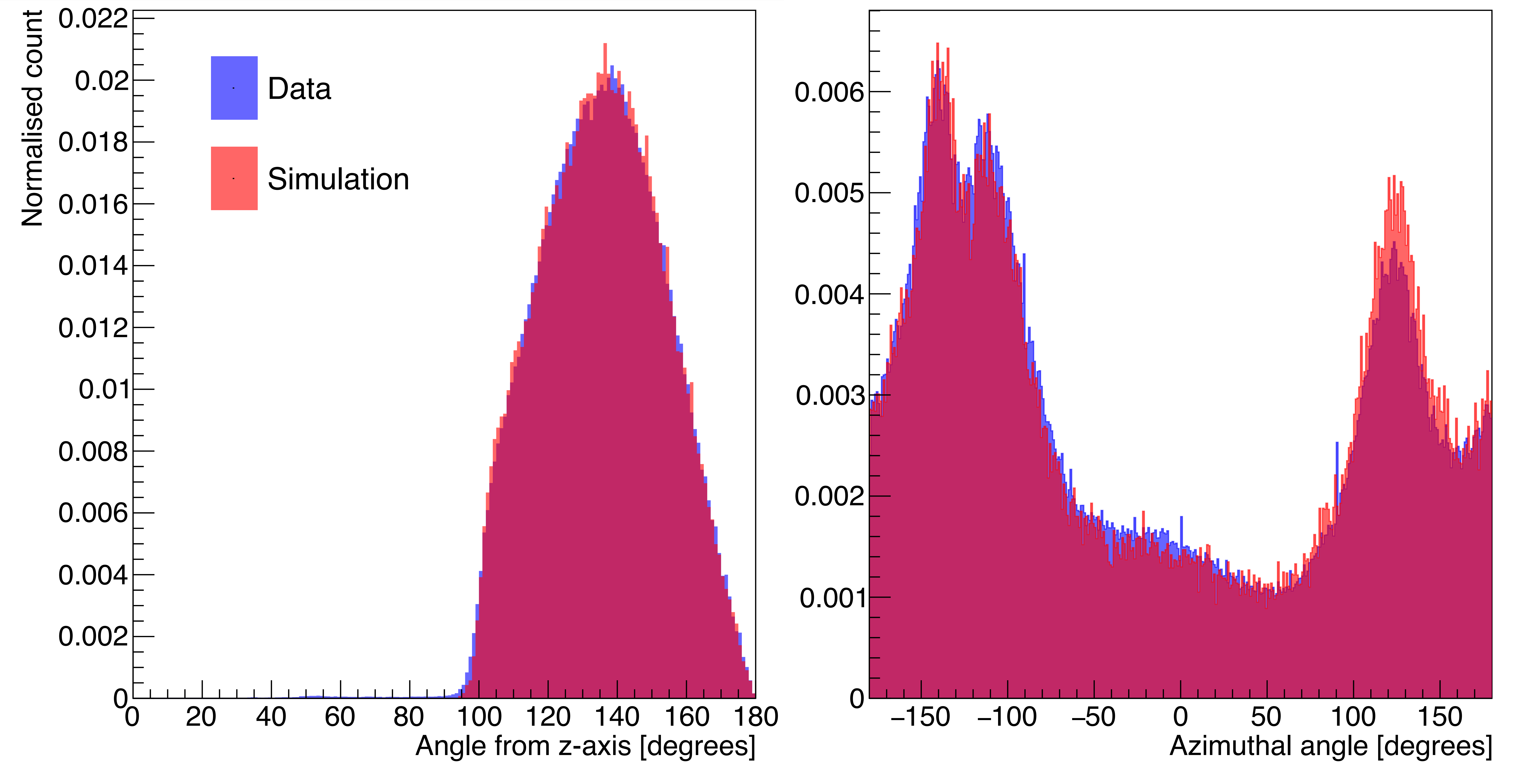
Interfacing with GHOST

- Need detector response to reconstruct location of shower in simulation.
- Use **FLUKA to simulate the hadronic processes, decays and muon-nuclear interactions**.
- **Disable all but EM physics** in GHOST to interface the two.
- Stitch together muon track, shower particles, neutron captures and decay products for a full spallation simulation.



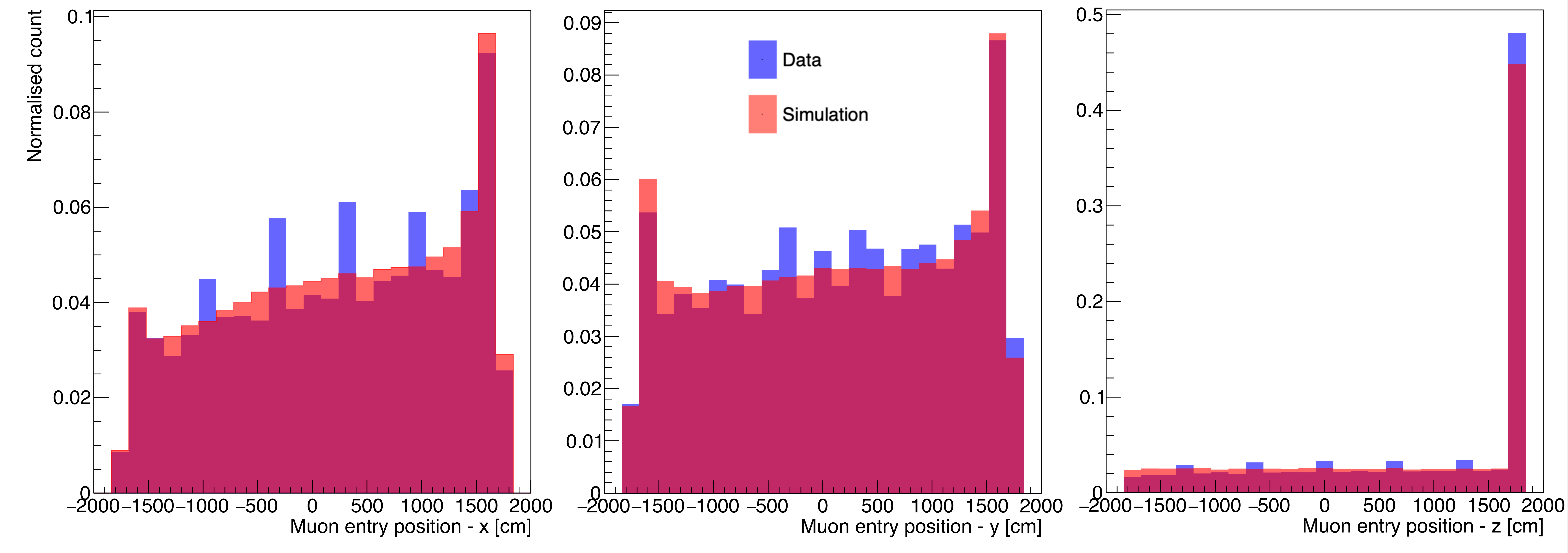
Data/MC comparison

University of Sheffield | Jack E.P. Fannon



- 10^6 simulated muons compared with ~46×10^6 reconstructed muon candidates from data

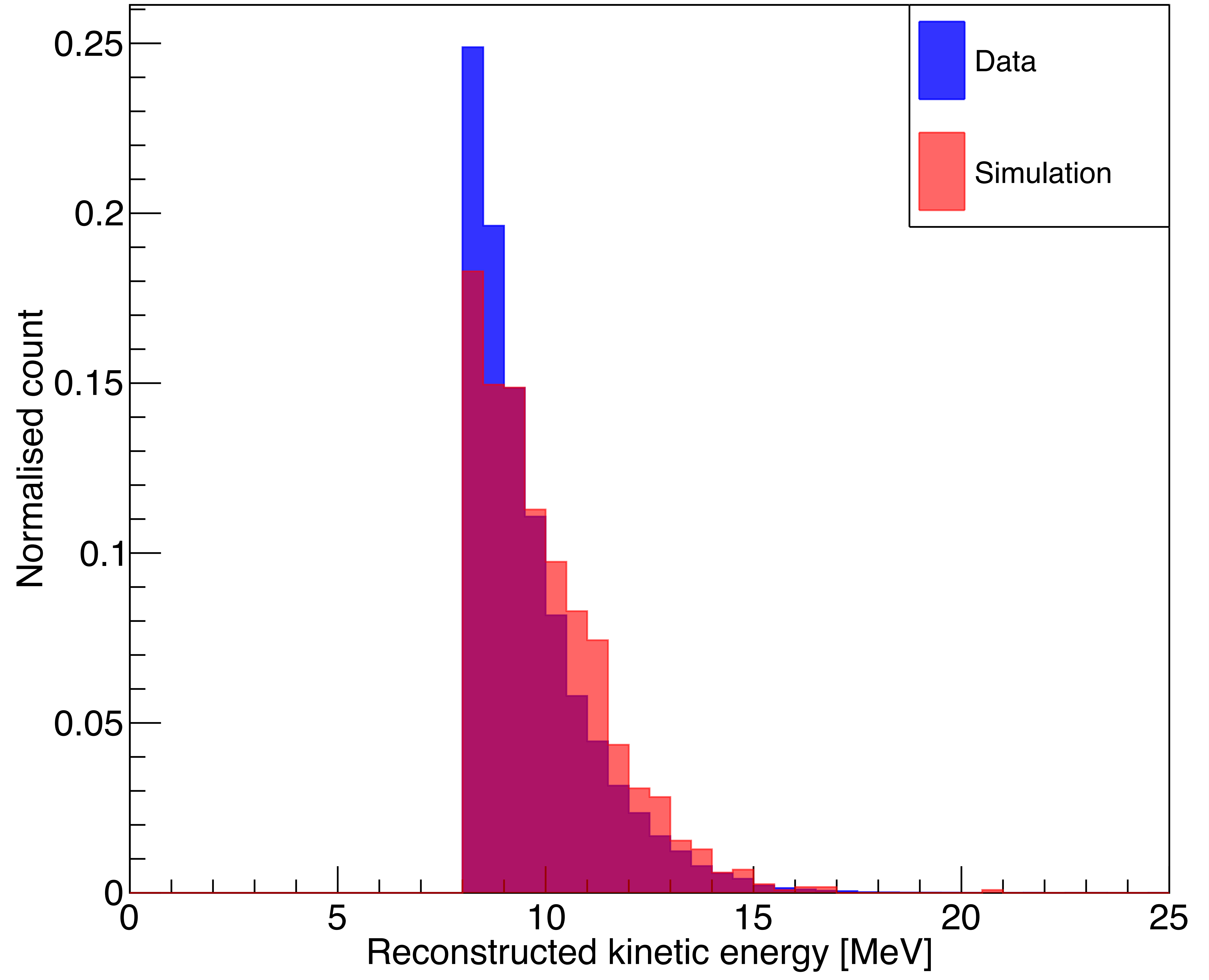
Data/MC comparison



- Good agreement also seen in entry positions.
- Peaks in data are expected to be caused by grid-based reconstruction — PMT snapping.

Data/MC comparison

- **Geant 4 simulated decay betas** where primaries were produced using FLUKA sims.
- Compare distributions once normalised to integral.
- **Excess in data expected** as simulation is only products from SK background decays and nothing else.
- Not comparing in < 8 MeV phase space due to **reactor neutrino cut**.

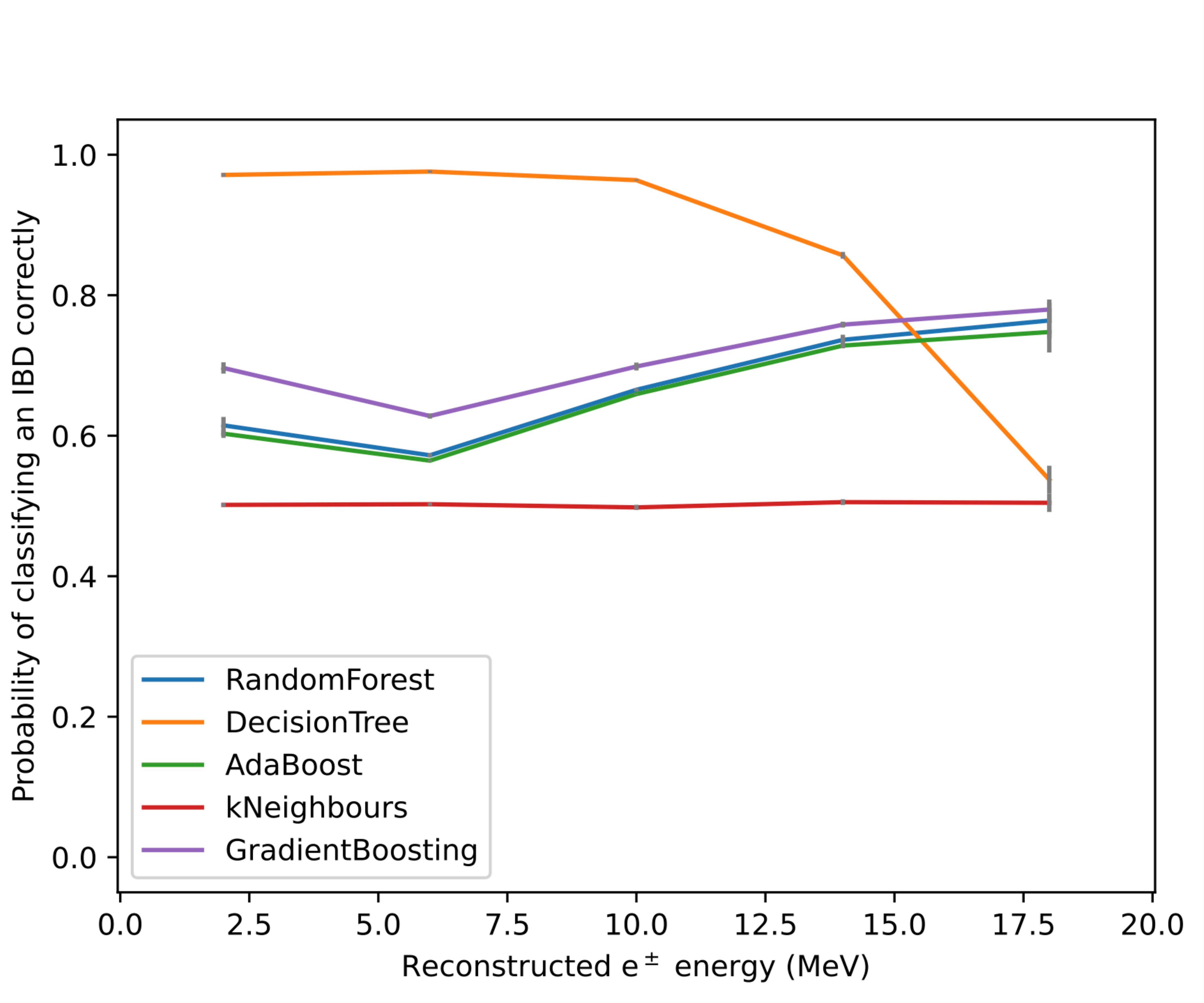


The classifiers

- Setup a pipeline for **training and testing the performance of five difference classifiers**:
  - BDT, random forest, AdaBoost, kNeighbours, and gradient boosted.

- Trained on a preliminary set of simulations created using results from A. Coffani's previous work.

- BDT shows good classification power: **above 90% probability of correctly identifying spallation and SRN events** in the < 15 MeV range.



Probability of correctly identifying an IBD shown against reconstructed positron/electron energy for five different classifiers.

Summary and next steps

University of Sheffield | Jack E.P. Fannon

☑ **FLUKA spallation simulation is setup:**

Show good agreement with data for muon entry point, muon direction and energy distribution of decay products.

☑ **Machine learning training/testing pipeline is in place:**

Discrimination power of five classifiers tested on a preliminary set of simulations.

BDT currently performs the best: >90% probability of classifying correctly

☐ **Simulate other backgrounds**

A more robust analysis requires other backgrounds to the DSNB search to be simulated: atmospheric & solar neutrinos.

Backups

Isotope table

University of Sheffield | Jack E.P. Fannon

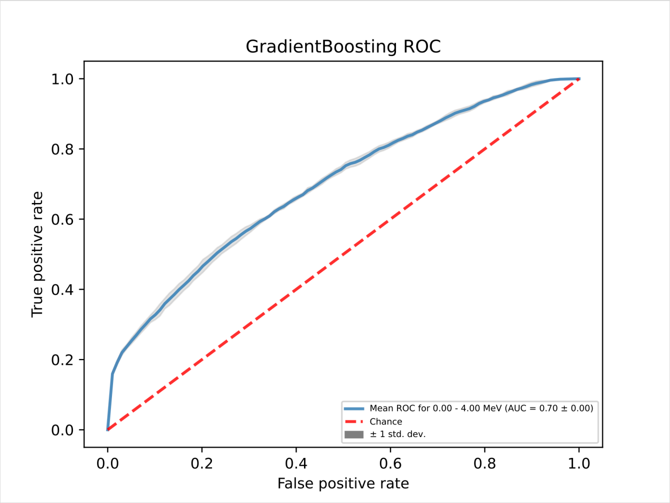
Isotope	Half-life[s]	E_{kin}[MeV]	Decay mode	Yields$\times 10^{-3}$ (μ^{-1}g^{-1}cm^3)	Yields[53]$\times 10^{-3}$ (μ^{-1}g^{-1}cm^3)
^{18}N	0.624		β^-	0.007 ± 0.004	0.006
^{17}N	4.173		$\beta^- n$	0.14 ± 0.02	0.19
^{16}N	7.13	4.27+6.13(γ)	$\beta^-\gamma$ (66%)	5.8 ± 0.1	5.8
		10.44	β^- (28%)		
^{16}C	0.747	~ 4	$\beta^- n$	0.01 ± 0.006	0.006
^{15}C	2.449	4.51+5.3(γ)	$\beta^-\gamma$ (63%)	0.27 ± 0.002	0.26
		9.77	β^- (37%)		
^{14}B	0.0138	14.55+6.09(γ)	$\beta^-\gamma$	0.01 ± 0.005	0.006
^{13}O	0.0086	$8 \sim 14$	β^+	0.06 ± 0.01	0.08
^{13}B	0.0174	13.44	β^-	0.59 ± 0.03	0.61
^{12}N	0.0110	16.38	β^+	0.36 ± 0.03	0.41
^{12}B	0.0202	13.37	β^-	3.8 ± 0.09	3.9
^{12}Be	0.0236	11.71	β^-	0.03 ± 0.008	0.03
^{11}Be	13.8	11.51)	β^- (55%)	0.21 ± 0.02	0.26
		9.41+2.1(γ)	$\beta^-\gamma$(31%)		
^{11}Li	0.0085	20.62	$\beta^- n$	0.007 ± 0.004	0.003
^{9}C	0.127	$3 \sim 15$	β^+	0.26 ± 0.02	0.29
^{9}Li	0.178	~ 10	$\beta^- n$ (51%)	0.59 ± 0.03	0.60
		13.6	β^-(49%)		
^{8}B	0.77	13.9	β^+	1.8 ± 0.06	1.9
^{8}Li	0.838	13.0	β^-	4.1 ± 0.09	4.2
^{8}He	0.119	9.67+0.98(γ)	$\beta^-\gamma$ (84%)	0.06 ± 0.01	0.07
		13.6	$\beta^- n$(16%)		

SK background isotopes taken from:

Alice Coffani. New studies on cosmogenic induced spallation background for Supernova relic neutrino search in the Super-Kamiokande experiment. Physics [physics]. Institut Polytechnique de Paris, 2021. English. ffNNT : 2021IPPAX112ff. fftel-03591741f

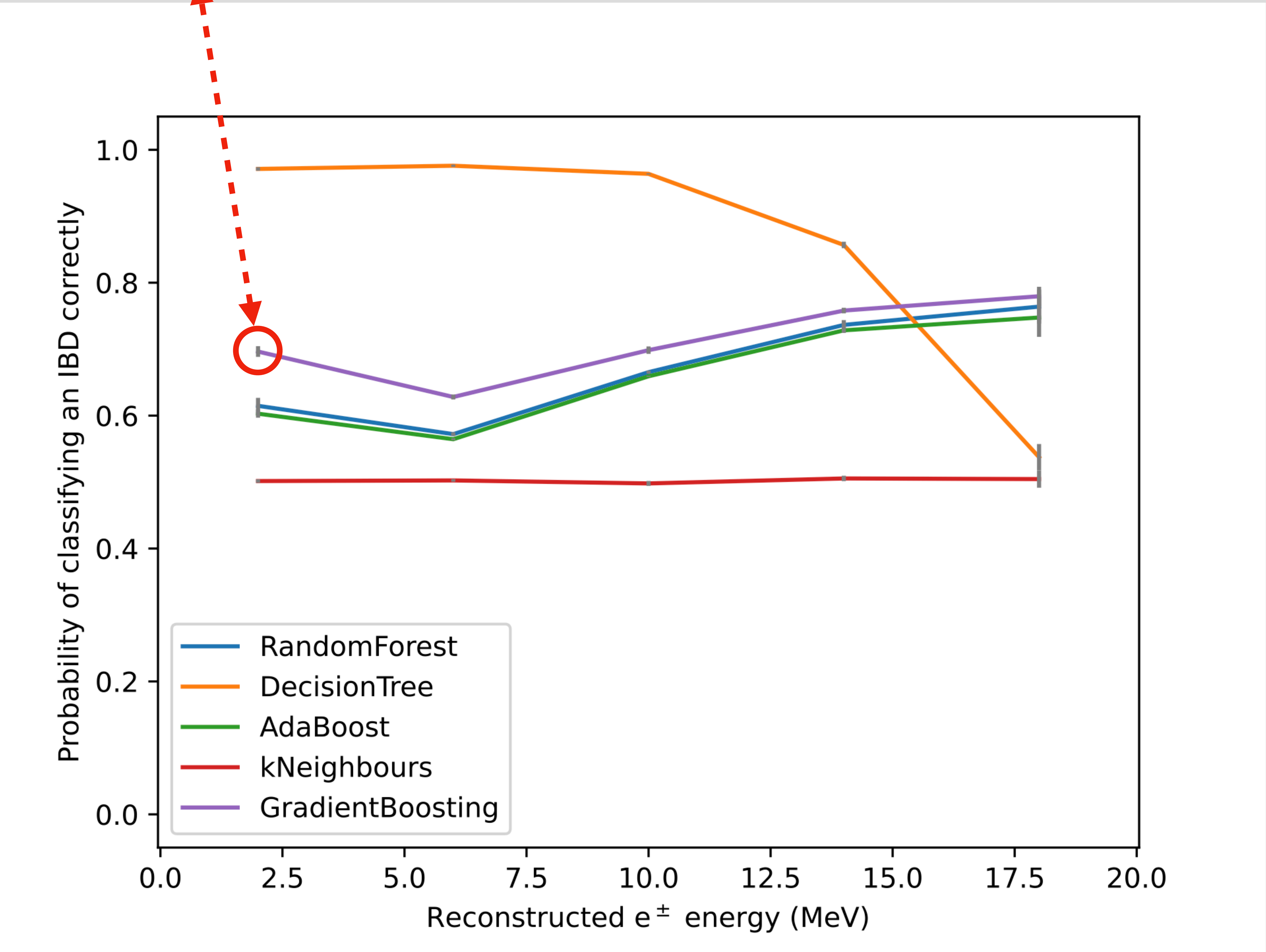
Classification performance - preliminary

University of Sheffield | Jack E.P. Fannon



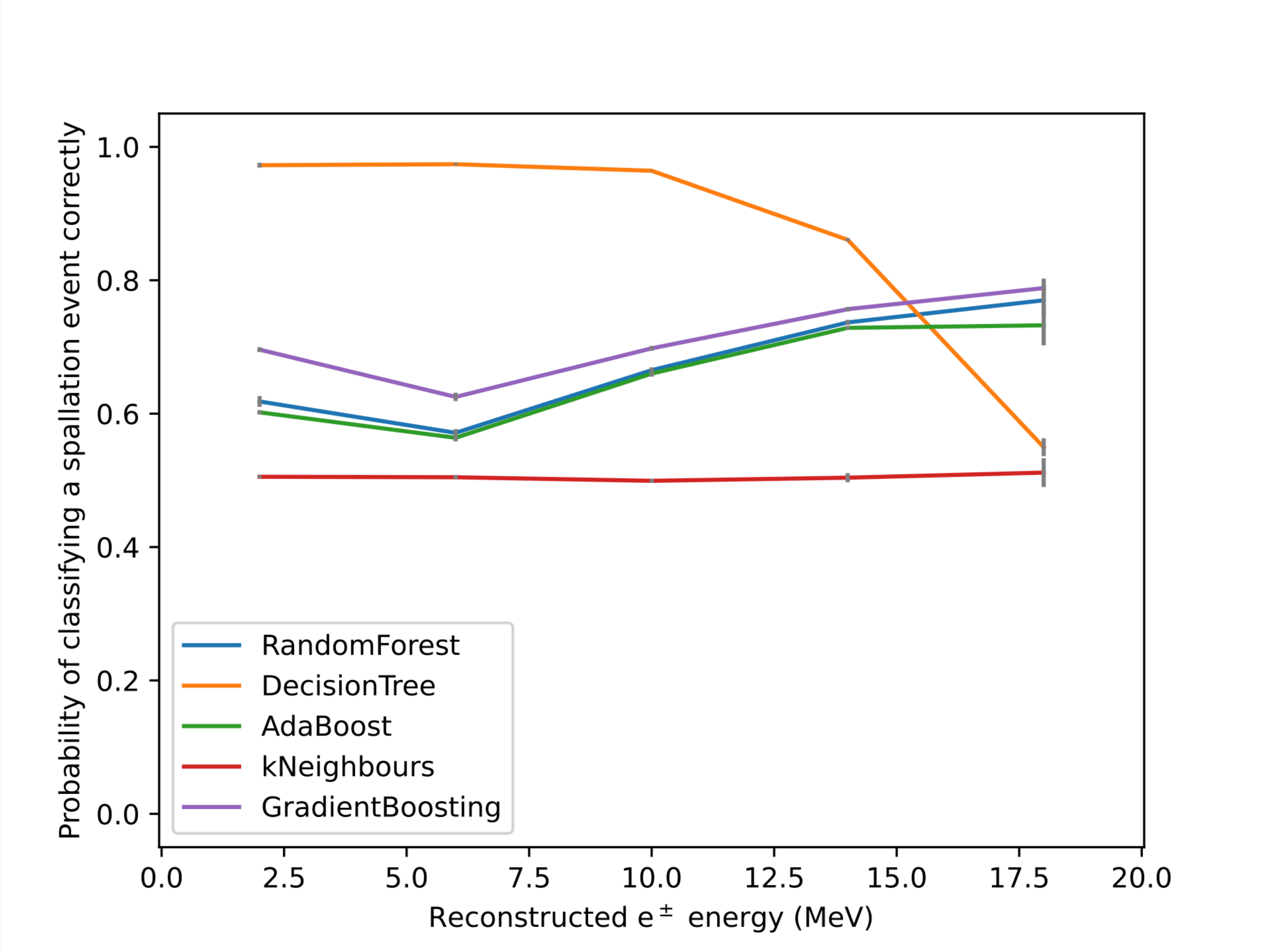
Area underneath ROC curve for five classifiers in five energy bins.
Bins are 4 MeV wide from 0 to 20 MeV.

IBD



Probability of correctly identifying an IBD shown against reconstructed positron/electron energy for five different classifiers.

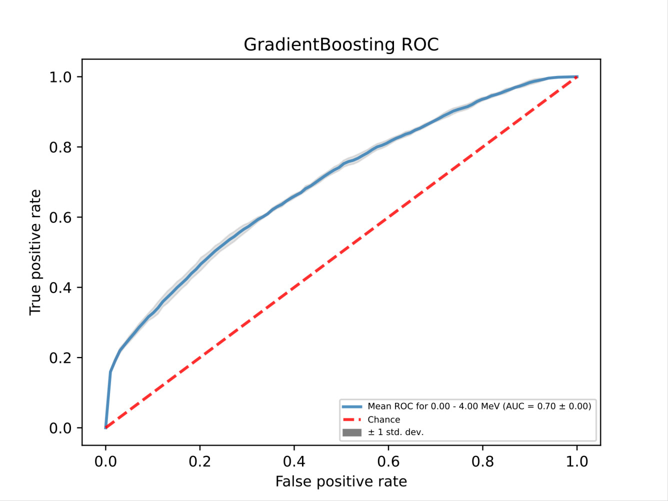
Spallation



Probability of correctly identifying an event caused by decay of a spallation product shown against reconstructed positron/electron energy for five different classifiers.

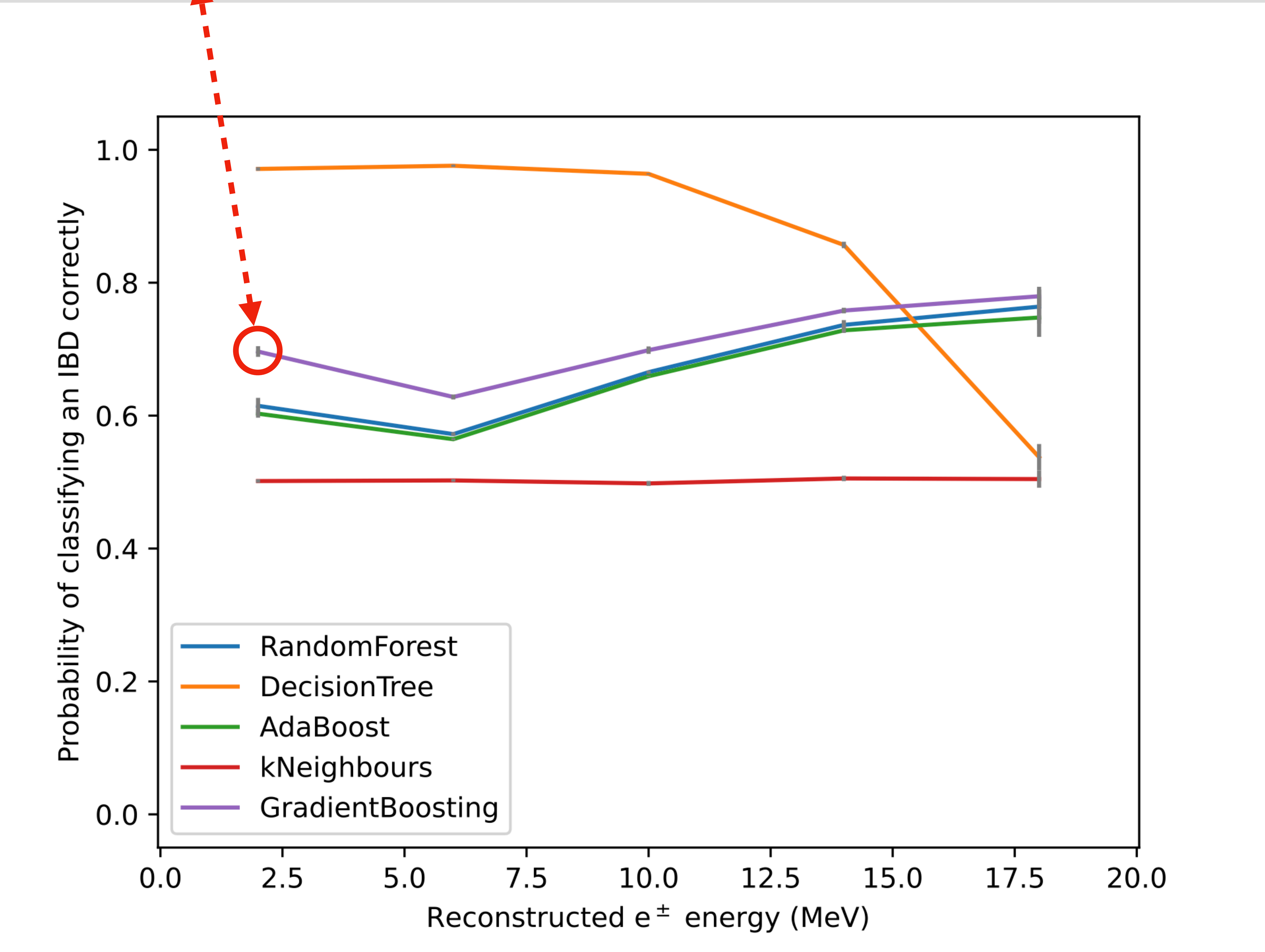
Classification performance - preliminary

University of Sheffield | Jack E.P. Fannon



Area underneath ROC curve for five classifiers in five energy bins.
Bins are 4 MeV wide from 0 to 20 MeV.

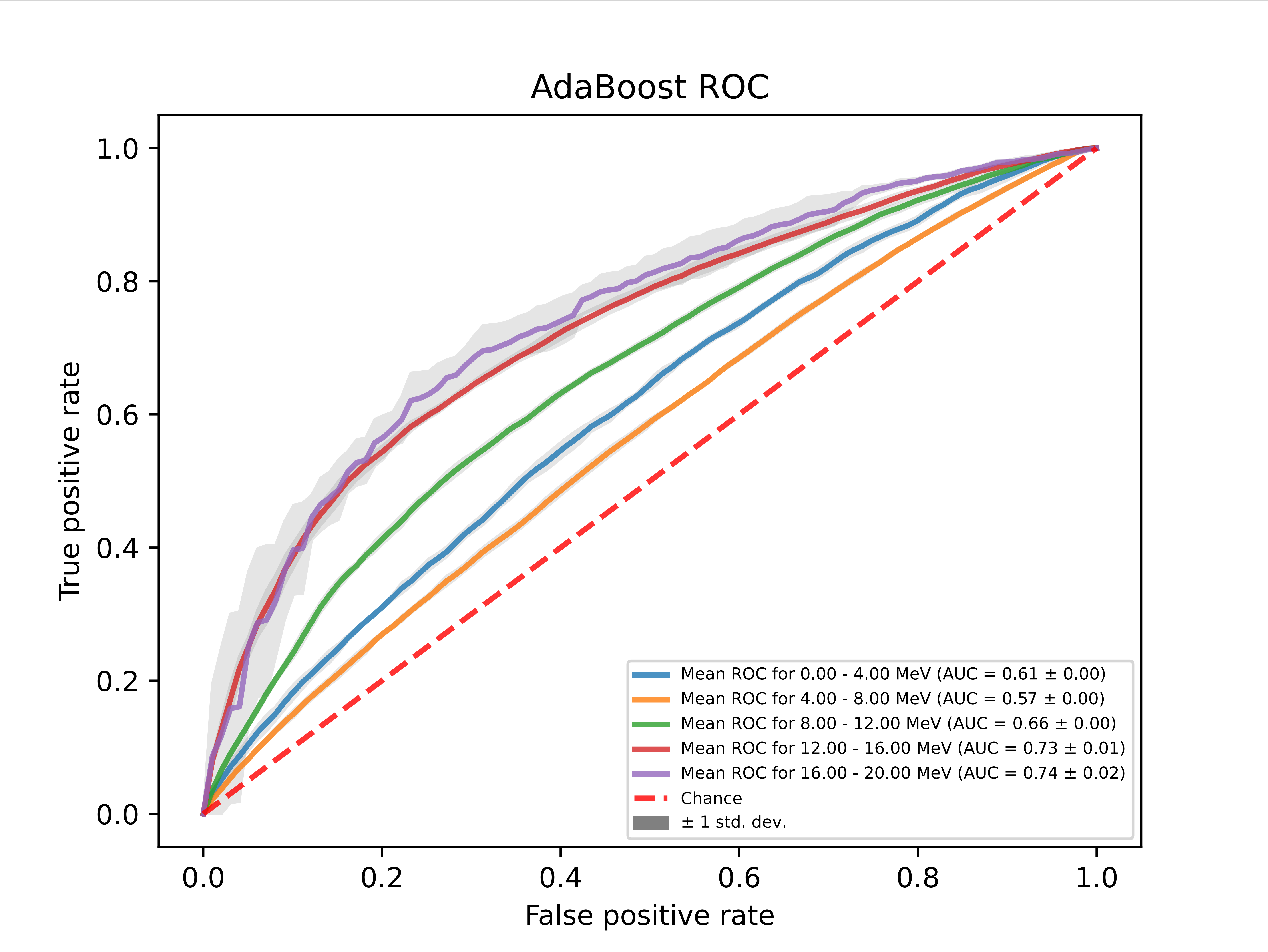
IBD



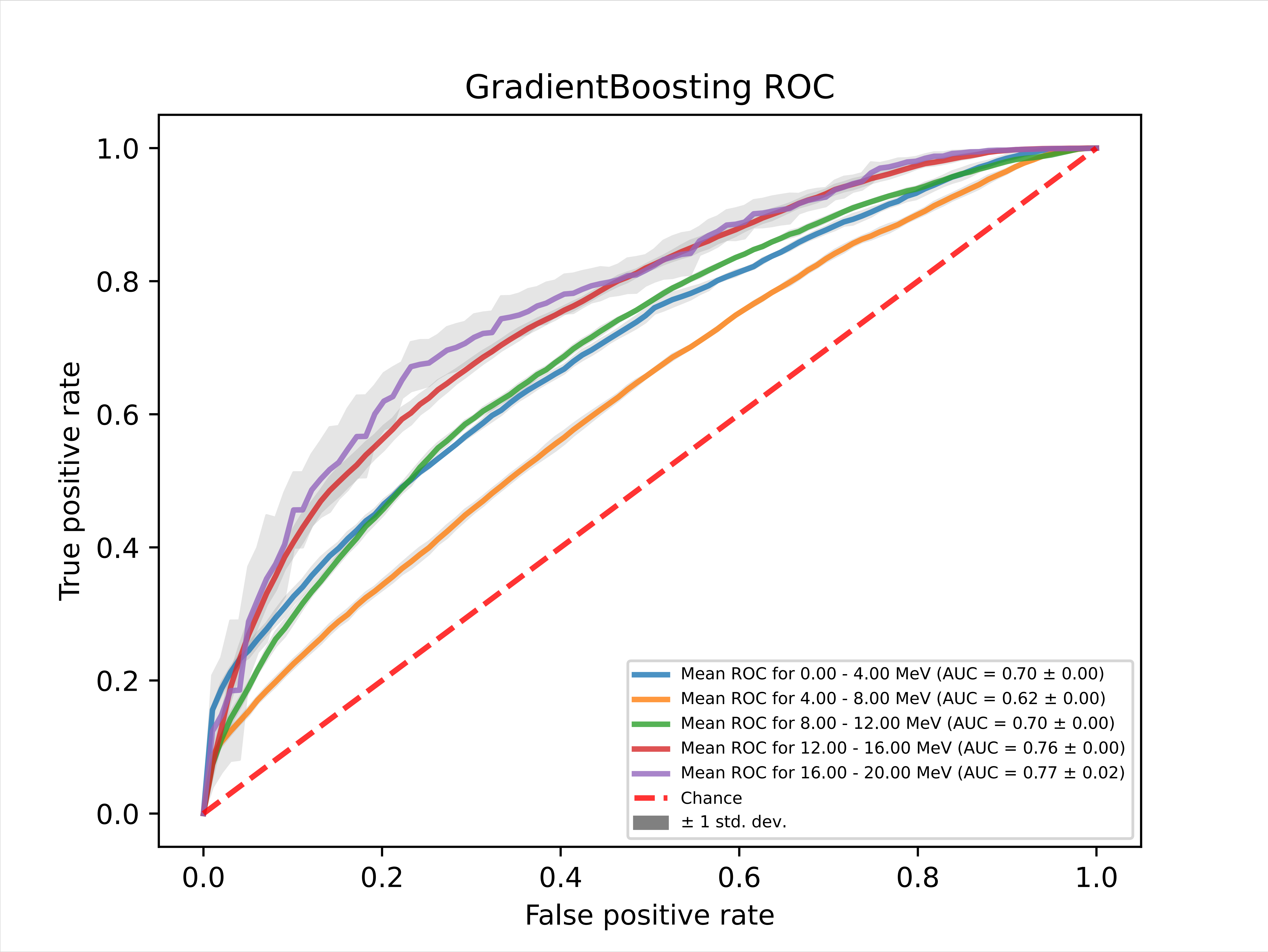
Probability of correctly identifying an IBD shown against reconstructed positron/electron energy for five different classifiers.

- Calculated by using a 5 split kFold validation and taking the area underneath the mean ROC curve.
- Boosted decision tree (yellow) shows the best classifying power.
- kNeighbours classifier is little better than random guessing (50% chance).
- AdaBoost, GradientBoosting and Random forest perform almost equally with a small dip in the 4-8 MeV bin.
- Reductions in classifying power could be related to classification of specific isotopes - requires further investigation.

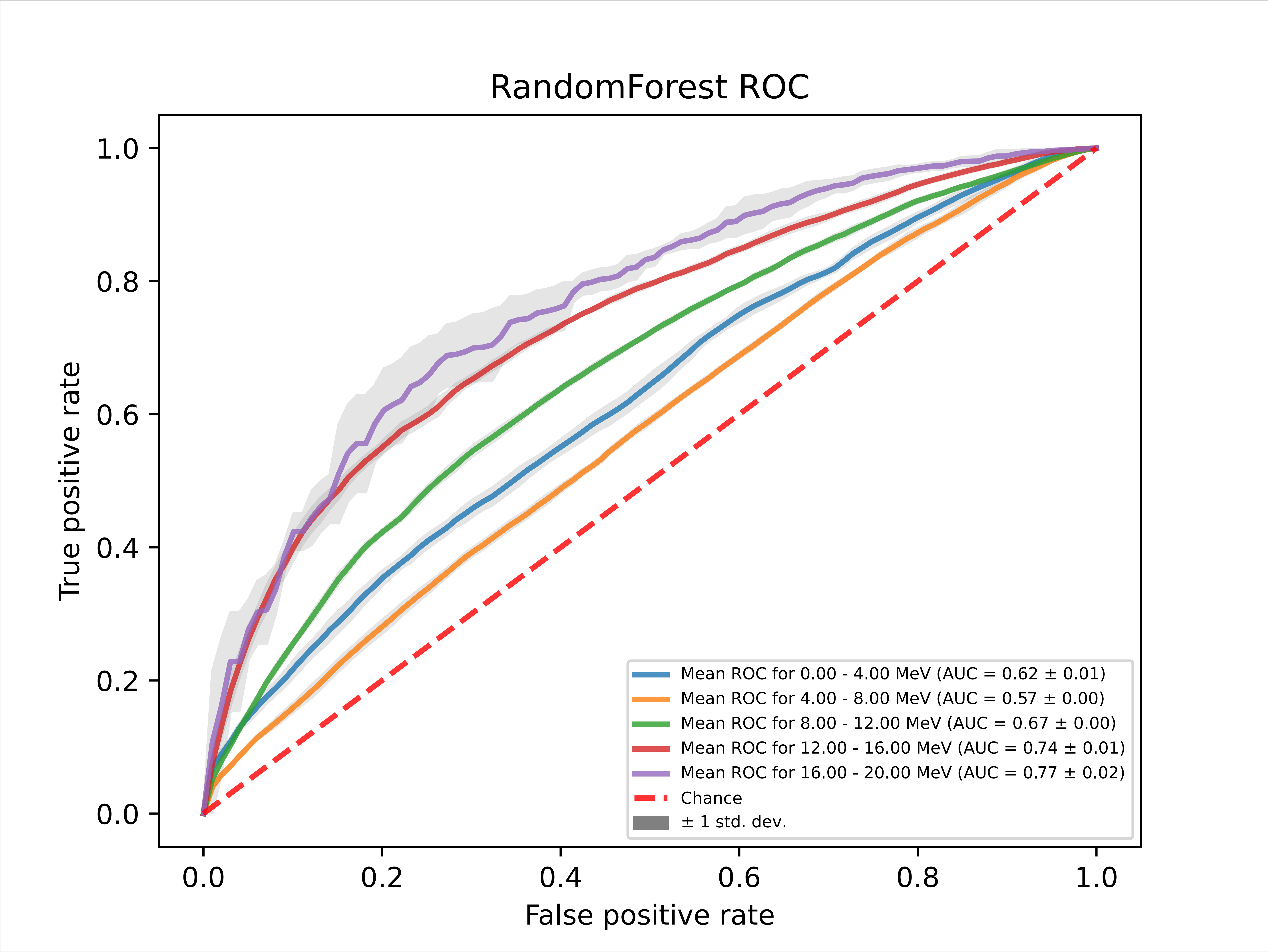
AdaBoost ROC



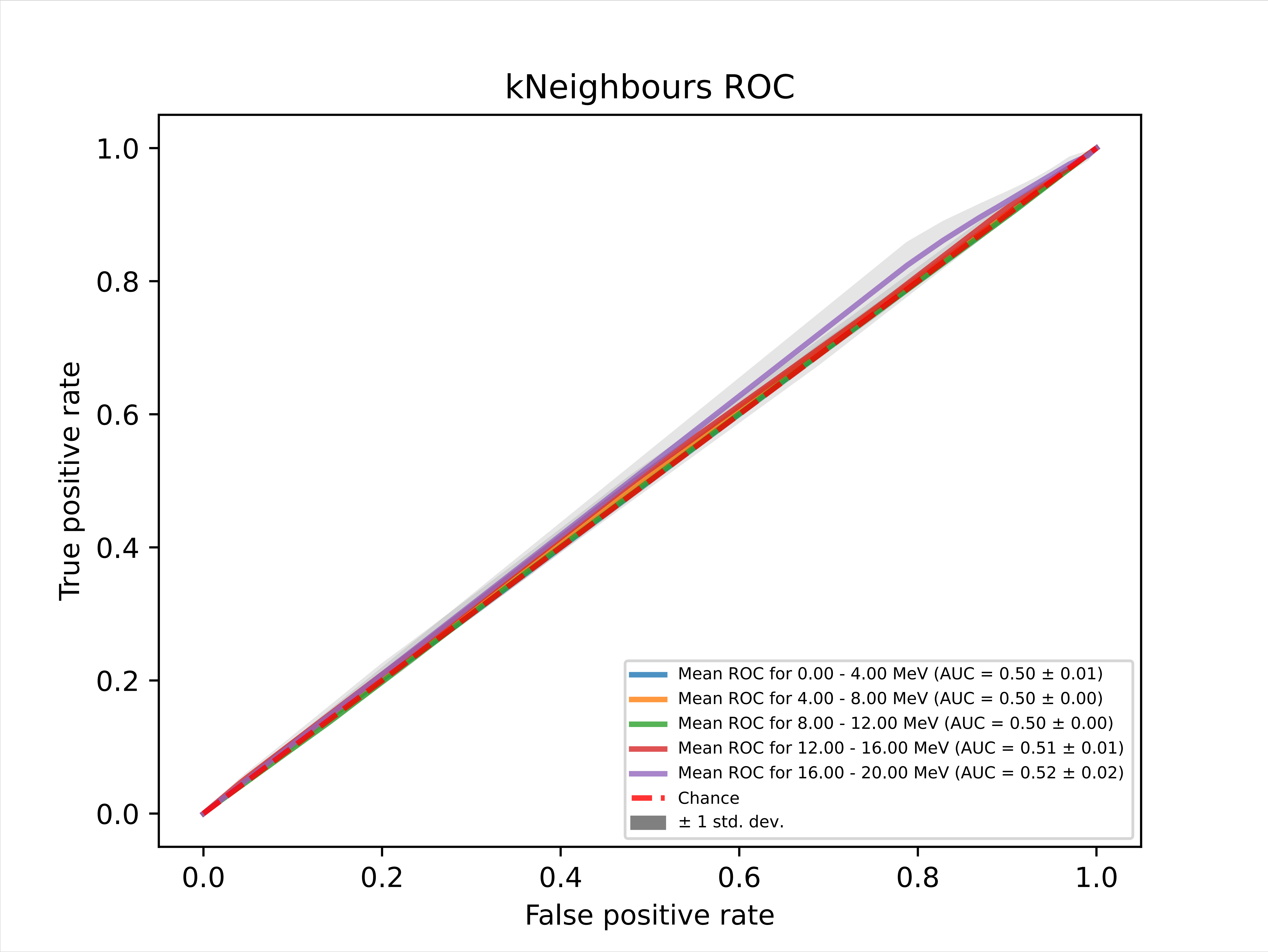
Gradient Boosting ROC



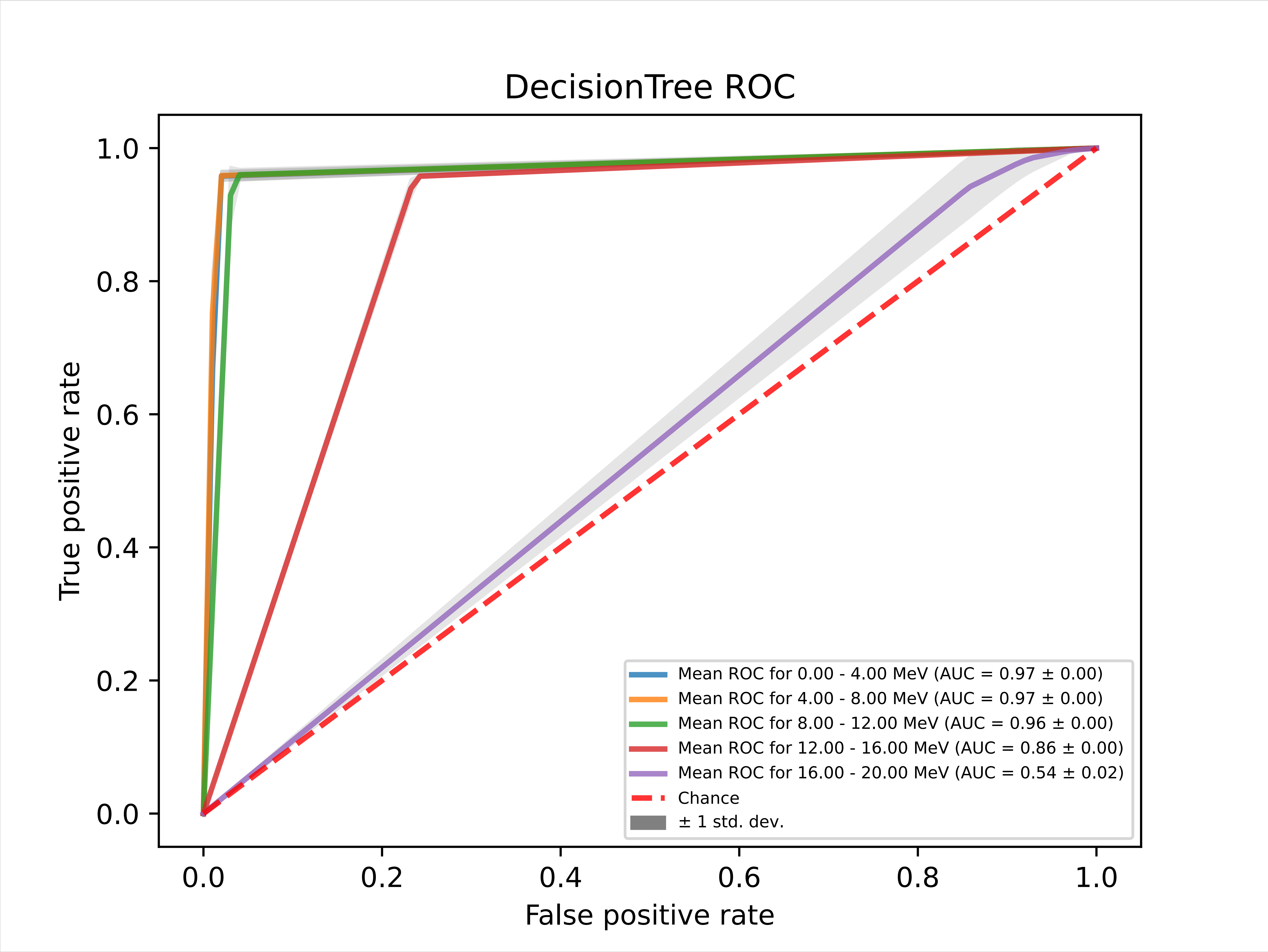
Random Forest ROC



kNeighbours ROC



BDT ROC


DecisionTree ROC
True positive rate
False positive rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Mean ROC for 0.00 - 4.00 MeV (AUC = 0.97 ± 0.00)
Mean ROC for 4.00 - 8.00 MeV (AUC = 0.97 ± 0.00)
Mean ROC for 8.00 - 12.00 MeV (AUC = 0.96 ± 0.00)
Mean ROC for 12.00 - 16.00 MeV (AUC = 0.86 ± 0.00)
Mean ROC for 16.00 - 20.00 MeV (AUC = 0.54 ± 0.02)
Chance
± 1 std. dev.