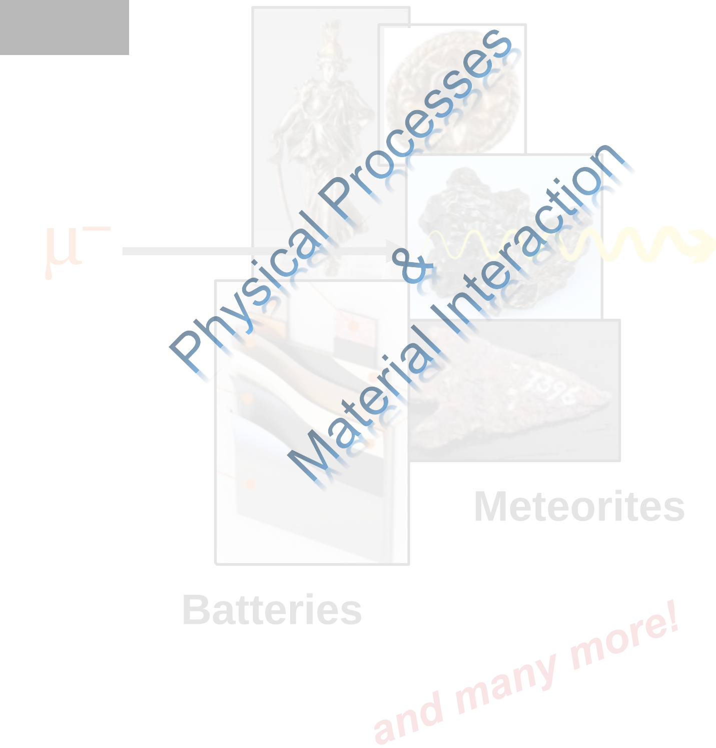
Michael Heiss :: Muon Induced X-ray Emission :: Paul Scherrer Institute

From Raw Data to Physics Results
and where Machine Learning enters the picture

SAMURAI Kick-Off Meeting, 6.9.2023

MIXE: Muon Induced X-ray Emission

Archeological artifacts

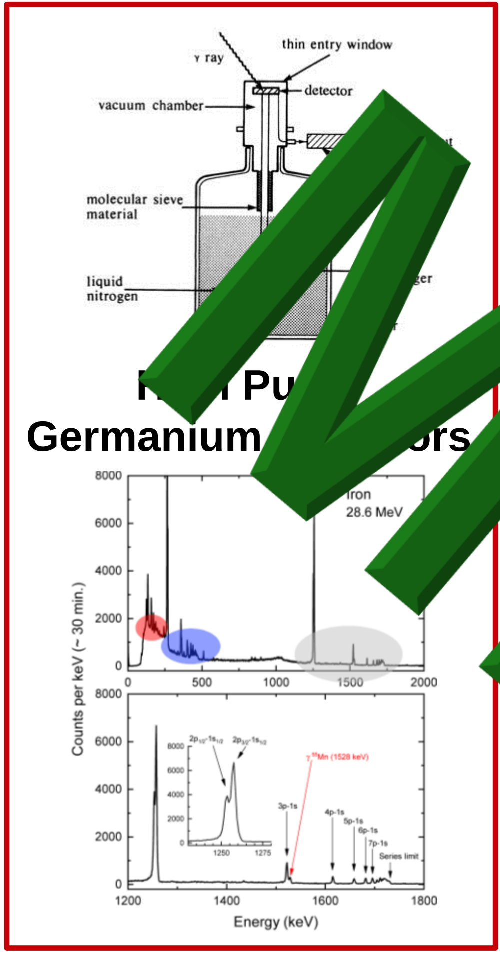


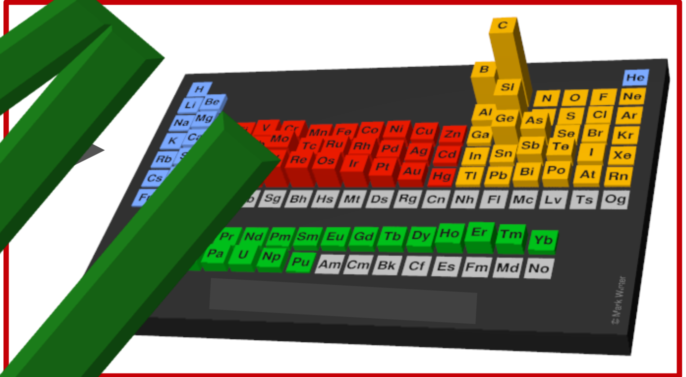
Meteorites

Batteries

and many more!

Raw Data

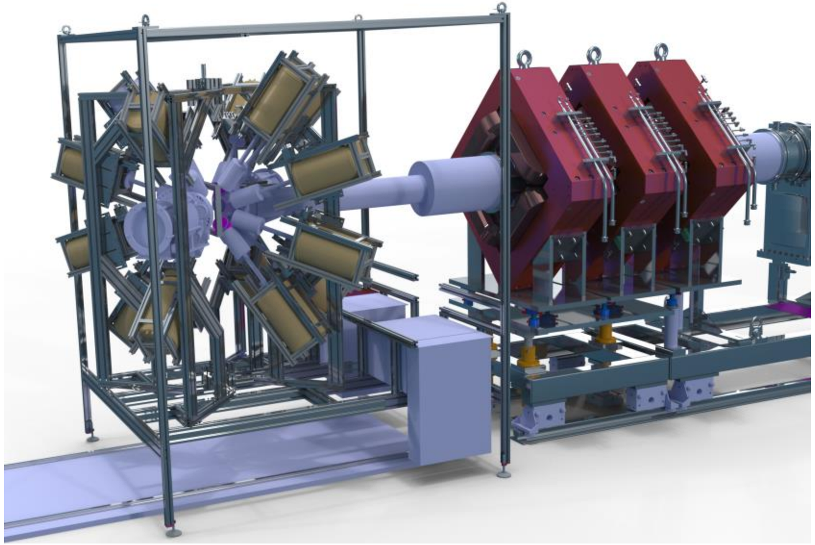




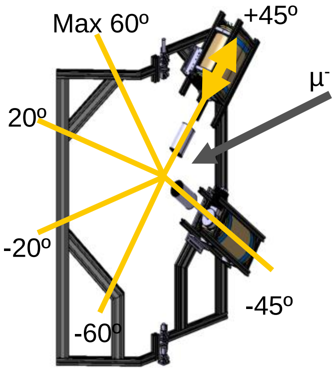
Physics Results

Raw Data

The GIANT Setup

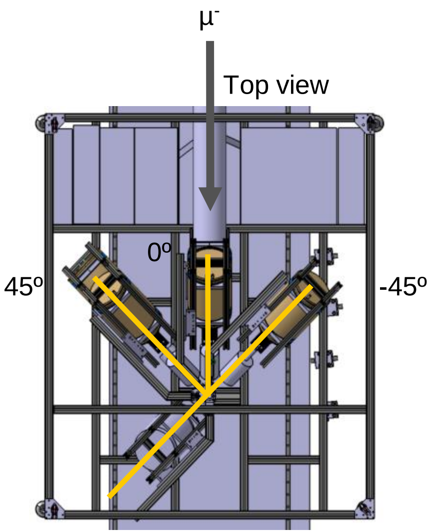


Max 4 HPGe per arm



Continuous -60º to +60º

Min Step 30º

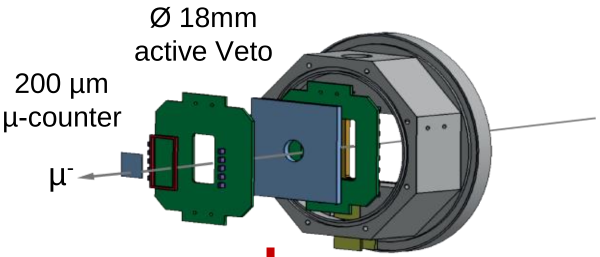


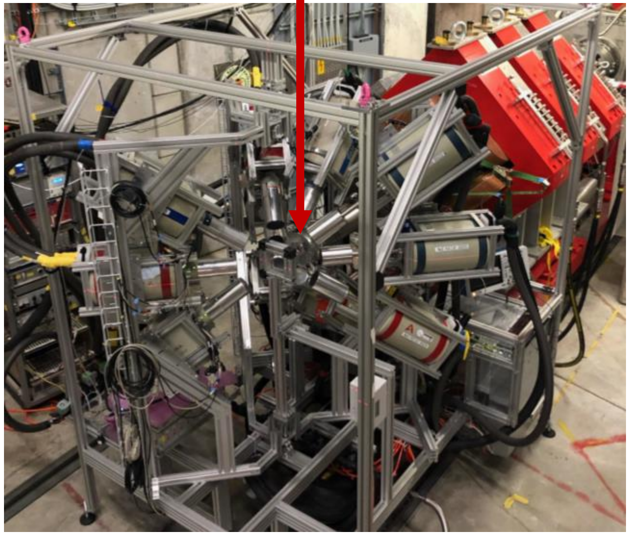
Continuous 0º - 340º
Min Step 35º

GIANT

German**I**um **A**rray for **N**on-destructive **T**esting

GIANT: Detector Setup





- Detectors
  - Scintillator with Silicon Photomultiplier readout
    - Muon Counter
    - Veto Detector

Start Time Only!

  - High Purity Germanium detectors
    - currently 11 detectors (up to 30 total)
    - various types (slightly different response)
    - shared with multiple other experiments
  - Silicon Drift Detectors
    - testing phase – currently 2 detectors
    - good candidates for low energy, low noise
    - similar signal structure

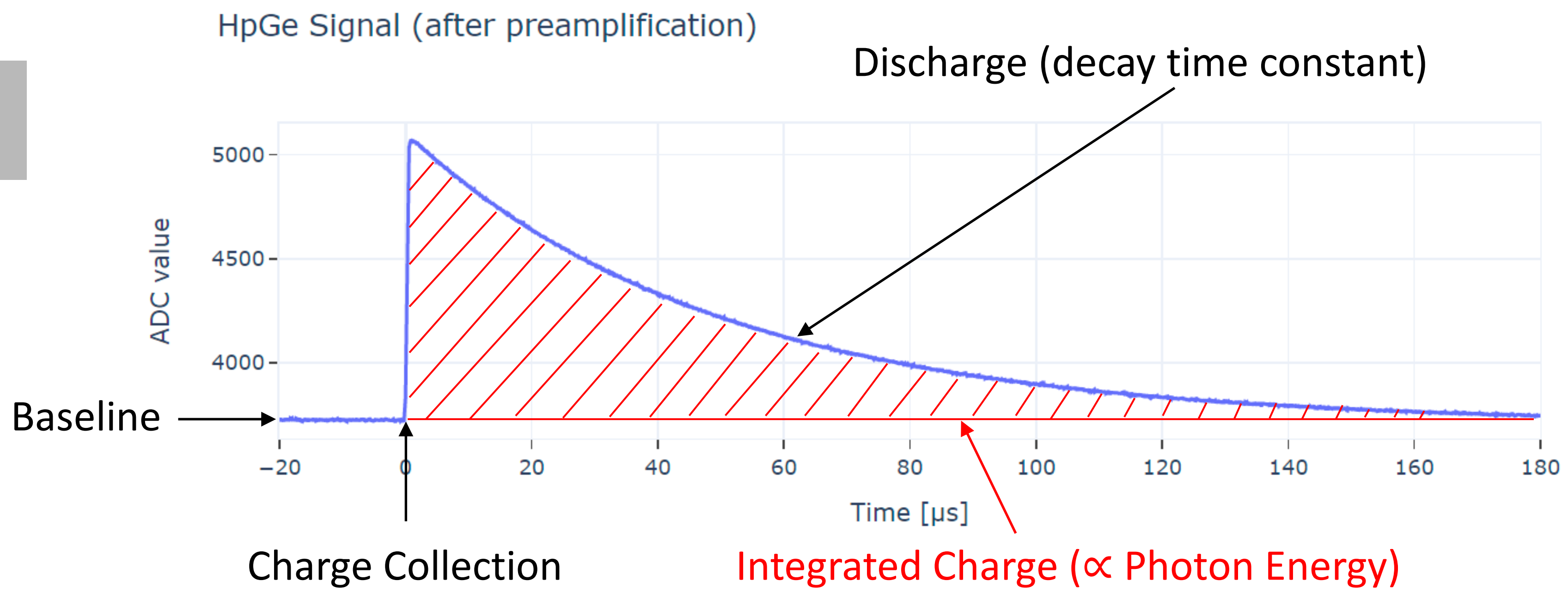
Future Addition to Project?

PAUL SCHERRER INSTITUT

GIANT: DAQ Setup

- MIDAS DAQ system (PSI development)
  - ingests and stores VME digitzer data
  - efficient binary format (easily convertible to more standard formats)
  - online preanalysis allows status monitoring
  - includes slow control (liquid nitrogen cooling, beam monitoring, etc.)
- Digitizer: Struck SIS3316-250-14 VME
  - 16 channels with 2V or 5V selectable dynamic range
  - 14-bit resolution with 250 MSps sampling rate
  - trapezoidal filter with decay time correction implemented on FPGA
  - fast optical readout
  - (partial) waveform readout for original and filtered waveforms
    - we usually store ~400 samples (~1.4µs per event)
    - allows for offline baseline correction
  - multiple devices chainable (external clock available)

High Purity Germanium detector signal

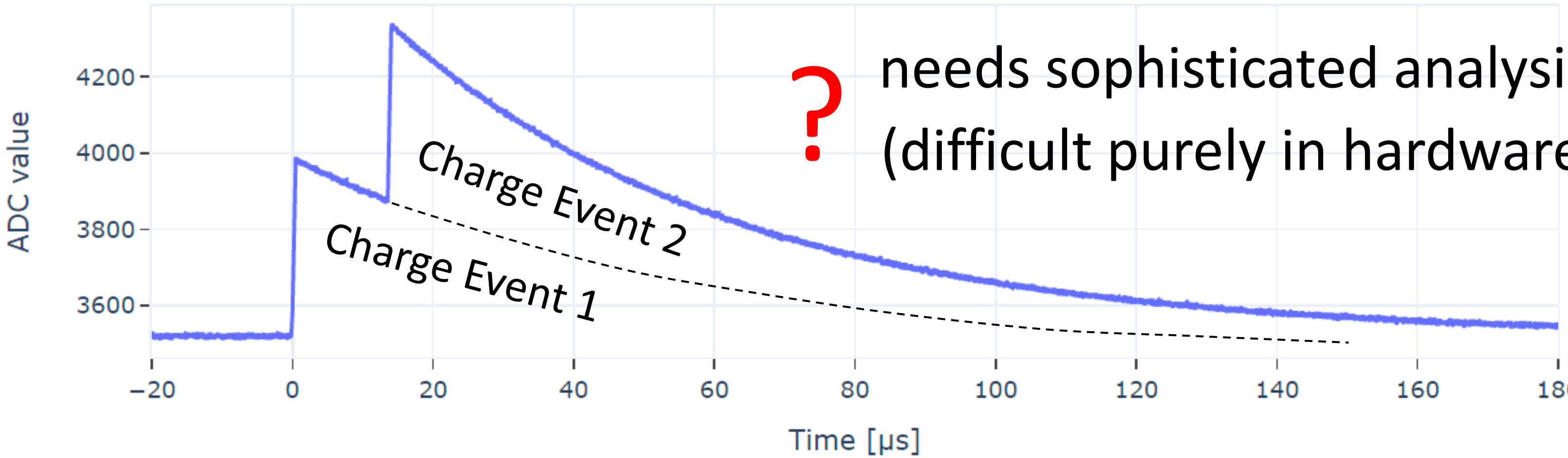


Potentially “best” method

at least in terms of energy resolution (measurement precision)

High Purity Germanium detector signal

But what about pileup – multiple photons at once?



needs sophisticated analysis
(difficult purely in hardware)

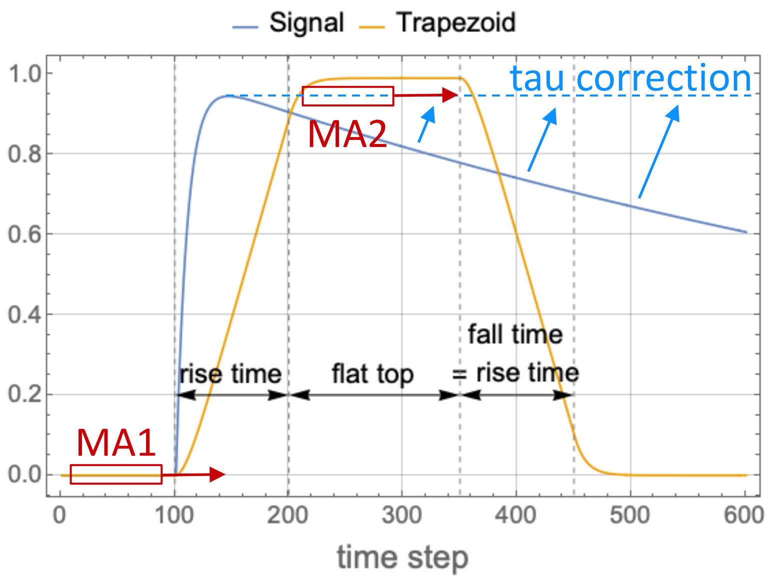
And what about storage?

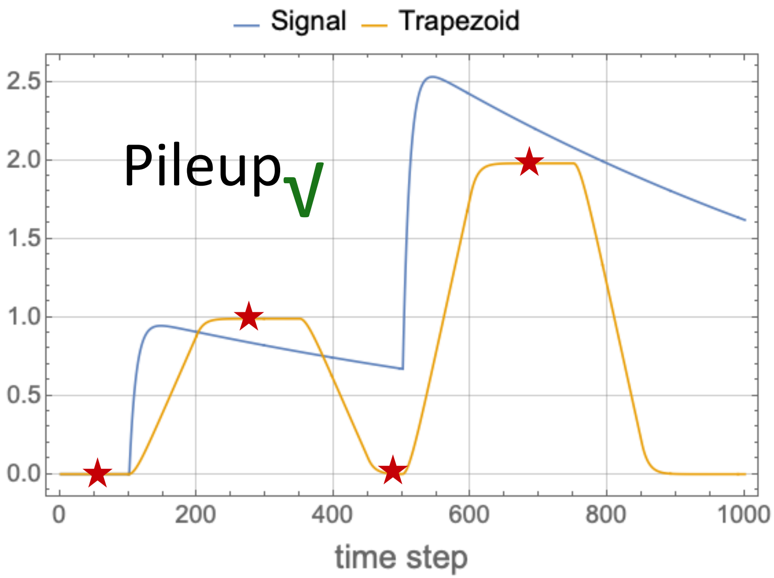
14 bit x 50.000 samples ≈ 1 MB per event

Tens of TB per Measurement!

High Purity Germanium detector signal

Good alternative – Trapezoidal Filter





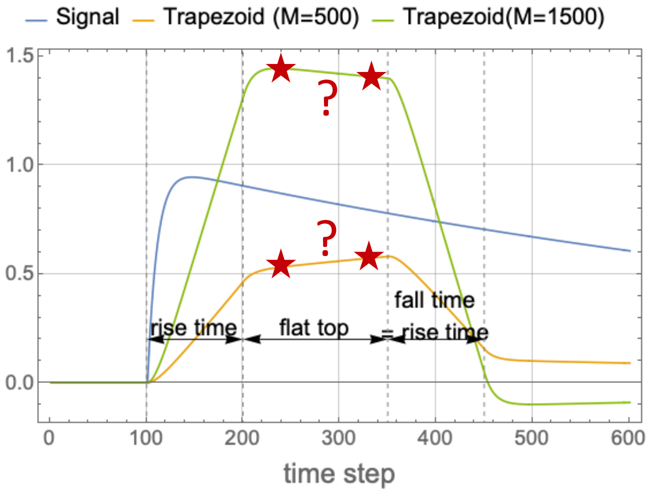
Easy to implement in hardware (FPGA)

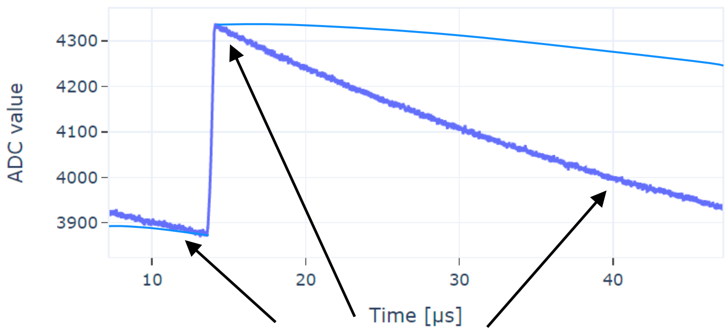
And what about storage?

14 bit x 4 values ➔ few GB per sample

High Purity Germanium detector signal

Disadvantages of Trapezoidal Filter





multiple decay constants

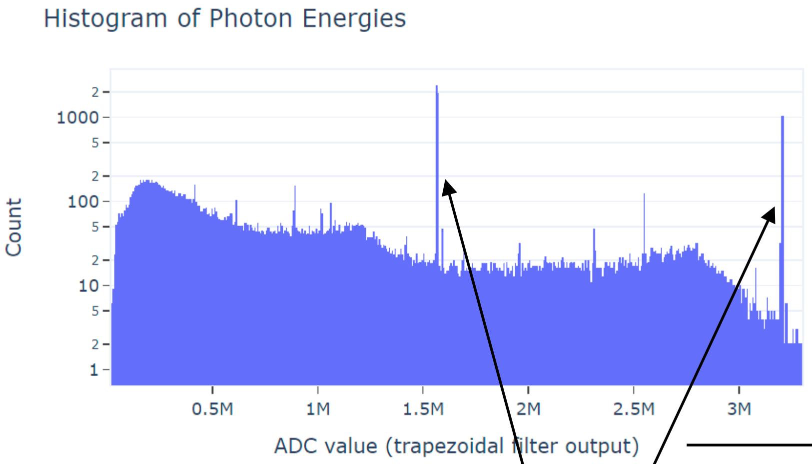
“Optimal” filter parameters strongly depend on detector

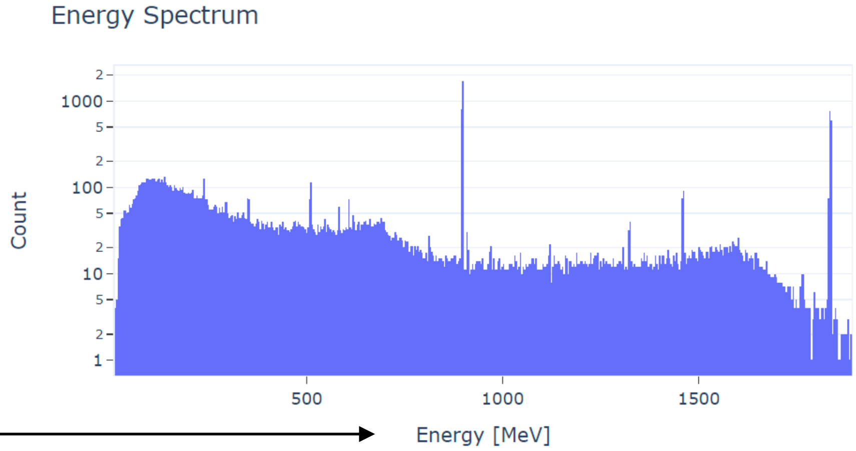
Incorrect tau correction can introduce baseline error

Physics Results

PAUL SCHERRER INSTITUT

Energy Spectrum (Y88 radioactive source)





Energy Calibration

Radiation: **Decay mode:** Electron Capture

Major Positrons:

Max E (MeV)	Avg E (MeV)	# per 100 dis
0.755	0.355	0.2

Max. Beta Range in air 290 cm or 9.51 ft
Max. Beta Range in water 0.31 cm

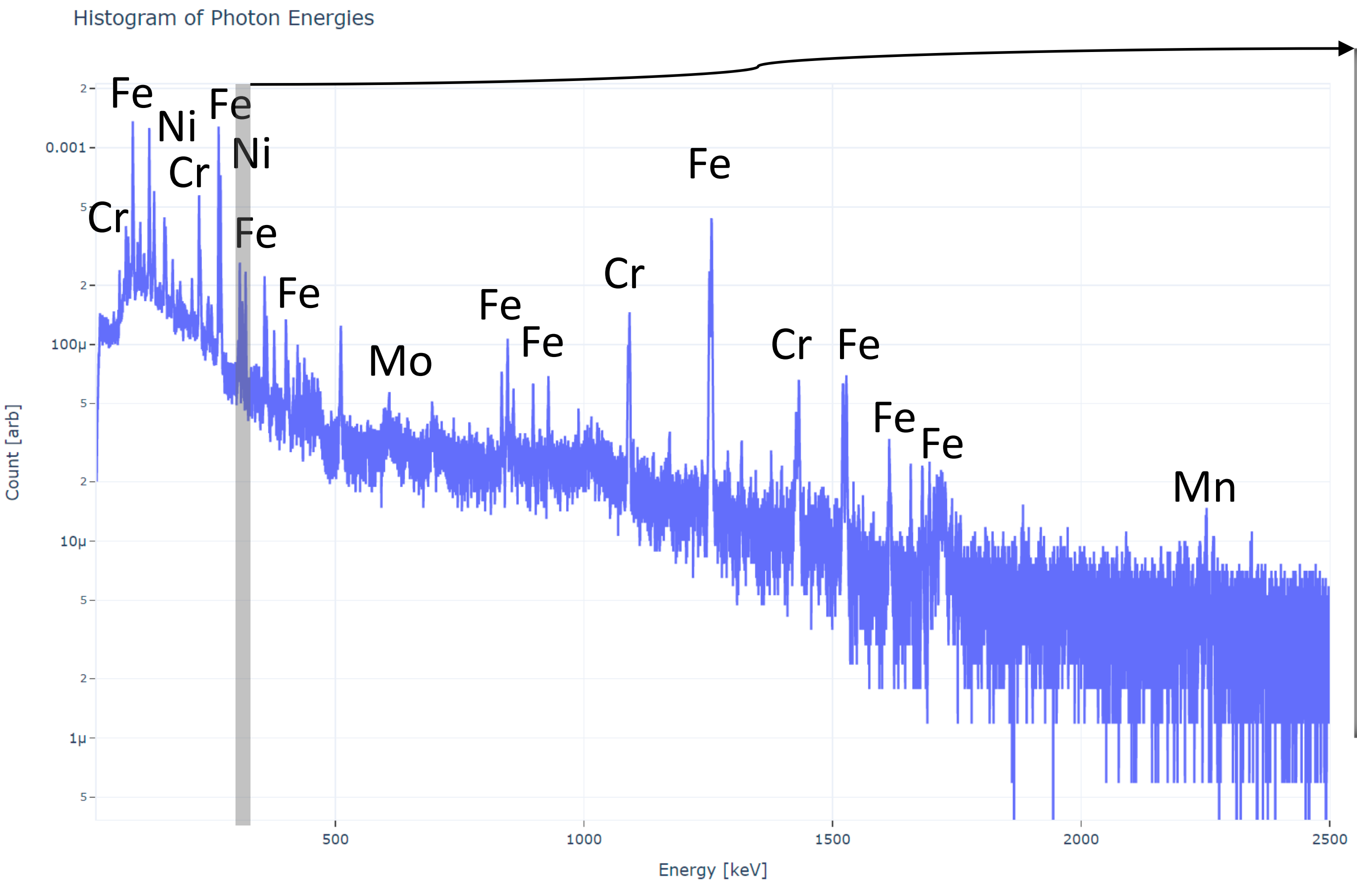
Major Gammas:

E (MeV)	# per 100 dis
0.898	93
1.836	99
2.734	0.6

Avg. gamma E = 1.060 MeV

Proper calibration for every detector:
many lines over wide range
using non-linear corrections

Muonic Energy Spectrum (Stainless Steel)



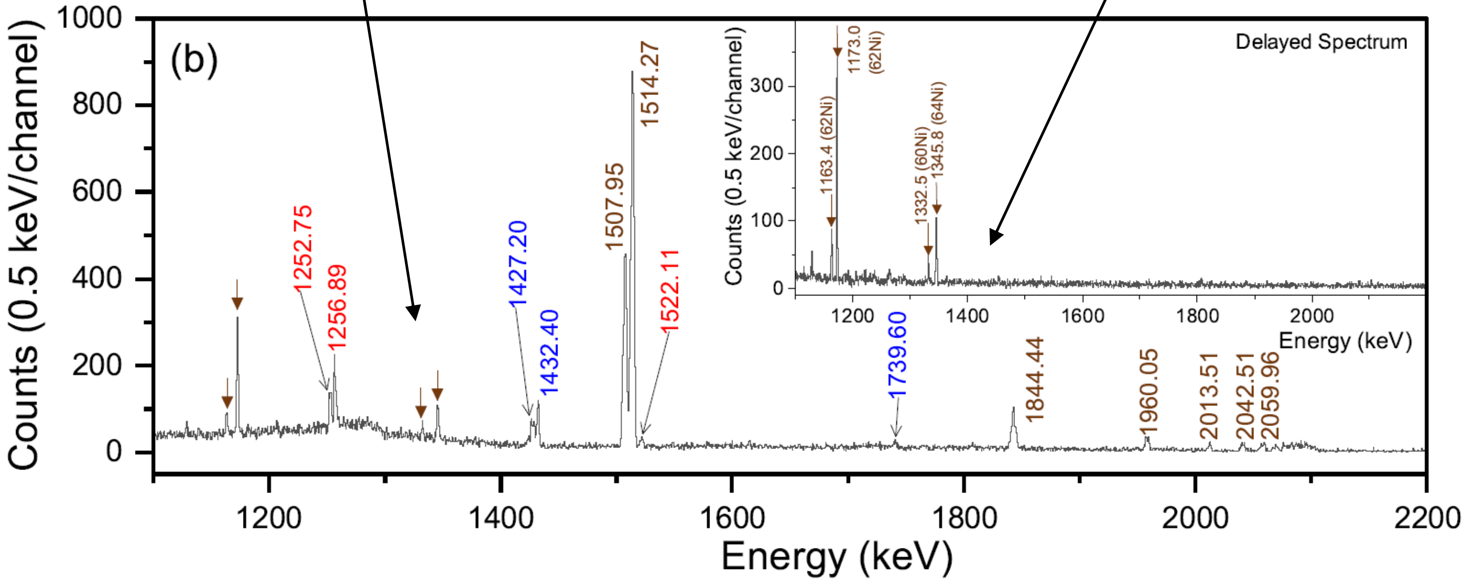
searching all values in between the energies:
301keV and 321keV

307.980	6g9/2-4f7/2	N7-P9	138Ba
318.000	4f7/2-3d5/2	M5-N7	114Cd
317.995	4f7/2-3d5/2	M5-N7	112Cd
317.992	4f7/2-3d5/2	M5-N7	111Cd
317.989	4f7/2-3d5/2	M5-N7	110Cd
305.082	4d5/2-2p3/2	L3-N5	52Cr
307.987	4d3/2-2p1/2	L2-N4	52Cr
305.094	4d5/2-2p3/2	L3-N5	53Cr
307.998	4d3/2-2p1/2	L2-N4	53Cr
318.432	8h11/2-5g9/2	O9-R11	166Er
318.434	8h11/2-5g9/2	O9-R11	168Er
314.4	Fe->56Mn		
309.099	8h11/2-5g9/2	O9-R11	165Ho
319.129	6g9/2-4f7/2	N7-P9	139La
319.127	6g9/2-4f7/2	N7-P9	138La
304.459	5g9/2-4f7/2	N7-O9	169Tm
307.632	5g7/2-4f5/2	N6-O8	169Tm
313.607	5d5/2-2p3/2	L3-O5	51V
316.120	5d3/2-2p1/2	L2-O4	51V
303.499	7h11/2-5g9/2	O9-Q11	184W
313.428	5g9/2-4f7/2	N7-O9	174Yb
316.792	5g7/2-4f5/2	N6-O8	174Yb
305.795	5f7/2-3d5/2	M5-O7	89Y
309.241	7g9/2-4f7/2	N7-Q9	121Sb
311.800	7h11/2-5g9/2	O9-Q11	187Re
303.459	6d5/2-2p3/2	L3-P5	48Ti
305.588	6d3/2-2p1/2	L2-P4	48Ti
313.489	7d5/2-2p3/2	L3-Q5	48Ti
315.629	7d3/2-2p1/2	L2-Q4	48Ti
304.9	Pb->206Tl		
304.783	4f7/2-3d5/2	M5-N7	107Ag
308.455	4f5/2-3d3/2	M4-N6	107Ag
313.949	4p3/2-1s1/2	K1-N3	23Na
308.341	3d5/2-2p3/2	L3-M5	58Ni
313.299	3d3/2-2p1/2	L2-M4	58Ni
320.216	7h11/2-5g9/2	N7-P9	192Os

Muonic Energy Spectrum including timing

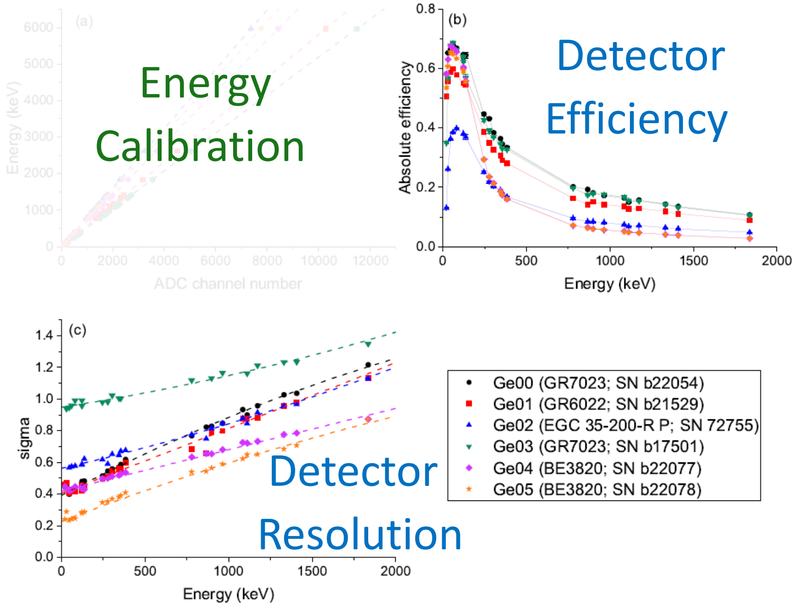
X-ray cascade:
prompt, <100ns delay

Nuclear Capture and
subsequent decay:
delayed, >100ns delay

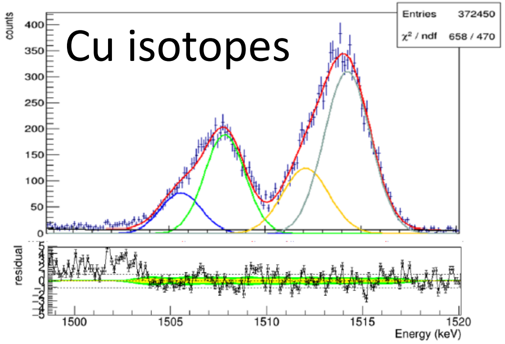


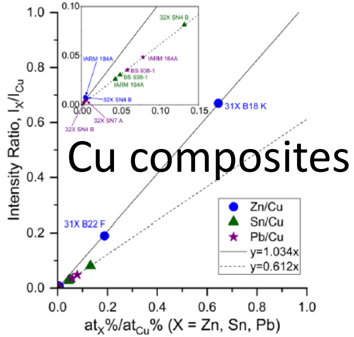
Quantitative Results

- So far only qualitative results ➔ Which elements / isotopes do we find in a sample?
- Quantitative results are much harder to obtain!
  - *Theory results on relative intensities are currently very unreliable*



Reference Samples allow for relative measurements





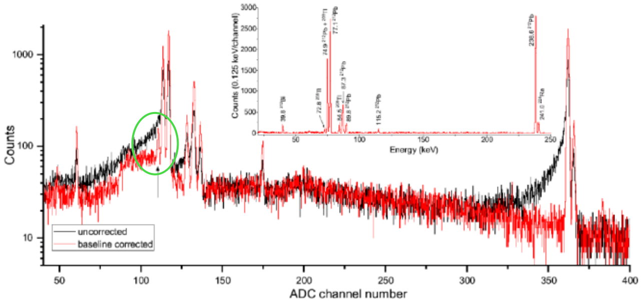
Attenuation of X-rays and gammas depends on exact material composition, density and penetration depth of the muon!

Part III

Machine Learning

Baseline correction

- Unoptimized trapezoidal filter settings and pile-up can lead to reduced energy / timing resolution and “missing” peaks

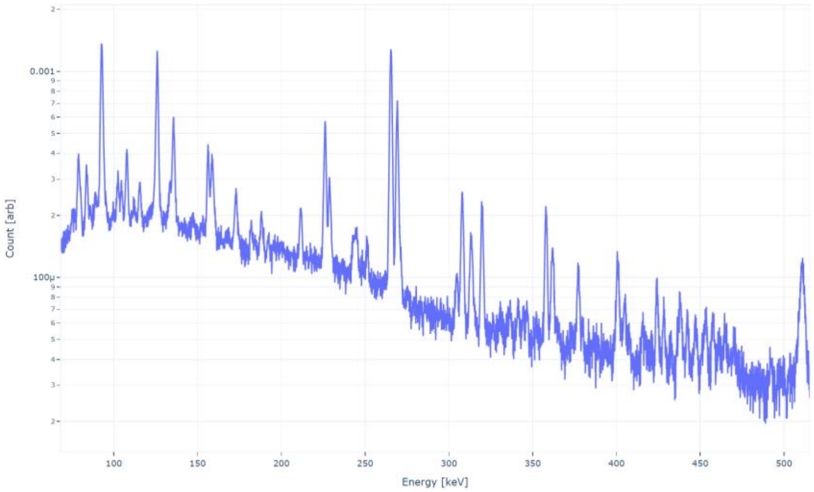


- Machine Learning Model could be used to
  - efficient optimization of trapezoidal filter parameters, e.g. reinforcement learning
  - replace trapezoidal filter by advanced algorithm based on neural network model
    - train using clean set of events with full waveforms
    - implement resulting model directly on FPGA?

low to medium priority

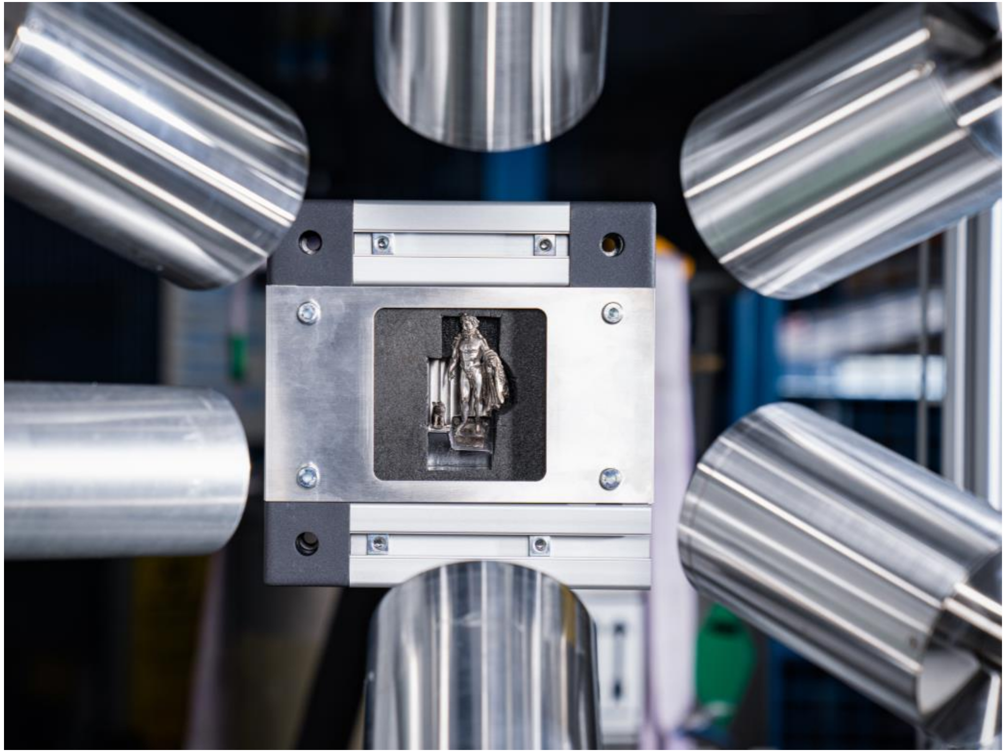
Peak identification

- Currently, the peak Identification is most serious analysis bottleneck!
  – has to done by experienced MIXE scientist (usually not by the user)
  – very time-consuming manual labor
  – overlapping peaks can make ID ambiguous
  – can not be done online → usually no ad-hoc decisions regarding the measurement
- Machine Learning Model could be used to
  – directly identify relevant peak candidates based on a dictionary learning approach
  – give probabilities for candidates based on full spectrum (not only single peaks)
  – empower users to analysize data without domain knowledge
  – allow for “online” (within minutes) analysis of spectra



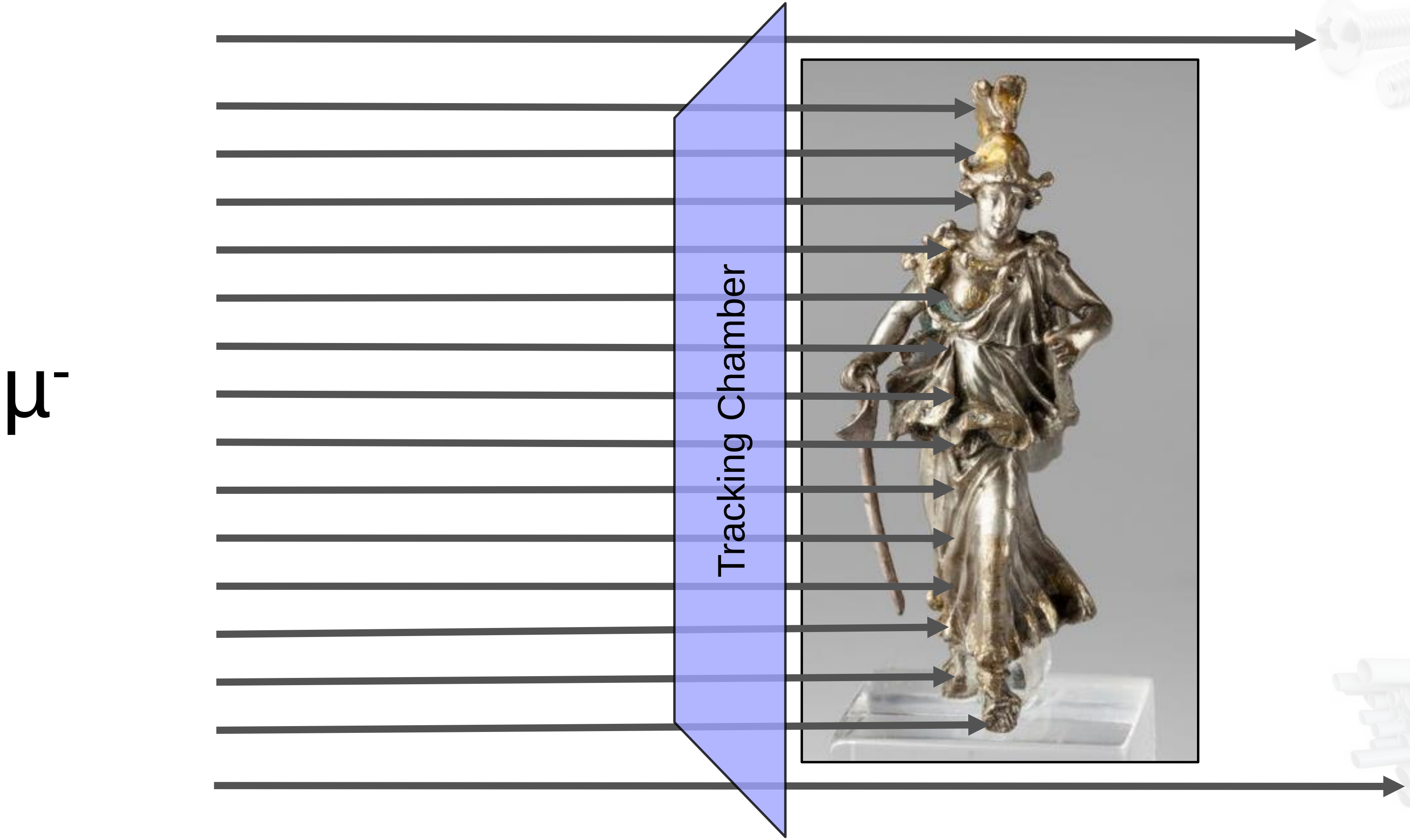
very high priority

Thank you for your attention!



Any Questions, Comments or Suggestions?

Bonus: Tracking



Twin GEM-TPC Tracking chamber

- Twin Time-Projection-Chamber
  - 1D strip readout – 1024 channels in total
  - X (cluster on strips) & Y (drift time)
- Currently track is given by 2 points
  - simple average of spatial charge distribution (X)
  - simple average of temporal charge distribution (Y)
- Machine Learning Model could be used to
  - recover angle information within each TPC
    - determine quality of the track
    - increase detector resolution
  - could be implemented e.g. as (convolutional) neural network, trained with simulation (existing)

low priority