Histogram tools: NEST and HistoFit

Oleksandr Volynets

Max-Planck-Institute for Physics

MaGe Workshop 2010, Munich

January 19, 2010

Outline

- Introduction
- NEST(New Energy Spectrum Toolkit):
  - Motivation
  - Implementation
  - Results
- HistoFit
  - Motivation
  - Implementation, method
  - Creating dummy data for check
  - Results of decomposing spectras

Introduction

- MaGe/Geant4 is a powerful tool to get the expected energy spectrum of the experiment.
- However, all we get is a set of files with spectras from each holder, cable, detector etc.
- But the Data only contains one single part which is naturally a sum of all contributions.
- We have to do offline histogram analysis to get all those files together to see how good we understand our detector.
- In this presentation two histogram tools for spectra analysis will be shown

NEST(New Energy Spectrum Toolkit)

Motivation

- The detector has many parts, we simulate (dominant) radioactive sources in separate parts;
- In order to plot MC spectrum we have to add all background contributions with its' weights according to activities/masses/time;
- To deal with all these histograms a tool with simple control of masses, activity etc. of all sources is needed;

Implementation

The tool called NEST(New Energy Spectrum Toolkit) was created by Jens Schubert. It has the following features:

- Draw separetely *Total energy in all crystals/Total energy in every crystal/Energy in the individual segment*;
- Draw separately *All/Single-crystal/Single-segment events*;
- All parameters of the sources are read from a single text file; all the parameters of the geometry are read from a single XML-file;
- Flexible control of output (counts in energy window, expected counts of every single background contribution etc.);
- Tag the contributions and get numbers (e.g. background index) for only needed tags;
- ...

Implementation

Example list of parameters:

```
# +-----------------------------------------------------------------------------+
# | +-------------------------------------------------------------------------+ |
# | | LIFE TIMES                                                              | |
# | +-------------------------------------------------------------------------+ |
# +-----------------------------------------------------------------------------+
#HALFlifetime___Agm108                  418 years   "the HALFlifetime of Agm108"
lifetime___Agm108                   1.9032785e10 s       "the lifetime of Agm108"

#HALFlifetime___Agm110                  249.9  days    "the HALFlifetime of Agm110"
lifetime___Agm110                   3.11497e7 s       "the lifetime of Agm110"

# +-----------------------------------------------------------------------------+
# | CONTAMINATION LEVELS  -- in Phase-1-Crystals                                |
# +-----------------------------------------------------------------------------+
contamination_Th232___Xtal_ANGx                          5.5e-8      Bq/kg   "The contamination of
contamination_Th232_MeasurementDate___Xtal_ANGx/date  2008-01-01  none    "The contamination of
```

Implementation

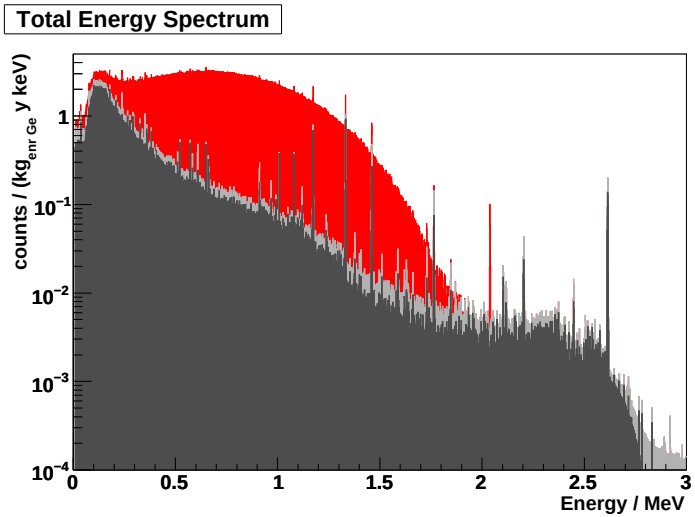
Example geometry description (PhaseIIideal.xml):

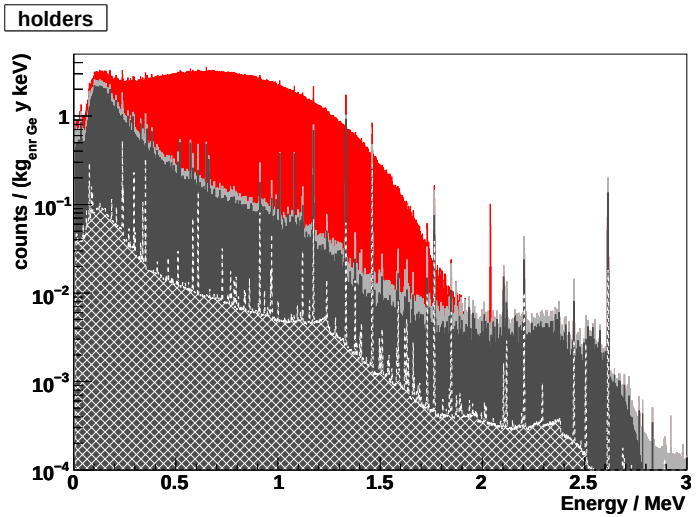
```
<ListOfElBaCos>
  <Replacer>
   <ReplaceExpressionName>    <RM>ProcBaseName</RM><RM>Nnu</RM>   </ReplaceExpressionName>
   <ReplaceExpressionValue>   <RM>DBD_2nu2beta</RM><RM>2   </RM>   </ReplaceExpressionValue>
   <ReplaceExpressionValue>   <RM>DBD_0nu2beta</RM><RM>0   </RM>   </ReplaceExpressionValue>
  </Replacer>

  <!-- loop over all phase-1 crystals -->
  <ListOfElBaCos switch="on">

   <Replacer>
    <ReplaceExpressionName>                <RM>XtalName</RM><RM>GeXXFrac</RM><RM>PhysVolIndex </R
    <ReplaceExpressionValue switch="on">   <RM>ANG1     </RM><RM>ANG1     </RM><RM>6            </R
    <ReplaceExpressionValue switch="on">   <RM>ANG2     </RM><RM>ANG2     </RM><RM>12           </R
    <ReplaceExpressionValue switch="on">   <RM>ANG3     </RM><RM>ANG3     </RM><RM>13           </R
    <ReplaceExpressionValue switch="on">   <RM>ANG4     </RM><RM>ANG4     </RM><RM>7            </R
    <ReplaceExpressionValue switch="on">   <RM>ANG5     </RM><RM>ANG5     </RM><RM>11           </R
    <ReplaceExpressionValue switch="on">   <RM>GTF1     </RM><RM>GTFx     </RM><RM>5            </R
    <ReplaceExpressionValue switch="on">   <RM>GTF2     </RM><RM>GTFx     </RM><RM>3            </R
    <ReplaceExpressionValue switch="on">   <RM>GTF3     </RM><RM>GTFx     </RM><RM>2            </R
    <ReplaceExpressionValue switch="on">   <RM>GTF4     </RM><RM>GTFx     </RM><RM>1            </R
    <ReplaceExpressionValue switch="on">   <RM>GTF5     </RM><RM>GTFx     </RM><RM>4            </R
    <ReplaceExpressionValue switch="on">   <RM>GTF6     </RM><RM>GTFx     </RM><RM>0            </R
    <ReplaceExpressionValue switch="on">   <RM>RG1      </RM><RM>RG1      </RM><RM>9            </R
    <ReplaceExpressionValue switch="on">   <RM>RG2      </RM><RM>RG2      </RM><RM>10           </R
    <ReplaceExpressionValue switch="on">   <RM>RG3      </RM><RM>RG3      </RM><RM>8            </R
   </Replacer>
   ....
```

Results


Total Energy Spectrum
counts / (kg$_{enr\ Ge}$ y keV)
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
0.5
1
1.5
2
2.5
3
Energy / MeV


holders
counts / (kg$_{enr\ Ge}$ y keV)
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
0.5
1
1.5
2
2.5
3
Energy / MeV

HistoFit

Motivation

Let's imagine the situation:

- Monte Carlo energy spectras calculated

Motivation

Let's imagine the situation:

- Monte Carlo energy spectras calculated
- Data energy spectrum obtained

Motivation

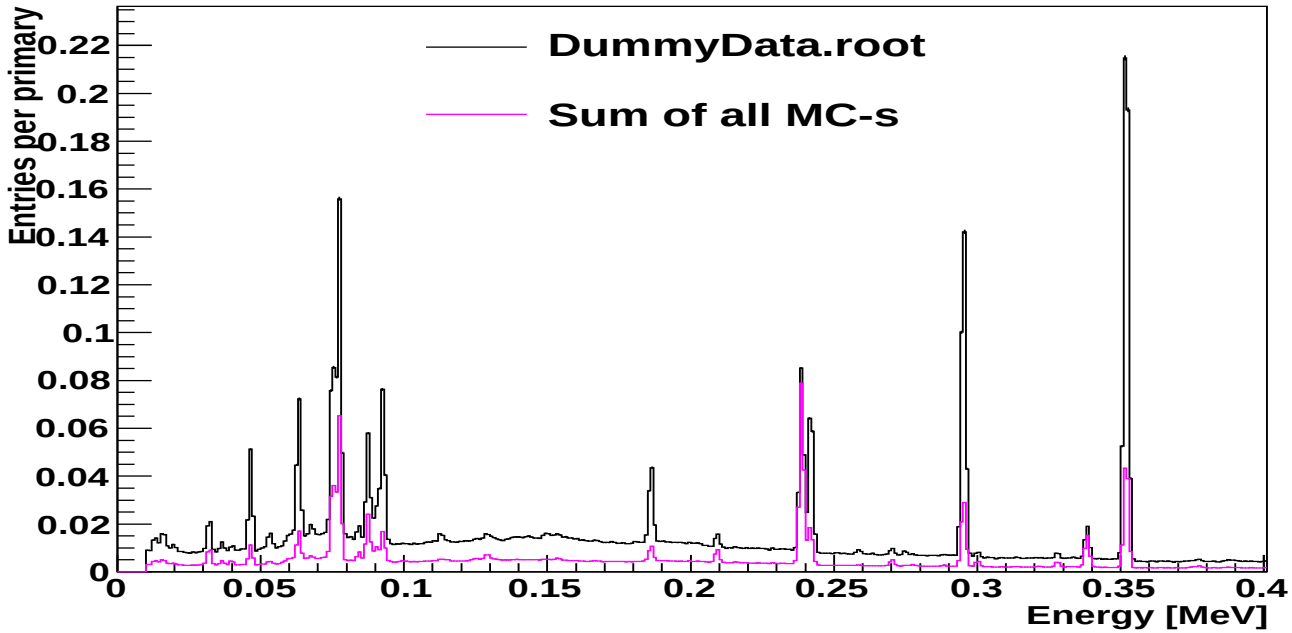
Let's imagine the situation:

- Monte Carlo energy spectras calculated
- Data energy spectrum obtained
- Monte Carlo describes the shape, but not the numbers

Motivation

Let's imagine the situation:

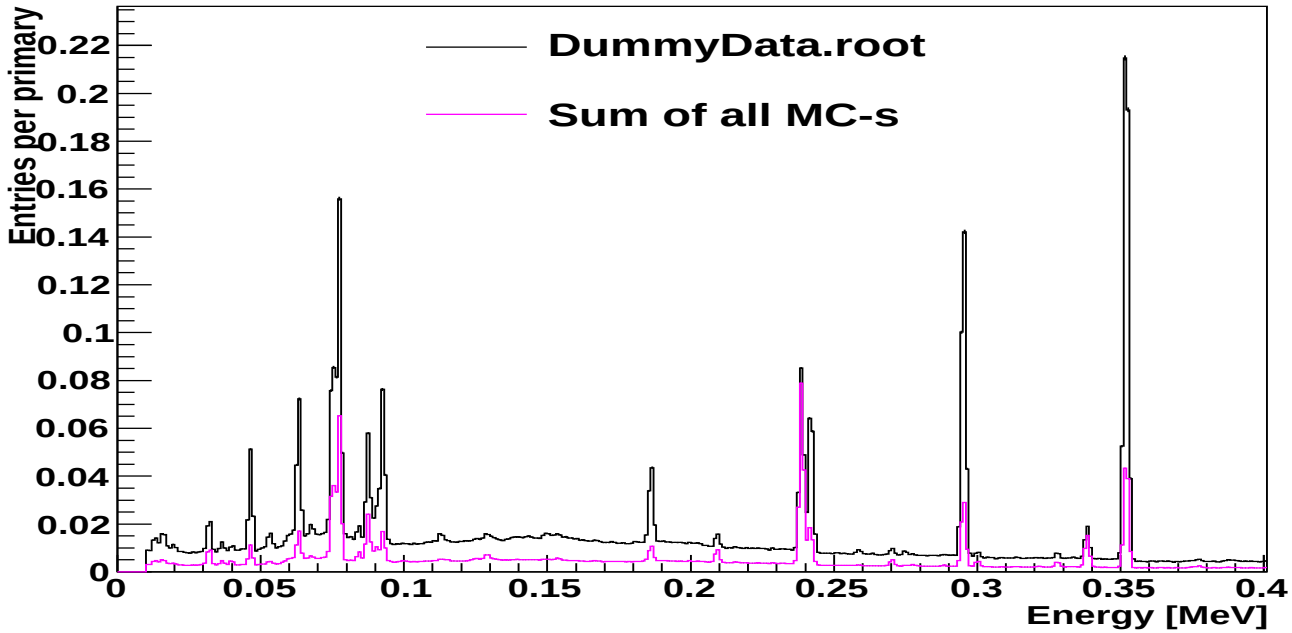
- Monte Carlo energy spectras calculated
- Data energy spectrum obtained
- Monte Carlo describes the shape, but not the numbers



Motivation

Let's imagine the situation:

- Monte Carlo energy spectras calculated
- Data energy spectrum obtained
- Monte Carlo describes the shape, but not the numbers
- Background or Signal contribution estimated wrong?



Motivation

Tool developed for:

- Fitting several main MC contributions to data spectrum
- Getting correct estimates of the contribution

Motivation, implementation

Tool developed for:

- Fitting several main MC contributions to data spectrum
- Getting correct estimates of the contribution

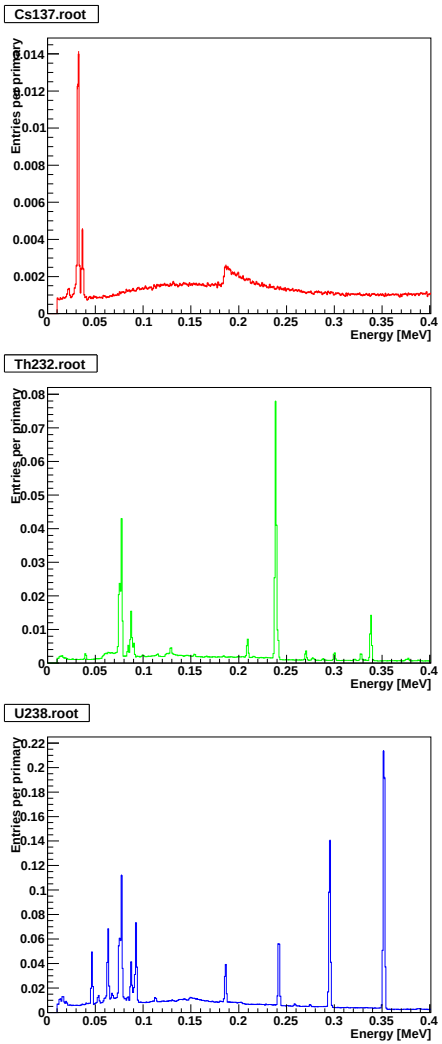
Method:

$$\chi^2 = \sum_{i,\text{bins}} \frac{\left(D_i - \alpha_0 MC_i^0 - \alpha_1 MC_i^1 - \ldots\right)}{\left(\sigma_{D_i}^2 + \alpha_0^2 \sigma_{MC_i^0}^2 + \alpha_1^2 \sigma_{MC_i^1} + \ldots\right)}$$

Use MINUIT to minimize and get α_i parameters

Creating dummy data spectrum

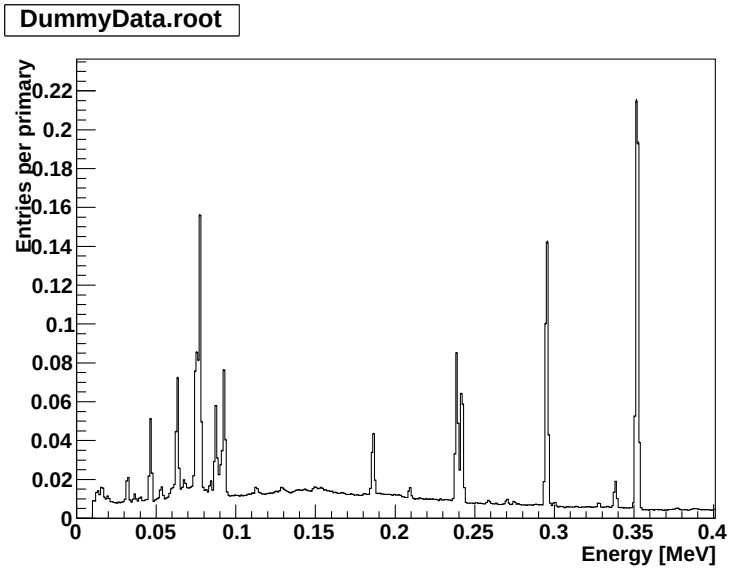
To check whether it works a dummy data was created:



$\times 2$

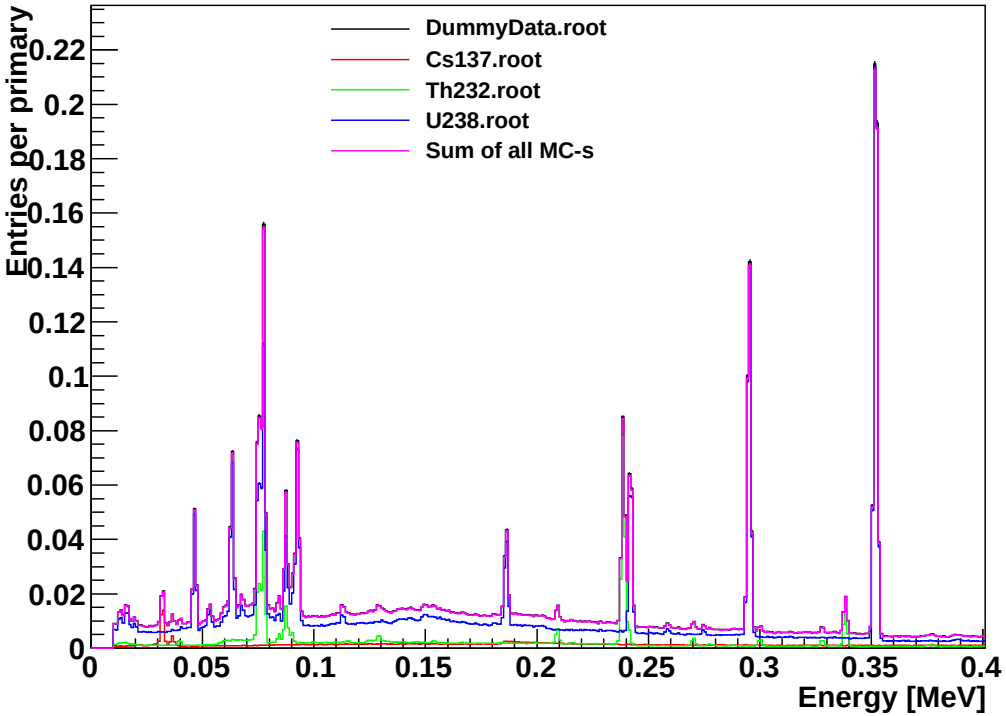
$\times 1 \Rightarrow$

$\times 5$



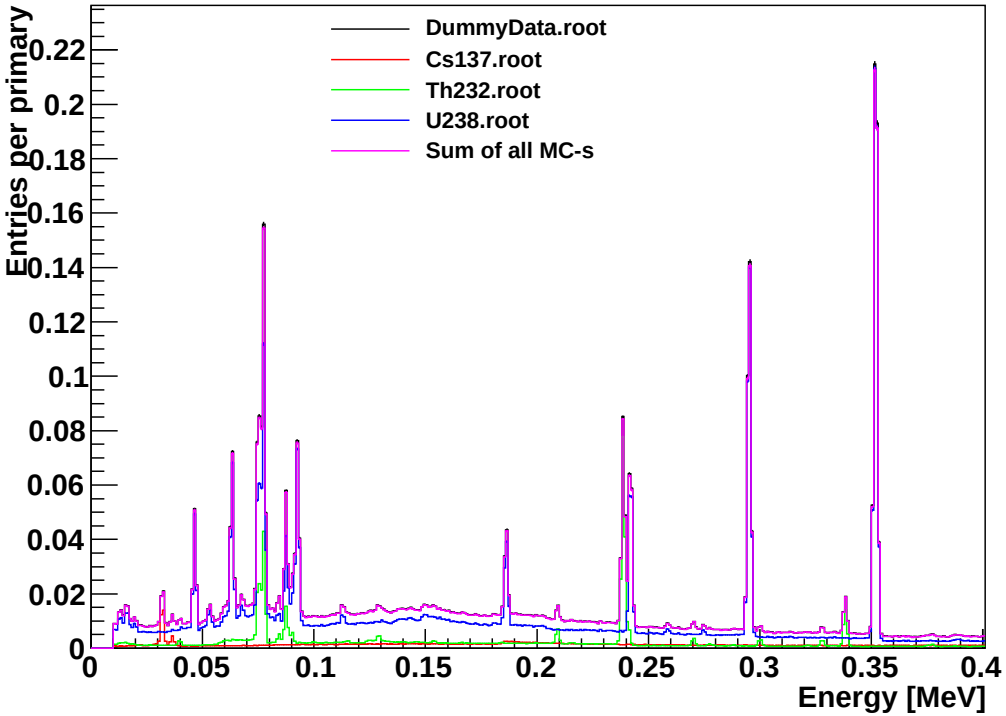
Self-check: Decomposing spectrum

Using the same 3 histograms to fit to "Dummy Data":



Self-check: Decomposing spectrum

Using the same 3 histograms to fit to "Dummy Data":



Parameter 0 $= 2 \pm 0.00679609$

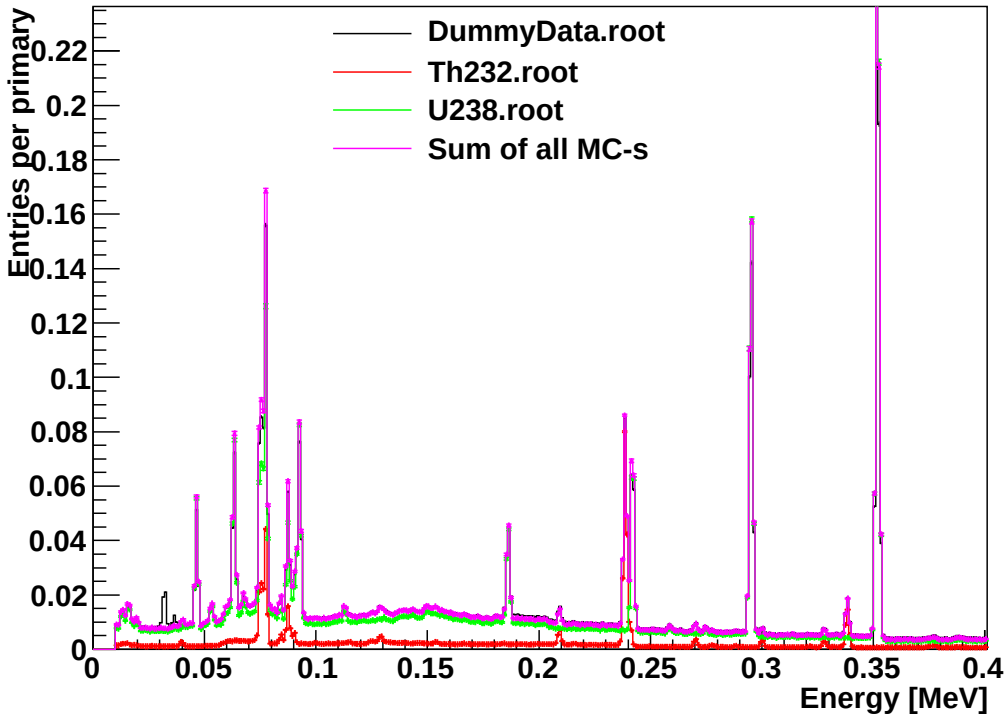
Parameter 1 $= 0.999997 \pm 0.00316915$ $\quad \chi^2 \approx 10^{-5}$

Parameter 2 $= 4.99999 \pm 0.00722358$

Self-check 2: Decomposing wrong spectrum

Same procedure, but now Dummy Data contains deposits from 3 sources: ^{137}Cs, ^{232}Th and ^{238}U, but set of MC-spectras contains only 2 of them. And it converges but with wrong results (using given spectras):

Self-check 2: Decomposing wrong spectrum



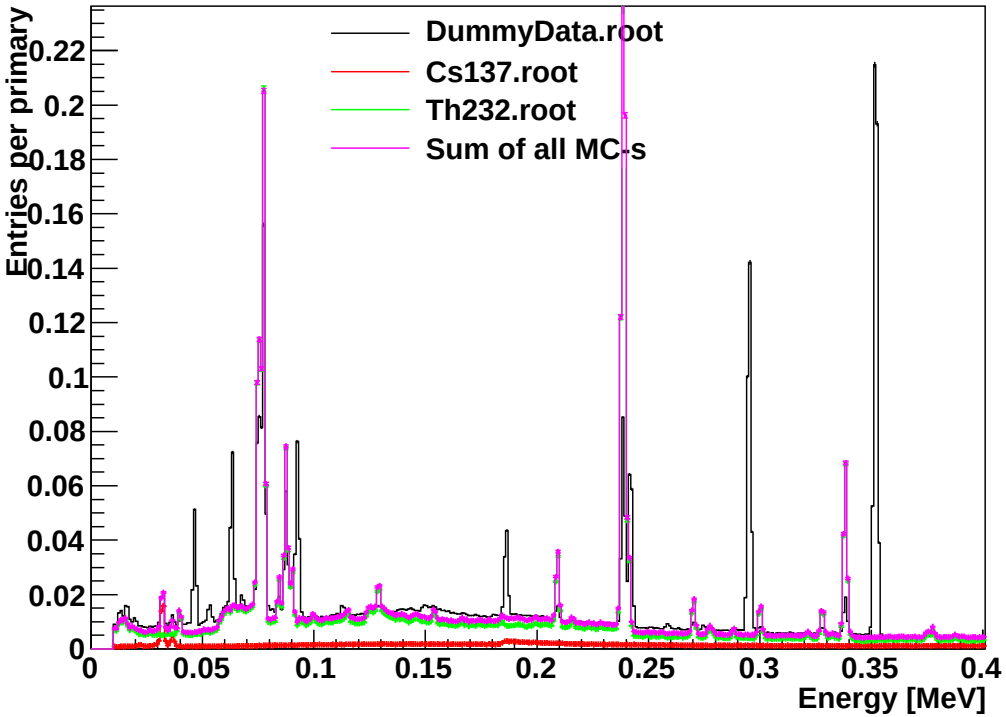
Parameter 0 = *not used*

Parameter 1 $= 1.029 \pm 0.003$

Parameter 2 $= 5.638 \pm 0.008$

$\chi^2 = 150000$

Self-check 2: Decomposing wrong spectrum



Parameter 0 $= 2.266 \pm 0.009$

Parameter 1 $= 4.794 \pm 0.008$

Parameter 2 $=$ *not used*

$\chi^2 = 530000$

Conclusions and discussion

Next steps:

- Wait for the data at GERDA, or
- Find some data with measuring several sources simultaneously

Is the following needed:

- Fitting MC-s within different regions (e.g. Sp1=0..1.5 MeV; Sp2=1..2.3 MeV) - complicated but possible
- Else?