Updates on the sensitivity study to BSM physics.

New and improved multi-dimensional frequentist limits & improved Bayesian models

Maros Petro

Outline

1. Background model

2. Methodology

1. **Feldman Cousins sensitivity**
2. **Bayesian sensitivity**

3. Results

1. $0\nu\beta\beta$
2. $0\nu\beta\beta\chi^0$
3. $0\nu\beta\beta\chi^0\chi^0$
4. $\nu_R\nu_L\beta\beta$

1. Background model

The first step in calculating the sensitvity is to have a proper background model, which defines how sensitive our experiment can be!

Obtaining the background model

To obtain the background model used in this analysis I performed a series of **simulations** of various background **sources** (identified mostly from NEMO-3).

Simulation setup:

1. Falaise version 5.1.5
2. Reconstruction pipeline: `MockCalibration -> ` **Cimrman** `-> ChargedParticleTracker -> GammaClusterizer -> SNCuts -> MiModule`
3. Magnetic field **off**
4. 8% FWHM energy resolution
5. Foil geometry: `RealisticFlat`

Data-cuts (using SNCuts)

- 2 tracks
- 2 foil vertices
- 2 **distinct** associated calo-hits
- $E_{sum} \in (300, 3500)$ keV
- max vertex distance on foil $r < 50$ mm
- ToF: $P_{int} \geq 4\%$ & $P_{ext} \leq 1\%$

Background sources

Simulated 100M events each:

1. $2\nu\beta\beta$ –> $T_{1/2}^{2\nu} = 9.4 \times 10^{19}$ yr
2. ^{214}Bi –> A = $10\ \mu Bq/kg$
3. ^{208}Tl –> A = $2\ \mu Bq/kg$
4. ^{40}K –> A = $58\ mBq/kg$
5. ^{234m}Pa –> A = $17\ mBq/kg$
6. Radon –> A = $150\mu Bq/m^3$

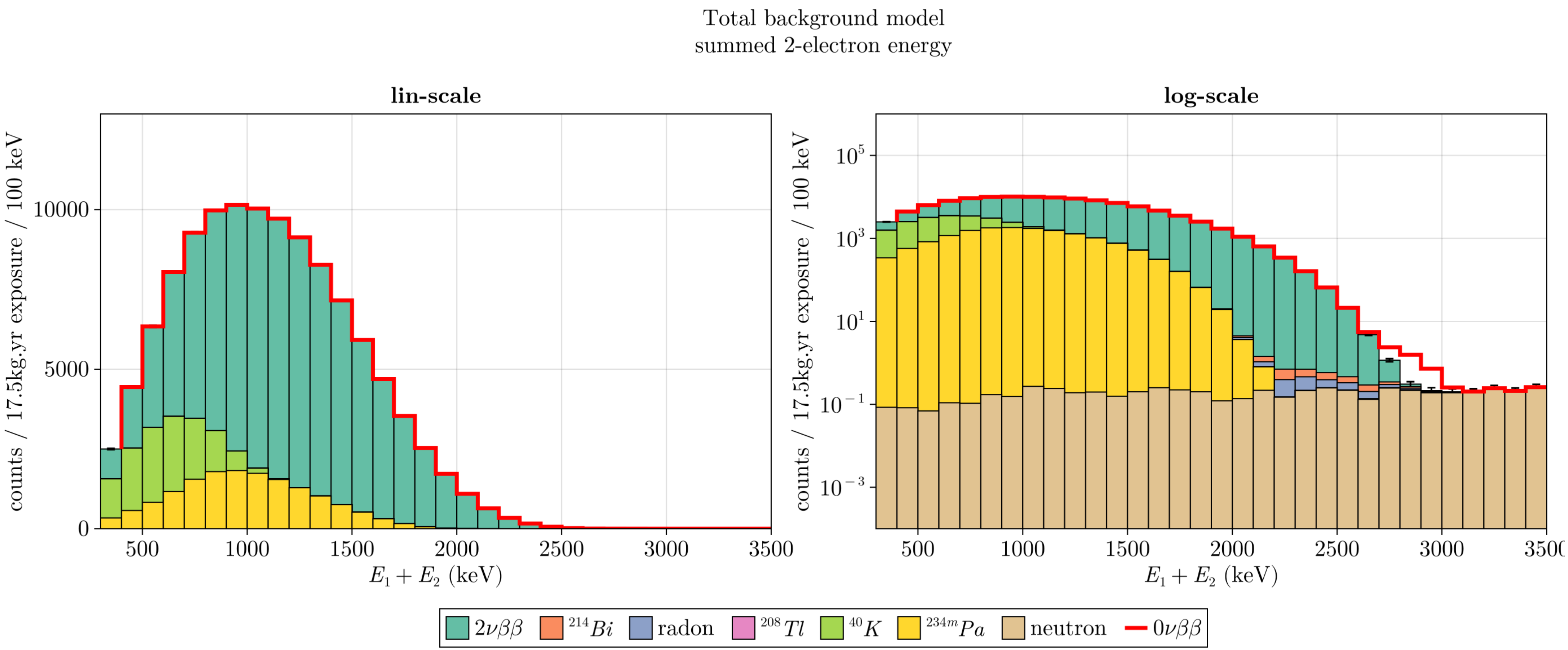
🤔 Neutron data is obtained from Sam but not used everywhere cause I wasn't able to do it properly in time. 🤭

Background Model

Emin = 300 keV

Emax = 3500 keV

Thalf = 4.5009999999999997e24 yr



Background by source

process	bkg counts
2nubb	87663.0±29.0
Bi214	8.903±0.022
Radon	10.62±0.15
Tl208	1.0227±0.0034
K40	12155.0±63.0
Pa234m	15578.0±39.0
neutron	5.91
total	115422.91±80.0

- Largest contribution due to $2\nu\beta\beta$ - almost 90k events (+10k compared to old tracking)
  - Good for 2nu physics
  - Bad for 0nu
- K40 and Pa234m relevant **only** at lower energies
- Radon will highly depend on whether or not we have anti-Rn
- For now, "simple" neutron data

2. Methodology

2.1. Feldman Cousins sensitvity

$$T^{1/2} = ln(2)\frac{N_A \cdot m \cdot t}{W}\frac{\varepsilon}{\mathcal{S}(\bar{b})}$$

The one you know: 1D approach!

I already gave several talks on this: DocDB#5943 and DocDB#5833

The general idea

1. Simulate relevant signal process
2. Chose 1 variable which is tracked (i.e.E_{sum}, E_i, ϕ..)
3. For each possible combination of ROI calculate signal ε and $\mathcal{S}(\bar{b})^*$ as signal-to-background ratio $r(ROI) = \frac{\varepsilon}{\mathcal{S}(\bar{b})}$
4. Find which ROI maximizes $r(ROI)$ –> from the $max(r)$ calculate sensitvity (simulation)
5. Once you've measured data –> fit in ROI to get $\bar{b}$ –> calculate sensitivity
   - use ε from simulation, $\bar{b}$ from data

* $\mathcal{S}(\bar{b})$ is a Fledman Cousins limit calculated at α CL for total expected number of background counts in ROI $\bar{b}$

Example for $0\nu\beta\beta$ and E_{sum}:

- ROI in this case is E^l_{sum}, E^u_{sum}, where l, u are the lower and upper boundaries of ROI
- Then sensitivity is a function of E_{sum} ROI:

$$T^{1/2}(ROI) = T^{1/2}(E^l_{sum}, E^u_{sum}) = const. \frac{\varepsilon(E^l_{sum}, E^u_{sum})}{\mathcal{S}(\bar{b}(E^l_{sum}, E^u_{sum}))}$$

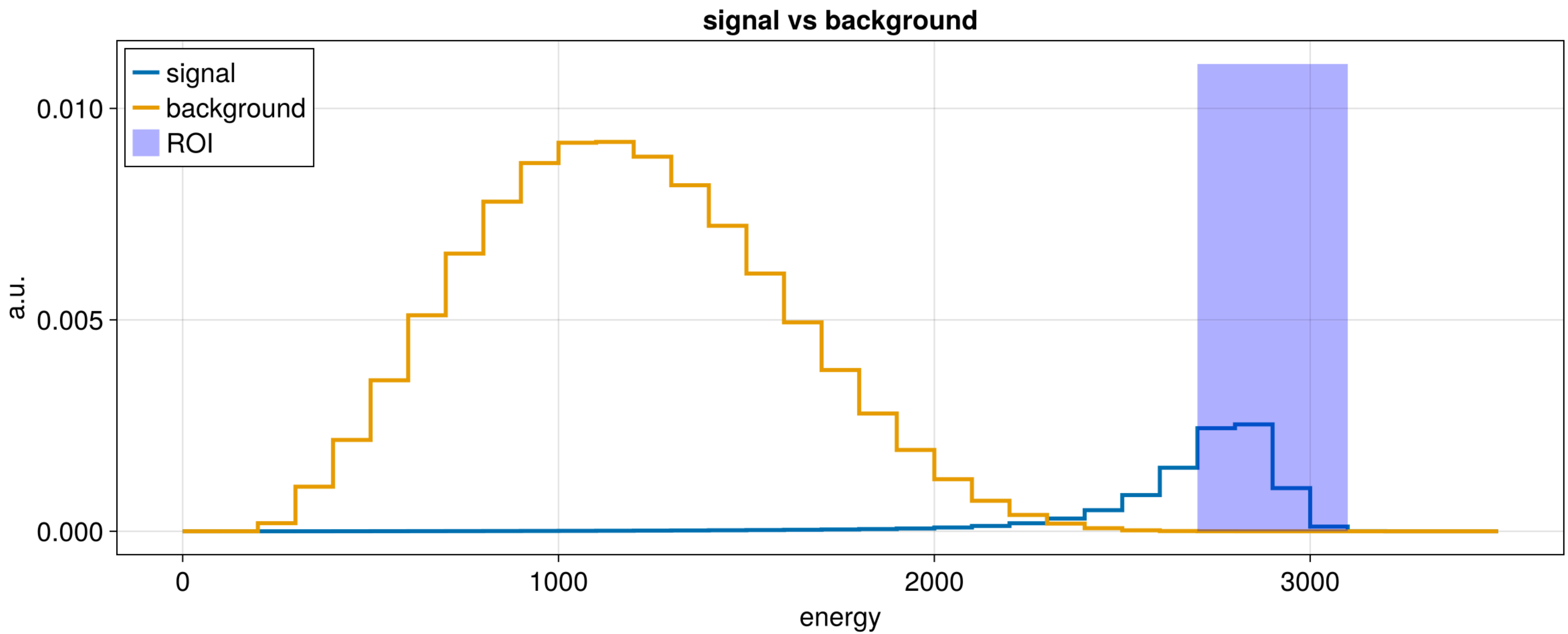
- We calculate $T^{1/2}$ for each combination and create a 2D map:

Example for $0\nu\beta\beta$ and E_{sum}:

Boundaries of ROI affect signal + background differently!

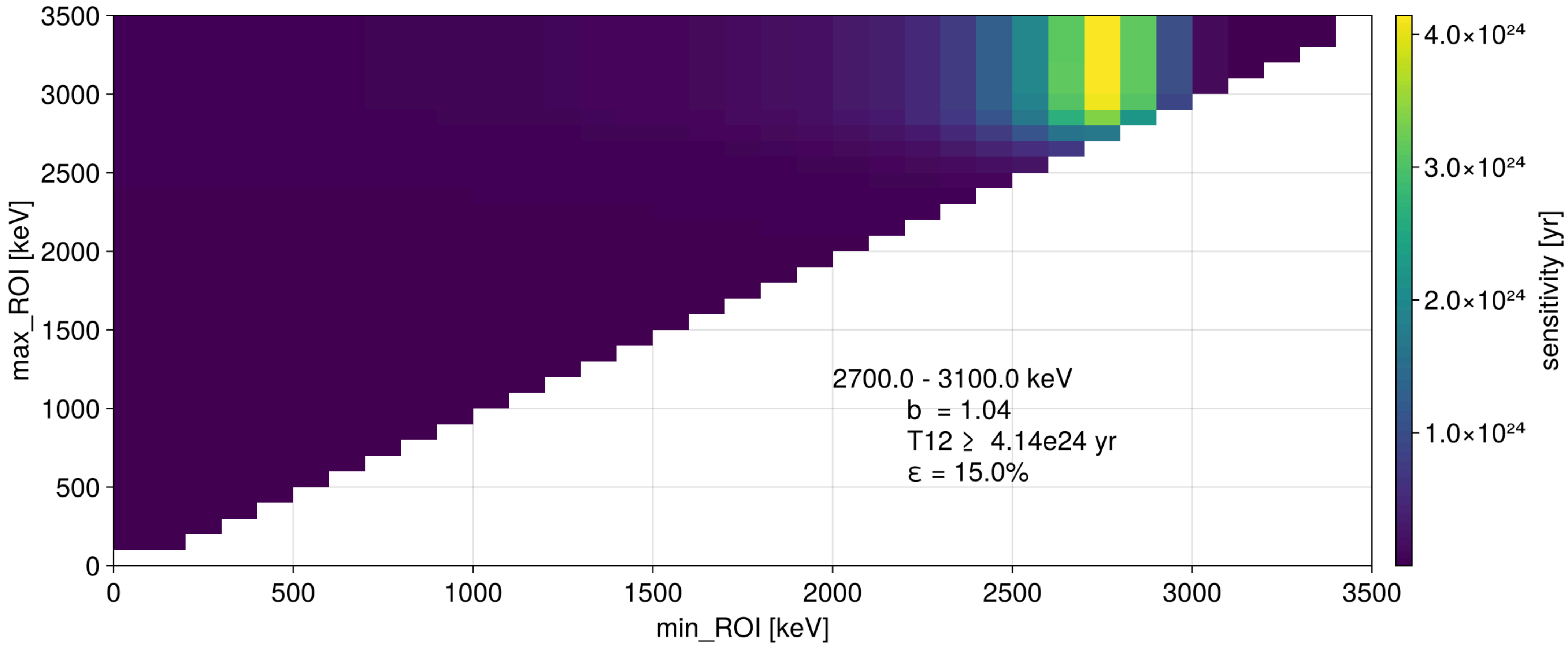
That we are interested in is:

- signal efficiency $\varepsilon(ROI)$
- expected background counts $\bar{b}(ROI)$



Example for $0\nu\beta\beta$ and E_{sum}:

For each combination of ROI (E^l_{sum}, E^u_{sum}) we get a different value of $T^{1/2}$ –> pick maximum




```
get_bkg_counts(): passed isotope bb0nu_foil_bulk is a signal process!!
```

- best ROI is $2700 - 3100$ keV
- with $\bar{b} = 1.04$ and $\varepsilon = 15$ we get sensitivity of $T^{1/2} \geq 4.14 \cdot 10^{24}$ yr

What if we look at another variable? Another signal process?

- For different signal processes (signal shapes!) we can be more sensitive in different channels: i.e. ϕ **should** be a better channel for RH spectra (will get to this later)
- Do we really need to chose only one? What if we can maximize **all of them at once!**

N-Dim ROI Search: the next step!

The general idea

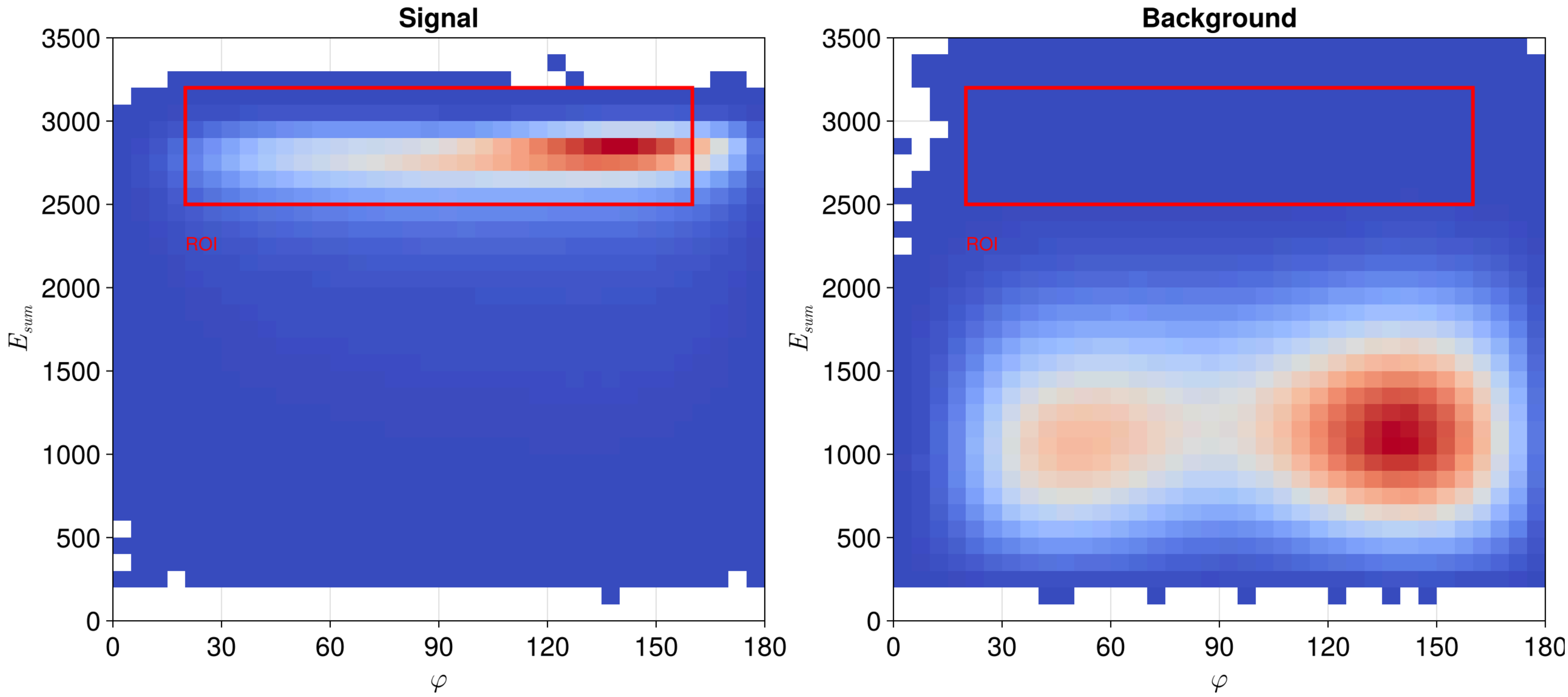
1. Simulate relevant signal process (again...)
2. Chose n variables, i.e. E_{sum}, ϕ, r
   - That means we have 6 parameters (2 for each ROI bound)
   - n can be essentially any number, but computational requirements grow fast (Can be used to study optimal vertex-distance, E_{max}, E_{min}, 2D angle etc.)
3. For each possible combination of ROI calculate signal ε and $\mathcal{S}(\bar{b})^*$ as signal-to-background ratio $r(ROI) = \frac{\varepsilon}{\mathcal{S}(\bar{b})}$
   - 💀 **This takes way too long and grows way too fast, when we add new dimensions** 💀
   - 🍰 **We instead use more advanced optimization algorithms that cut down the computing time!** 🍰
4. Find which ROI maximizes $r(ROI)$ –> from the $max(r)$ calculate sensitvity (simulation)
5. Once you've measured data –> fit in ROI to get $\bar{b}$ –> calculate sensitivity
   - use ε from simulation, $\bar{b}$ from data

Extending to multiple dimensions:

Let's take the example methodology above, but use 2 variables for best ROI: E_{sum} and φ

Again, we look at the spectrum for signal and background (now a 2D histogram):

$E^l_{sum} =$ 2500, $E^u_{sum} =$ 3200, $\varphi^l =$ 20, $\varphi^u =$ 160



We can see that taking a 2D ROI (basically a data-cut), we can keep a lot of signal while reducing the amount of background. In this example however, it is not exactly clear which is the **best** choice!

This approach can be extended to many dimensions!

Example for $0\nu\beta\beta$ N-Dim:

- Choose n = 3, with variables φ, E_{sum}, r
  - ROI consists of 6 parameters: (E^l_{sum}, E^u_{sum}), (φ^l, φ^u) and (r^l, r^u), where l, u are the lower and upper boundaries of ROI
  - let's collectively call them Θ
- Choose optimization algorithm
  - there's many libraries written for this purpose, the goal is to **choose one that works well for the particular problem**
  - I tried multiple, the best results seem to be found with **Evolutionary Centers Algorithm** ECA[1]
- Then sensitivity is a function of Θ:

$$T^{1/2}(\Theta) = const. \frac{\varepsilon(\Theta)}{\mathcal{S}(\bar{b}(\Theta))}$$

[1]https://doi.org/10.1007/978-981-13-0860-4_6

Example for $0\nu\beta\beta$ N-Dim:

Basics of ECA approach

1. The algorithm starts with N randomly placed initial guesses Θ_i
2. Calculates the values of $T^{1/2}(\Theta_i)$ for each
3. Identifies "best performing" Θ_i
4. Creates new set of N around the best performers
5. Repeat until convergence (or failure...)

Example for $0\nu\beta\beta$ N-Dim:

$$\text{Best } \Theta =$$

variable	ROI
E_{sum}	(2710, 3350) keV
φ	$(10, 180)^\circ$
r	(0, 50) mm

$$\text{Resulting sensitivity}$$

variable	value
ε	0.145
$\bar{b}$	0.73
$T^{1/2}$	$\geq 4.22 yr$

2.2 Bayesian sensitivity

$$T^{1/2} = ln(2)\frac{N_A \cdot m \cdot t}{W}\frac{\varepsilon}{\mu_S}$$

where μ_S is the posterior 90% CI on the signal!

2.2 Bayesian sensitivity

The general idea

1. Simulate relevant signal and background processes
    - form **truth** pdfs on signal and background processes
2. Chose 1 variable which is tracked (i.e.E_{sum}, E_i, ϕ..)
3. Create appropriate model which describes the data
    - One of the model's parameters is n_S -> the expected number of signal events
4. Generate a pseudo-data from the simulated pdfs
5. Extract **pdf** of the signal counts **given data**: $pdf(n_S|data)$
6. Extract the 90% CI from $pdf(n_S|data)$ –> Calculate $T^{1/2}$
7. Repeat steps 4.-6. N times to mitigate statistical fluctuations and obtain **median sensitivity**
8. Profit 🍾🍾🍾

More detailed description in: DocDB#5943

Example, generic Bayesian inference

Let's say we have some Gaussian signal and Exponential background and have measured some data.

We have:

- μ_{sig} mean of signal Gaussian
- Θ_{sig} proportion of signal in data
- λ_{bkg} decay rate of background
- Θ_{bkg} proportion of background in data

- For simplicity, let's say we know μ_{sig} and λ_{bkg}.
- We just want to find: Θ_{sig}, Θ_{bkg}.

* let's ignore σ for now..

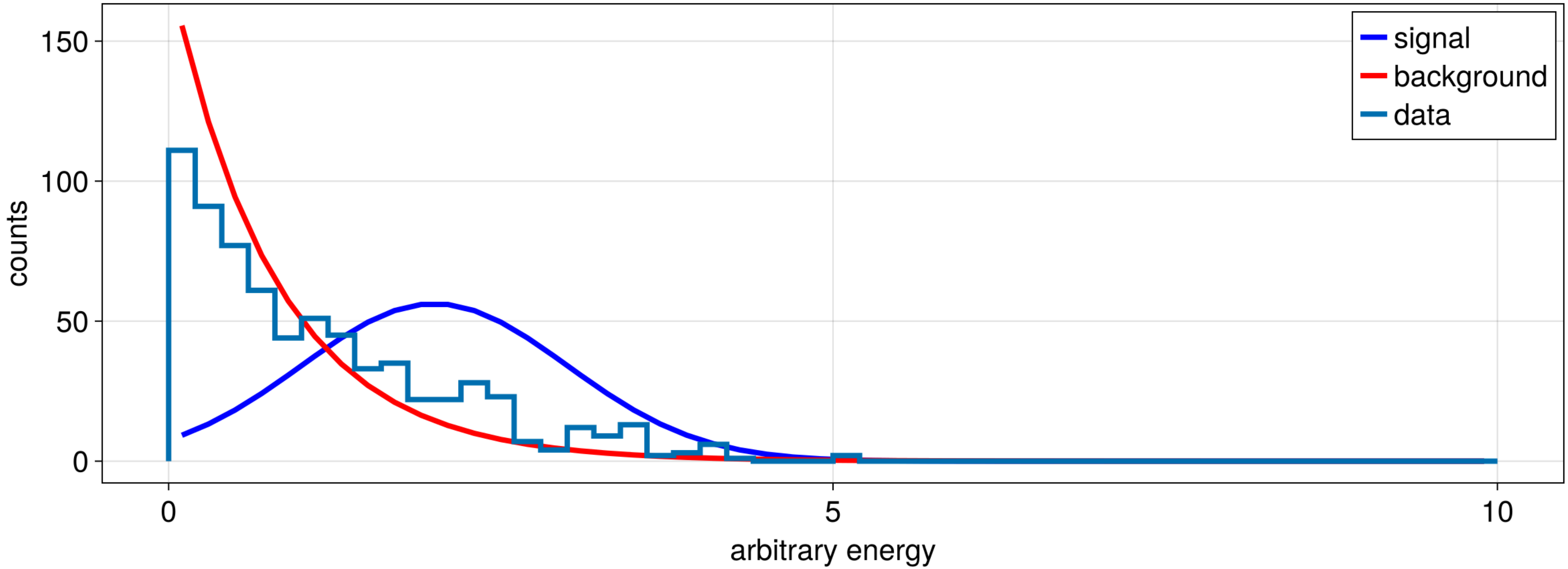
Example, generic Bayesian inference

Our likelihood:

$$\mathcal{L}(data|\Theta_{sig}, \Theta_{bkg}) = \prod_{i}^{N^{obs}} \left(\Theta_{bkg}\lambda_{bkg}e^{-\lambda_{bkg}E_i} + \Theta_{sig}\frac{1}{\sigma\sqrt{2\pi}}e^{-(\frac{E_i-\mu_{sig}}{\sigma})^2} \right)$$

Example experiment with n data-points:

n = 705



Parameters of the experiment:

μ_{sig} = 2.0

Θ_{sig} = 0.301

λ_{bkg} = 0.8

Θ_{bkg} = 1 - Θ_{sig} = 0.7

For a Bayesian inference we must chose the so-called **prior** (describes our knowledge of the paramaters)

$$pdf(\Theta) \sim \text{some distribution}$$

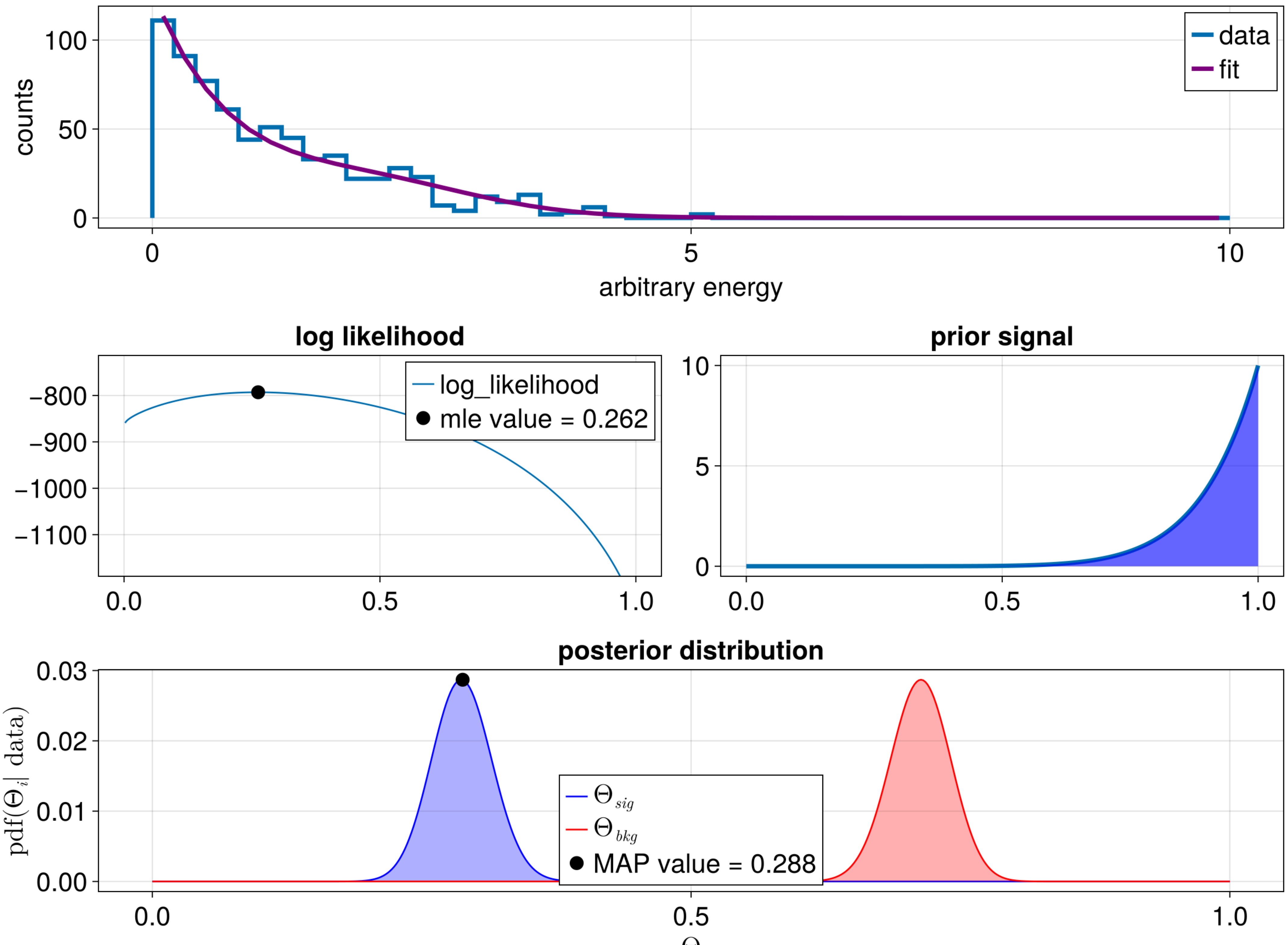
What we really want as the result is the so-called **posterior**: **The distribution of the parameter given data!**

Since our parameter of interest is Θ_{sig}:

$$pdf(\Theta_{sig}|data) \propto \mathcal{L}(data|\Theta_{sig}, \Theta_{bkg}) \cdot pdf(\Theta_{sig})$$

p1 = 10.0

p2 = 1.0



counts
arbitrary energy
data
fit
log likelihood
log_likelihood
mle value = 0.262
prior signal
posterior distribution
pdf(Θ_i| data)
Θ_{sig}
Θ_{bkg}
MAP value = 0.288

Example for $0\nu\beta\beta$ and E_{sum}:

What was...bad!

In DocDB#5943 I showed **general model Exponential background + gaussian signal**

$$\mathcal{L}(data|n_S, n_B, \lambda) = \frac{1}{n_S + n_B} \prod_{i}^{N^{obs}} \left(n_B \lambda e^{-\lambda E_i} + n_S \frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{E_i - \mu}{\sigma})^2} \right)$$

Where $\mathcal{L}(data|n_S, n_B, \lambda)$ is the likelihhod of data **given** parameters, n_S, n_B, λ are expected number of signal, background counts and decay constant. N_{obs} are the observed events.

This description works for most experiments **that are not SuperNEMO!!**

- Our ROI is very wide:
  - cannot be described well with and exponential, or flat (as is the case for CUPID and Gerda)
- Our Signal is not gaussian:
  - due to energy losses in gas, the signal is more Landau shaped!

Example for $0\nu\beta\beta$ and E_{sum}:

What is...better! 🥁

Detailed likelihood model with each process having its own pdf!

$$\mathcal{L}(data|\vec{\Theta}) = \frac{1}{\sum \Theta_i} \prod_i^{N^{obs}} \left(\Theta_1 pdf(sig, E_i) + \sum_{j=2}^{n_{pars}} \Theta_j pdf(bkg_j, E_i) \right)$$

Where $pdf(sig, E_i)$ and $pdf(bkg_j, E_i)$ are the signal and background normalized spectra evaluated at E_i.

- This way we have better description of the shapes of the spectra (if we consider simulation to be correct...)

Priors and posetrior

Priors

- the parameters are set up as a "proportion of the total spectrum", so each process contributes some percentage of the total spectrum: $\Sigma_i \Theta_i = 1$
- signal $p(\Theta_1|data) \sim Uniform(0, 10^{-4})$
    - 10^{-4} represents roughly $T^{1/2} \geq 10^{24} yr$ for $17.5 kg.y$

- backgrounds $\Theta_i \sim Uniform(0, 1)$

The only condition is that $\sum \Theta_i = 1.0$.

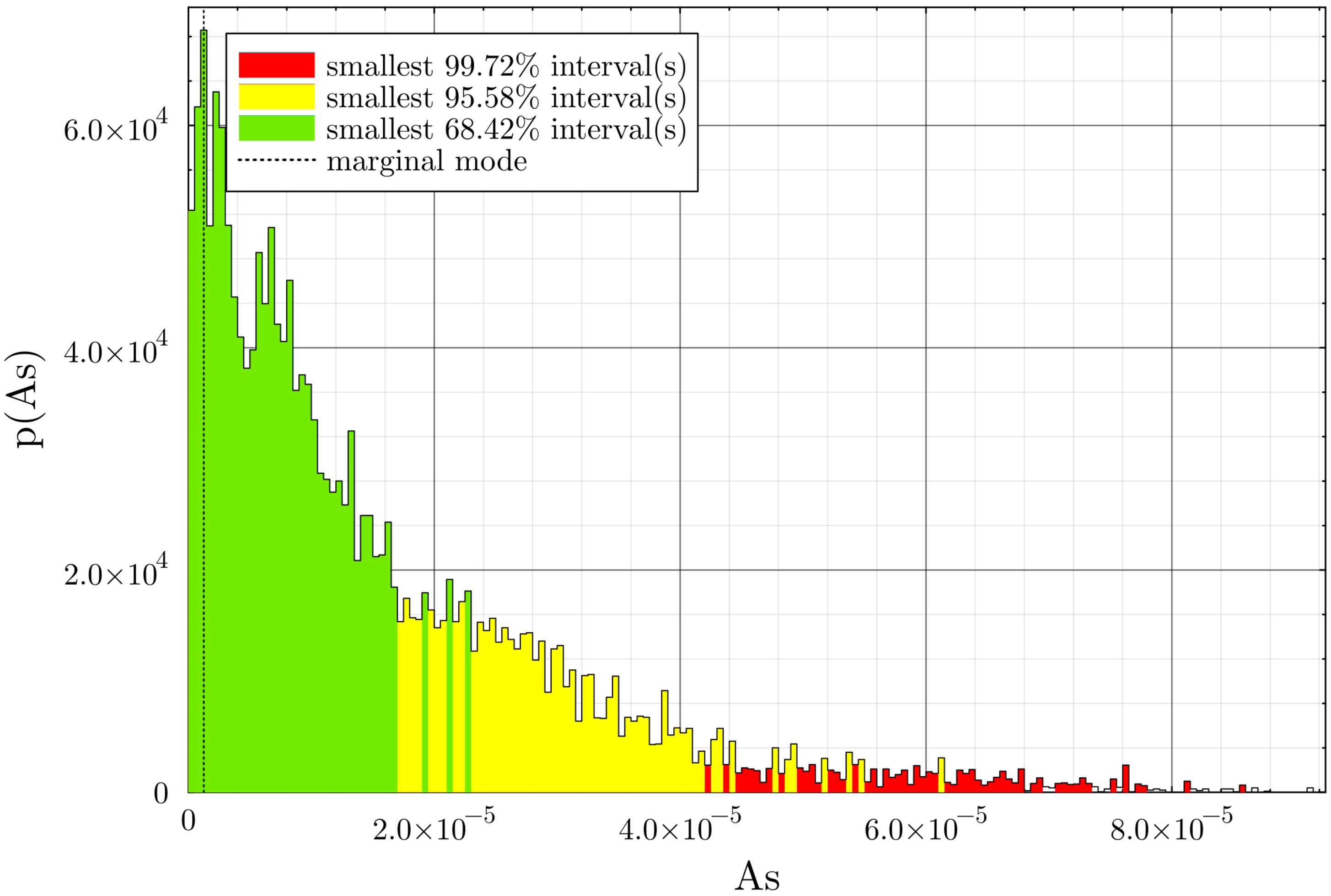
Example for $0\nu\beta\beta$ and E_{sum}:

Posterior

The posterior distributions represent the **possible distributions of the parameter** given data. That means, that for each parameter (proportion of the spectrum), we get a pdf. For example, if $0\nu\beta\beta$ represented say 30% of our measured data, we'd get a pdf with peak around 30%. If there's no signal present, the peak would be expected at 0. The shape of the pdf is how we calculate confidence interval.

Example for $0\nu\beta\beta$ and E_{sum}:

Results: For a single Bayesian inference we get posterior distribution of Θ_{sig}:



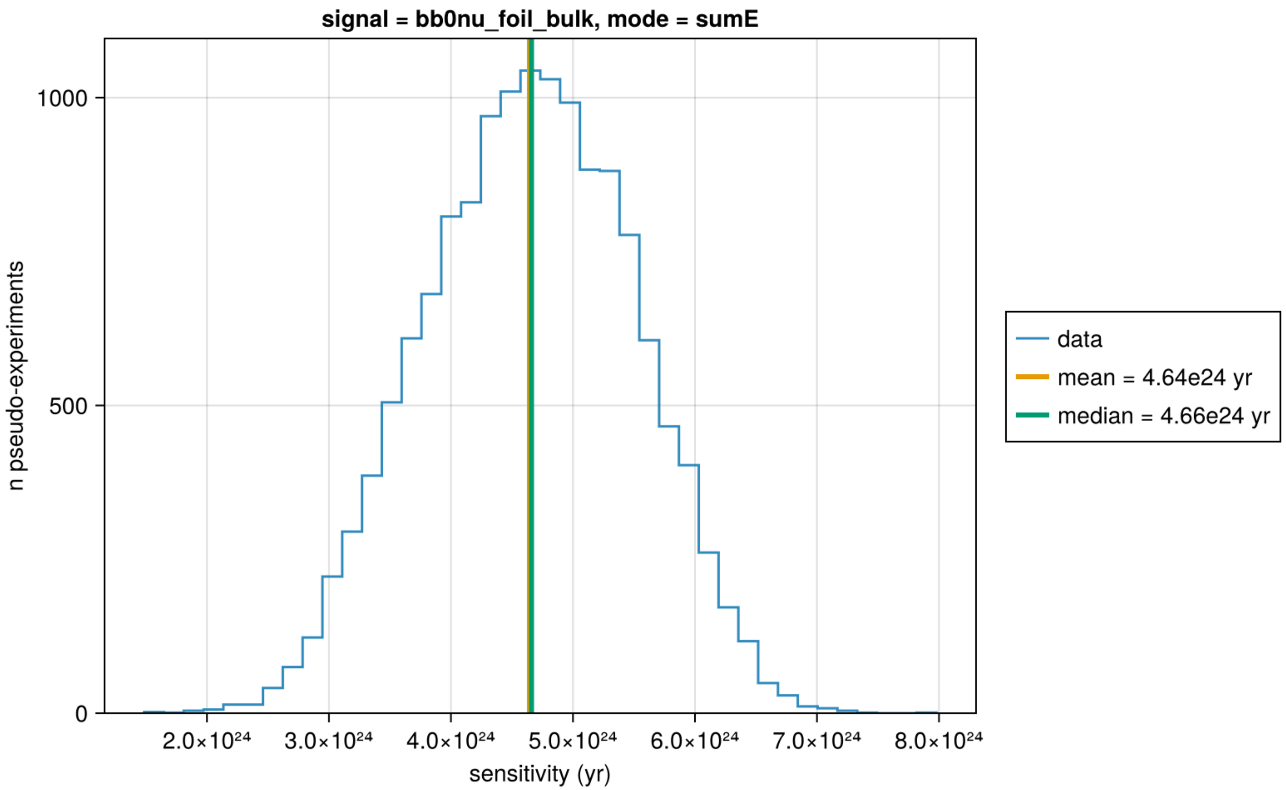
This means, the most probable value for Θ_{sig} is very close to zero, with 90% CI around 3×10^{-5}, which corresponds to 3.5 events. This is the 90% CI on the number of signal events that should be plugged into the sensitivity equation.

Example for $0\nu\beta\beta$ and E_{sum}:

Danger!

The Bayesian inference method shown can fluctuate a lot! Both due to the statistical nature of the Bayesian MCMC sampling and due to how the data fluctuates (this is true for Frequentist as well). To limit the effects, we create **n pseudo experiments** to gather a large enough statistics for drawing conclusions.

Example for $0\nu\beta\beta$ and E_{sum}



I ran this algorithm in 20 parallel jobs on CC-LYON cluster for 2 days each. The following steps were within each job:

1. For each iteration **generate** pseudo-data based on the background model
2. Perform bayesian inference and extract **Sensitivity**
3. Repeat for 2 days
4. Save all sensitivities into a file

The resulting median sensitivity for $0\nu\beta\beta$ is:

$$T_{1/2}^{0\nu} \geq 4.66 \times 10^2 4yr$$

Results

Numbers in parenthasis are with **manually** added "total neutrons in given ROI" (without performing a dedicated ND/Bayes analysis)

signal	1D	ND	Bayes
$0\nu\beta\beta$	$4.14 \times 10^{24} y$	$4.22(3.20) \times 10^{24} y$	$4.66 \times 10^{24} y$
$0\nu\beta\beta\chi^0$	$1.45 \times 10^{23} y$	$1.48(1.45) \times 10^{23} y$	$2.38 \times 10^{23} y$
$0\nu\beta\beta\chi^0\chi^0$	$2.31 \times 10^{22} y$	$2.32(2.32) \times 10^{22} y$	$1.43 \times 10^{22} y$
$\nu_L\nu_R\beta\beta$	$1.30 \times 10^{22} y^*$	$1.30 \times 10^{22} y$	$1.09 \times 10^{21} y^{**}$

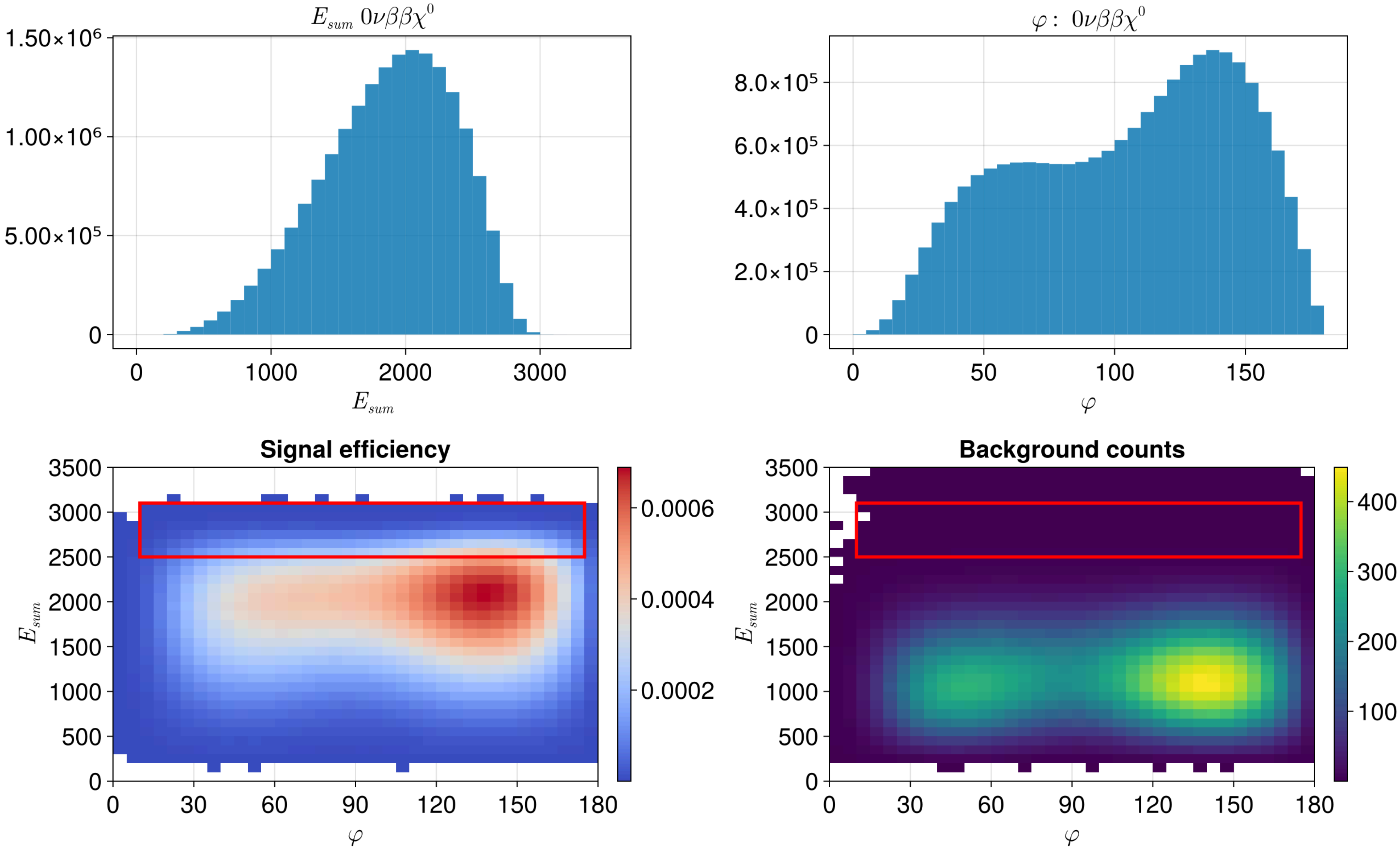
* for angular distribution

** for angular distribution, need to investigate why this is so low.

Results: $0\nu\beta\beta\chi^0$

ROI: $\varphi \in (5, 175)^\circ, E_{sum} \in (2500, 3100) keV, r \in (0, 50) mm$

$T^{1/2} \geq 1.48 \times 10^{23}$ y; $\varepsilon = 0.017$; $\bar{b} = 25.944\ (+1.2)$

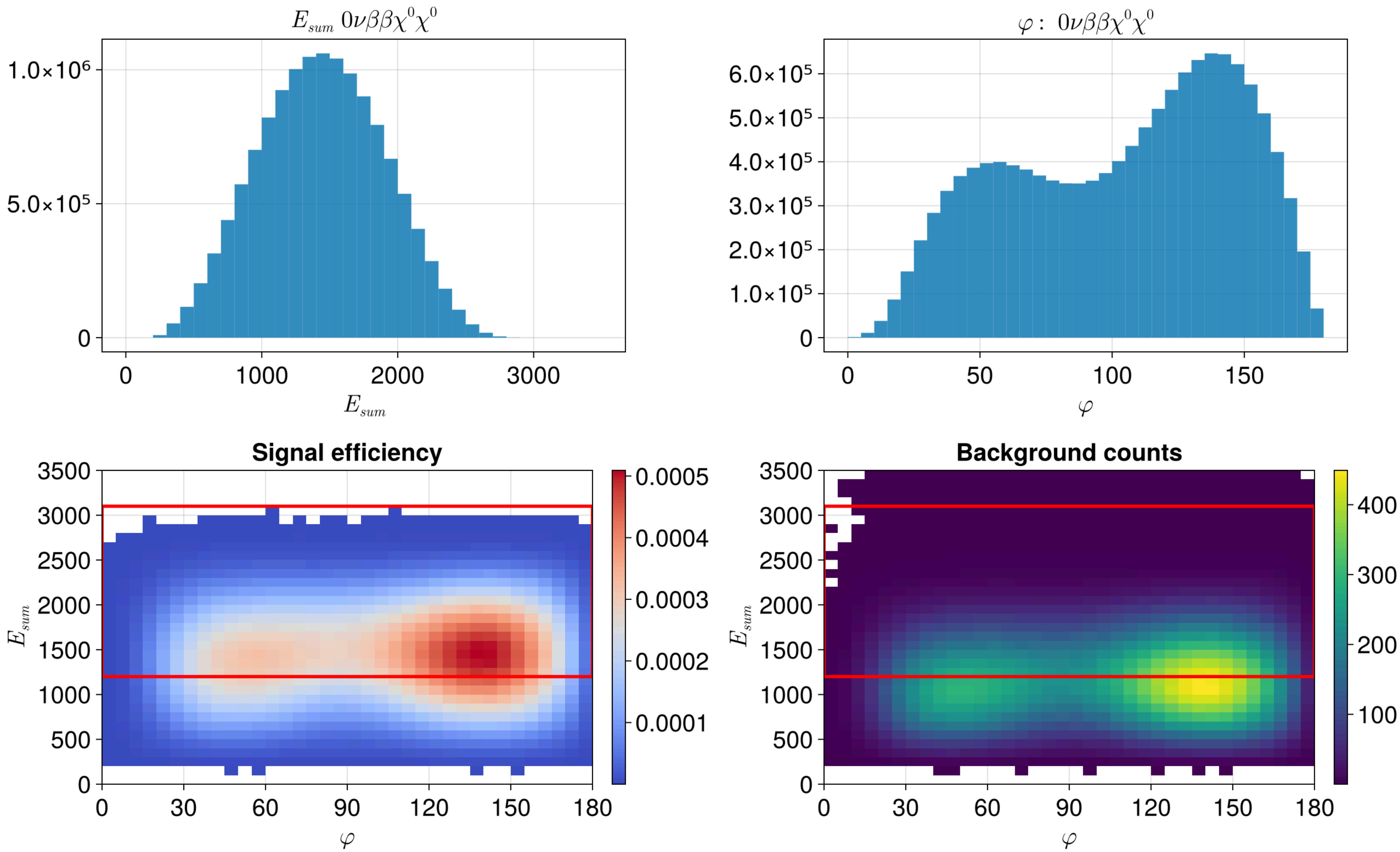


E_{sum} $0\nu\beta\beta\chi^0$
$\varphi:$ $0\nu\beta\beta\chi^0$
E_{sum}
φ
Signal efficiency
Background counts
E_{sum}
φ

Results: $0\nu\beta\beta\chi^0\chi^0$

ROI: $\varphi \in (0, 180)^\circ, E_{sum} \in (1200, 3100) keV, r \in (0, 50) mm$

$T^{1/2} \geq 2.32 \times 10^{22}$ y; $\varepsilon = 0.09$; $\bar{b} = 44880.87\ (+3.66)$

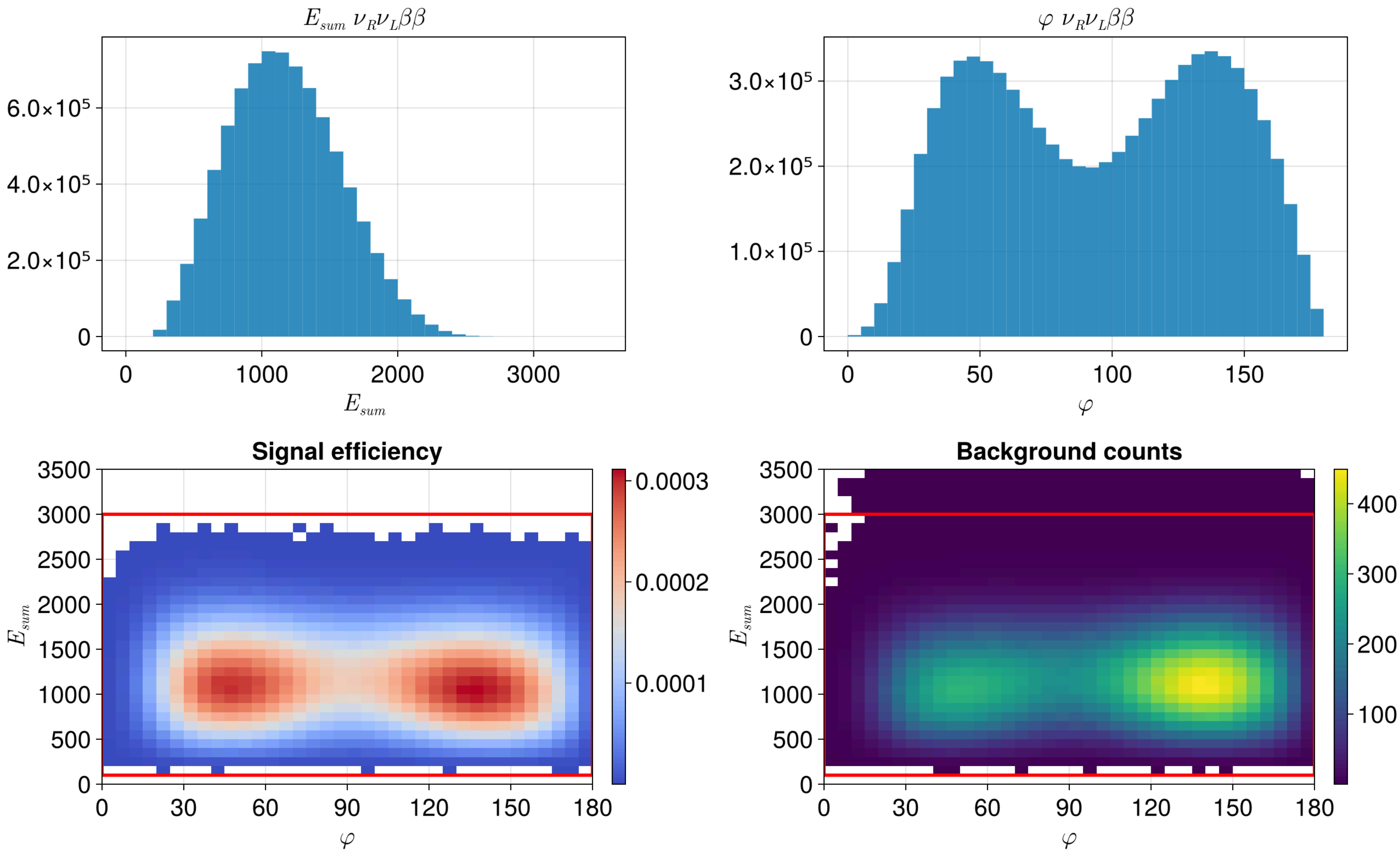


E_{sum} $0\nu\beta\beta\chi^0\chi^0$
$\varphi:$ $0\nu\beta\beta\chi^0\chi^0$
E_{sum}
φ
Signal efficiency
Background counts
E_{sum}
φ

Results: $\nu_R \nu_L \beta\beta$

ROI: $\varphi \in (5, 175)^\circ, E_{sum} \in (100, 3000) keV, r \in (0, 50) mm$. **Basically take everything?**

$T^{1/2} \geq 1.30 \times 10^{22}$ y; $\varepsilon = 0.08$; $\bar{b} = 115520.50$



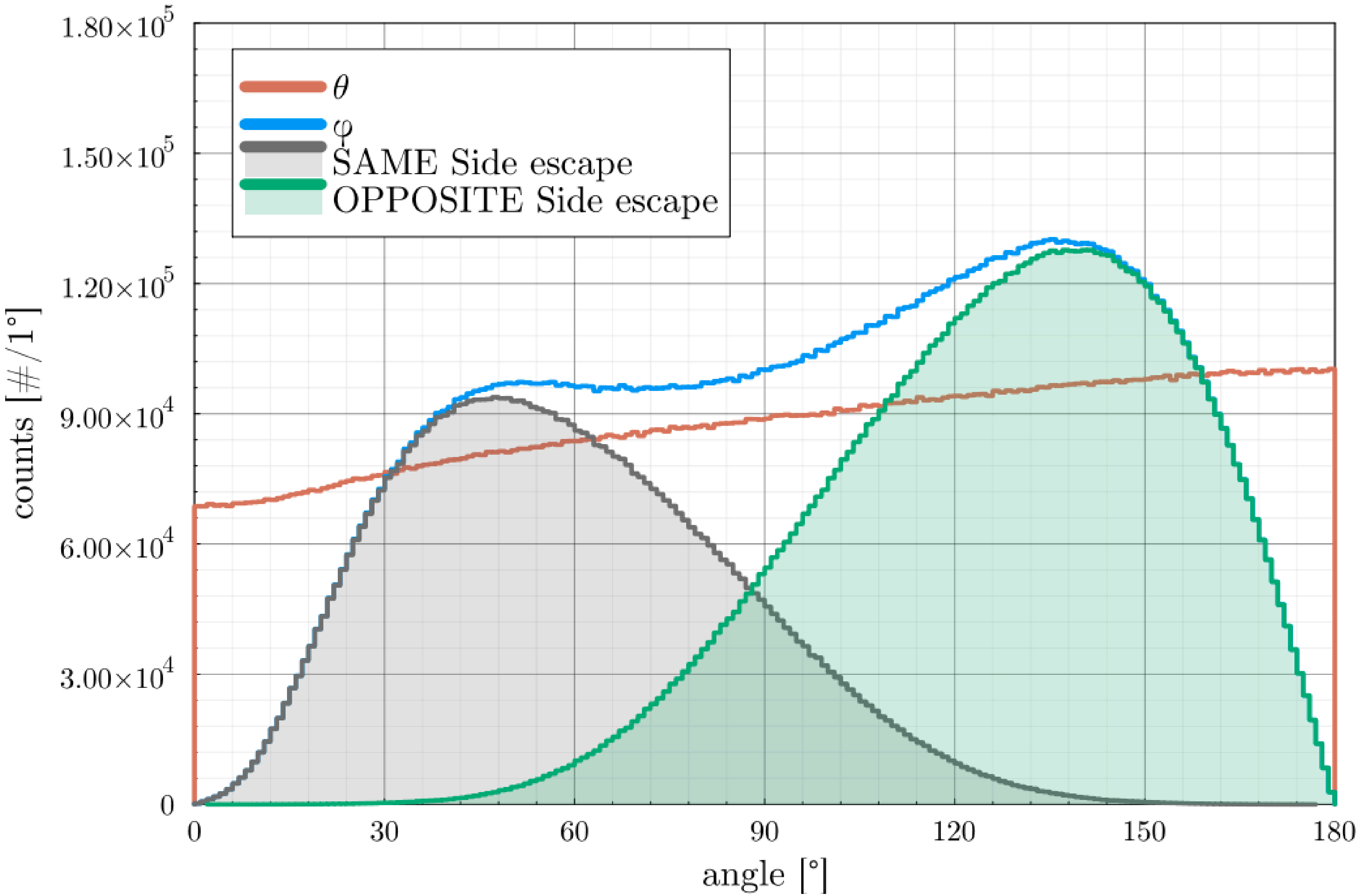
E_{sum} $\nu_R\nu_L\beta\beta$
6.0×10⁵
4.0×10⁵
2.0×10⁵
0
0
1000
2000
3000
E_{sum}
φ $\nu_R\nu_L\beta\beta$
3.0×10⁵
2.0×10⁵
1.0×10⁵
0
0
50
100
150
φ
Signal efficiency
3500
3000
2500
2000
1500
1000
500
0
E_{sum}
0
30
60
90
120
150
180
φ
0.0003
0.0002
0.0001
Background counts
3500
3000
2500
2000
1500
1000
500
0
E_{sum}
0
30
60
90
120
150
180
φ
400
300
200
100

Conclusions

- Using n-dimensional approach (sometimes) leads to *slight* increase in sensitivities
  - there are issues that need investigation
  - better optimization should improve the results even further
    - The choice of data-cut on $r < 50mm$ **is way too strict!**
    - Will try to optimize ToF values as well
  - Choice of variables can be extended, at the cost of computing time
  - Machine learning techniques for MVA could improve it even further
- Bayesian approach is more sensitive than frequentist when signal shape is more different from backgrounds!
  - not sure what's wrong with RH spectra, but this could be it?
- Need to implement neutron data
- We have an idea to use data-partitioning for angular distribution

Future

Based on research of angular distribution for the electrons, depending on which side of foil they are emitted, the reconstruction is differently efficient. We could use this fact!




```
Activating project at `~/Work/Phd/SNSensitivityEstimate`
```