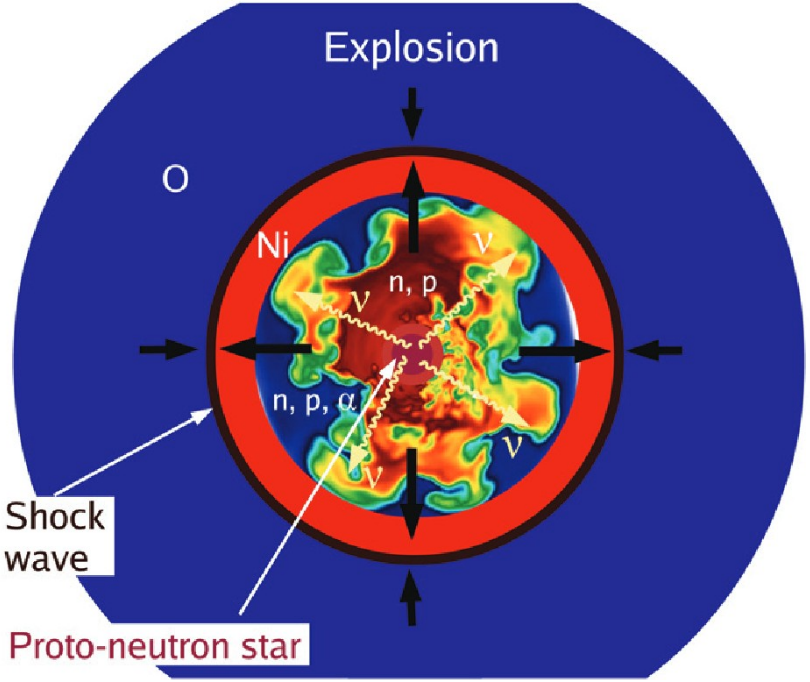
Unfolding of the supernova burst neutrino spectrum

Thilo Birkenfeld
Physics Institute III B, RWTH Aachen University
Europe Meeting, 24.10.2022

Core-Collapse Supernova



- Core collapse supernovae produce a short (~**10s**), strong (**L~10^{53} erg/s**) neutrino burst of **all flavor**
- Neutronization

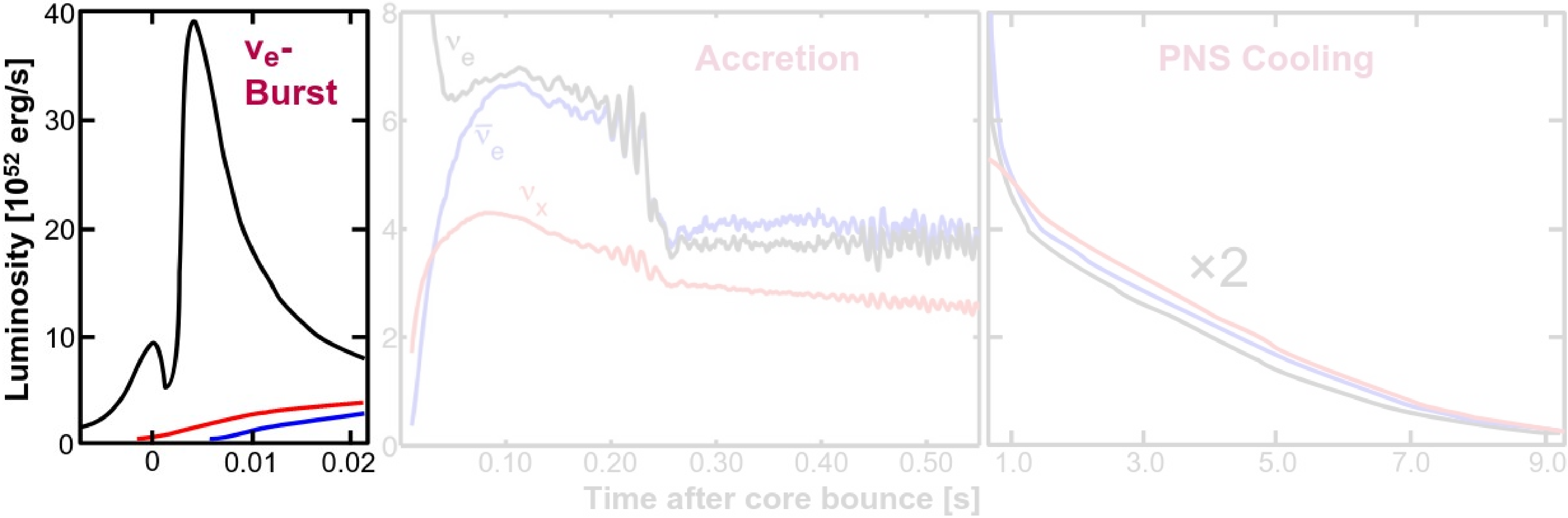
$$e^- + p \to n + \nu_e$$

- Pair-production

$$e^- + e^+ \to \nu_f + \bar{\nu}_f$$

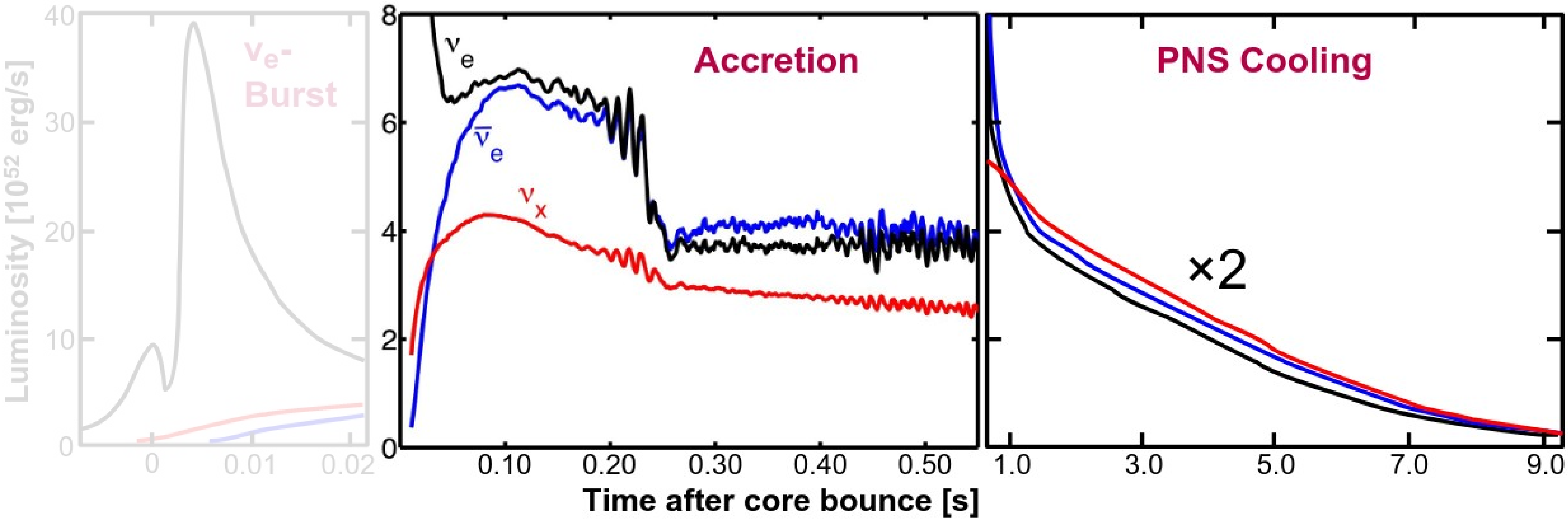
- And more …
- $\langle E_\nu \rangle \sim 10 - 15\,\mathrm{MeV}$

Neutronization burst



- Within ~ ms a neutronization burst emits a large number of electron neutrinos
  - Caused by shock wave becoming transparent for neutrinos
  - Highly **non-thermal** spectrum

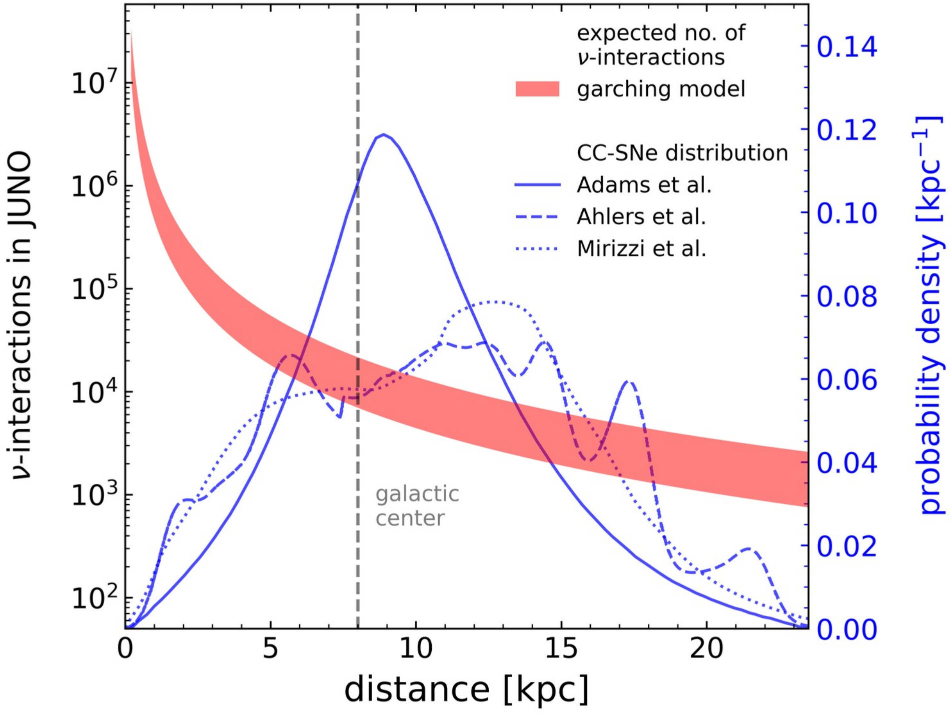
Accretion and cooling



- Accretion: Luminosity and $\langle E_\nu \rangle$ rapidly changing
- Cooling phase: **thermal spectra** with decreasing temperature

Cumulative Spectra

- @ 10 kpc distance
  ~ O(10^4) Events



- MSW leads to flavor conversion
  - NH

$$\begin{bmatrix} F_{\nu_e} \\ F_{\bar{\nu}_e} \\ F_{\nu_x} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \cos^2\theta_{12} & \sin^2\theta_{12} \\ \frac{1}{4} & \frac{1}{4}\sin^2\theta_{12} & \frac{1}{4}(2+\cos^2\theta_{12}) \end{bmatrix} \begin{bmatrix} F^0_{\nu_e} \\ F^0_{\bar{\nu}_e} \\ F^0_{\nu_x} \end{bmatrix}$$

  - IH

$$\begin{bmatrix} F_{\nu_e} \\ F_{\bar{\nu}_e} \\ F_{\nu_x} \end{bmatrix} = \begin{bmatrix} \sin^2\theta_{12} & 0 & \cos^2\theta_2 \\ 0 & 0 & 1 \\ \frac{1}{4}\cos^2\theta_{12} & \frac{1}{4} & \frac{1}{4}(2+\sin^2\theta_{12}) \end{bmatrix} \begin{bmatrix} F^0_{\nu_e} \\ F^0_{\bar{\nu}_e} \\ F^0_{\nu_x} \end{bmatrix}$$

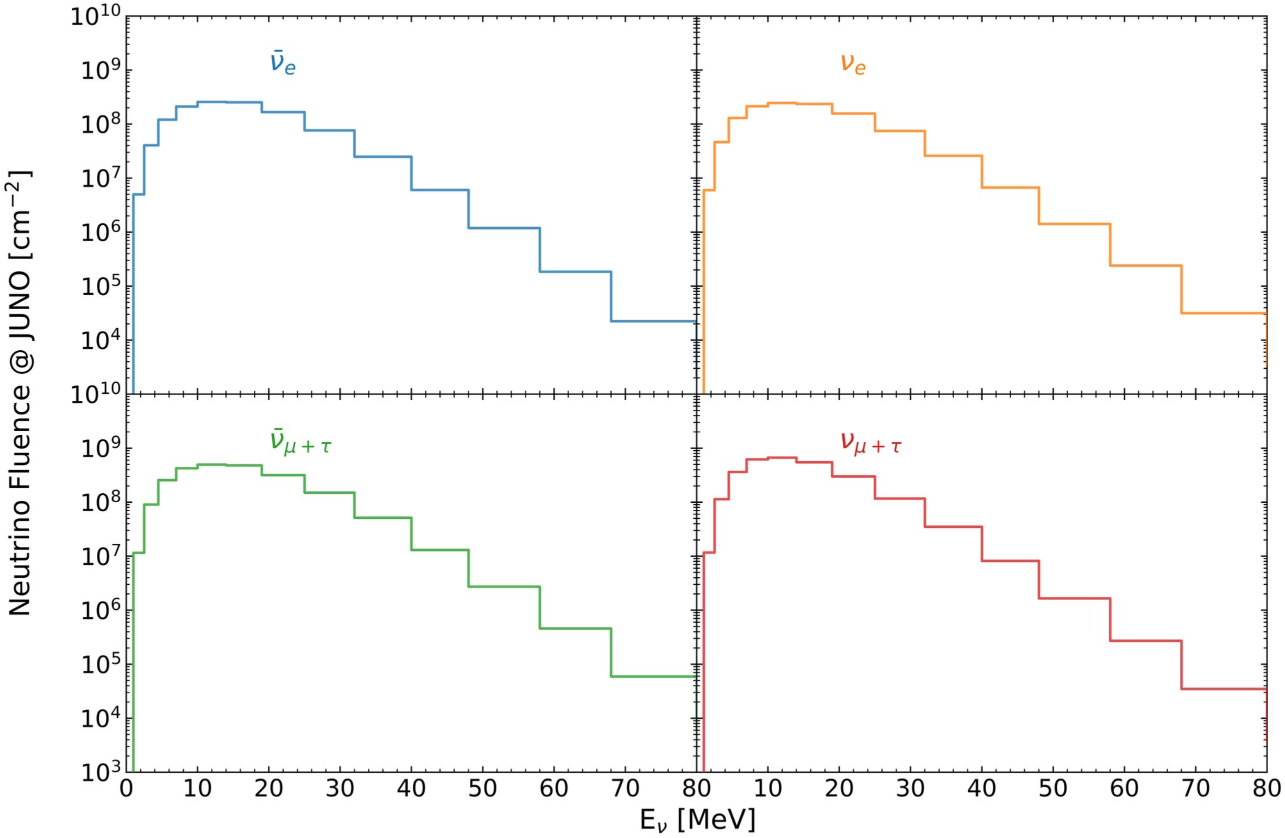
- Time integrated ν_{f}-spectra are mixture of different compositions (thermal and non-thermal)
  - Dependence on details of the supernova

Motivation

- Model-independent reconstruction of spectra

→ Unfolding

Example: 11 M⊙ SN @ 10 kpc distance



Event Typology

Supernova Neutrinos in JUNO

- Multiple channels
  - Inverse Beta Decay (IBD)
    - $e^+ + n$, $\Delta t \sim 220\,\mu s$
  - Proton elastic scattering (PES)
  - Electron elastic scattering (EES)
  - ^{12}C – interactions
    - $^{12}C^*$ excited state $\rightarrow \gamma$ (15.1 MeV)
    - $^{12}N/^{12}B$ production
      - $e^{-(+)} + e^{+(-)}$, $\Delta t \sim 10\,ms$

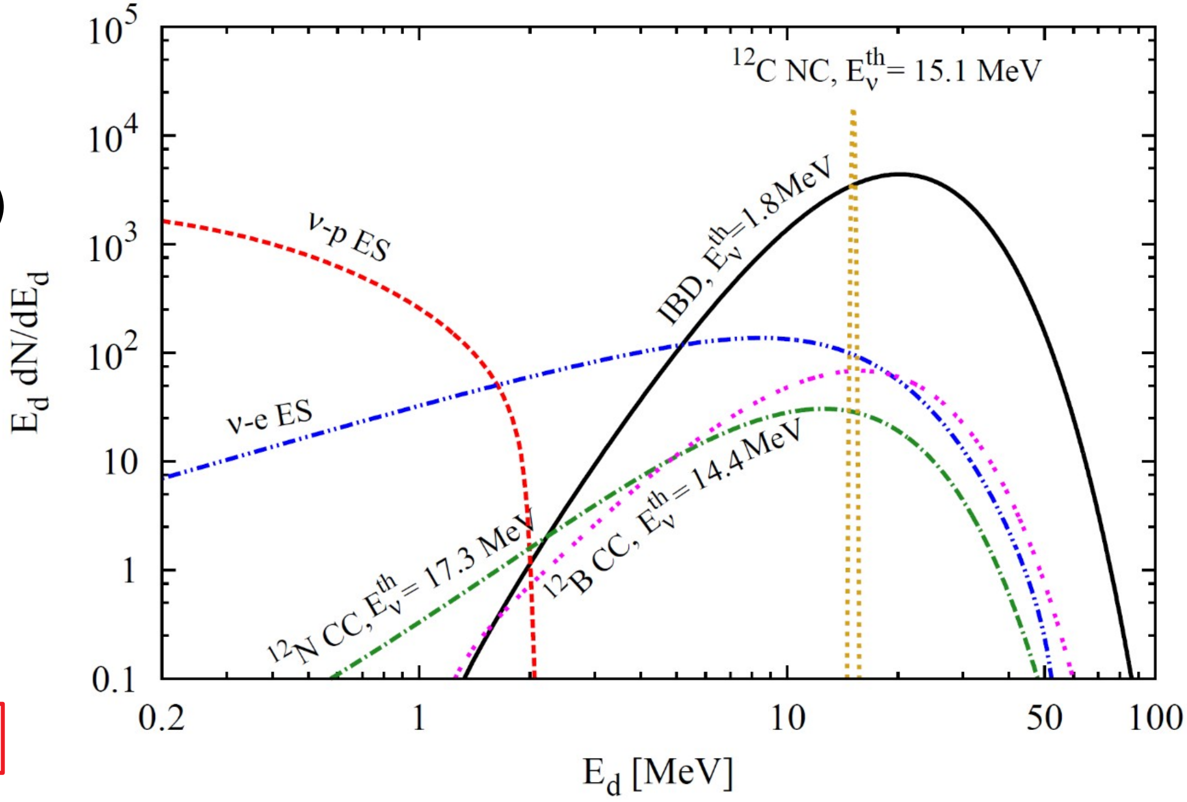


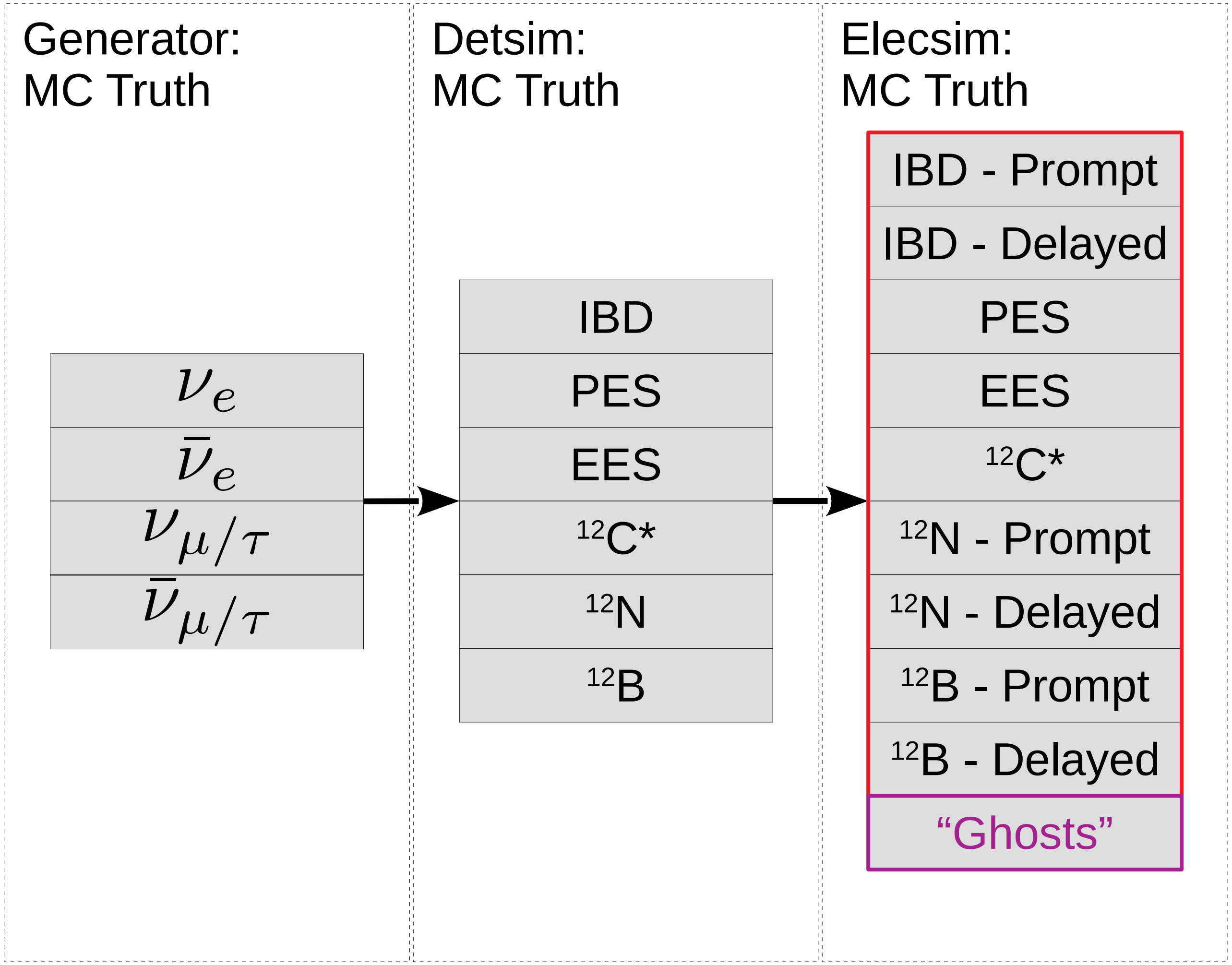
Fig 1. Expected energy spectrum for 10 kpc core-collapse Supernova

$$\sigma_f^{channel} \neq \sigma_{f'}^{channel} \quad \text{in general}$$

Channel identification required for multi flavor analysis

MC Truth

- What is MC truth type information? (Without cluster algorithm)



Generator:
MC Truth

- ν_e
- $\bar{\nu}_e$
- $\nu_{\mu/\tau}$
- $\bar{\nu}_{\mu/\tau}$

Detsim:
MC Truth

- IBD
- PES
- EES
- ^{12}C*
- ^{12}N
- ^{12}B

Elecsim:
MC Truth

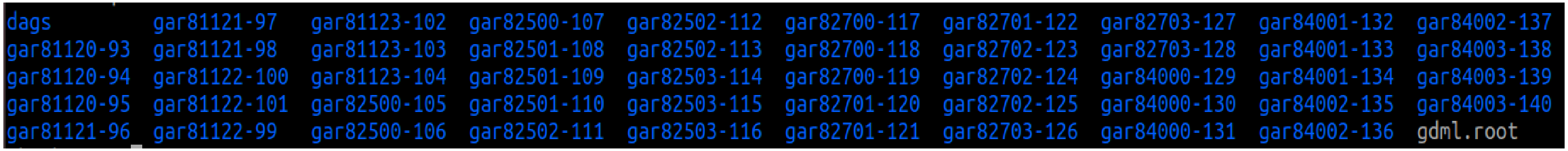
- IBD - Prompt
- IBD - Delayed
- PES
- EES
- ^{12}C*
- ^{12}N - Prompt
- ^{12}N - Delayed
- ^{12}B - Prompt
- ^{12}B - Delayed
- “Ghosts”

- Timing
- Secondary particle type
- Each readout window will be assigned to biggest contributor (1 MC Truth)
- “Ghosts” from **afterpulses** or splitted events
- **For event selection**

Event Selection

Data Set

- 48 CC-SN @ 10 kpc



```
dags         gar81121-97   gar81123-102  gar82500-107  gar82502-112  gar82700-117  gar82701-122  gar82703-127  gar84001-132  gar84002-137
gar81120-93  gar81121-98   gar81123-103  gar82501-108  gar82502-113  gar82700-118  gar82702-123  gar82703-128  gar84001-133  gar84003-138
gar81120-94  gar81122-100  gar81123-104  gar82501-109  gar82503-114  gar82700-119  gar82702-124  gar84000-129  gar84001-134  gar84003-139
gar81120-95  gar81122-101  gar82500-105  gar82501-110  gar82503-115  gar82701-120  gar82702-125  gar84000-130  gar84002-135  gar84003-140
gar81121-96  gar81122-99   gar82500-106  gar82502-111  gar82503-116  gar82701-121  gar82703-126  gar84000-131  gar84002-136  gdml.root
```

- 3 different seed for each model

- ~ 500’000 events

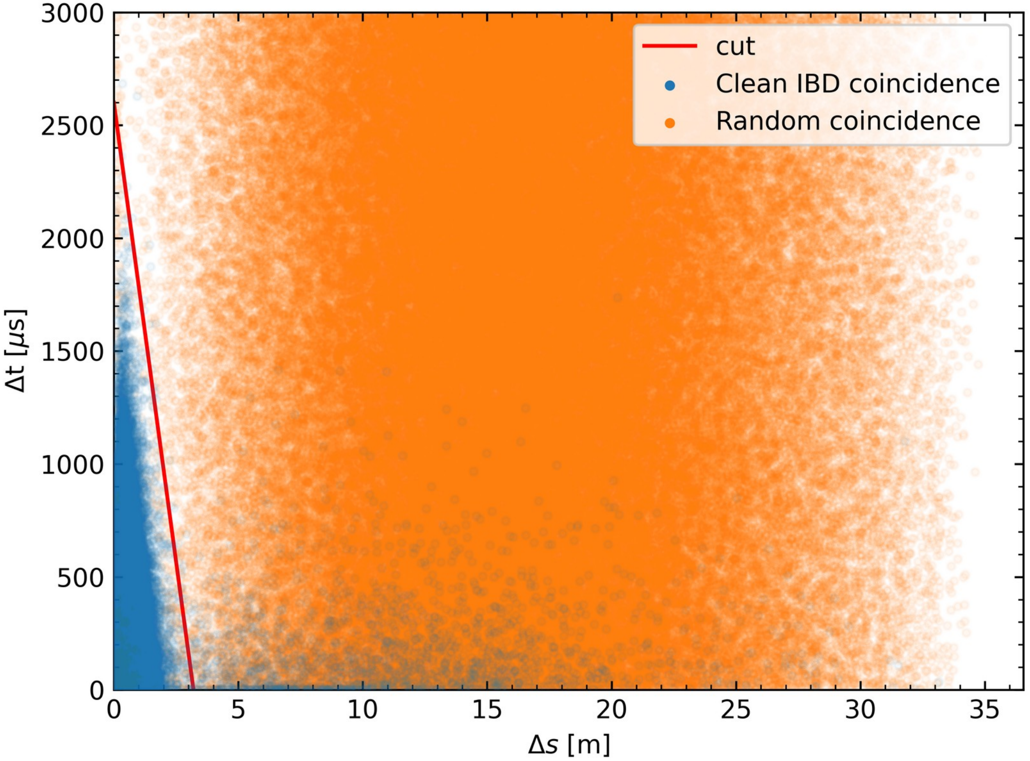
- **Full Chain** (including elecsim)

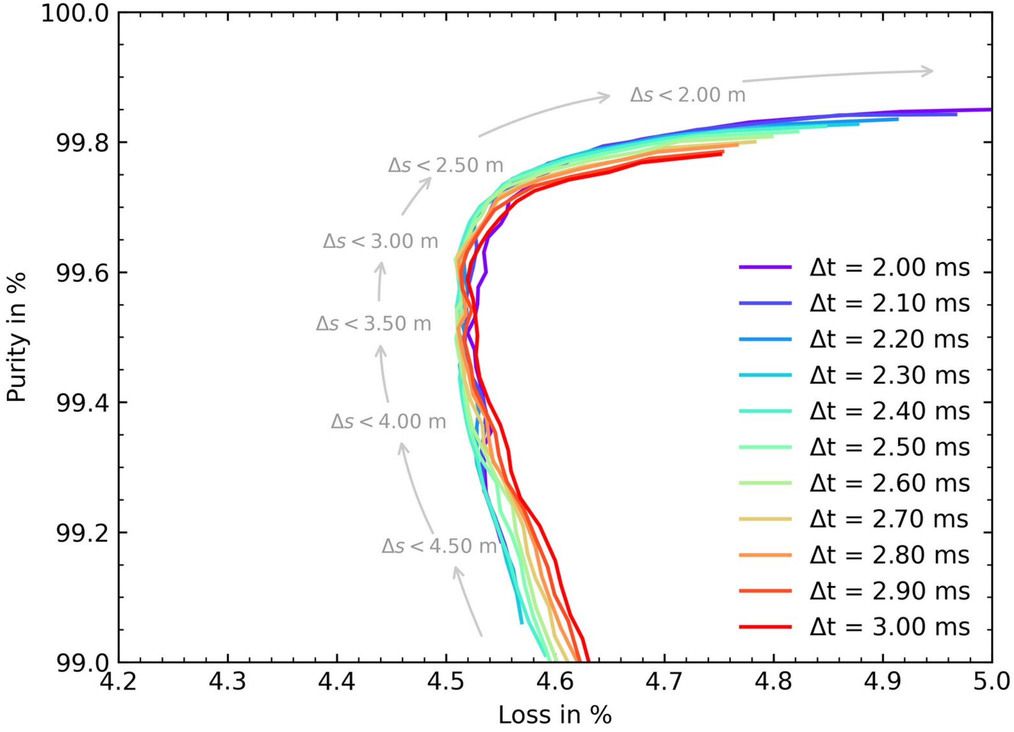
IBD selection

- Delayed energy: $E_{max} < 5.25\,\mathrm{MeV}$
- Prompt energy: $E_{min} > 1.0\,\mathrm{MeV}$

Purity = # correct selection / # total selection
Loss = # missed events / # all events

- Optimized by minimal loss → (loss will contaminate other channels)

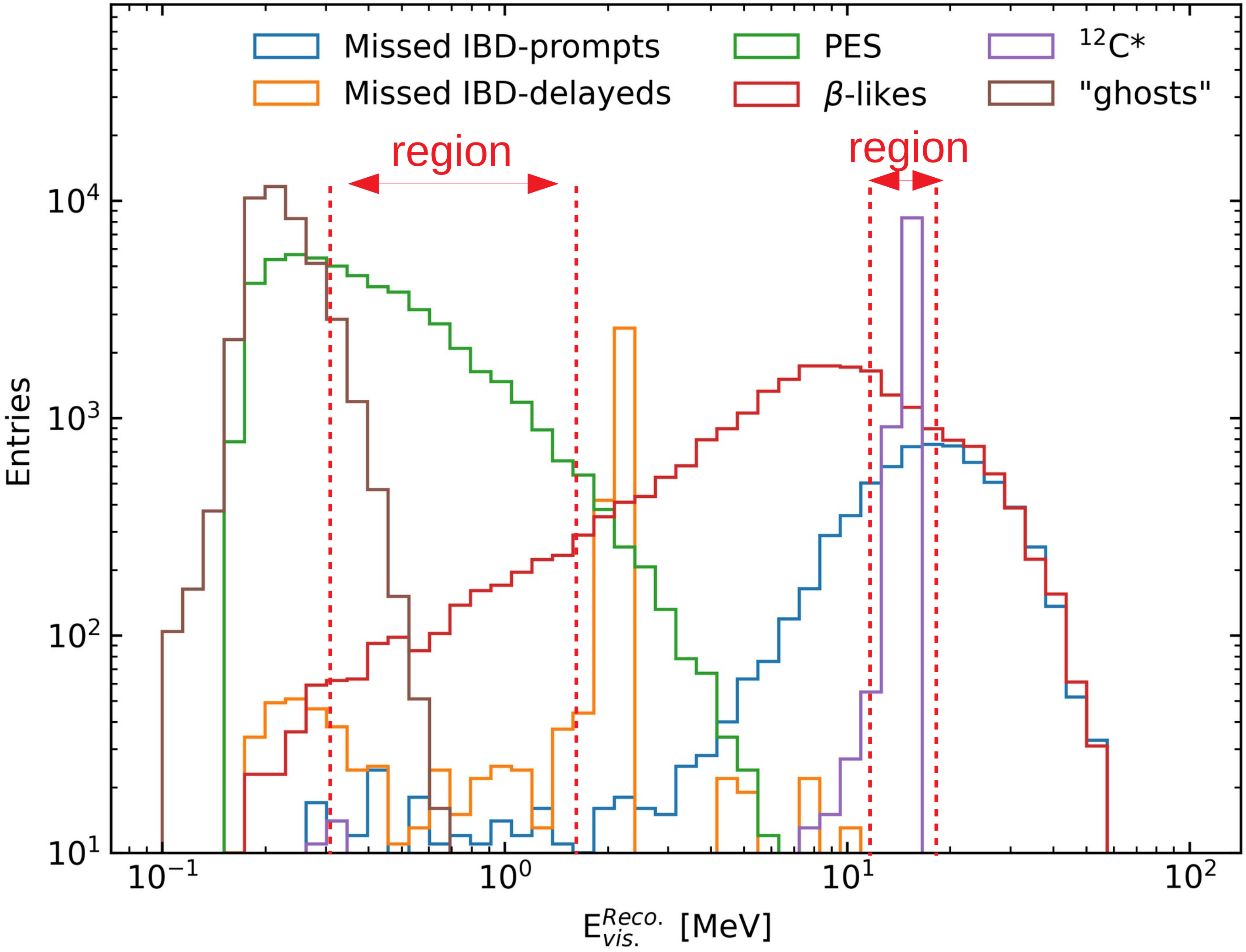




Distribution after IBD Selection

- Select PES and $^{12}C^*$ events from energy region

- Cut optimized by generic trade-off between loss and purity



β-like and "Ghosts"

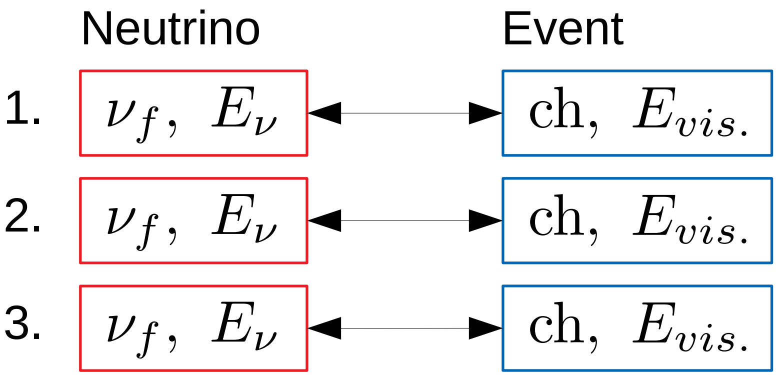
- "ghosts" are in lowest energy region
  - Everything below min. proton energy regarded as "ghosts"

- Everything is selected except β-likes
  - Left over is β-like channel

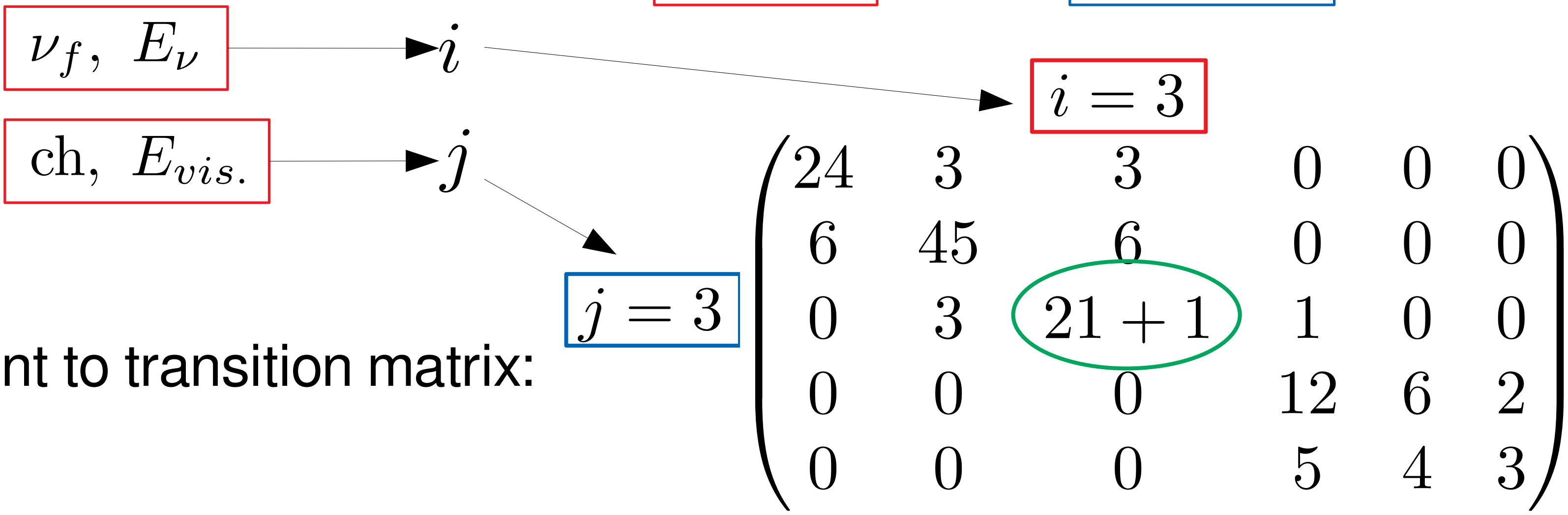
Spectra Unfolding

Response Matrix Construction

- With MC Truth definition
- With neutrino flavor simulation



- Get bin IDs:



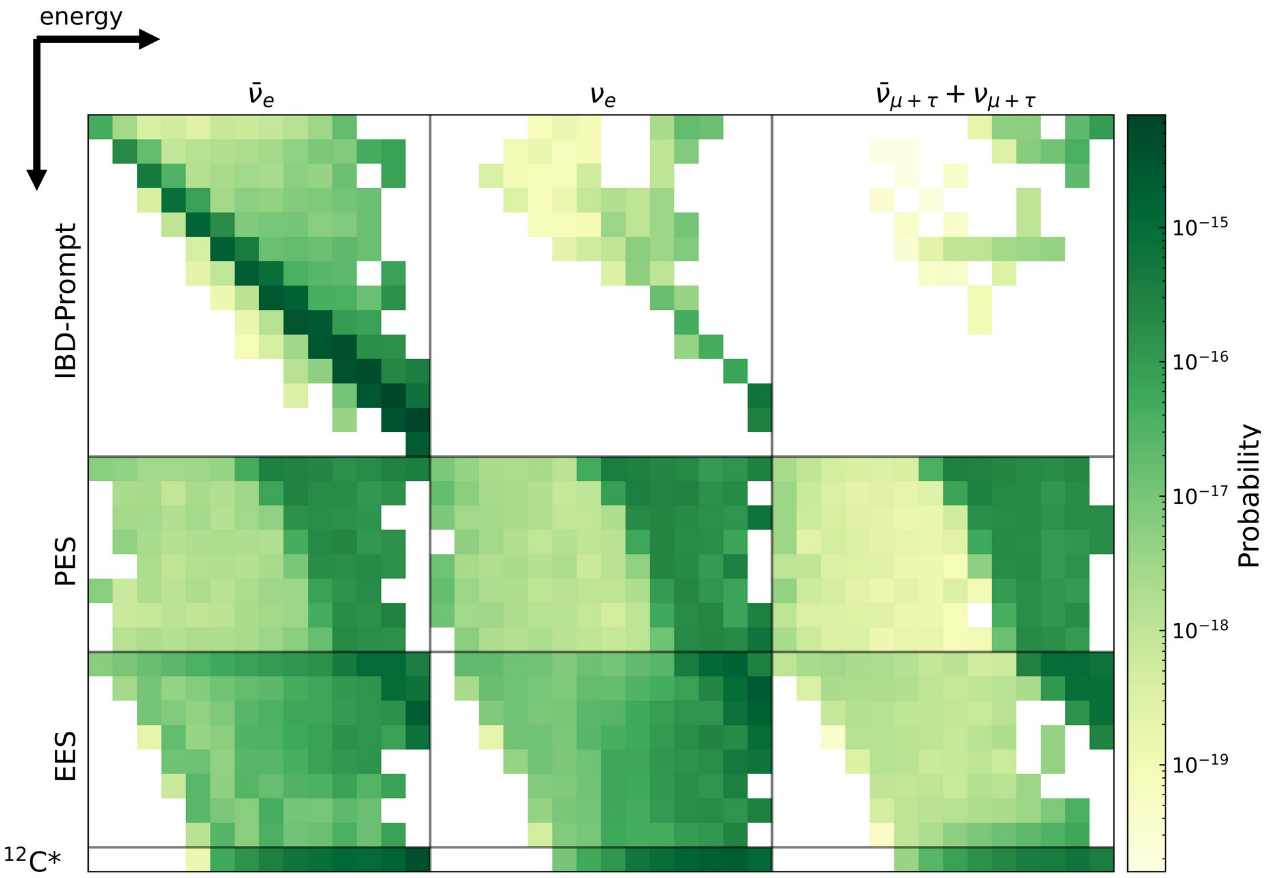
- Add each event to transition matrix:

$$\begin{pmatrix} 24 & 3 & 3 & 0 & 0 & 0 \\ 6 & 45 & 6 & 0 & 0 & 0 \\ 0 & 3 & 21+1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 12 & 6 & 2 \\ 0 & 0 & 0 & 5 & 4 & 3 \end{pmatrix}$$

- Normalize each column by no. ν's that **pass through JUNO** → probabilities (M)

Direct Unfolding

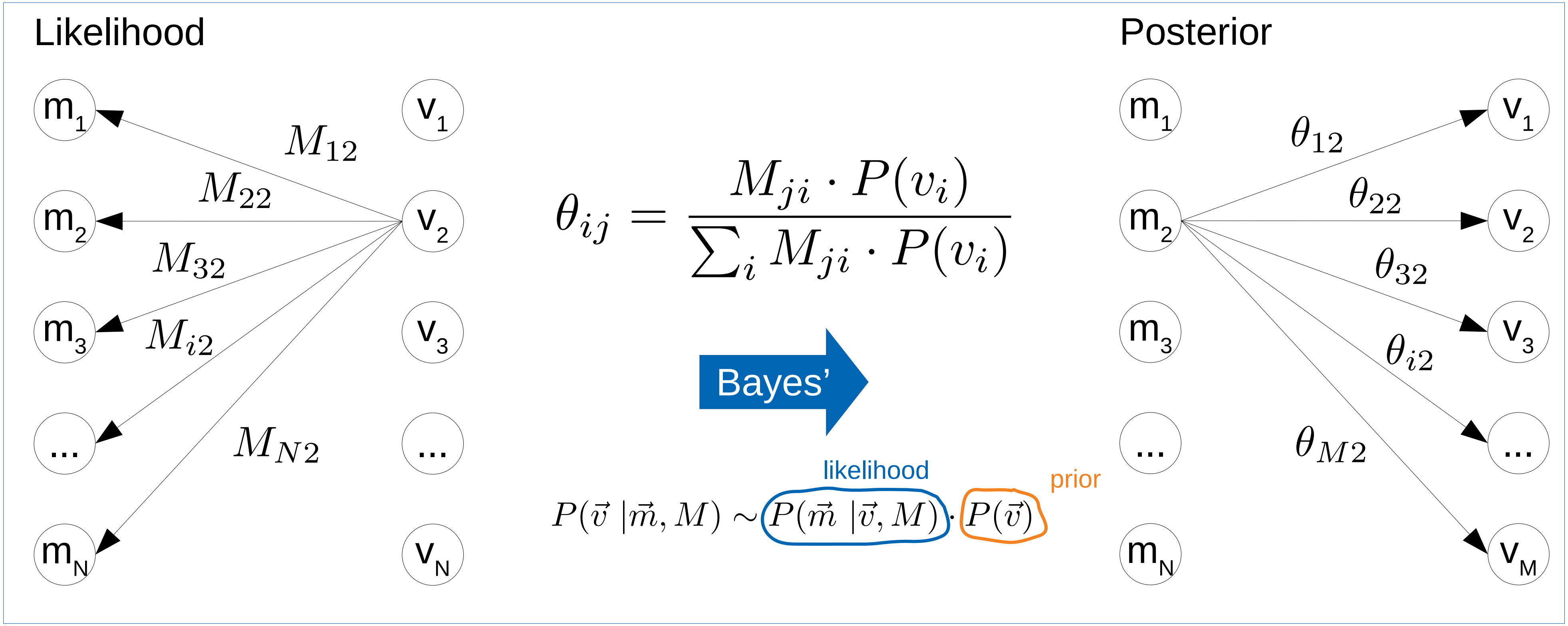
Response Matrix



- **Important**:

  Physical spectrum is original neutrino energy distribution
- Includes cross-section, trigger and selection efficiencies and energy smearing

Inference Step



$$\theta_{ij} = \frac{M_{ji} \cdot P(v_i)}{\sum_i M_{ji} \cdot P(v_i)}$$

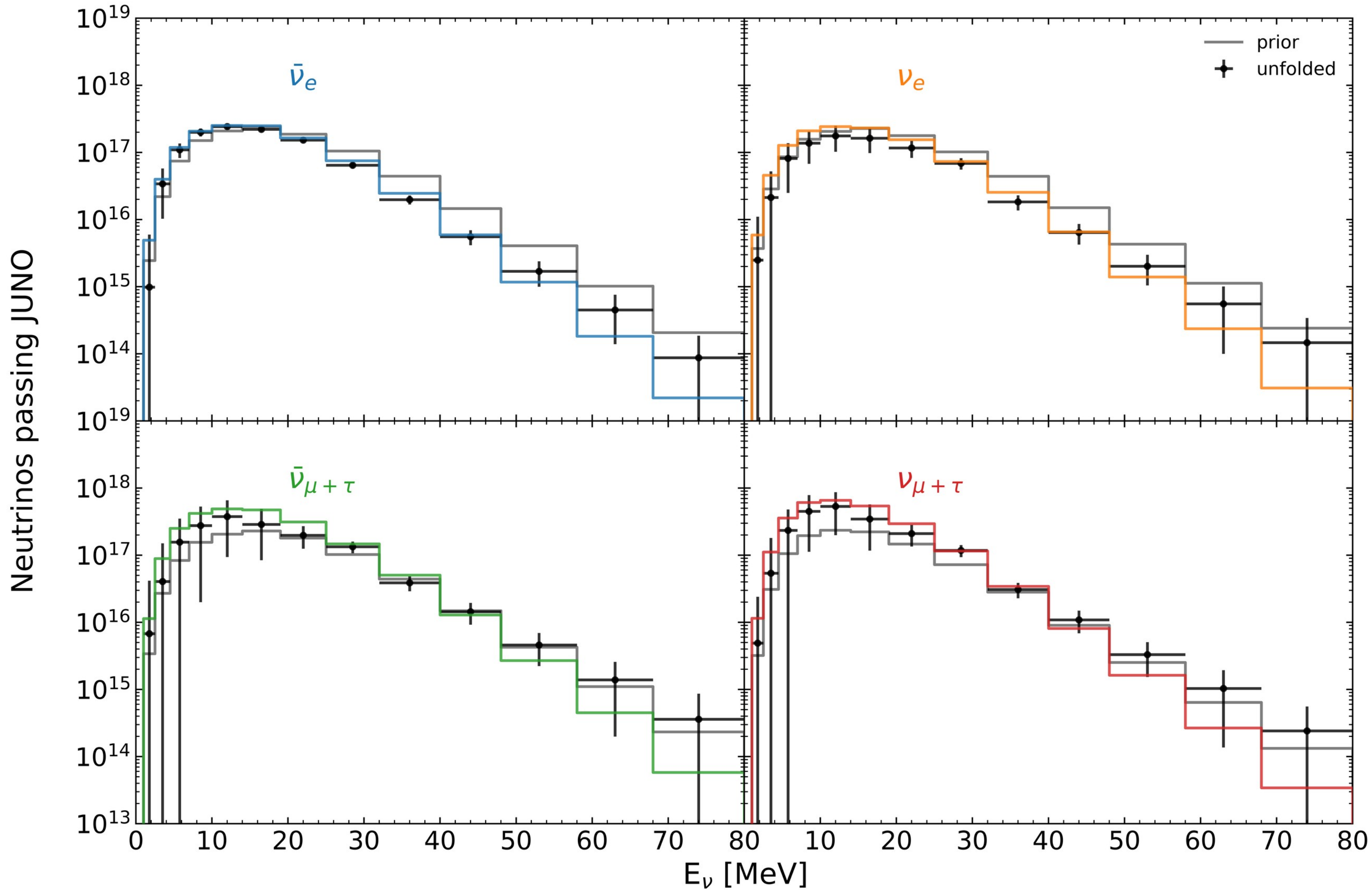
$$P(\vec{v}\,|\vec{m}, M) \sim \overset{\text{likelihood}}{\boxed{P(\vec{m}\,|\vec{v}, M)}} \cdot \overset{\text{prior}}{\boxed{P(\vec{v})}}$$

Full physical spectrum:

$$\boxed{\tilde{v}_i \approx \frac{1}{\epsilon_i} \sum_{j=1}^{n_{\vec{m}}} \theta_{ij} \cdot m_j}$$

$$\epsilon_i = \text{efficiency}$$

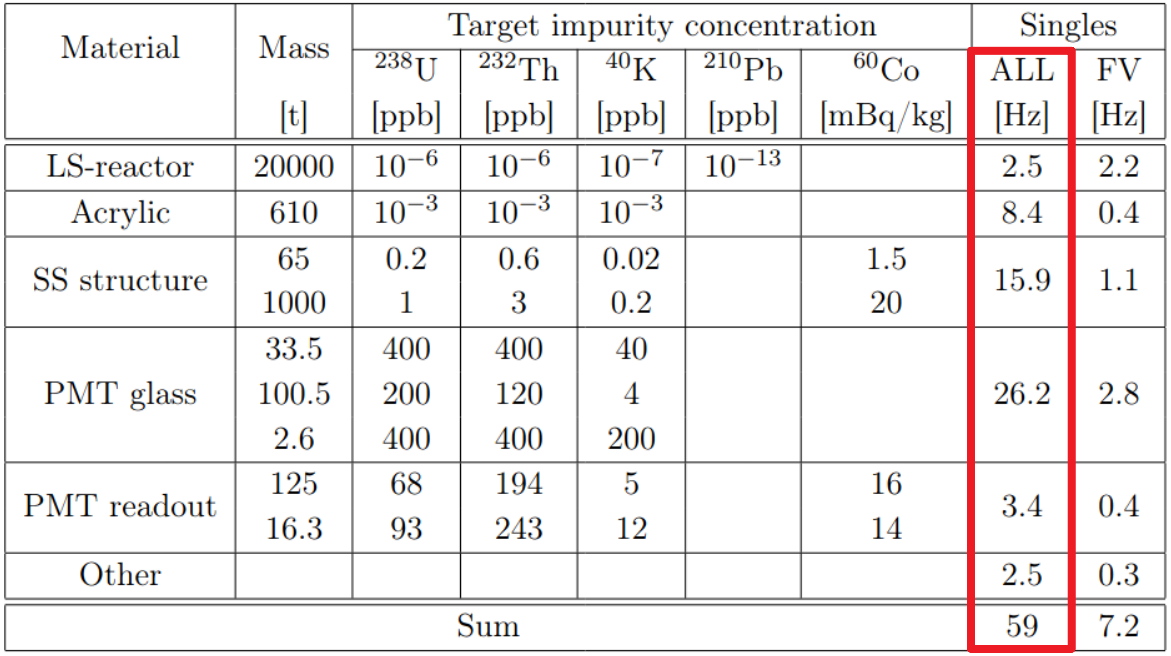
Example Unfolding



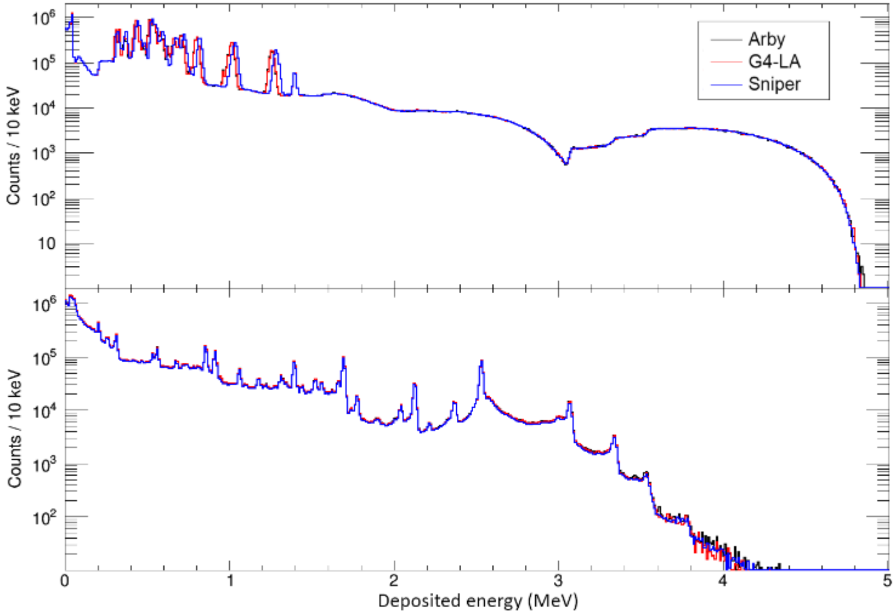
Backgrounds

Single backgrounds

https://juno.ihep.ac.cn/cgi-bin/Dev_DocDB/ShowDocument?docid=7057



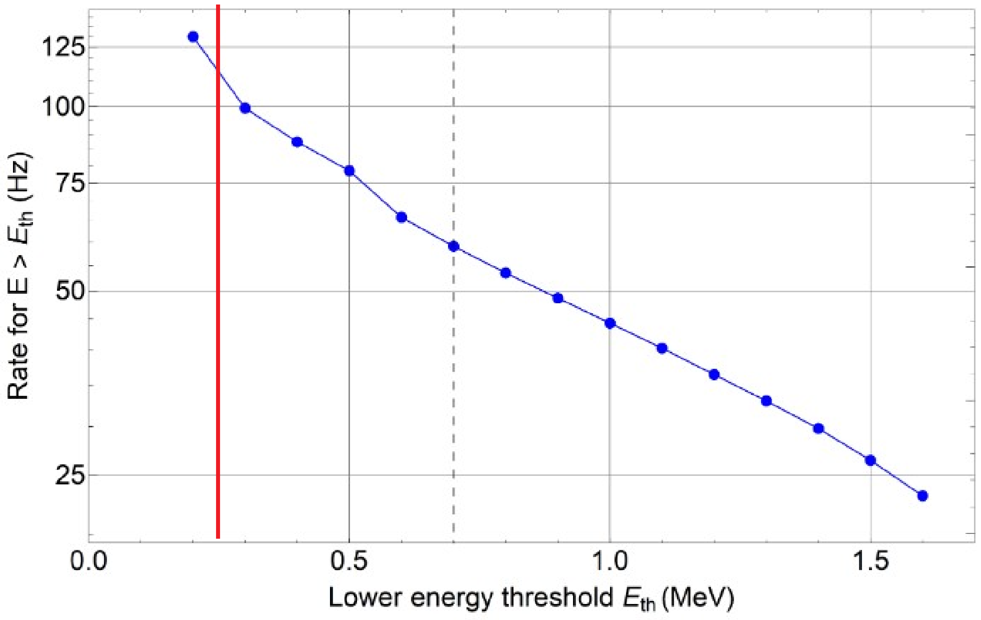
Material	Mass	Target impurity concentration					Singles	
		^{238}U	^{232}Th	^{40}K	^{210}Pb	^{60}Co	ALL	FV
	[t]	[ppb]	[ppb]	[ppb]	[ppb]	[mBq/kg]	[Hz]	[Hz]
LS-reactor	20000	10^{-6}	10^{-6}	10^{-7}	10^{-13}		2.5	2.2
Acrylic	610	10^{-3}	10^{-3}	10^{-3}			8.4	0.4
SS structure	65 1000	0.2 1	0.6 3	0.02 0.2		1.5 20	15.9	1.1
PMT glass	33.5 100.5 2.6	400 200 400	400 120 400	40 4 200			26.2	2.8
PMT readout	125 16.3	68 93	194 243	5 12		16 14	3.4	0.4
Other							2.5	0.3
Sum							59	7.2



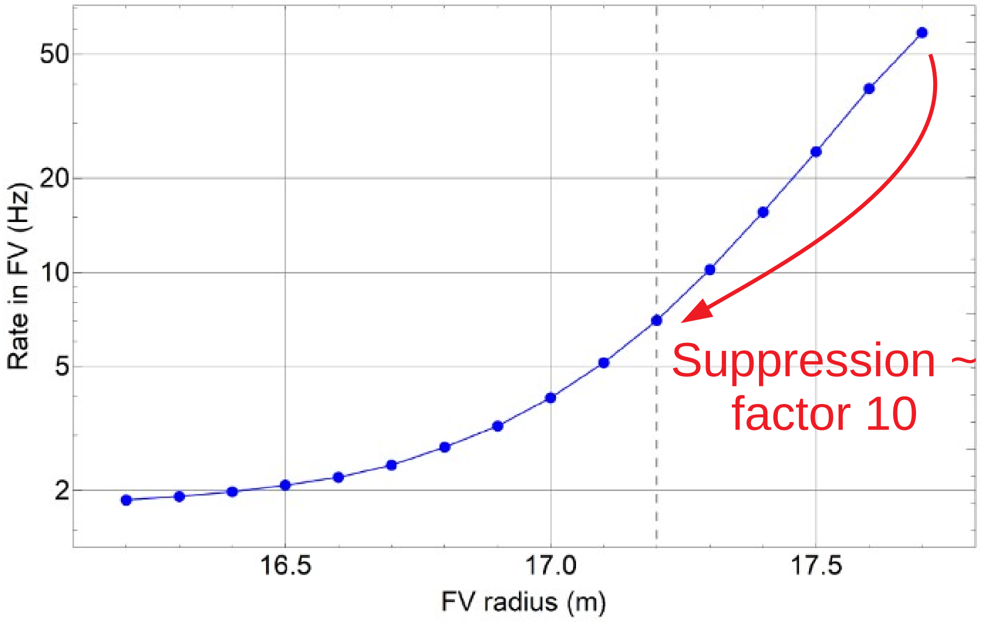
- Without any cuts ~ 590 singles in triggered (~ 10 second SN signal)
- BUT: This is with a Energy threshold of 700 keV

Energy & Fiducial cuts

1st Cut (Energy)



2nd Cut (Fiducial Volume)



Suppression ~ factor 10

- With my energy cut and FV cut → ~110 events in 10 secs.
- But loose ~ 4% data (with vertex reco.)

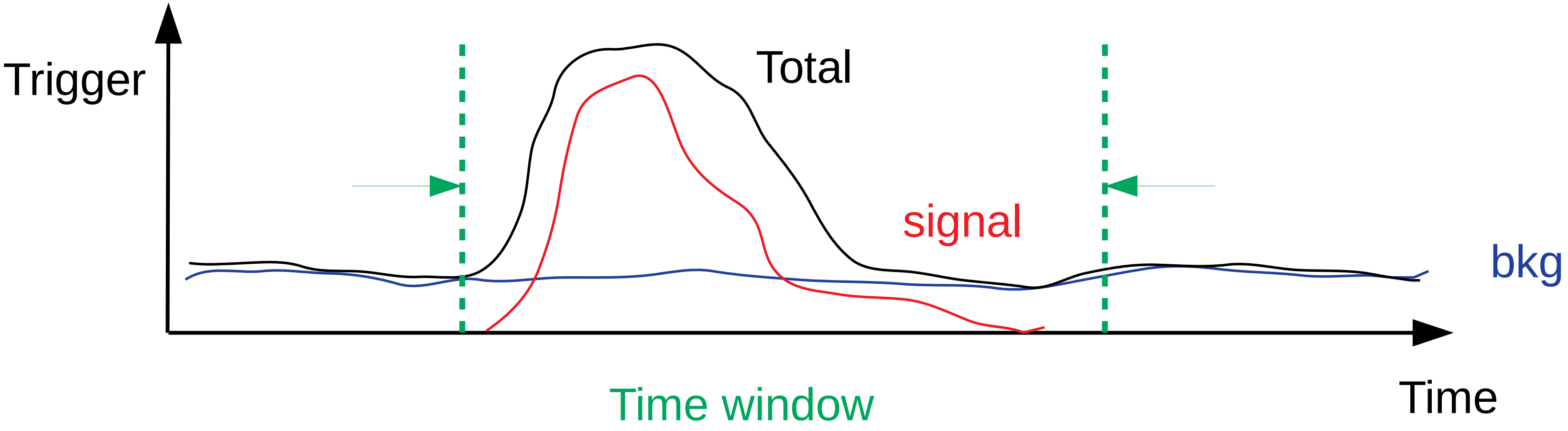
Background selection

- ~ 110 with FV cut (10 second)
- In low energy region:
    - Contaminates PES
    - Contaminates EES

- More sophisticated selection should improve selection
    - Pulse Shape discrimination for PES
    - Characteristic energies for certain decays

Include in Simulation

- Again timing is crucial for SN



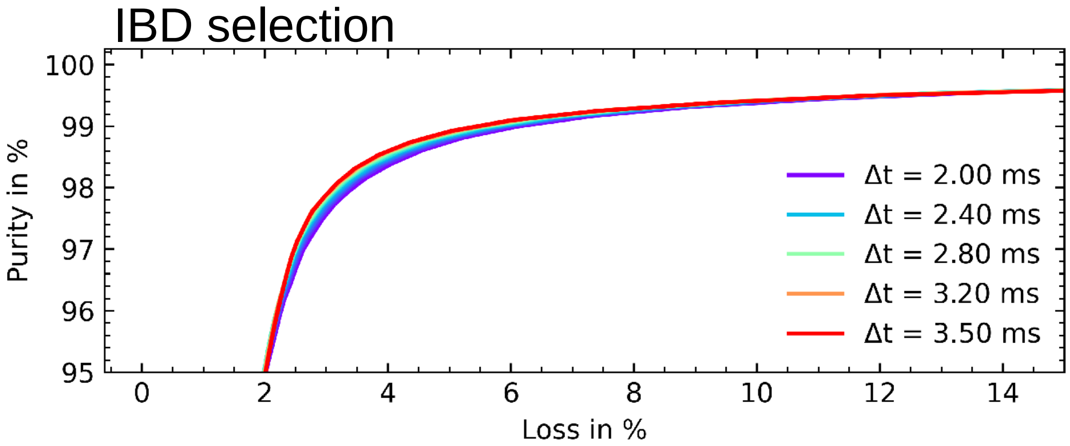
- Time window (start, end) should be optimized for each SN (or any bursts)
  - Different duration + different distance gives different signal to bkg ratio

- There will be one or two **online** burst alerts from global trigger and MM trigger
  - Fixed length (60 secs)
  - Offline analysis might yield better results

- This effect needs to be simulated properly → to be implemented

Summary / Outlook

- Established working unfolding scheme
    - Direct unfolding (regularization on neutrino spectra)
- Primary results very promising
- Estimated small contribution from radioactive backgrounds
- Produced larger data set ~ 55 models with more variety & updated JUNO simulation



Backups

Reformulated inference

- Define:

$$\theta_{ij} = \frac{M_{ji} \cdot P(v_i)}{\sum_i M_{ji} \cdot P(v_i)}$$

- Inference step:

$$\boxed{\tilde{v}_i \approx \frac{1}{\epsilon_i} \sum_{j=1}^{n_{\vec{m}}} \theta_{ij} \cdot m_j}$$

  - $\tilde{v}_i$ reconstructed bin content in physical distribution (collection of $\tilde{v}_i$ gives reconstructed distribution)
  - ϵ_i in inefficiency of physical bin i
  - $n_{\vec{m}}$ number of bins in measured distribution

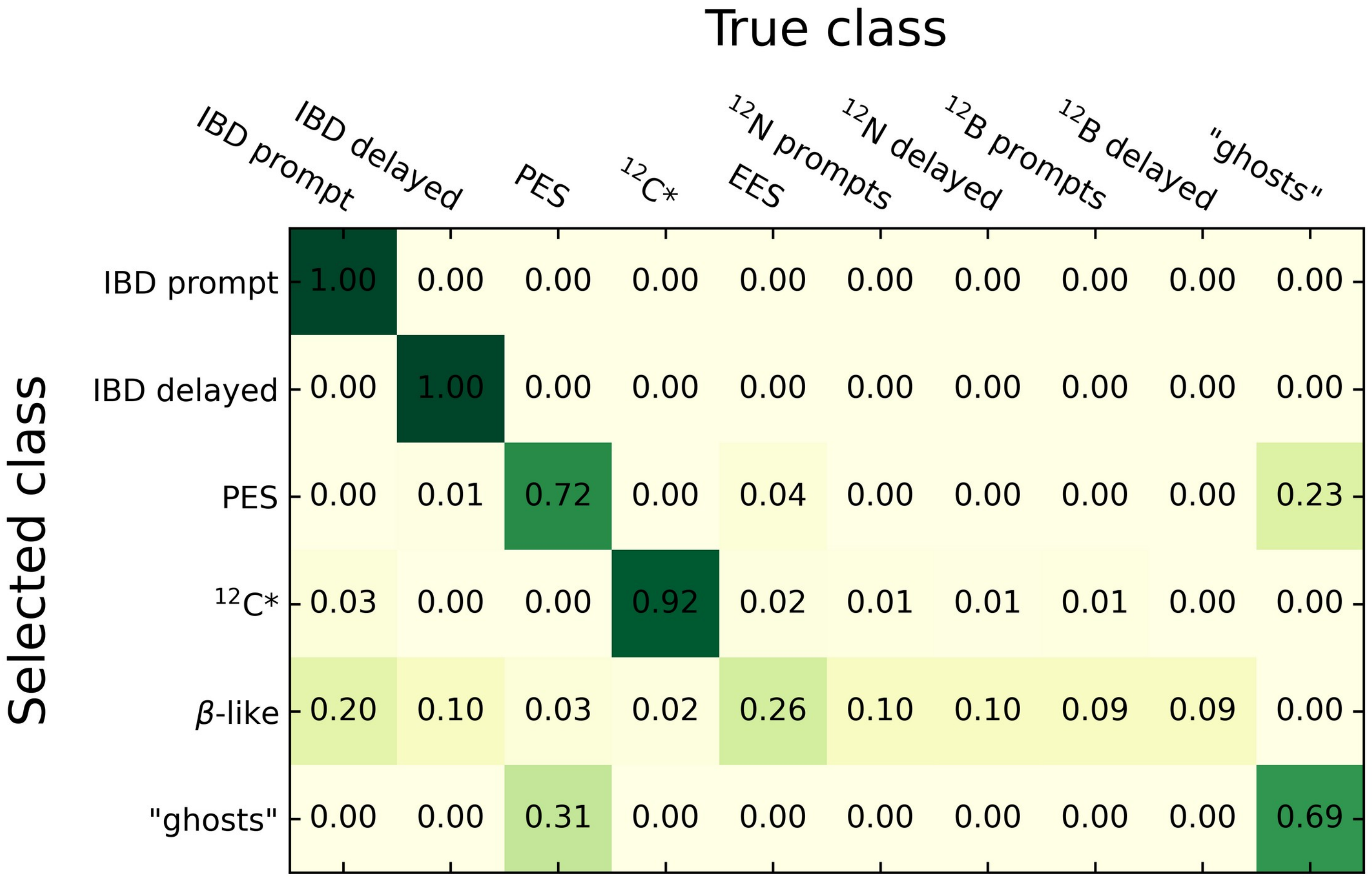
Response construction

- With MC simulations we get the physical spectrum $\vec{v}$ and the observed spectrum $\vec{m}$
  - In our case we actually have a collection of simulations and spectra

- Using the MC truth to measured relations on event level we can construct our response matrix
  - Which is an **estimator** for the true response matrix
    - Statistical fluctuations from limited amount of MC
    - Systematics coming from the model choices

$$\overline{M}_{MC} \approx M$$

Performance – Contamination Matrix

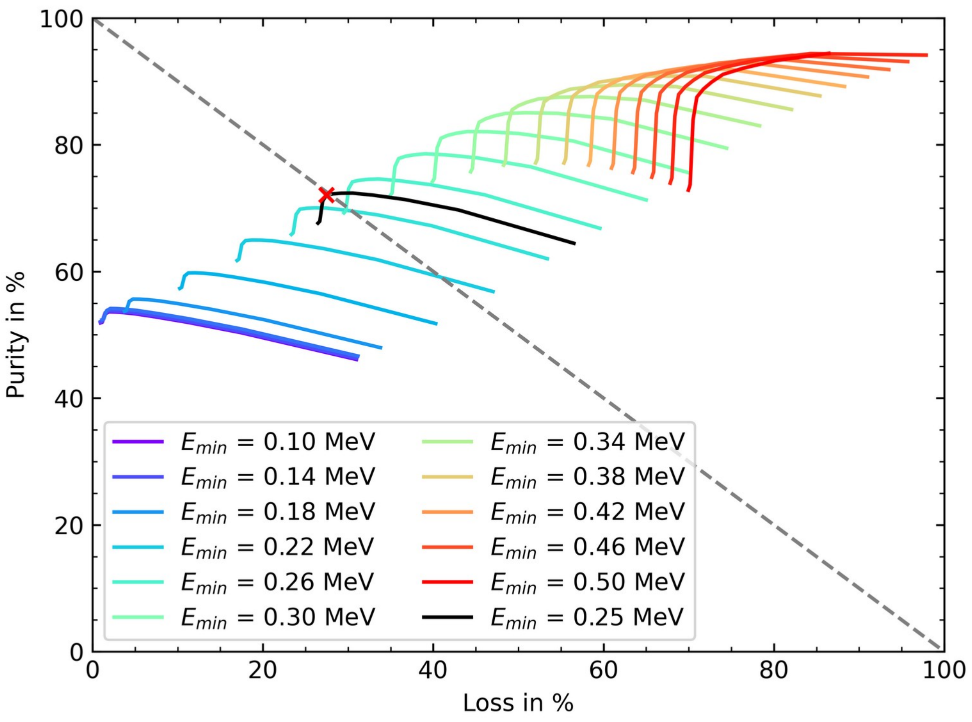


	IBD prompt	IBD delayed	PES	$^{12}C*$	EES	^{12}N prompts	^{12}N delayed	^{12}B prompts	^{12}B delayed	"ghosts"
IBD prompt	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IBD delayed	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PES	0.00	0.01	0.72	0.00	0.04	0.00	0.00	0.00	0.00	0.23
$^{12}C*$	0.03	0.00	0.00	0.92	0.02	0.01	0.01	0.01	0.00	0.00
β-like	0.20	0.10	0.03	0.02	0.26	0.10	0.10	0.09	0.09	0.00
"ghosts"	0.00	0.00	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.69

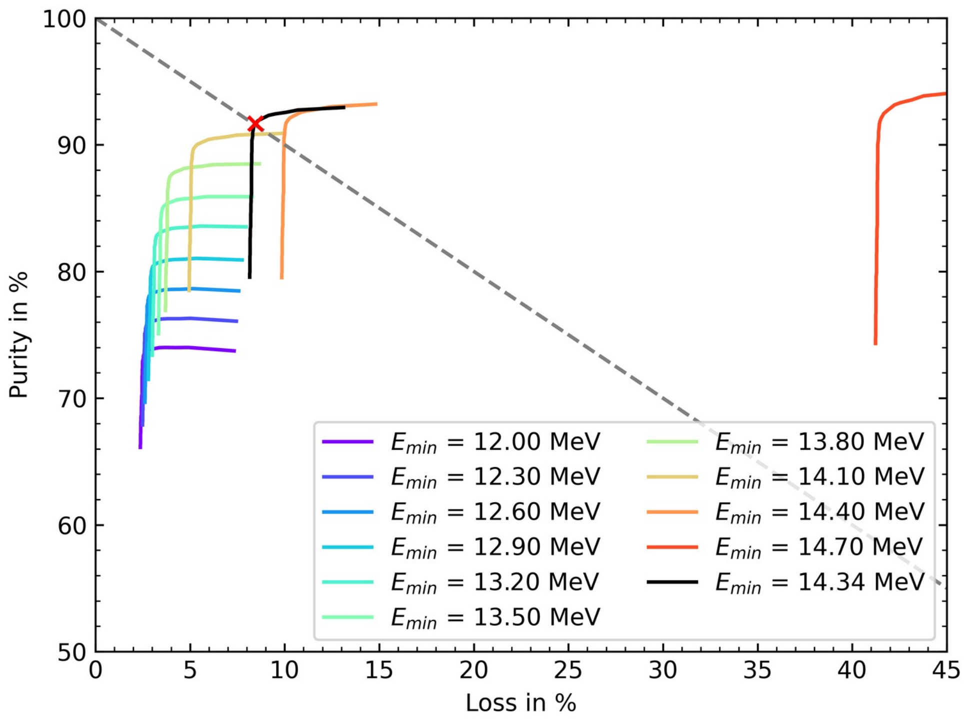
- Entries = # events selected / # events in selected class (fraction of selected events from true class...)

PES & ^{12}C* - Optimization

PES



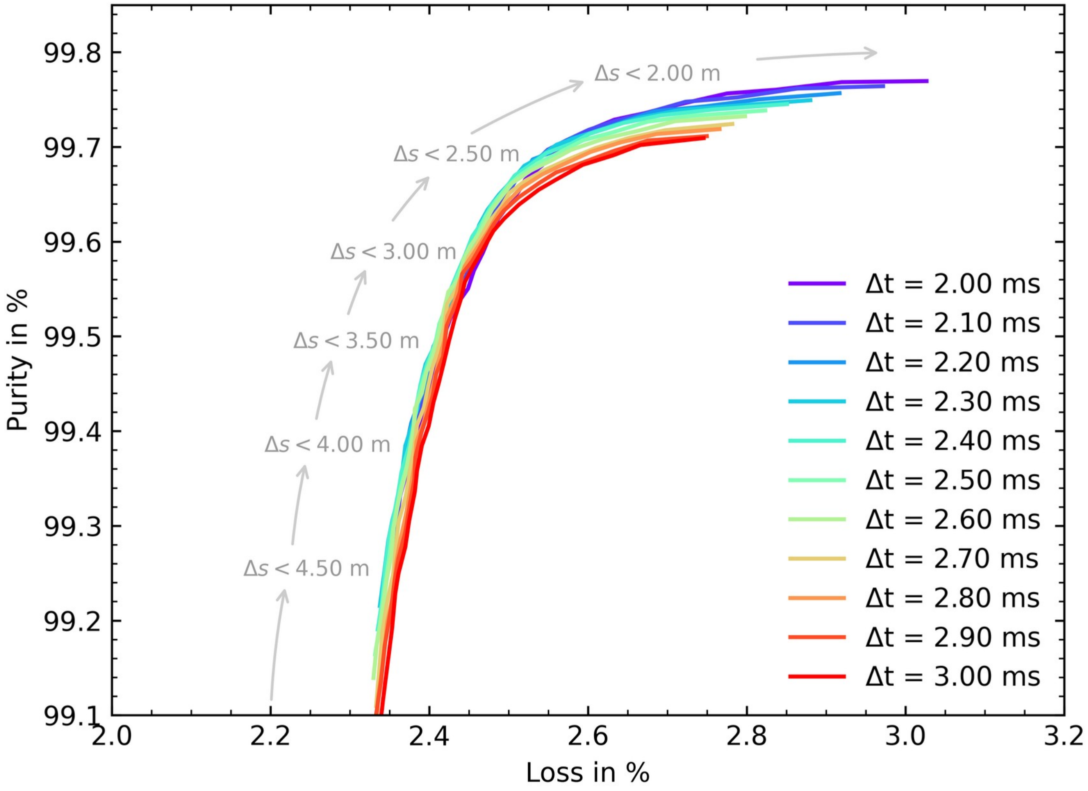
^{12}C*



- **Gray** line is diagonal between max purity and max loss
- Define optimum as interception of diagonal with highest purity

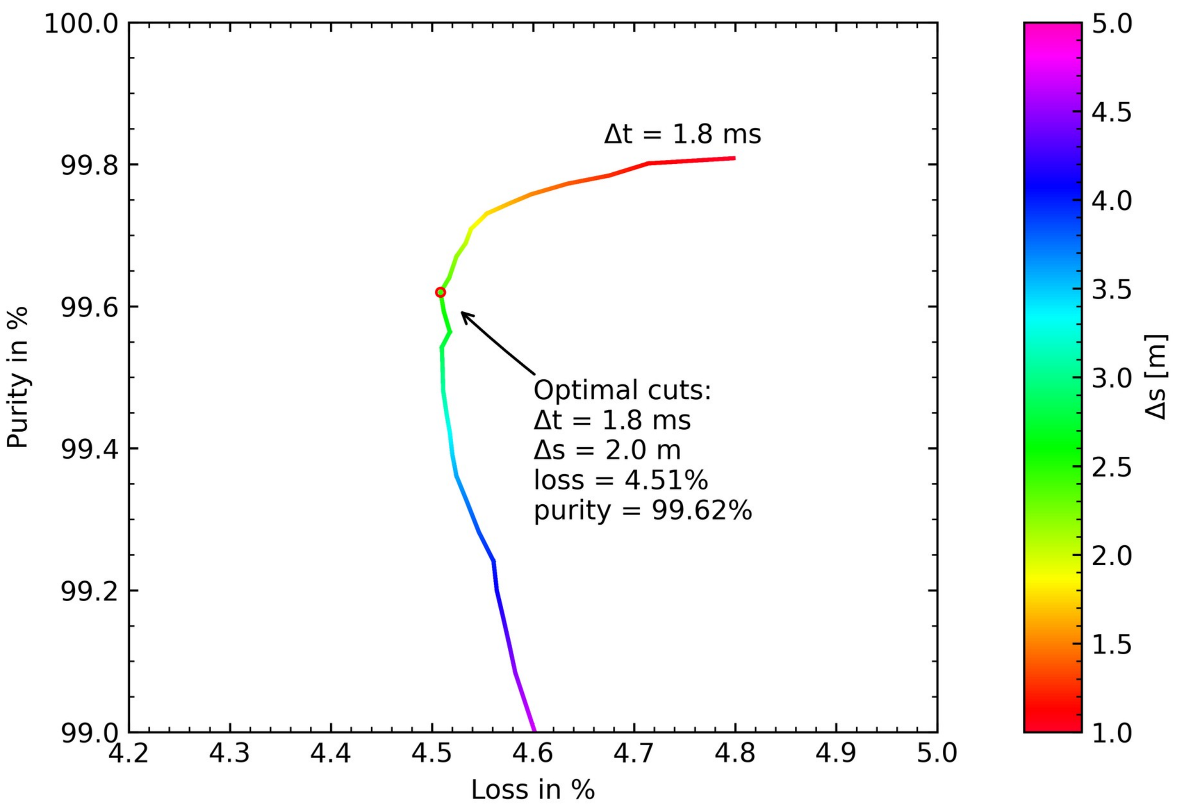
Generic trade-off between **purity** and **efficiency**

IBD delayed

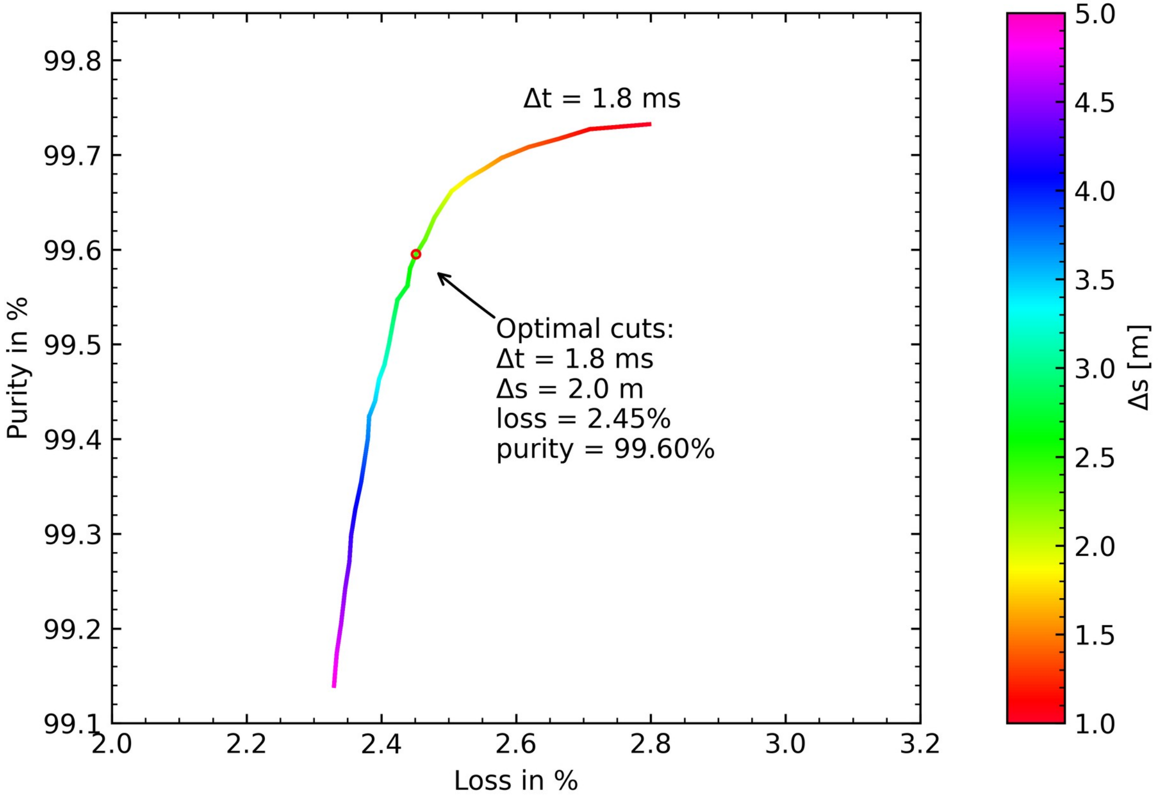


Purity in %
Loss in %
99.8
99.7
99.6
99.5
99.4
99.3
99.2
99.1
2.0
2.2
2.4
2.6
2.8
3.0
3.2
$\Delta s < 2.00$ m
$\Delta s < 2.50$ m
$\Delta s < 3.00$ m
$\Delta s < 3.50$ m
$\Delta s < 4.00$ m
$\Delta s < 4.50$ m
Δt = 2.00 ms
Δt = 2.10 ms
Δt = 2.20 ms
Δt = 2.30 ms
Δt = 2.40 ms
Δt = 2.50 ms
Δt = 2.60 ms
Δt = 2.70 ms
Δt = 2.80 ms
Δt = 2.90 ms
Δt = 3.00 ms

IBD prompt optimal curve



IBD delayed optimal curve

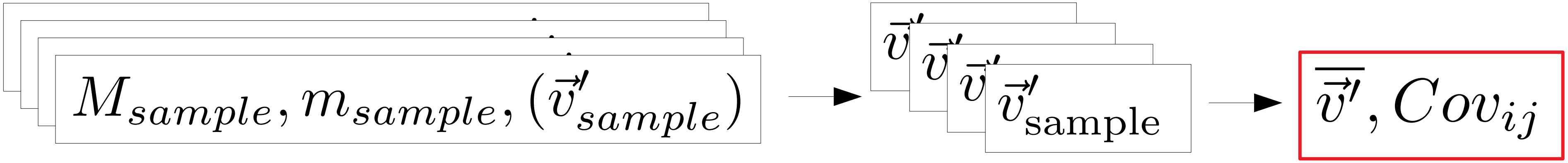


Uncertainty handling – statistical

- Original d’Agostini method used gaussian error propagation
  - Did not take poisson distribution into account
  - Did not take error propagation from updated priors into account

  Does **not** describe cases with **small numbers** well!

- Improved method (Improved iterative Bayesian unfolding)
  - Added poisson fluctuations
  - Estimates uncertainties from sampling underlying distributions

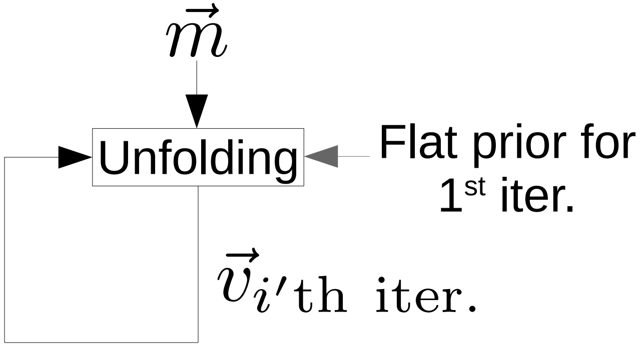


$$M_{sample}, m_{sample}, (\vec{v}'_{sample}) \rightarrow \vec{v}'_{\text{sample}} \rightarrow \boxed{\overline{\vec{v}'}, Cov_{ij}}$$

Bayesian inference and prior update

- Starting from a **flat prior,** we basically updated our knowledge (in this case no knowledge) of the physical spectrum with our unfolding
- Now that we have an “updated” version of the physical distribution $\vec{v}'$ we can rerun our unfolding by assuming the shape of $\vec{v}'$ as prior:

  $\longrightarrow \vec{v}''$ 2nd iteration
- This procedure can be redone iteratively until the updated distribution converges (*Iterative bayesian prior update*)
  - To much iterations blow up statistical uncertainties
  - In some cases the unfolded distribution does not converge



$\vec{m}$

Unfolding

Flat prior for 1st iter.

$\vec{v}_{i'\mathrm{th\ iter.}}$

- Starting from a **non-flat prior** (prior knowledge), usually the physical distribution $\vec{v}'$ is only updated once
  - Multiple iterations can lose initial prior knowledge

Bayesian inference and prior update

- Starting from a **flat prior,** we basically updated our knowledge (in this case no knowledge) of the physical spectrum with our unfolding
- Now that we have an “updated” version of the physical distribution $\vec{v}'$ we can rerun our unfolding by assuming the shape of $\vec{v}'$ as prior:
- This pro converges (*Iterative*
  - To m statis or for er.
  - In so distri

The use of priors acts as **regularization** for this method.
(as smooth priors are usually assumed)

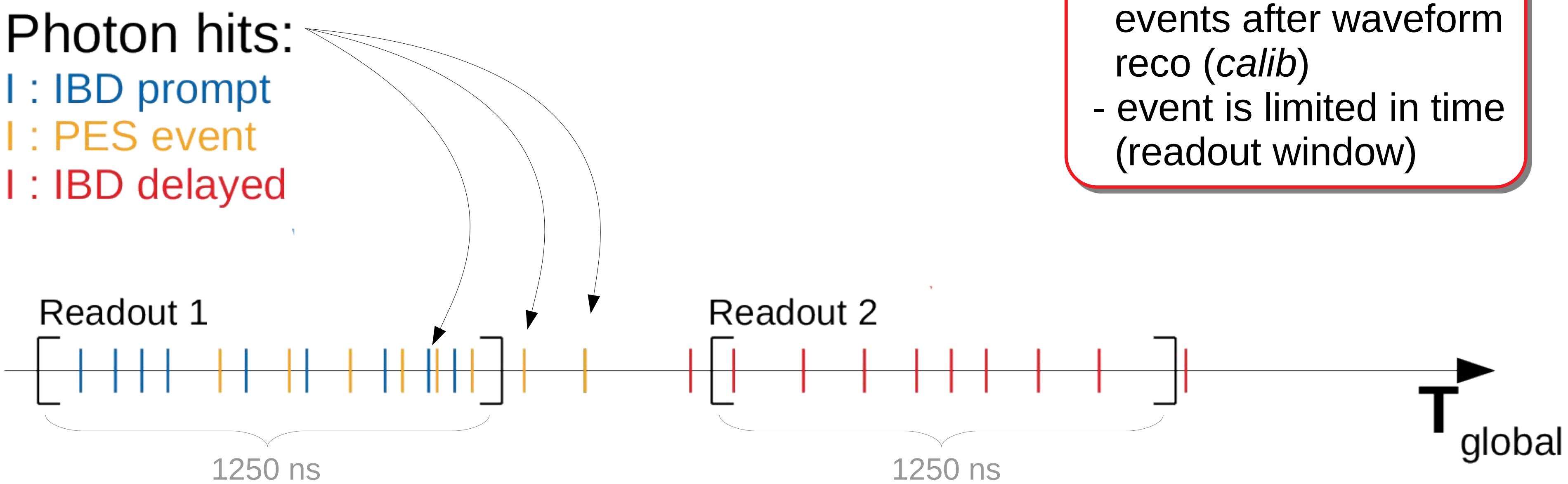
Other unfolding methods (e.g. SVD unfolding)
use manual regularizations
to constrain the blow up of statistical fluctuations.

- Starting from a **non-flat prior** (prior knowledge), usually the physical distribution $\vec{v}'$ is only updated once
  - Multiple iterations can lose initial prior knowledge

Pile-Up

- Photon hits from N neutrino events (usual just two) in one readout window (after *elecsim*)
- Low signal rate:
  - Pile-up of signal + background
- High signal rate:
  - Pile-up of signal + signal

Readout event:
- “raw” data output of detector
- in software given by *elecsim*
- but usually work on events after waveform reco (*calib*)
- event is limited in time (readout window)



Photon hits:

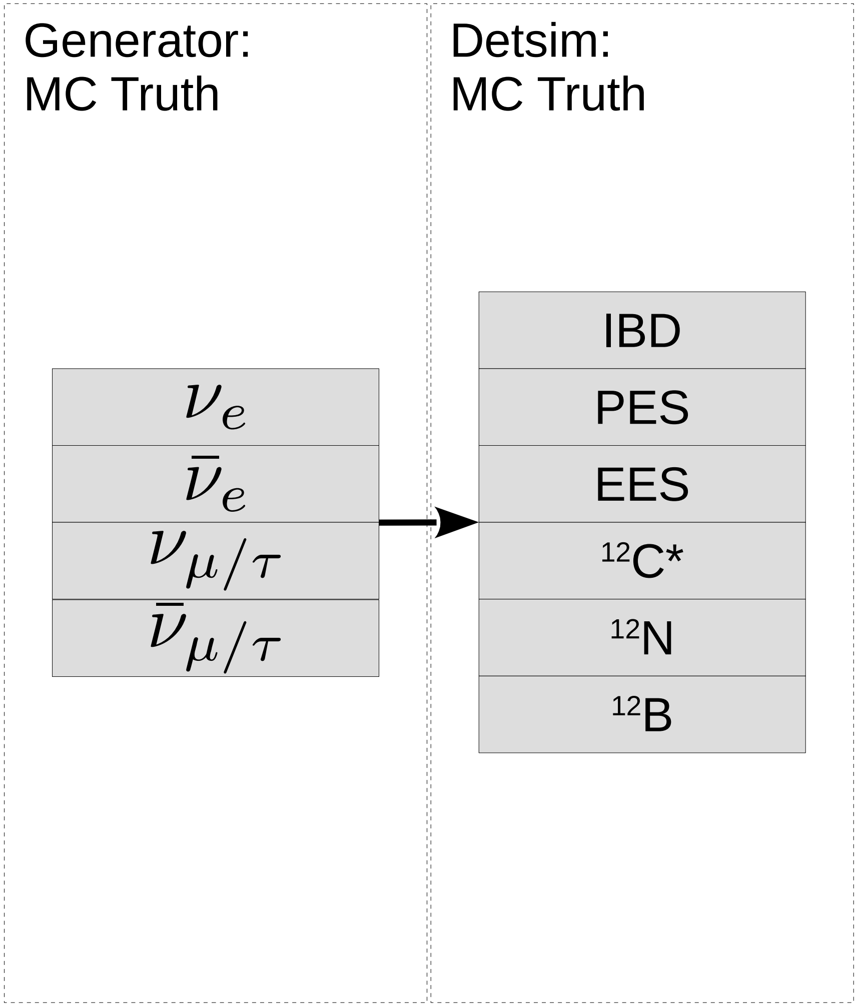
| : IBD prompt

| : PES event

| : IBD delayed

MC Truth

- What is MC truth type information?



Generator:
MC Truth

ν_e

$\bar{\nu}_e$

$\nu_{\mu/\tau}$

$\bar{\nu}_{\mu/\tau}$

Detsim:
MC Truth

IBD

PES

EES

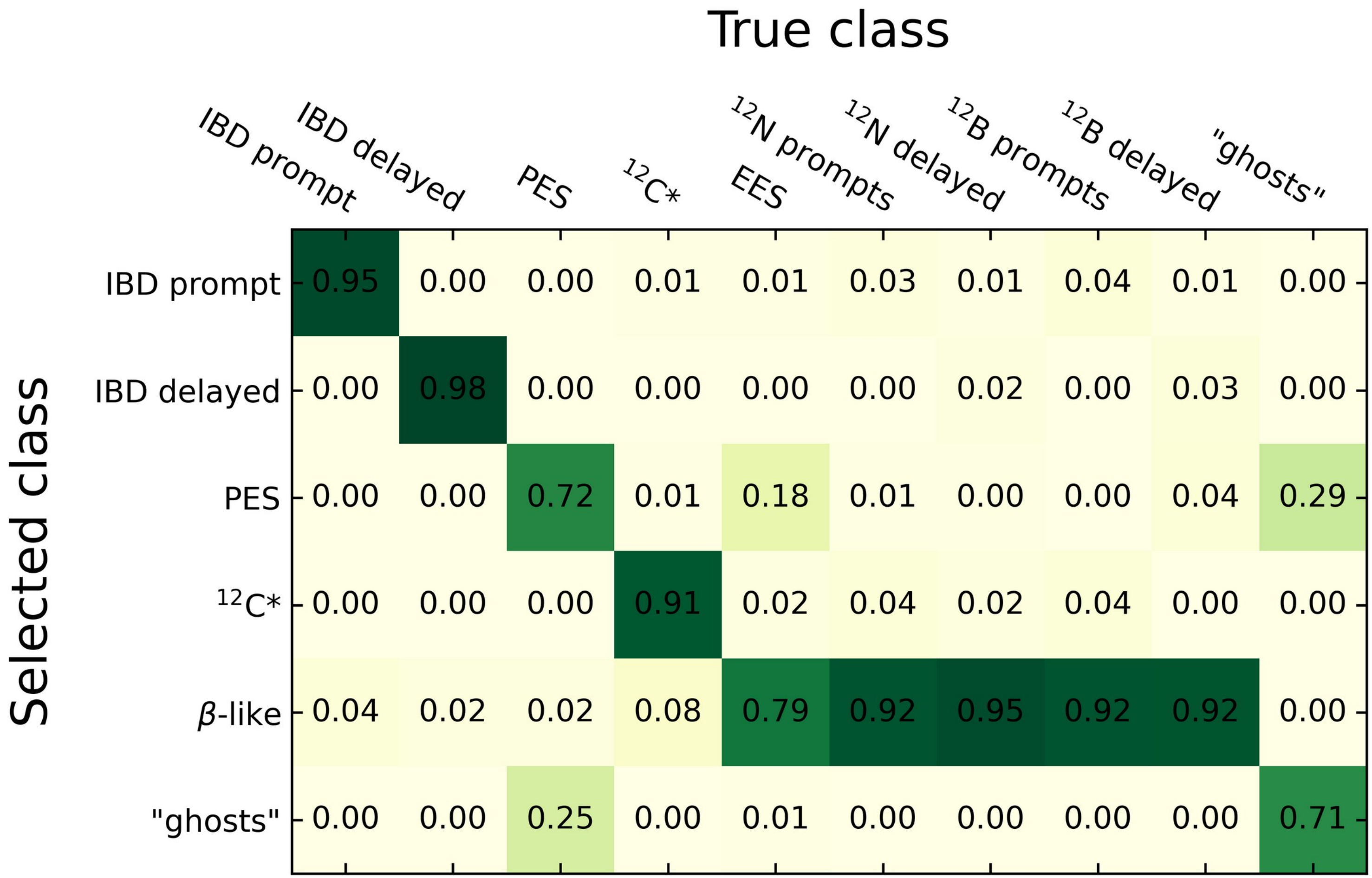
^{12}C*

^{12}N

^{12}B

- Interaction channel
- Secondary particle type
- Secondary energy
- Secondary position
- **For event reconstruction (energy + vertex)**

Total Confusion Matrix



	IBD prompt	IBD delayed	PES	$^{12}C*$	EES	^{12}N prompts	^{12}N delayed	^{12}B prompts	^{12}B delayed	"ghosts"
IBD prompt	0.95	0.00	0.00	0.01	0.01	0.03	0.01	0.04	0.01	0.00
IBD delayed	0.00	0.98	0.00	0.00	0.00	0.00	0.02	0.00	0.03	0.00
PES	0.00	0.00	0.72	0.01	0.18	0.01	0.00	0.00	0.04	0.29
$^{12}C*$	0.00	0.00	0.00	0.91	0.02	0.04	0.02	0.04	0.00	0.00
β-like	0.04	0.02	0.02	0.08	0.79	0.92	0.95	0.92	0.92	0.00
"ghosts"	0.00	0.00	0.25	0.00	0.01	0.00	0.00	0.00	0.00	0.71

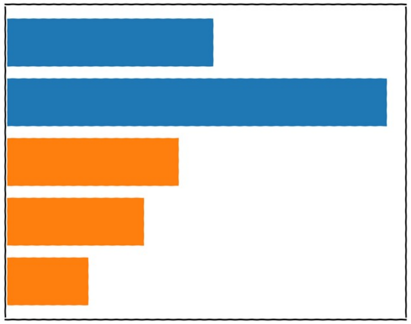
- Entries = # events selected / # events in true class
  (fraction of true events selected into class...)

Unfolding Multiple Spectra

- For binned distributions problem can be described by matrix multiplication

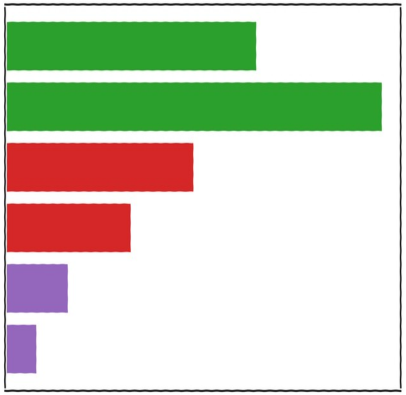
$$\vec{m} = M \cdot \vec{v}$$

Measured distribution



$$\left(\begin{array}{c}30\\55\\\hline 25\\20\\12\end{array}\right) = \left(\begin{array}{cc|cc|cc}0.6 & 0.05 & 0.1 & 0.0 & 0.0 & 0.0\\0.1 & 0.75 & 0.2 & 0.0 & 0.0 & 0.0\\\hline 0.0 & 0.05 & 0.7 & 0.05 & 0.0 & 0.0\\0.0 & 0.0 & 0.00 & 0.6 & 0.6 & 0.4\\0.0 & 0.0 & 0.00 & 0.25 & 0.4 & 0.6\end{array}\right) \cdot \left(\begin{array}{c}40\\60\\\hline 30\\20\\\hline 10\\5\end{array}\right)$$

Physical distribution



- Can concatenate multiple physical distributions (e.g. neutrino flavor)
- Can concatenate multiple measured distributions (e.g. interaction channels)
- Response consist of different blocks
  - Block describes bin migration from one physical class to one measured class

Inference problem

- Now consider a true measurement (or a different MC sample)
- Simple matrix inversion (M^{-1}) not possible (response might not be invertible, but even if, the inversion only holds for expectation values **not for sample measurement**)

$$\begin{pmatrix}25\\58\\21\\25\\7\end{pmatrix} = \begin{pmatrix}0.6 & 0.05 & 0.1 & 0.0 & 0.0 & 0.0\\ 0.1 & 0.75 & 0.2 & 0.0 & 0.0 & 0.0\\ 0.0 & 0.05 & 0.7 & 0.05 & 0.0 & 0.0\\ 0.0 & 0.0 & 0.00 & 0.6 & 0.6 & 0.4\\ 0.0 & 0.0 & 0.00 & 0.25 & 0.4 & 0.6\end{pmatrix} \cdot \begin{pmatrix}?\\?\\?\\?\\?\\?\end{pmatrix}$$

$$\begin{matrix}0.3 & 0.15 & 0.00 & 0.1 & 0.0 & 0.0\end{matrix}$$

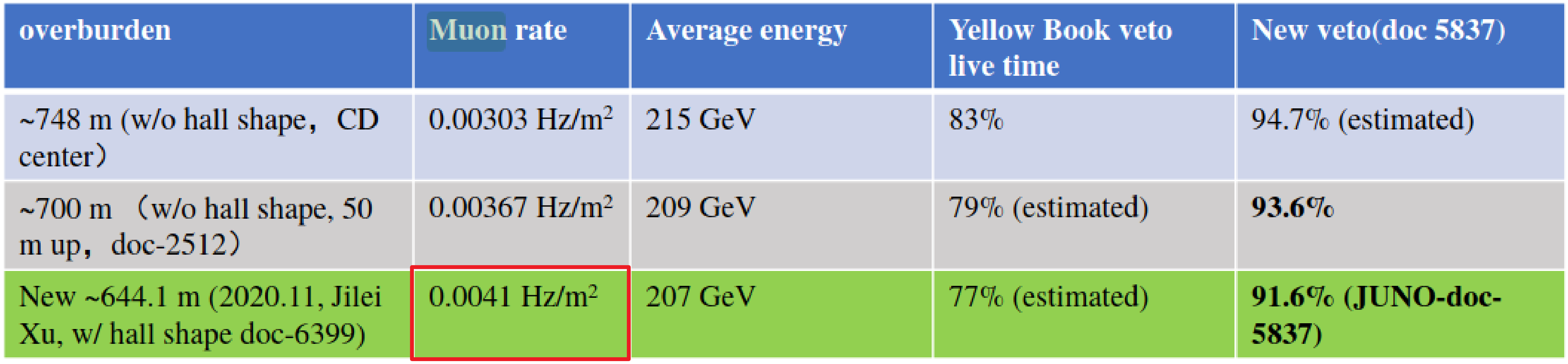
- How can we infer the physical spectrum using our knowledge on the response?
  - Keep in mind we don't know the true number of lost events (inefficiency bin), but we still know the mean inefficiency

- **Bayes' theorem** can help:

$$P(\vec{v}\,|\vec{m}, M) \sim \overbrace{P(\vec{m}\,|\vec{v}, M)}^{\text{likelihood}} \cdot \overbrace{P(\vec{v})}^{\text{prior}}$$

- Method of d'Agostini Improved iterative Bayesian unfolding

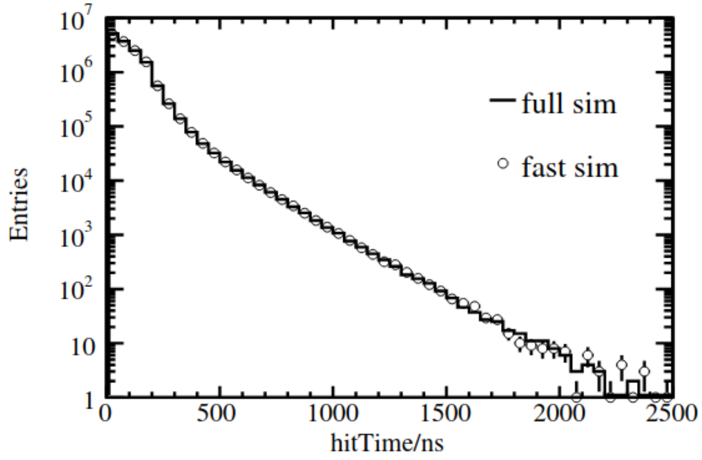
How about muons?



overburden	Muon rate	Average energy	Yellow Book veto live time	New veto(doc 5837)
~748 m (w/o hall shape，CD center）	0.00303 Hz/m^2	215 GeV	83%	94.7% (estimated)
~700 m （w/o hall shape, 50 m up，doc-2512）	0.00367 Hz/m^2	209 GeV	79% (estimated)	**93.6%**
New ~644.1 m (2020.11, Jilei Xu, w/ hall shape doc-6399)	0.0041 Hz/m^2	207 GeV	77% (estimated)	**91.6% (JUNO-doc-5837)**

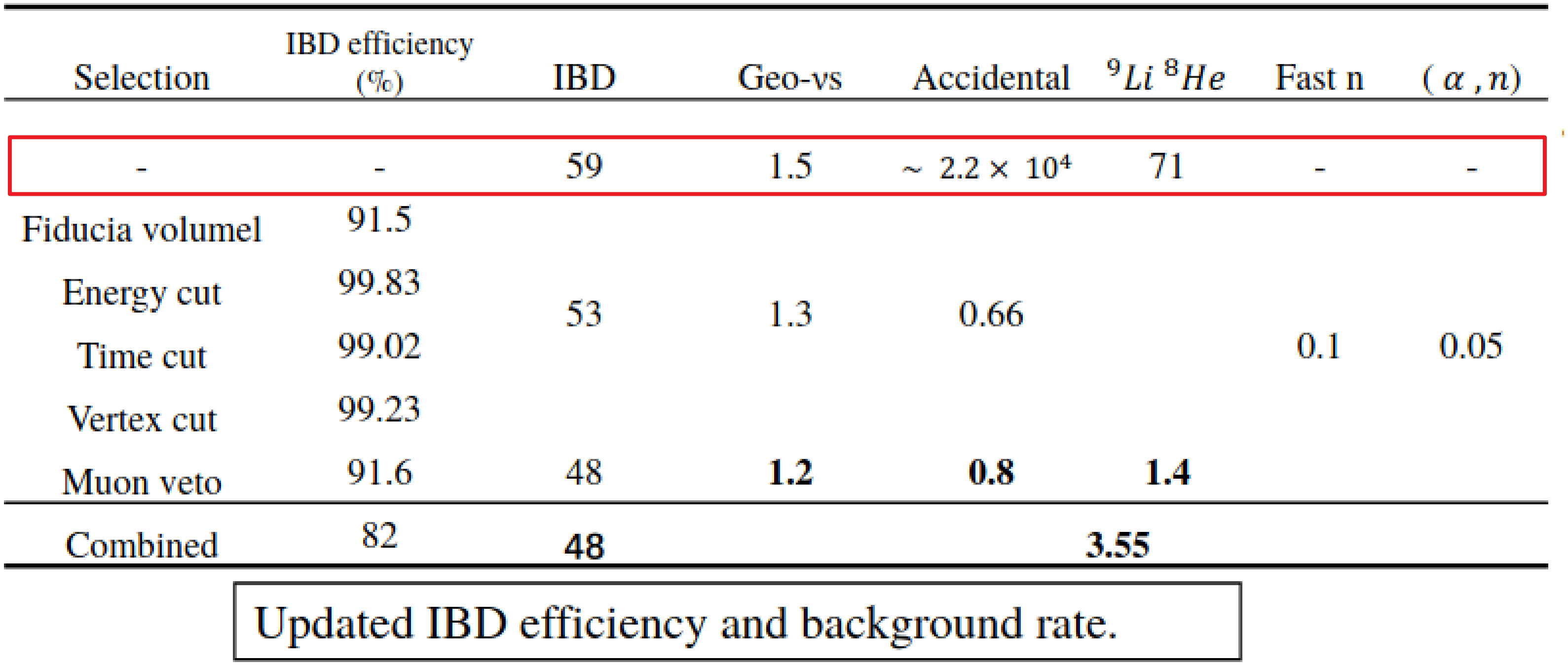
- With detector cross-section → ~4 Hz muon rate
- Muon pulse ~ 1.5 to 2 μs
- Small fraction of downtime
  - 10 kpc SN: pile up with signal very

Old software



Backgrounds considerations

- How big is actual bkg contribution (with global trigger >200 MeV)?
  - New paper about to be published



Selection	IBD efficiency (%)	IBD	Geo-vs	Accidental	$^{9}Li\ ^{8}He$	Fast n	(α, n)
-	-	59	1.5	$\sim 2.2 \times 10^4$	71	-	-
Fiducia volumel	91.5	53	1.3	0.66		0.1	0.05
Energy cut	99.83						
Time cut	99.02						
Vertex cut	99.23						
Muon veto	91.6	48	**1.2**	**0.8**	**1.4**		
Combined	82	**48**	**3.55**				

Updated IBD efficiency and background rate.

- In readout of ~ 10 sec. reactor, Geo-vs & ^{9}Li. ^{8}He neglectable
- How about accidental? (here accidental coincidences)

Selected Energy Distributions