Core-to-Core Program

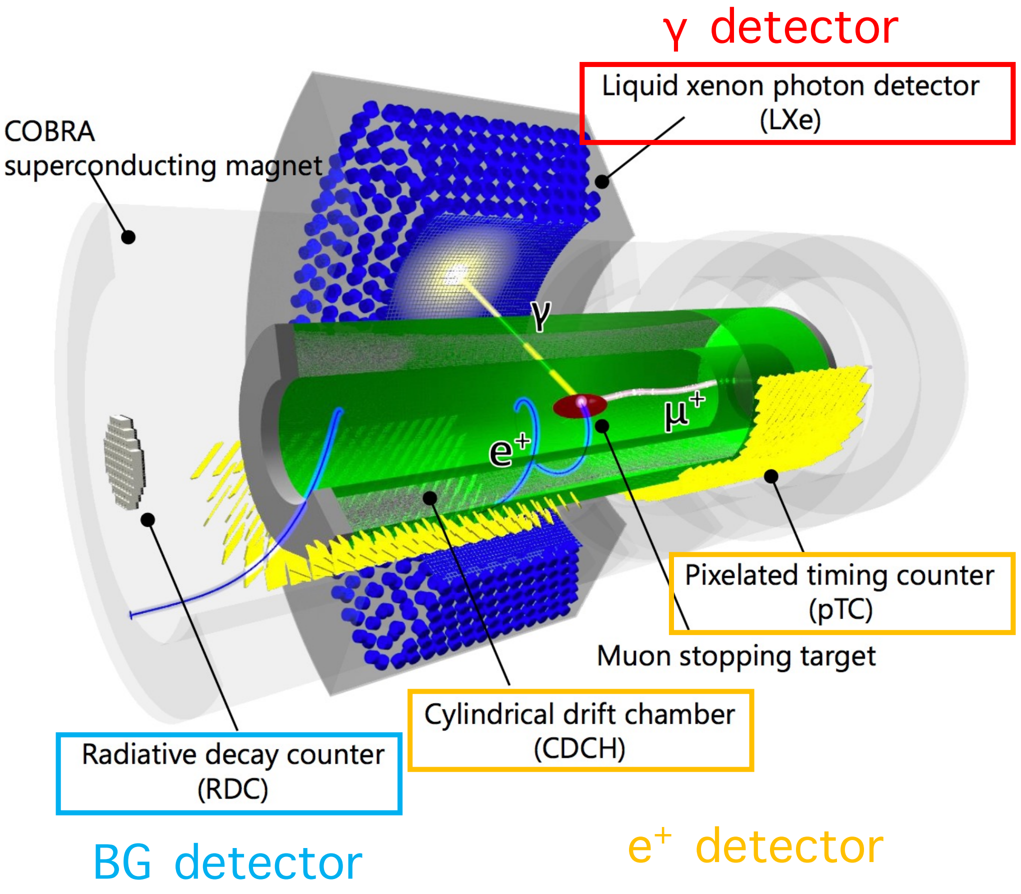
Pileup Analysis for the Liquid Xenon Detector of the MEG II experiment

15/09/2021, 15pT3-7

Rina Onda on behalf of MEG II Collaboration

The University of Tokyo

γ Detector of MEG II Experiment



γ detector

Liquid xenon photon detector
(LXe)

COBRA
superconducting magnet

Pixelated timing counter
(pTC)

Muon stopping target

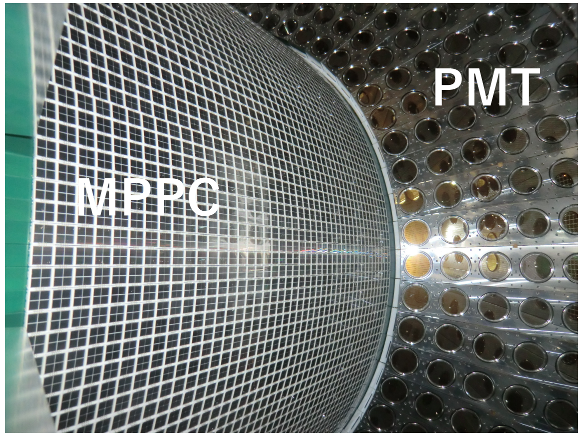
Cylindrical drift chamber
(CDCH)

Radiative decay counter
(RDC)

BG detector

e^+ detector

Inside LXe

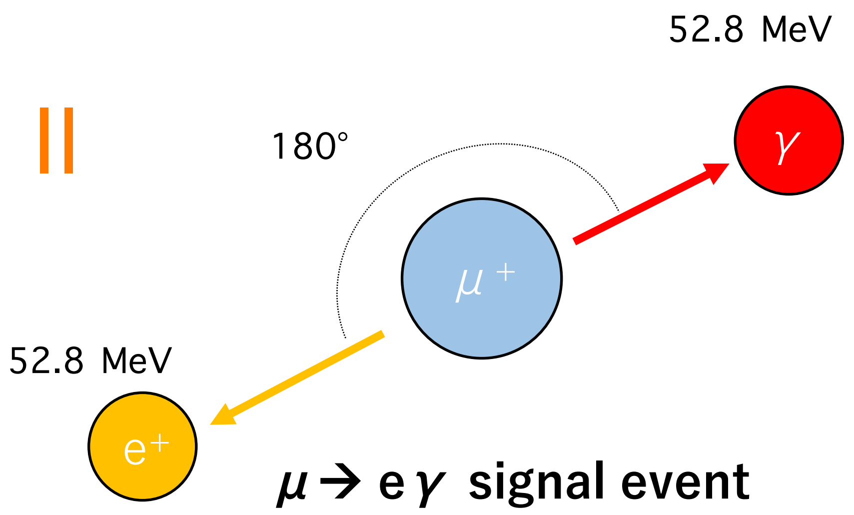


- MEG II experiment searches μ→ eγ decay, which is one of charged Lepton Flavor Violation.
- Liquid xenon photon detector (LXe) detects energy, position and timing of γ.
- Scintillation lights from liquid xenon are detected with PMTs and MPPCs.
- In this talk, the pileup analysis for the LXe detector will be reported.

Signal & BG in MEG II

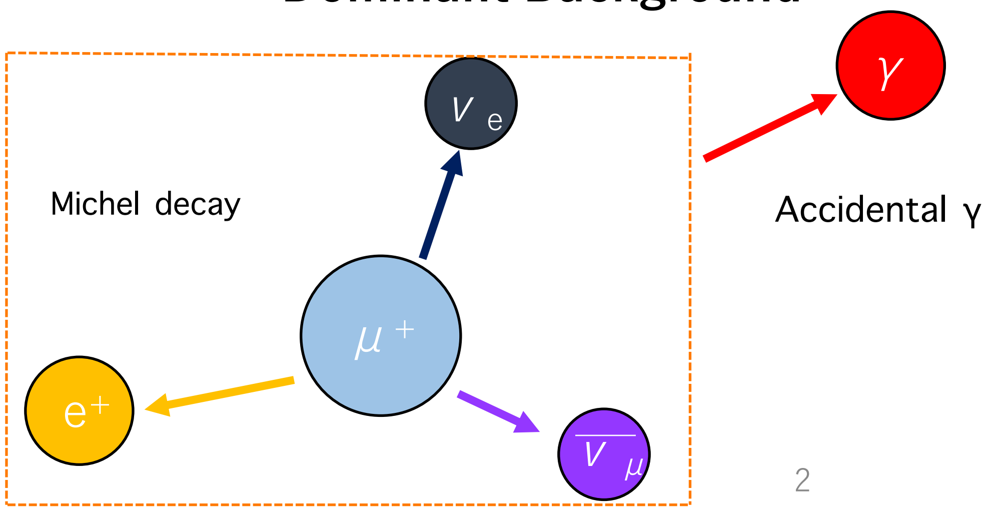
- μ→ eγ signal event can be characterized by
  - $E_e = E_\gamma = 52.8$ MeV
  - back to back
  - coincident in time
- **The dominant background derives from the accidental coincidence of e^+ and γ-ray from different μ decays.**
- The number of the accidental background is proportional to the square of the beam rate R_μ:

  $N_{bg} \propto R_\mu^2$
- Background γ-rays hit to LXe at 0.7 MHz.



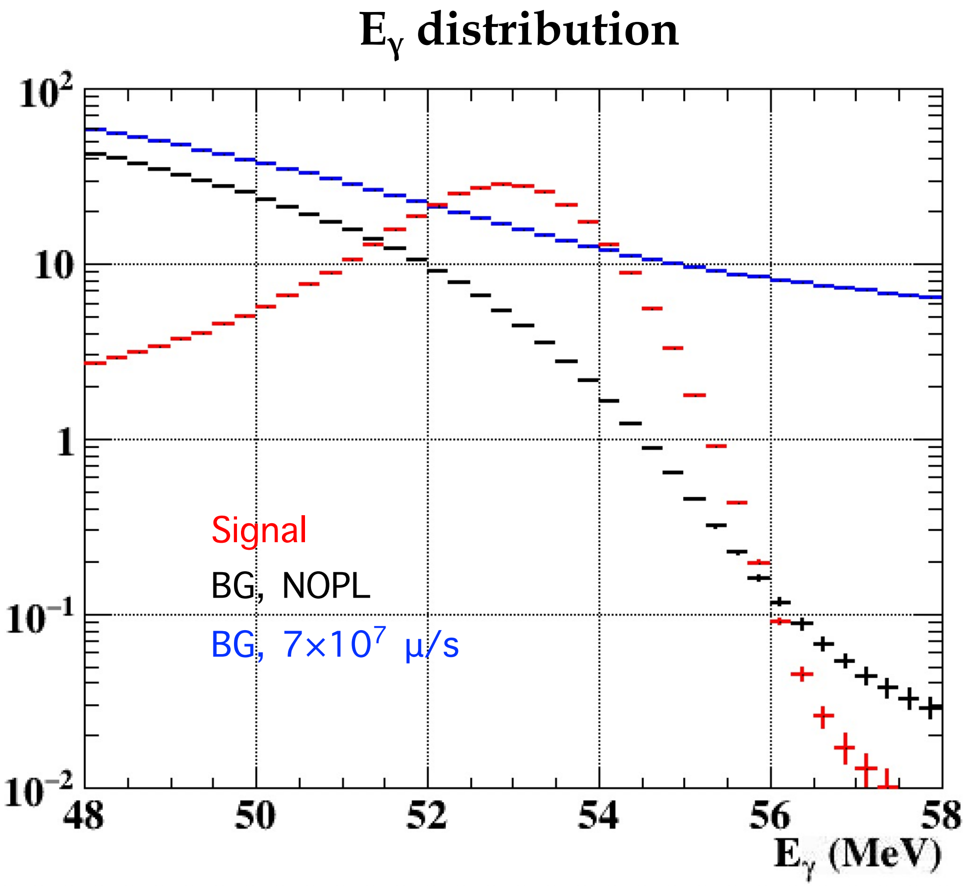
μ→ eγ signal event

Dominant Background



Pileup γ-ray

- The energy, position and timing of a γ-ray are reconstructed using information measured by LXe.
- The pileup γ-rays can **greatly affect the energy reconstruction** since it uses information of all channels.
  ↔ The effects on the position and the timing are limited since they are reconstructed using local information.
- The existence of the pileups **increases the number of background events** in the signal region:
  w/o pileup: 42 Hz → w/ pileup: 131 Hz for 52-54 MeV
- Therefore, the pileup elimination is crucial for the better sensitivity.

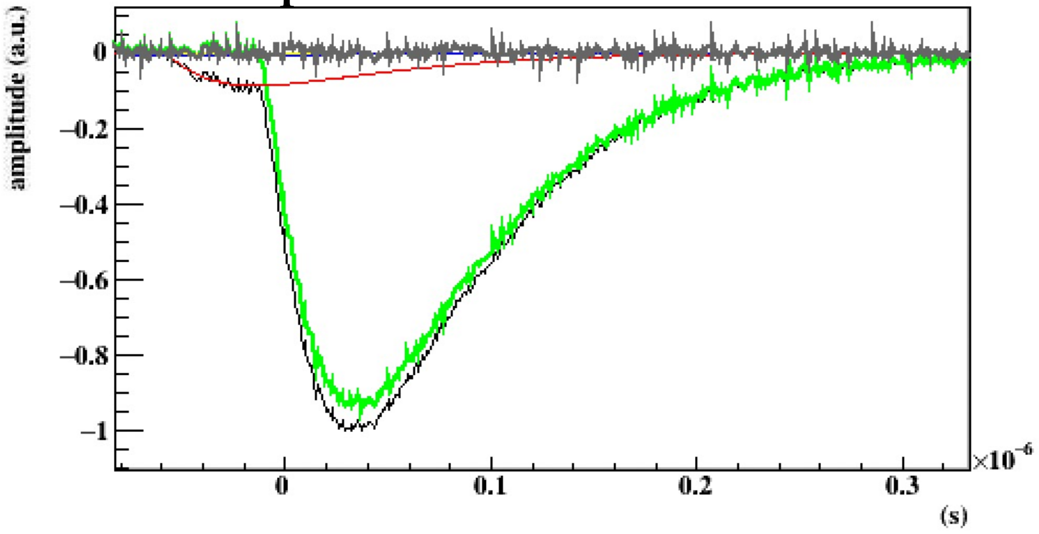


*Signal distribution is scaled for visibility.

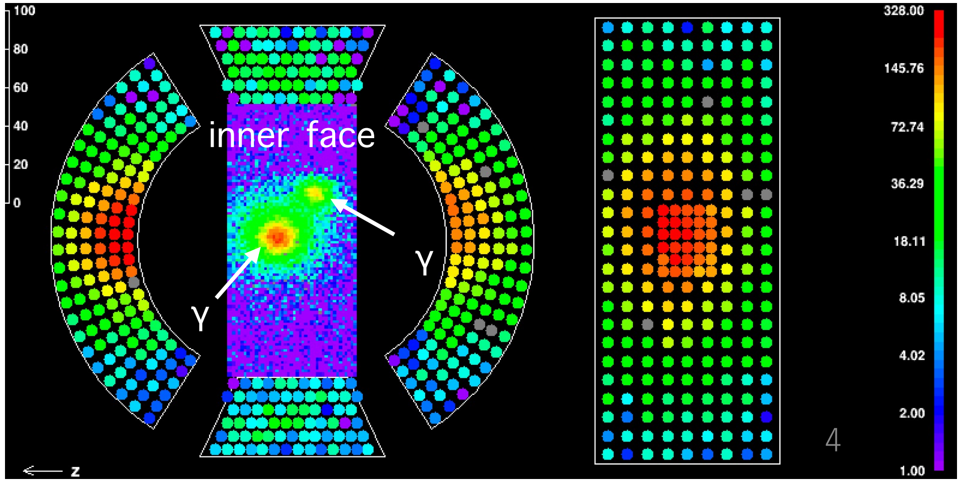
Previous Algorithm for Pileup Elimination

- There were algorithms already implemented.
- It consists of two steps:
  1. **Unfolding with sum waveform fitting**
     - Take sums of MPPC and PMT channels
     - Fit a template waveform
     - The waveforms are unfolded.
     - Sensitive to off-timing pileups
  2. **Rejection with peak search in charge distribution**
     - Search peaks whose charges are larger than a threshold on inner face.
     - The events with pileups are rejected.
     - Sensitive to on-timing pileups
- They are processed independently.

Example of MPPC Sum Waveform

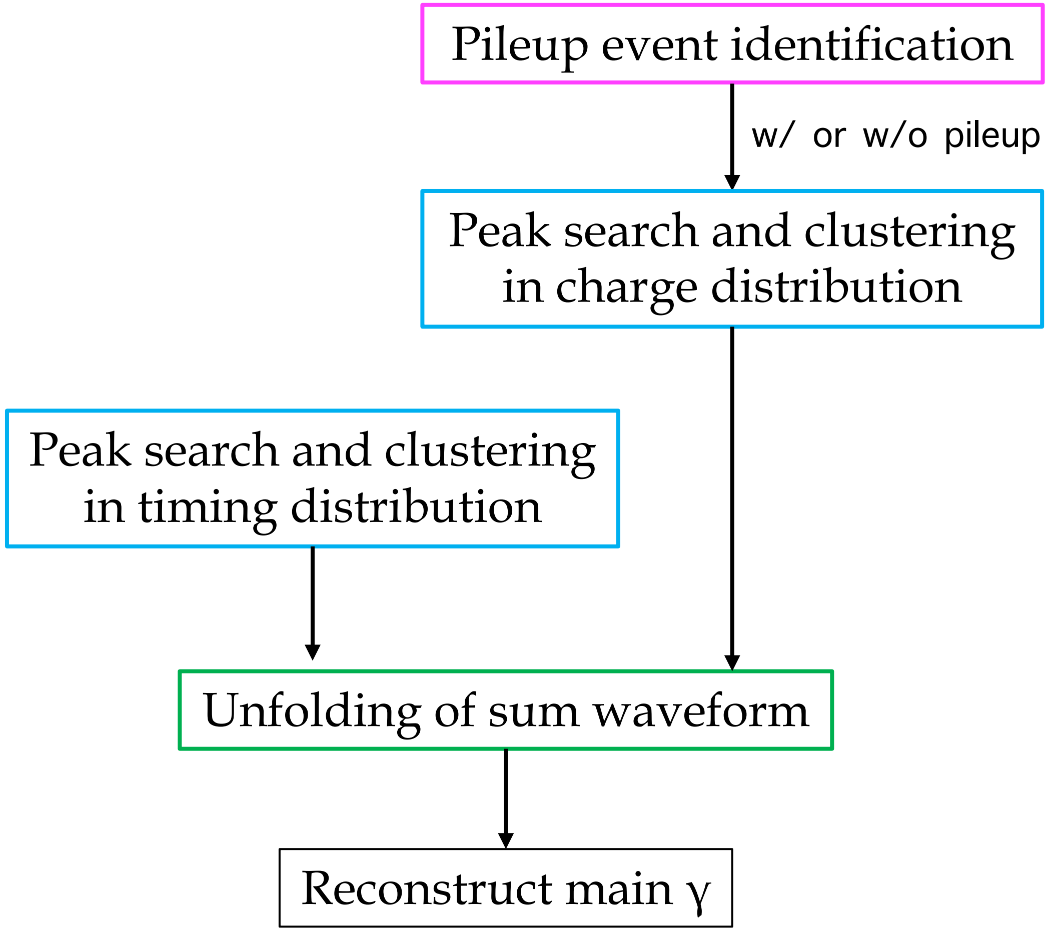


Example of Charge Distribution



New Algorithm for Pileup Analysis

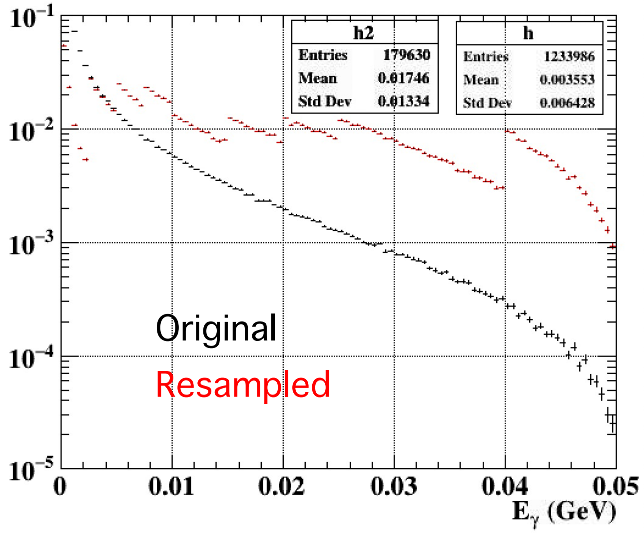
- A new algorithm was developed to improve the performance.
- It consists of three steps:
  1. Pileup event identification with DL-based algorithm
  2. Peak search and clustering of channels in charge and timing distributions
  3. Unfolding of sum waveform



Step1: DL-based Pileup Identification

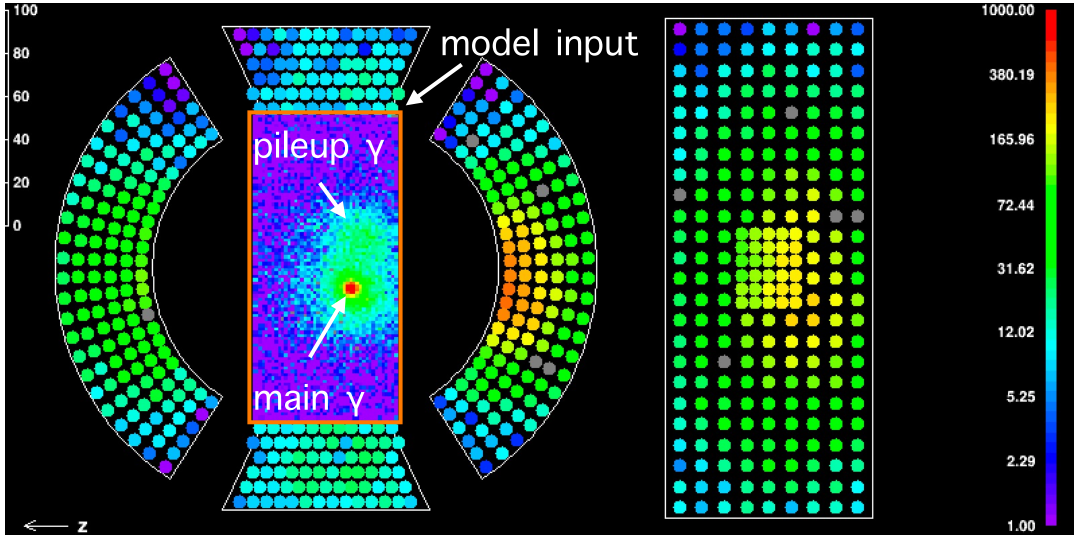
- The deep learning-based pileup identification method was implemented.
- The DL model judge whether the event likely has pileup γ-rays.
- Model architecture
  - Based on EfficientNet (https://arxiv.org/pdf/1905.11946.pdf)
    ← CNN with efficiently scaled model architecture
  - Inputs: Charge distribution of inner face (93 × 44 pixels)
  - Outputs: Probability to include pileup γ-rays
- Dataset
  - Generated with MC
  - Main γ (uniform 20-100 MeV, 1.6×10^5 events)
  - Pileup γ (resampled from the original pileup γ, 1.2×10^5 events)
- Implemented with Pytorch and converted to ONNX after training on Google Colaboratory

E_γ Distribution of Pileup γ

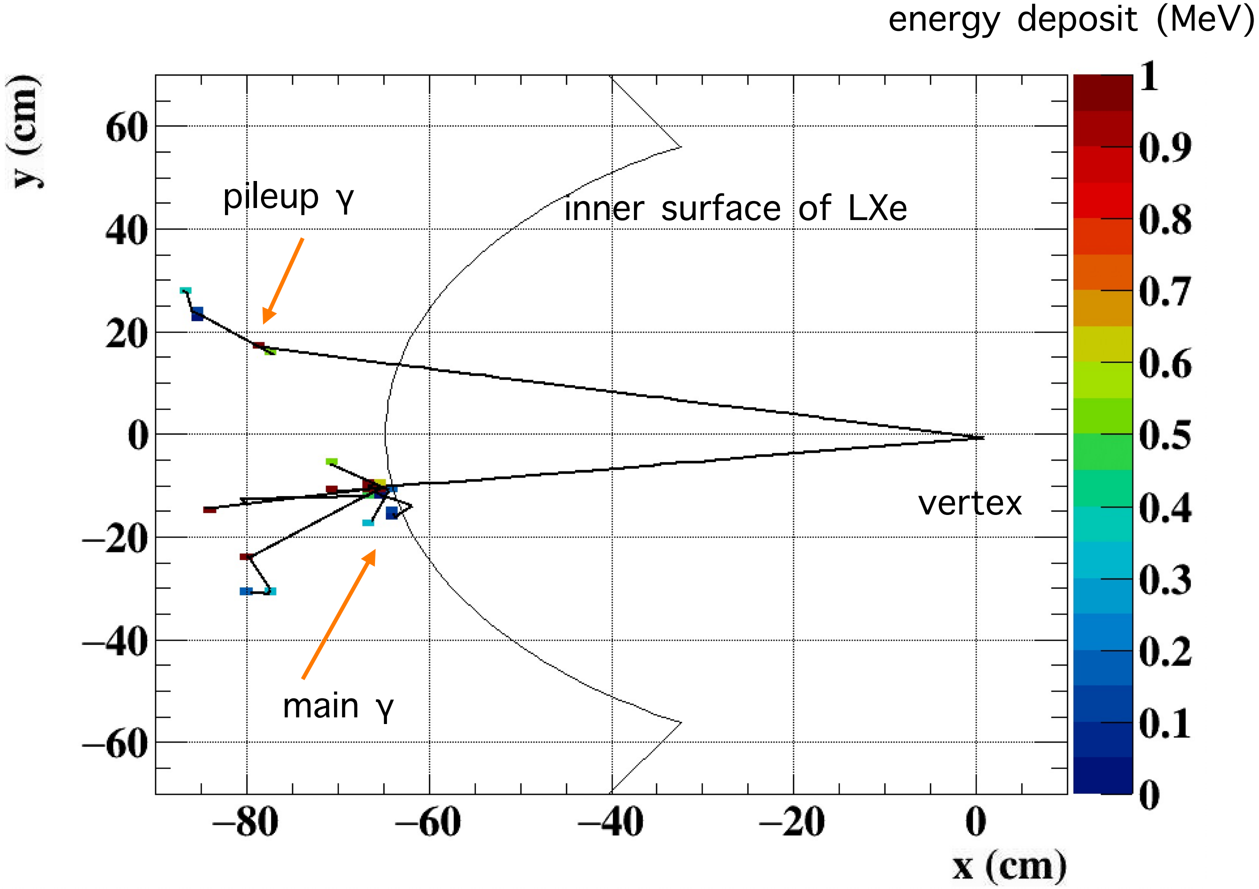


Example of Event

Main γ + 1 pileup γ

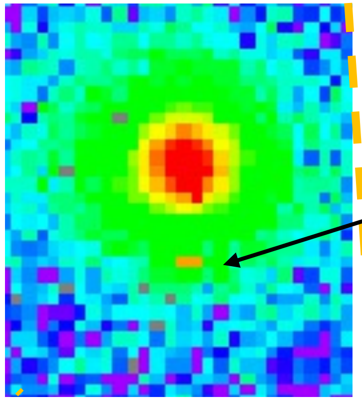


- The peak search cannot find the pileup γ-ray
  ← Smaller than the threshold due to the deep conversion position.
- The DL model estimates the probability to include pileups as 0.83.



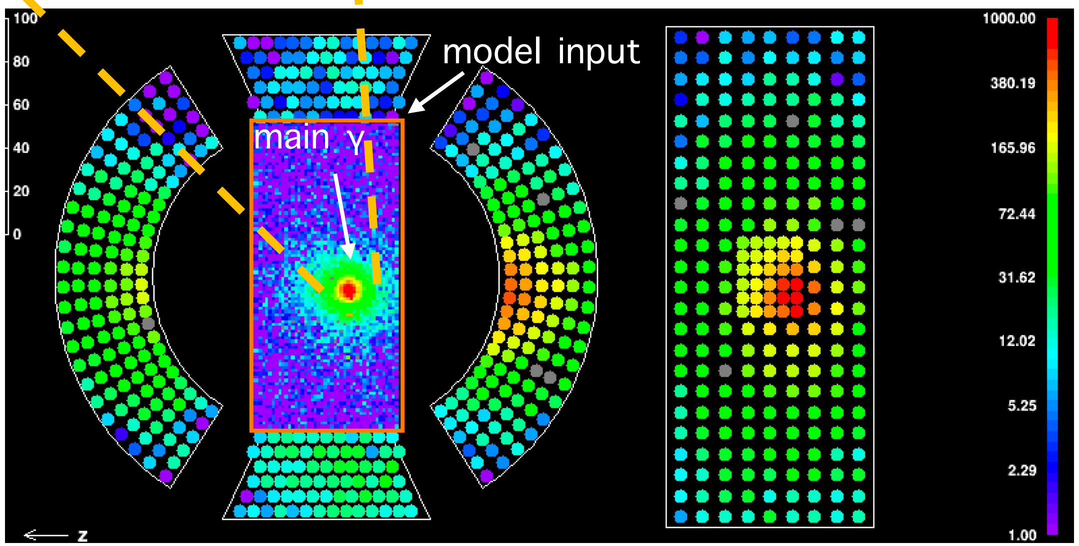
iiner面とかの説明

Example of Event

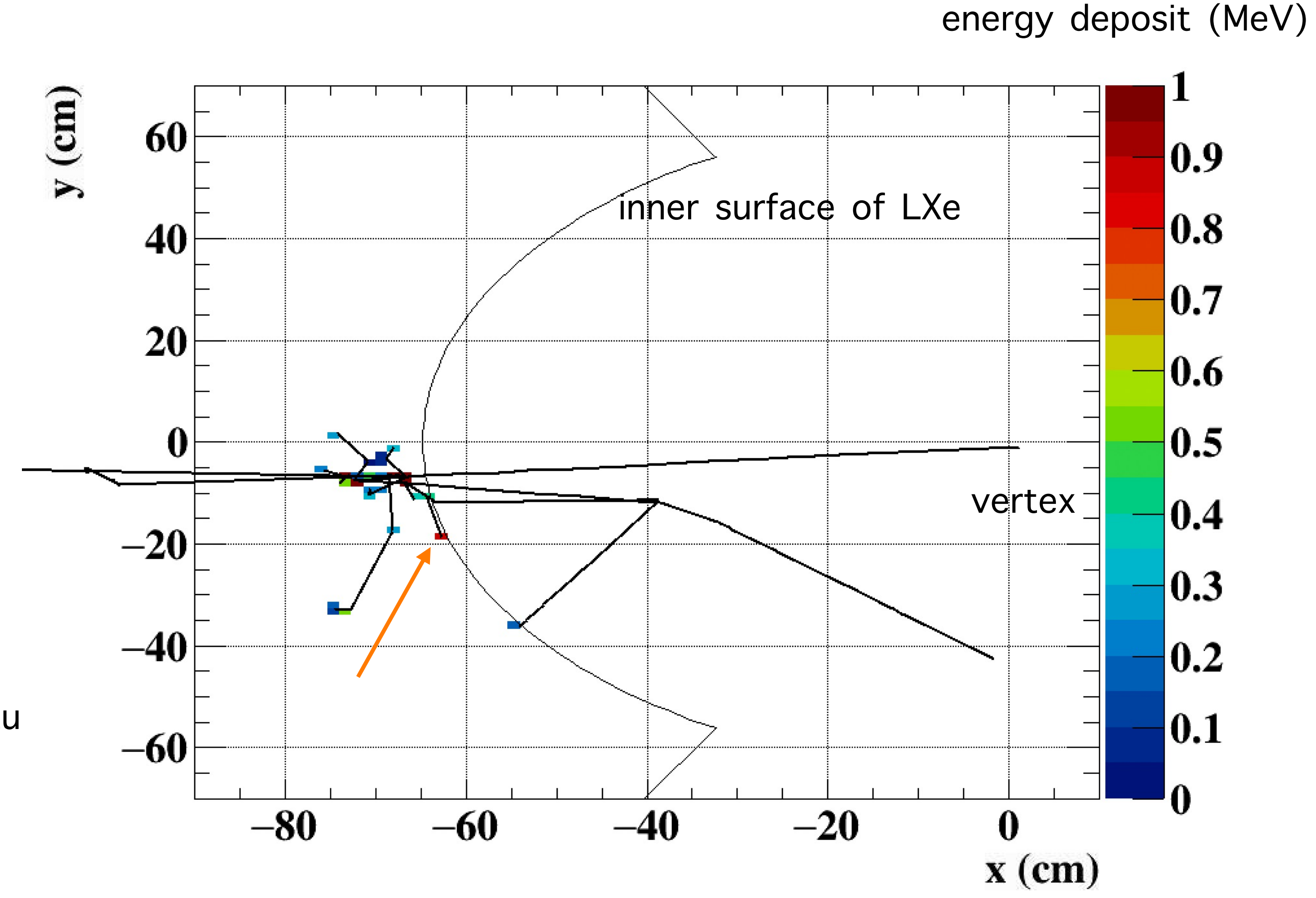


recognized as a pileup γ-ray

Only Main γ



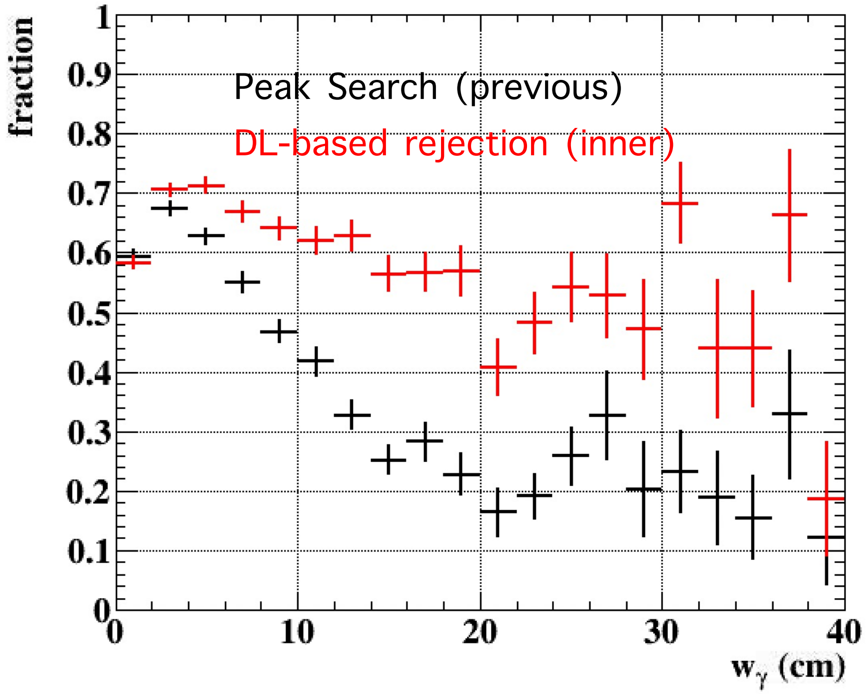
- The peak search regards the fake peak due to a shower fluctu or noise as the pileup γ-ray.
  ← Larger than the threshold.
- The DL model estimates the probability to include pileups as 0.11.



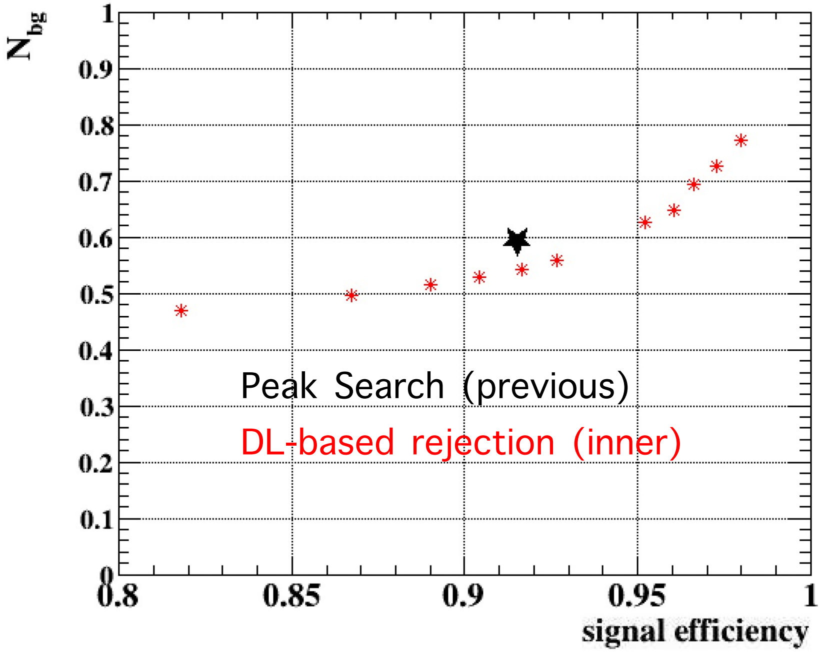
A gamma-ray escaped from the shower converted around the fake peak.

Performance of DL-based Pileup Identification

Depth Dependence of Found Pileup Fraction



signal efficiency v.s. N_{bg} (52-54 MeV)



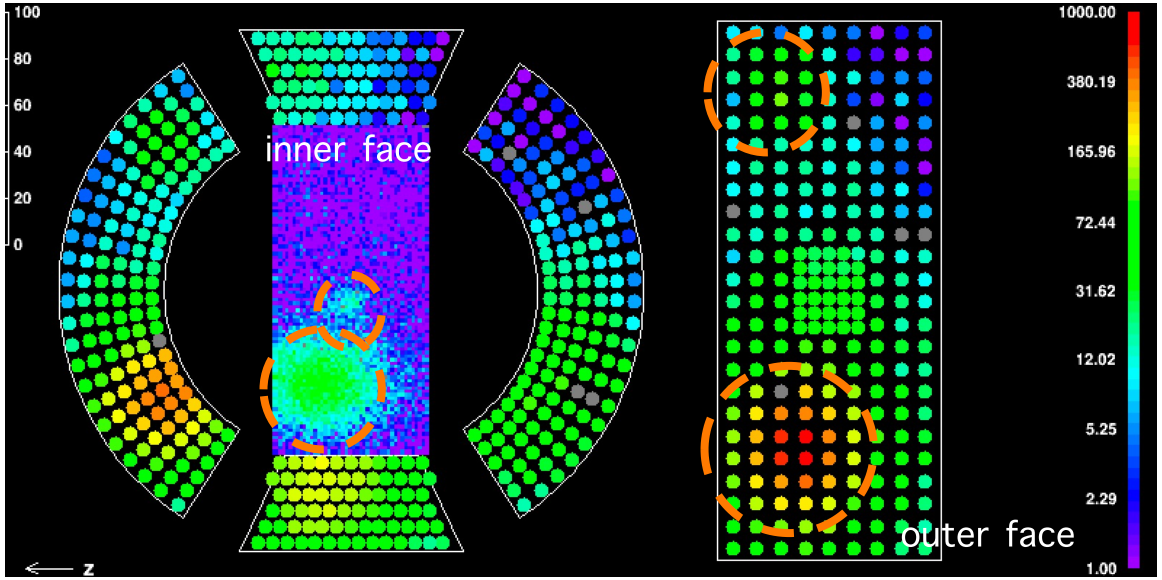
*Different points correspond to different thresholds.

- The DL-based algorithm achieved the higher pileup rejection efficiency especially in deeper region.
  ←A peak structure is not required by utilizing the global distribution.
- The number of backgrounds N_{bg} decrease by 5% at the same signal efficiency.
  ←Higher detection efficiency and tolerance to the fake peak.

Step2: Peak Search & Clustering in Charge Distribution

- Two peak search and clustering methods are implemented.
- One is based on a charge distribution.
    1. Peak search is performed on the inner/outer face.
    2. The channel at the center of the found peak is assigned to a cluster.
    3. The neighboring channels whose charges are larger than a threshold are added to the same cluster.
- The on-timing pileup γ-rays entering can be found.

Example of Charge Distribution



Step2: Peak Search & Clustering in Charge Distribution

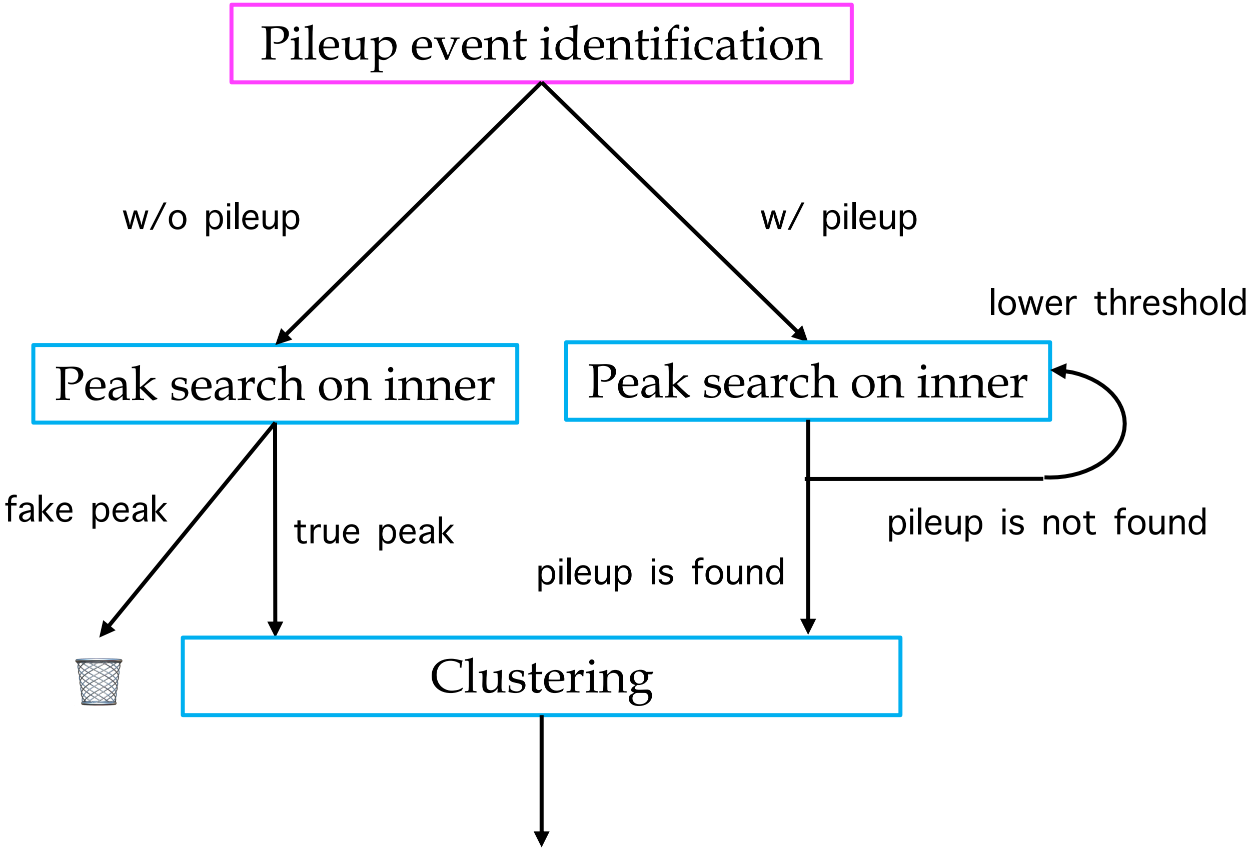
The information whether the event likely has pileup γ-rays from the DL model was used to **switch the peak search method** in the charge distribution on the inner face.

- **”w/o pileup”**:
  the peak search with the nominal threshold, and **peaks with small energies are discarded**.

← Tolerant to the fake peaks.

- **“w/ pileup”**:
  the peak search **reducing the threshold until a pileup γ-ray is found**.

← The deeper events can be found with the lower threshold.



Step2: Peak Search & Clustering in Timing Distribution

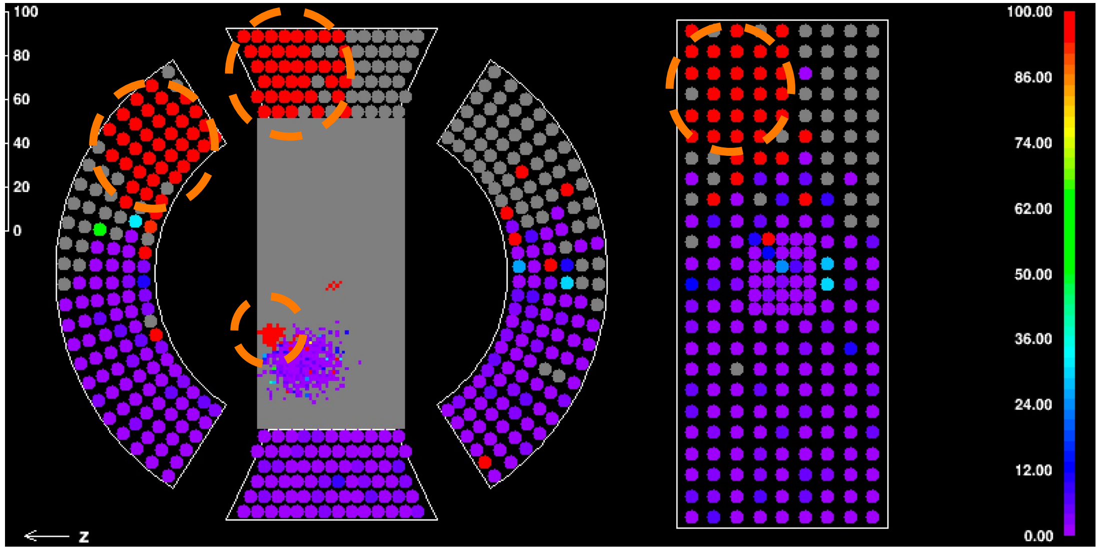
- The other is based on a timing χ^2 distribution.
- The χ^2 of i-th channel is defined as

$$\chi_i^2 = \frac{(t_\gamma - t_i)^2}{\sigma_i^2},$$

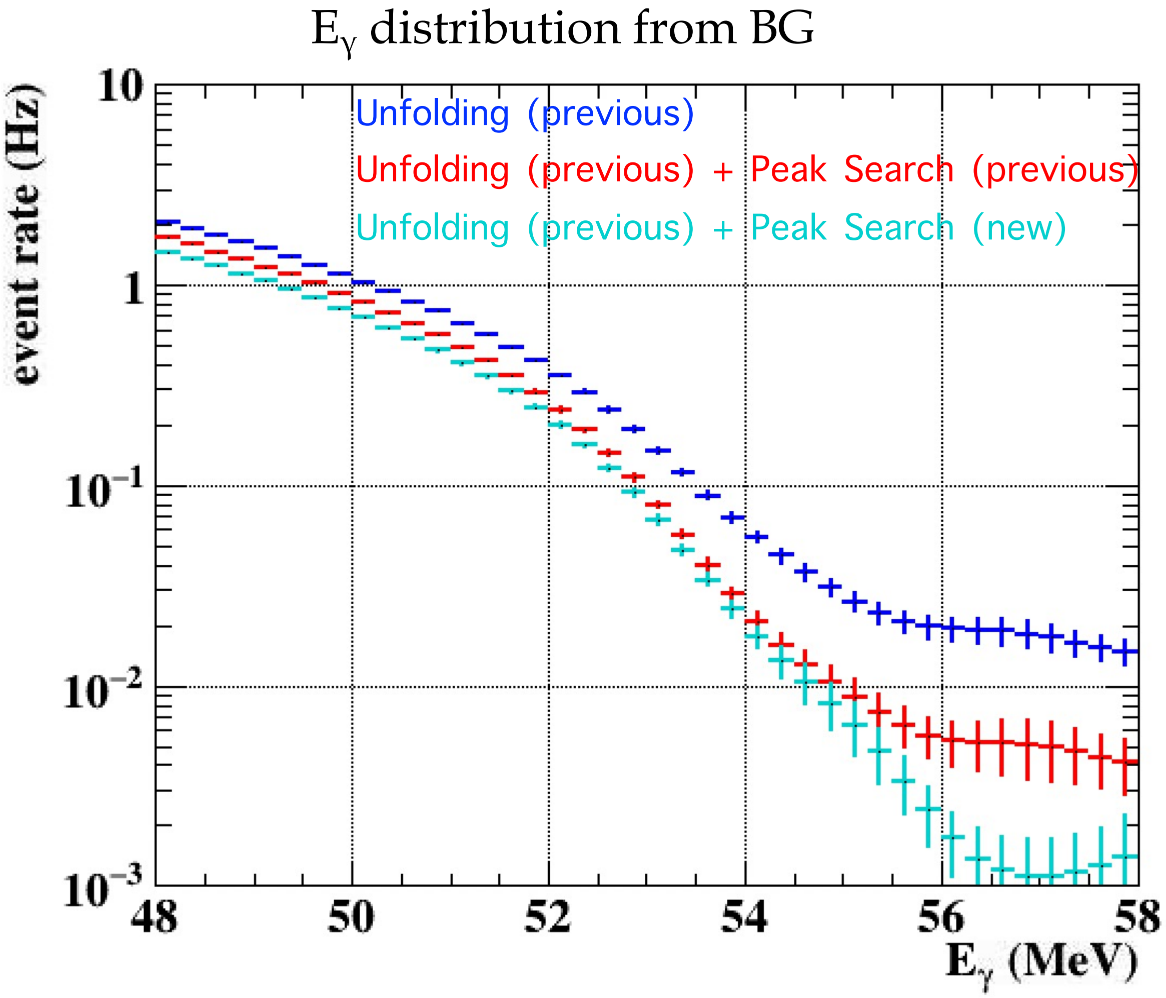
  where t_γ is the reconstructed γ timing, and t_i and σ_i are the timing and its uncertainty of the channel.

- The clustering is performed as follows:
  1. **Find channels whose χ^2 are larger than a threshold**.
  2. One of the found channels is assigned to a cluster.
  3. The neighboring channels whose χ^2 are larger than a threshold are added to the same cluster.
- The off-timing pileup γ-rays entering can be found.

Timing χ^2



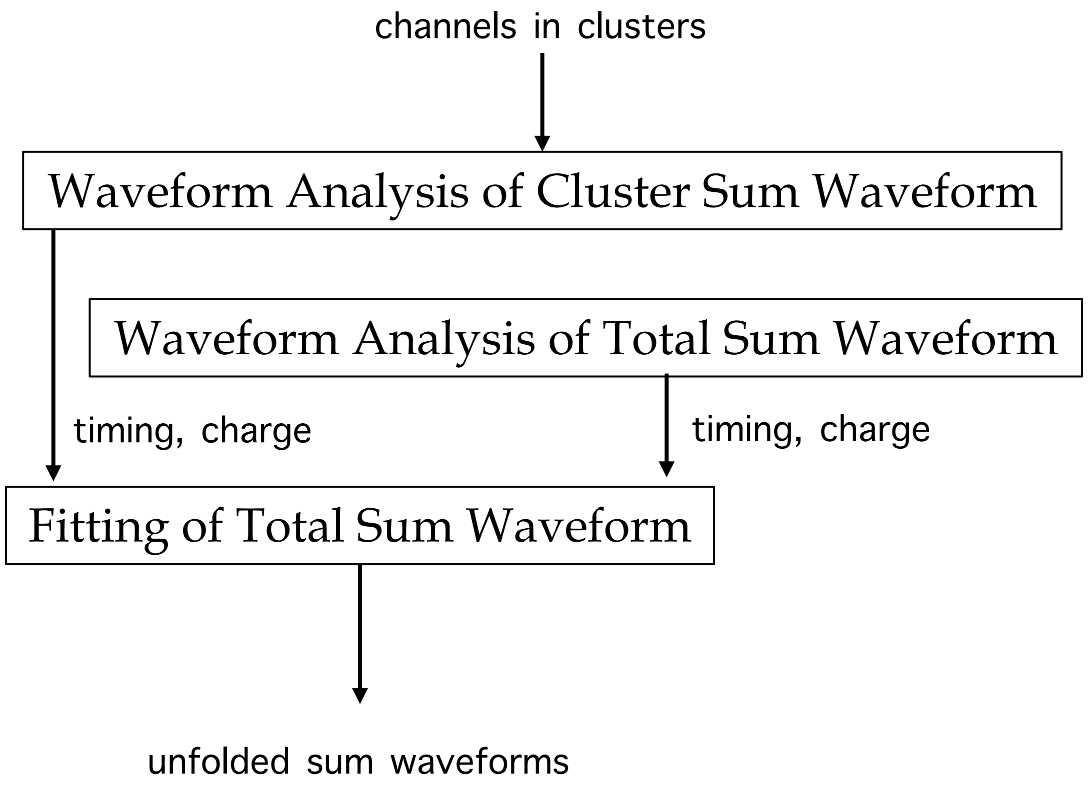
Performance of Peak Search & Clustering



- The events the peak search methods find more than one γ are rejected.
- **The new peak search algorithm can find more pileup γ-rays** compared to the previous one.

Step3: Unfolding of Sum Waveform

- The pileup unfolding in sum waveform was developed.
- Two types of sum waveforms are generated:
  - **Total sum waveform**: All MPPCs/PMTs
  - **Cluster sum waveform**: MPPCs/PMTs belonging to each cluster generated by the clustering
- Pulse timings and charges are extracted from the sum waveforms.
- **Template waveforms are fit to the total sum waveforms** using the timings and charges as initial values.



channels in clusters

Waveform Analysis of Cluster Sum Waveform

Waveform Analysis of Total Sum Waveform

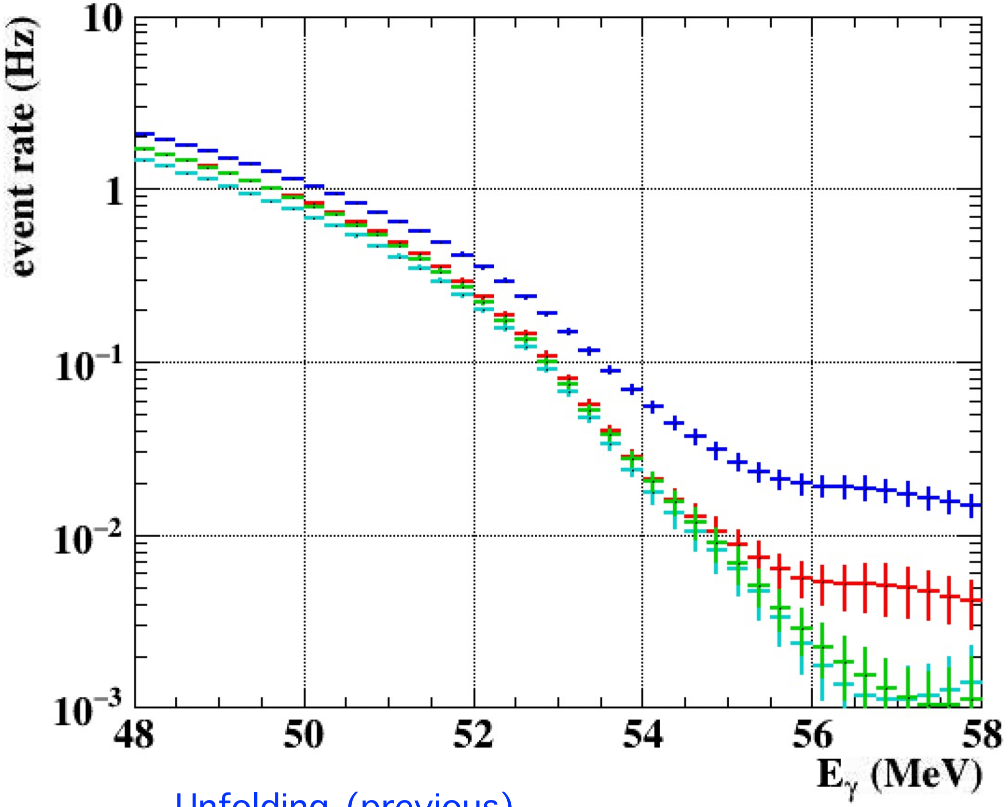
timing, charge

timing, charge

Fitting of Total Sum Waveform

unfolded sum waveforms

Performance of Unfolding



E_γ distribution from BG

Unfolding (previous)

Unfolding (previous) + Peak Search (previous)

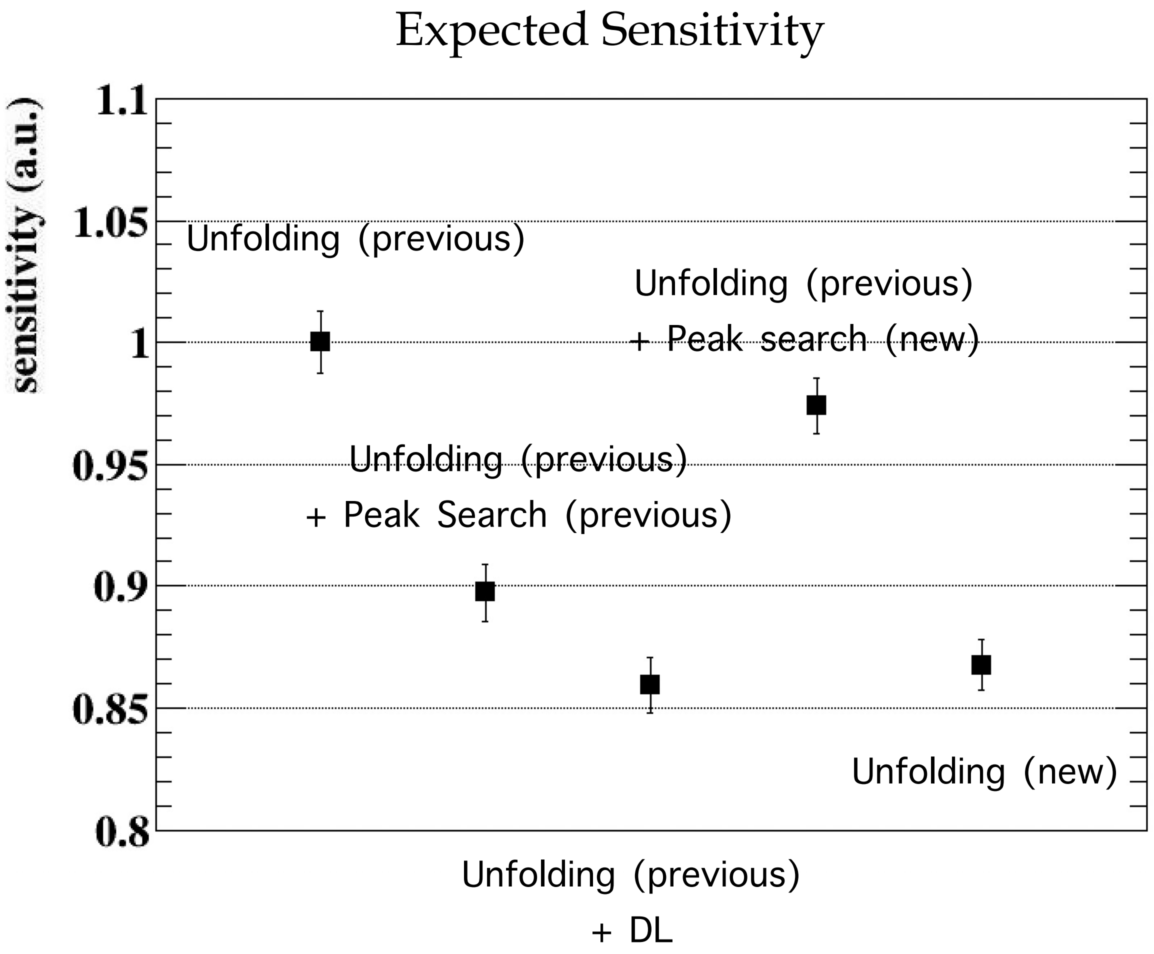
Unfolding (previous) + Peak Search (new)

Unfolding (previous)

- The unfolding **recovers the signal efficiency by 10%.**
  ↔ **backgrounds increases by 5%.**
- As a result, **4% less backgrounds at the same signal efficiency** was achieved compared to the previous one.

E_γ = 52-54 MeV	N_{bg}	Signal efficiency
Unfolding (previous)	1	1
Unfolding (previous) + Peak Search (previous)	0.59	0.91
Unfolding (previous) + Peak Search (new)	0.50	0.81
Unfolding (new)	0.55	0.91

Effect on Sensitivity



- The effect of the sensitivity was investigated.
- The new algorithm **improves the sensitivity by 3%** compared to the unfolding + peak search (previous).
- The sensitivity of DL-based rejection is equivalent to that of the new unfolding method.

Summary

- The MEG II experiment searches μ→ eγ decay.
- The pileup analysis for the LXe detector is important to reduce the γ-ray background events in the signal region.
- The new algorithm for the pileup analysis was developed.
- It consists of three steps:
    1. Pileup event identification with DL-based algorithm
    2. Peak search and clustering of channels in charge and timing distributions
    3. Unfolding of sum waveform
- The new algorithm was found to **improve the sensitivity by 3%** compared to the previous algorithm.

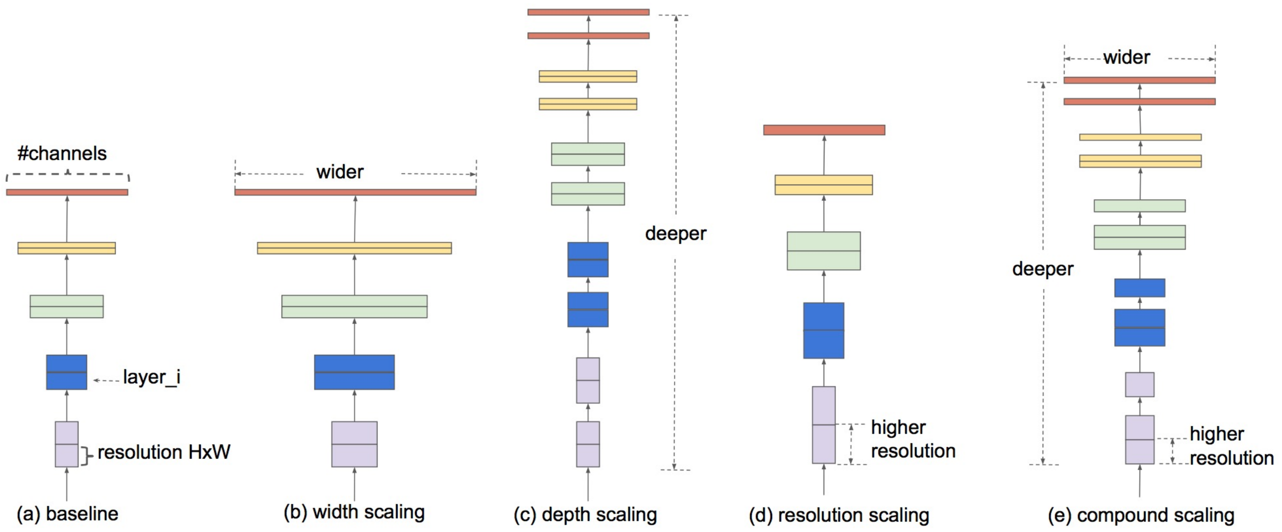
Prospect

- The performance must be evaluated in more realistic situation.
  - Decreasing MPPC PDE due to radiation damage
  - Coherent noise among channels
  - Existence of dead channels
  - Precision of calibration
- It also must be evaluated with data.

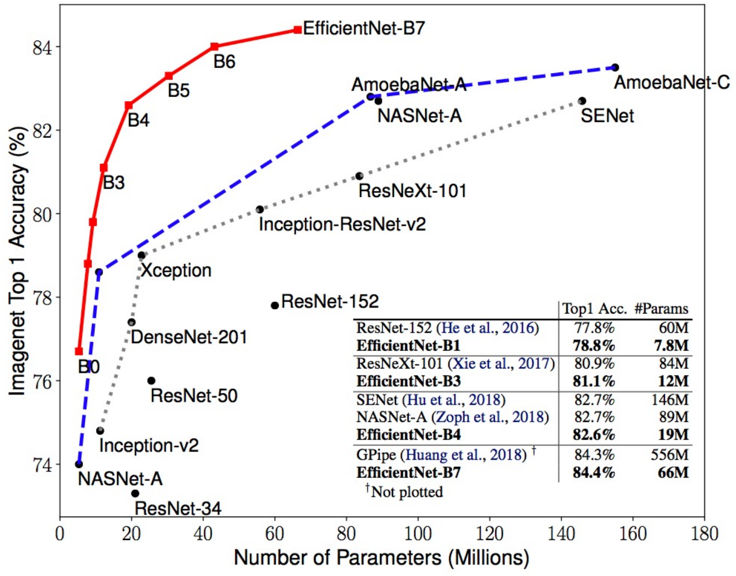
Backup Slides

EfficientNet

- A type of CNN
- https://arxiv.org/pdf/1905.11946.pdf
- The performance of DL models can be improved by scaling up the original model.
- The optimal scaling method was investigated, and they introduced the efficiently scaled models.
  → A better performance with less parameters was achieved compared to other models.



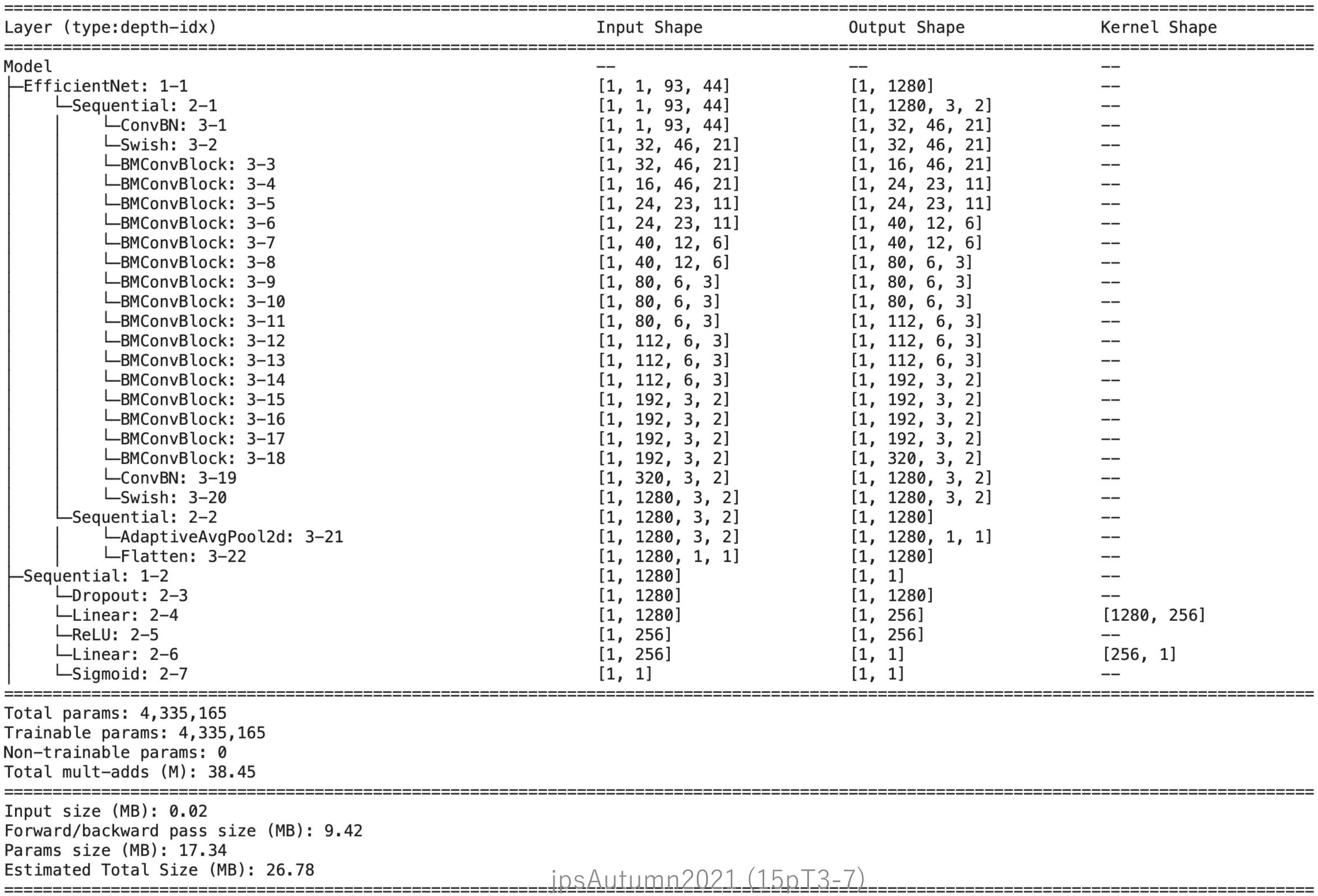
(a) baseline (b) width scaling (c) depth scaling (d) resolution scaling (e) compound scaling



	Top1 Acc.	#Params
ResNet-152 (He et al., 2016)	77.8%	60M
EfficientNet-B1	**78.8%**	**7.8M**
ResNeXt-101 (Xie et al., 2017)	80.9%	84M
EfficientNet-B3	**81.1%**	**12M**
SENet (Hu et al., 2018)	82.7%	146M
NASNet-A (Zoph et al., 2018)	82.7%	89M
EfficientNet-B4	**82.6%**	**19M**
GPipe (Huang et al., 2018) †	84.3%	556M
EfficientNet-B7	**84.4%**	**66M**

†Not plotted

Model Architecture



```
======================================================================================================================================
Layer (type:depth-idx)                                      Input Shape               Output Shape              Kernel Shape
======================================================================================================================================
Model                                                       --                        --                        --
├─EfficientNet: 1-1                                         [1, 1, 93, 44]            [1, 1280]                 --
│    └─Sequential: 2-1                                      [1, 1, 93, 44]            [1, 1280, 3, 2]           --
│    │    └─ConvBN: 3-1                                     [1, 1, 93, 44]            [1, 32, 46, 21]           --
│    │    └─Swish: 3-2                                      [1, 32, 46, 21]           [1, 32, 46, 21]           --
│    │    └─BMConvBlock: 3-3                                [1, 32, 46, 21]           [1, 16, 46, 21]           --
│    │    └─BMConvBlock: 3-4                                [1, 16, 46, 21]           [1, 24, 23, 11]           --
│    │    └─BMConvBlock: 3-5                                [1, 24, 23, 11]           [1, 24, 23, 11]           --
│    │    └─BMConvBlock: 3-6                                [1, 24, 23, 11]           [1, 40, 12, 6]            --
│    │    └─BMConvBlock: 3-7                                [1, 40, 12, 6]            [1, 40, 12, 6]            --
│    │    └─BMConvBlock: 3-8                                [1, 40, 12, 6]            [1, 80, 6, 3]             --
│    │    └─BMConvBlock: 3-9                                [1, 80, 6, 3]             [1, 80, 6, 3]             --
│    │    └─BMConvBlock: 3-10                               [1, 80, 6, 3]             [1, 80, 6, 3]             --
│    │    └─BMConvBlock: 3-11                               [1, 80, 6, 3]             [1, 112, 6, 3]            --
│    │    └─BMConvBlock: 3-12                               [1, 112, 6, 3]            [1, 112, 6, 3]            --
│    │    └─BMConvBlock: 3-13                               [1, 112, 6, 3]            [1, 112, 6, 3]            --
│    │    └─BMConvBlock: 3-14                               [1, 112, 6, 3]            [1, 192, 3, 2]            --
│    │    └─BMConvBlock: 3-15                               [1, 192, 3, 2]            [1, 192, 3, 2]            --
│    │    └─BMConvBlock: 3-16                               [1, 192, 3, 2]            [1, 192, 3, 2]            --
│    │    └─BMConvBlock: 3-17                               [1, 192, 3, 2]            [1, 192, 3, 2]            --
│    │    └─BMConvBlock: 3-18                               [1, 192, 3, 2]            [1, 320, 3, 2]            --
│    │    └─ConvBN: 3-19                                    [1, 320, 3, 2]            [1, 1280, 3, 2]           --
│    │    └─Swish: 3-20                                     [1, 1280, 3, 2]           [1, 1280, 3, 2]           --
│    └─Sequential: 2-2                                      [1, 1280, 3, 2]           [1, 1280]                 --
│    │    └─AdaptiveAvgPool2d: 3-21                         [1, 1280, 3, 2]           [1, 1280, 1, 1]           --
│    │    └─Flatten: 3-22                                   [1, 1280, 1, 1]           [1, 1280]                 --
├─Sequential: 1-2                                           [1, 1280]                 [1, 1]                    --
│    └─Dropout: 2-3                                         [1, 1280]                 [1, 1280]                 --
│    └─Linear: 2-4                                          [1, 1280]                 [1, 256]                  [1280, 256]
│    └─ReLU: 2-5                                            [1, 256]                  [1, 256]                  --
│    └─Linear: 2-6                                          [1, 256]                  [1, 1]                    [256, 1]
│    └─Sigmoid: 2-7                                         [1, 1]                    [1, 1]                    --
======================================================================================================================================
Total params: 4,335,165
Trainable params: 4,335,165
Non-trainable params: 0
Total mult-adds (M): 38.45
======================================================================================================================================
Input size (MB): 0.02
Forward/backward pass size (MB): 9.42
Params size (MB): 17.34
Estimated Total Size (MB): 26.78
======================================================================================================================================
```

Data Pre-processing

- Dead channel recovery

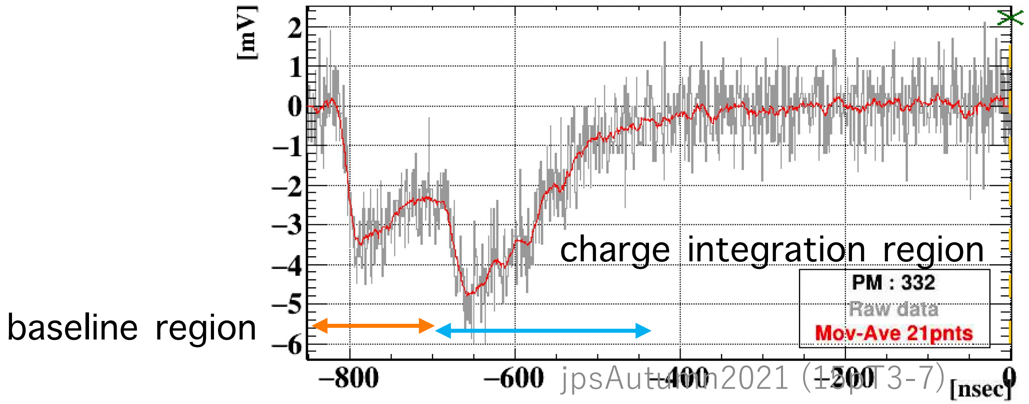
  Values of dead channels are estimated by the mean of surroundings.

- Normalization

  Normalized by the maximum value, i.e. all input values are no more than 1.
  ← Suppress the energy dependence

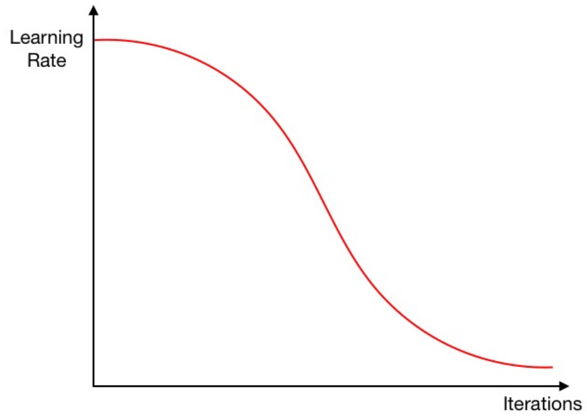
- Cut off

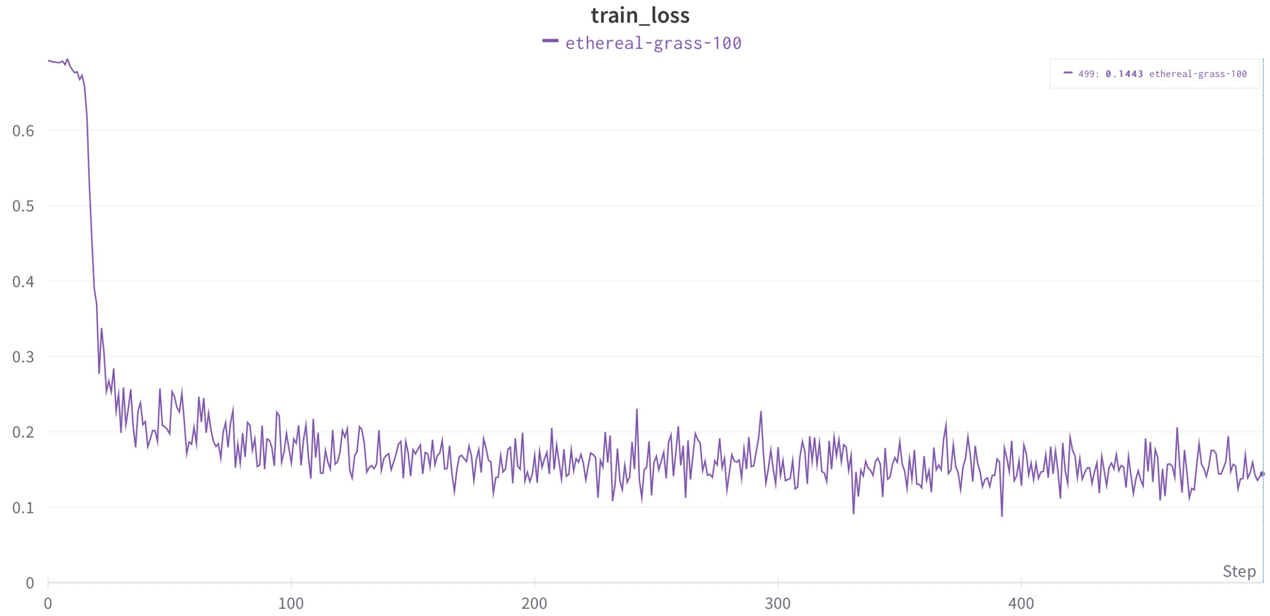
  Negative charges are set to 0, i.e. all input values are no less than 0.
  ← Due to a failure of the baseline calculation



Training Details

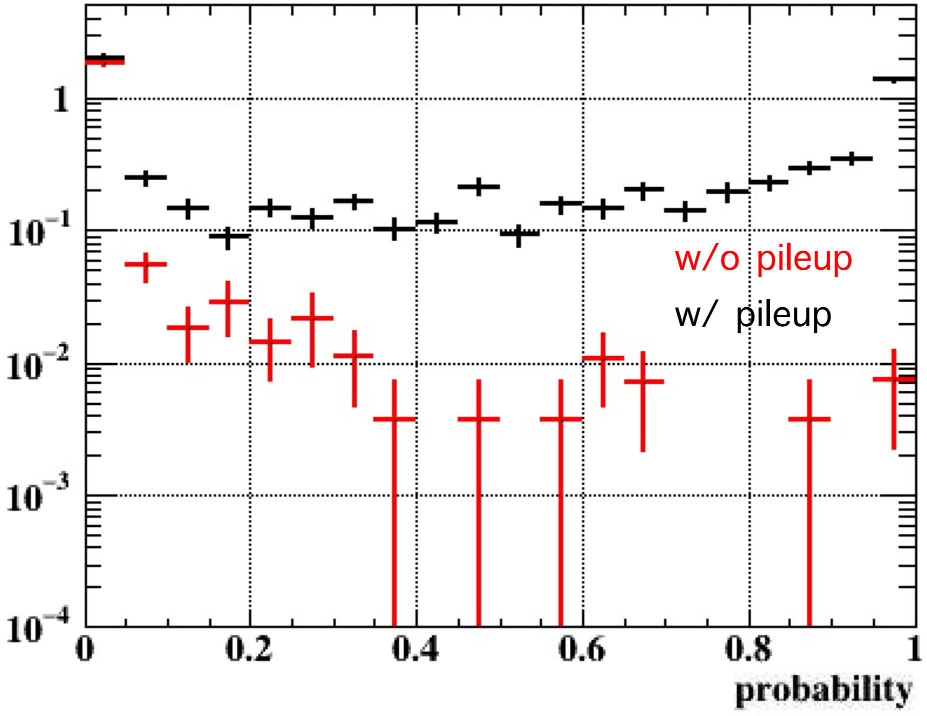
- Optimizer: SGD, lr=0.01
- Loss: Binary cross entropy
- Scheduler: CosineAnnealing(max_T=500)
- Batch size: 200
- n_epochs: 500


CosineAnnealing

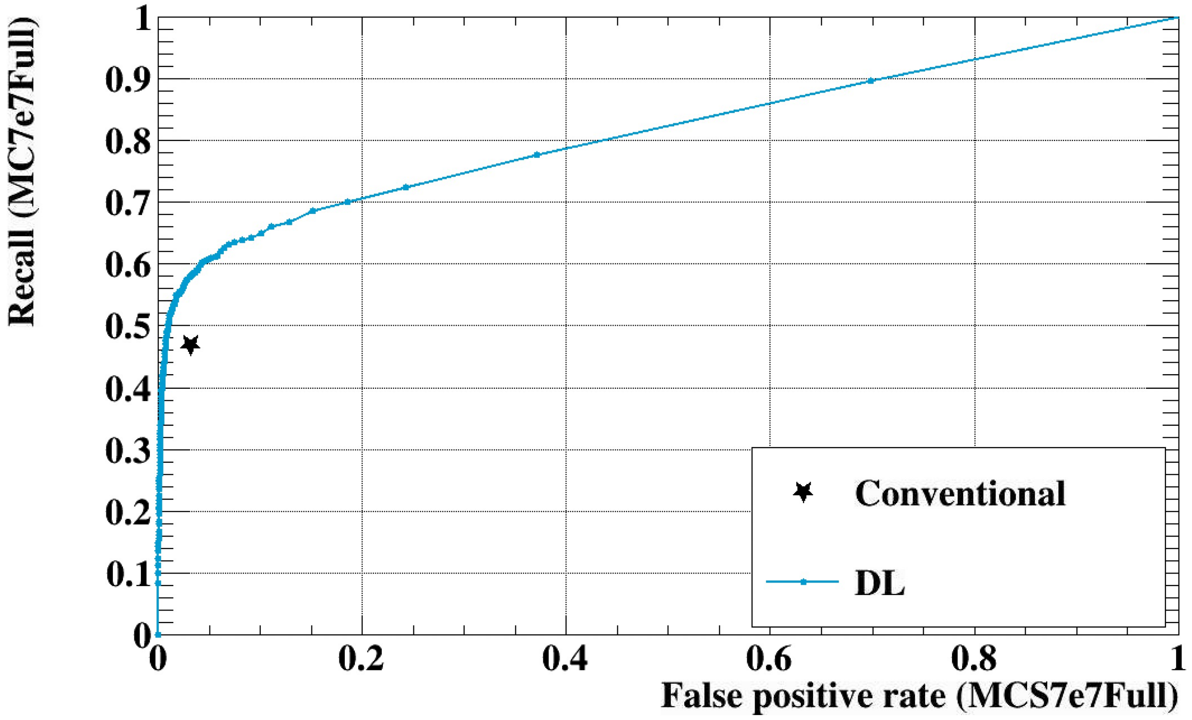


DL Model Performance

Output (BG)



ROC

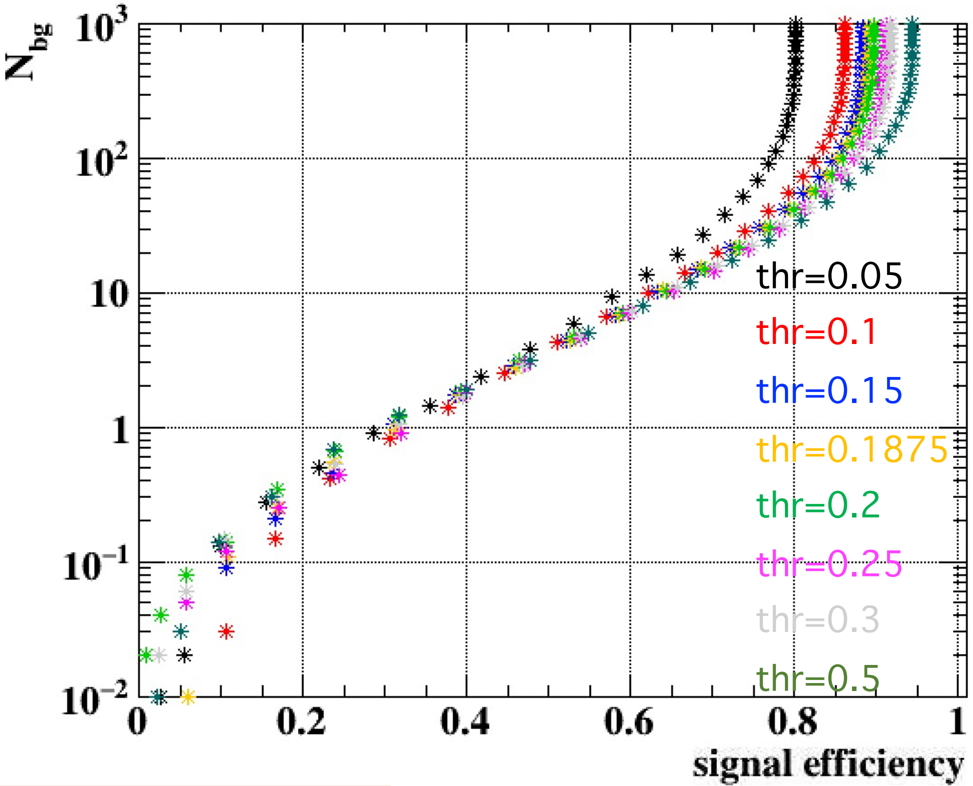


- The outputs of the DL model are well separated.
- The DL model (inner) has better recall at the same FPR point.

Optimization of Threshold

- Threshold scan was performed by defining “signal box” with R_{sig}.
- R_{sig} is defined for each event as:

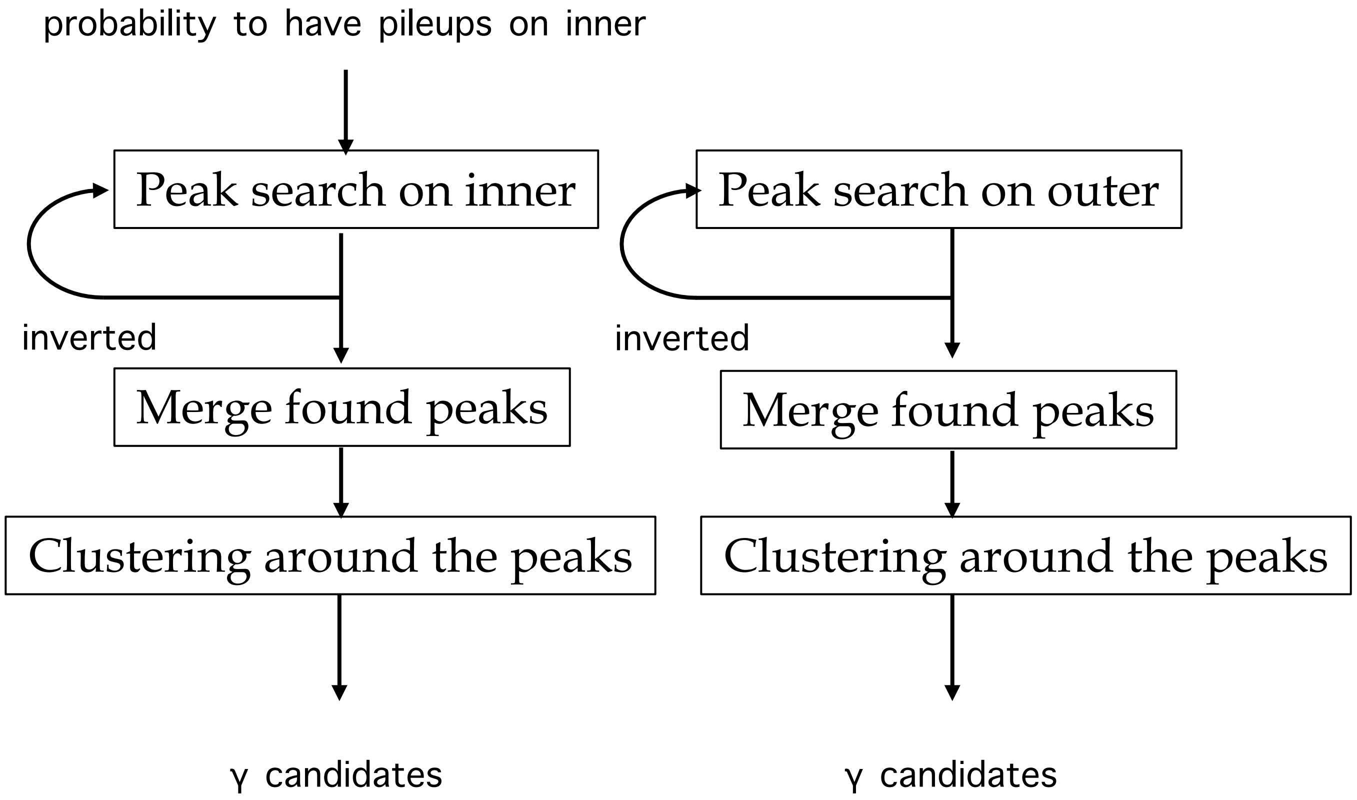
  $R_{sig} = \log(L_{signal}(x)/L_{bg}(x))$, where x is MEG observables.
- N_{bg} is the least at thr=0.25 up to signal efficiency of 70% except for 40% point.



Signal Efficiency	N_{bg}							
	thr=0.05	thr=0.10	thr=0.15	thr=0.1875	thr=0.20	thr=0.25	thr=0.3	thr=0.5
0.3	1.01	0.79	0.95	0.89	1.08	0.79	0.94	1.11
0.4	2.13	1.77	1.98	1.88	2.00	1.82	1.80	1.91
0.5	4.65	3.95	3.86	3.73	4.01	3.62	3.70	3.73
0.6	11.53	8.42	8.07	7.77	7.66	7.22	7.22	7.27
0.7	31.77	19.00	18.00	17.44	16.36	14.40	15.29	14.64
0.8	357.68	61.66	48.48	44.77	41.84	37.61	36.44	32.86

Algorithm Flow in Peak Search & Clustering

Search by light distribution


probability to have pileups on inner

Peak search on inner

inverted

Merge found peaks

Clustering around the peaks

γ candidates

Peak search on outer

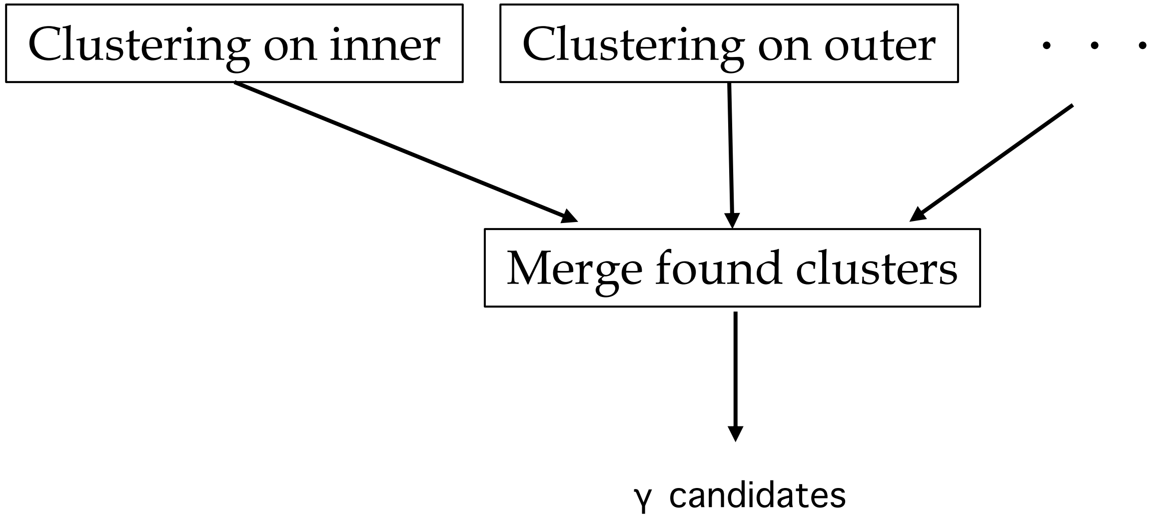
inverted

Merge found peaks

Clustering around the peaks

γ candidates

Search by time distribution


Clustering on inner

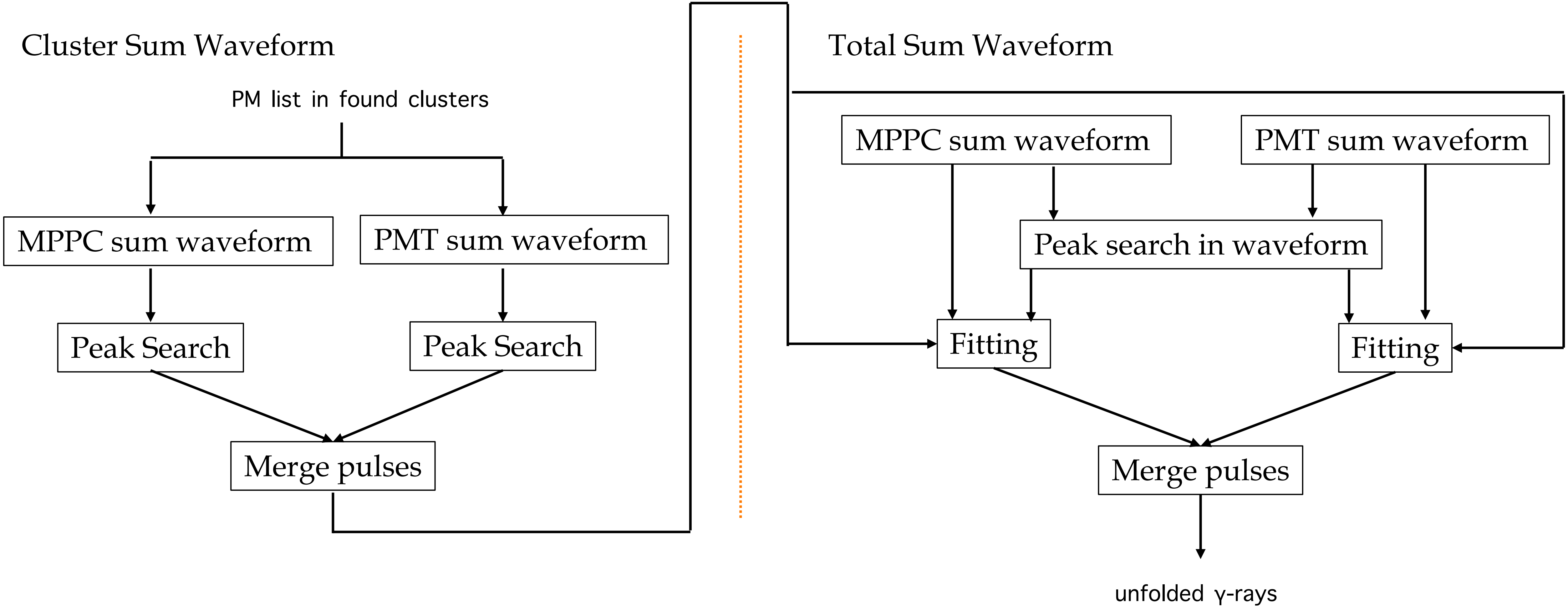
Clustering on outer

・・・

Merge found clusters

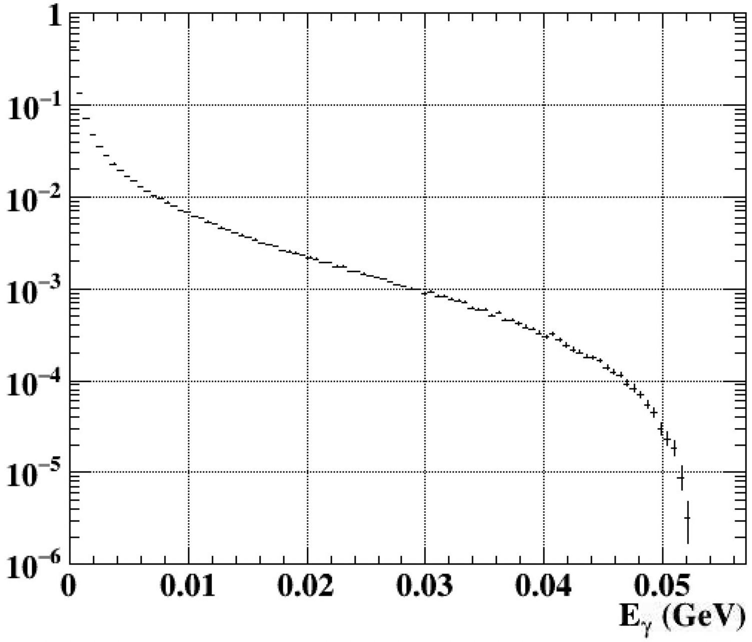
γ candidates

Algorithm Flow in Sum Waveform Unfolding

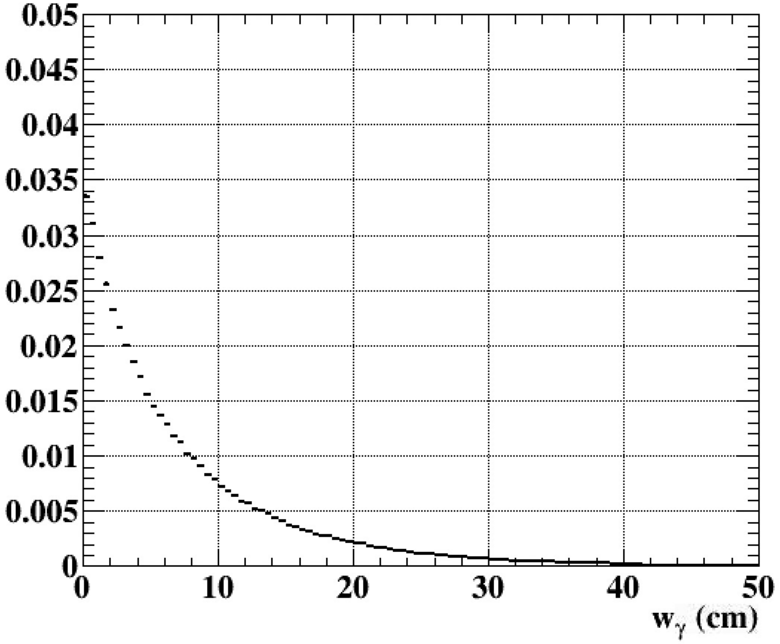


Pileup γ-rays

E_γ distribution of Pileup

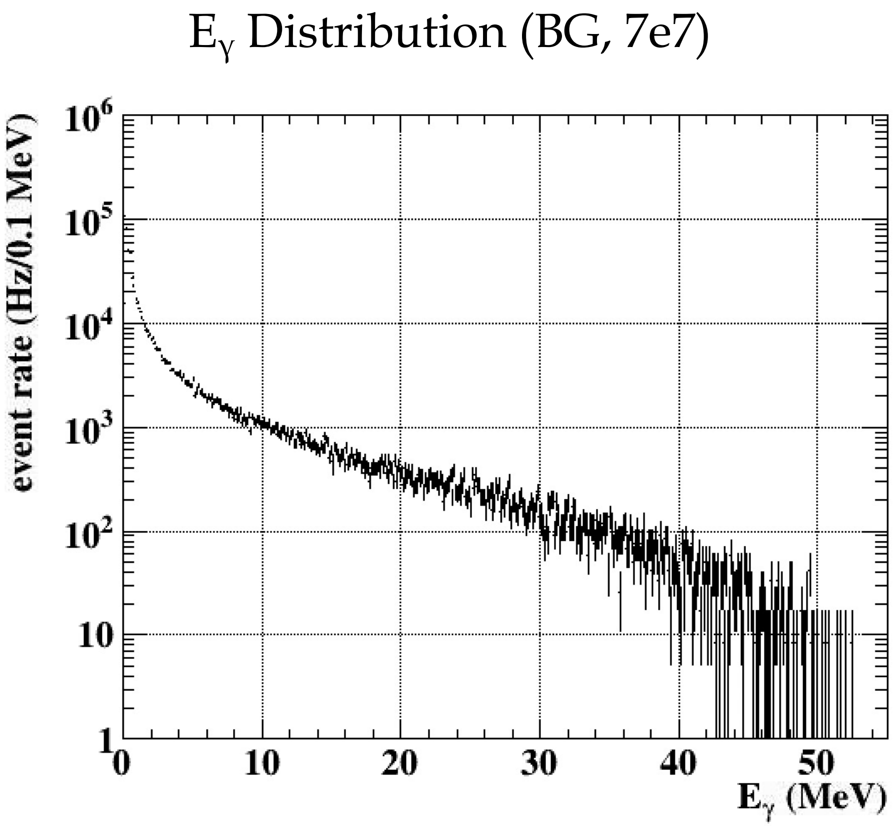


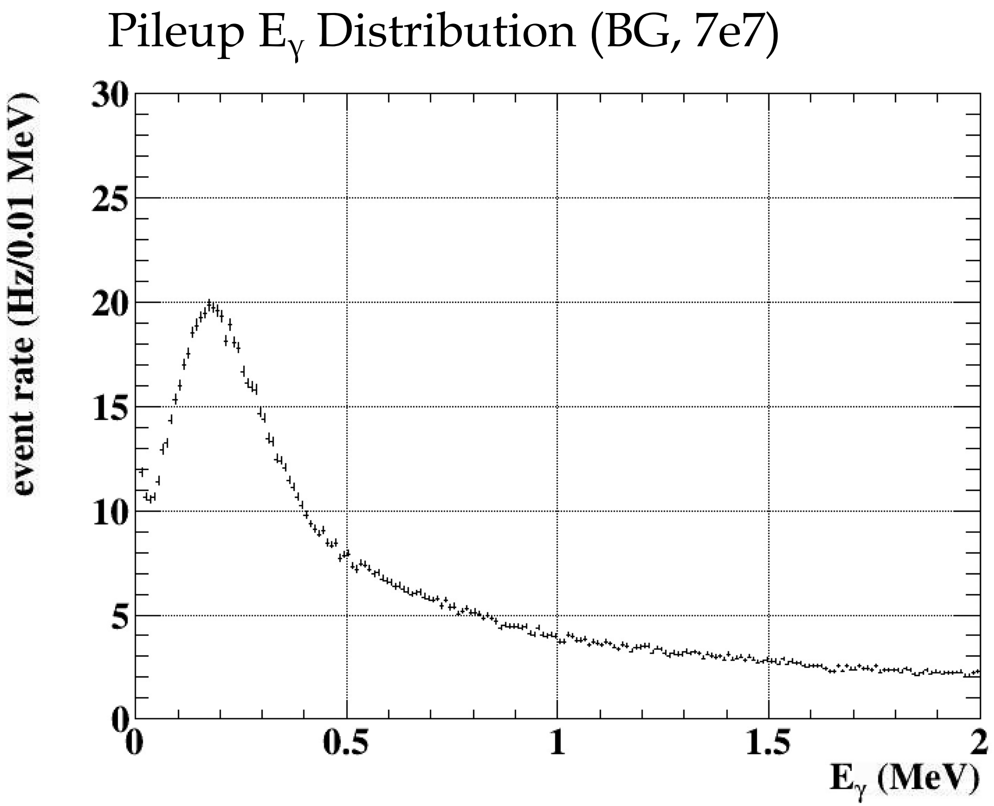
w_γ distribution of Pileup



- A small energy and a shallow conversion point are dominant.

Pileup γ-rays



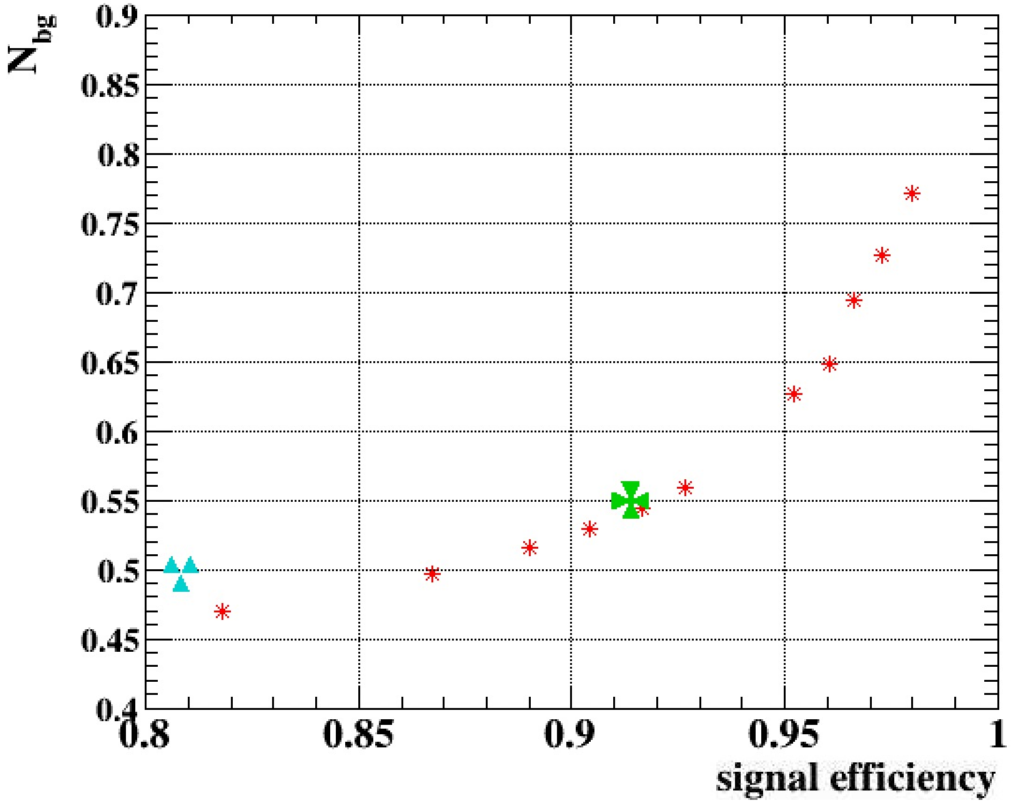


The pileup analysis can find pileups whose energies are more than 0.2 MeV.

The event rate of γ-ray hits for $E_\gamma > 0.2$ MeV is 0.7 MHz.

$(|u| < 25 \text{ cm}) \wedge (|v| < 71 \text{ cm}) \wedge (0 < w < 38.5 \text{ cm})$

- a


N_{bg}
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.8
0.85
0.9
0.95
1
signal efficiency