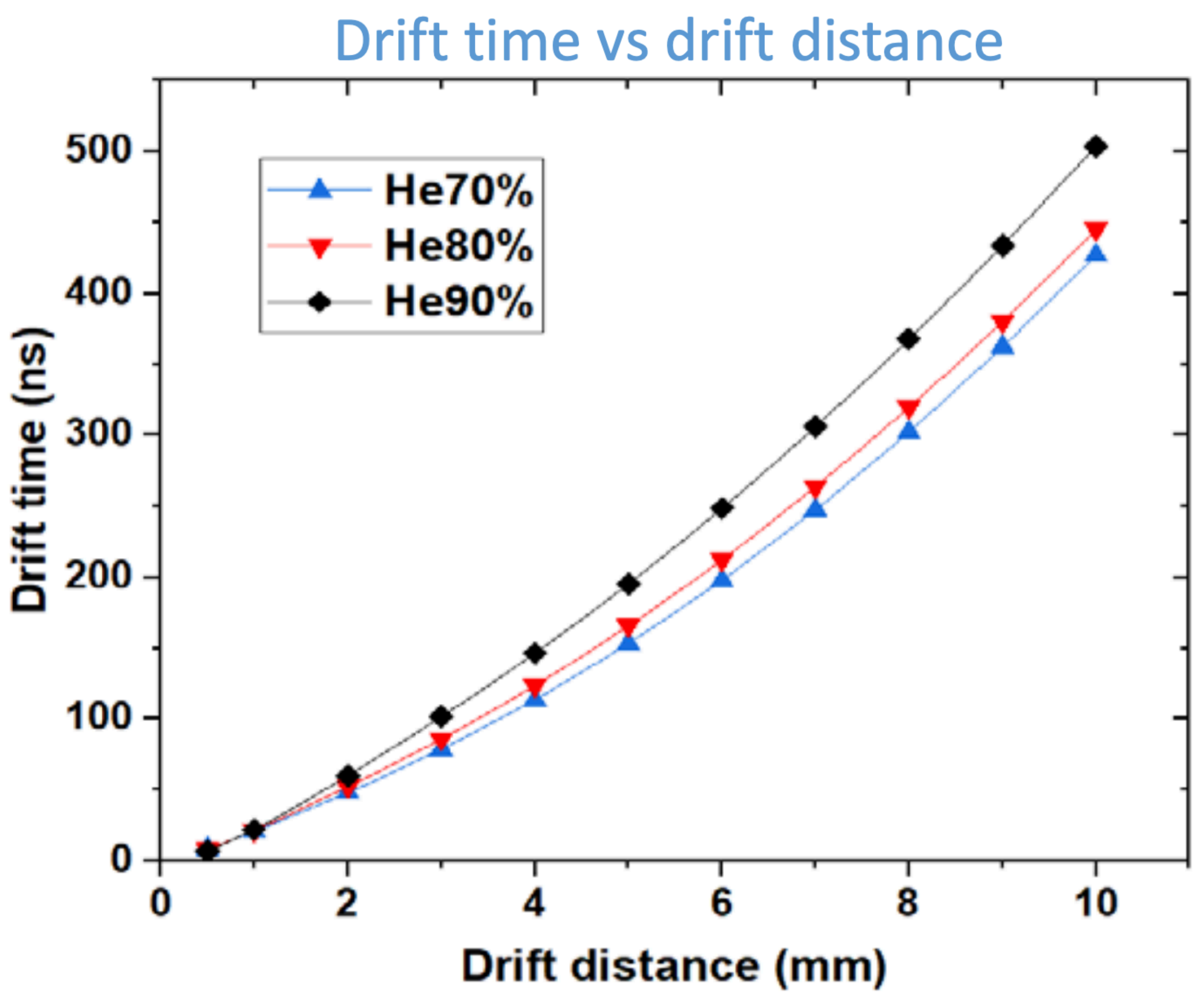
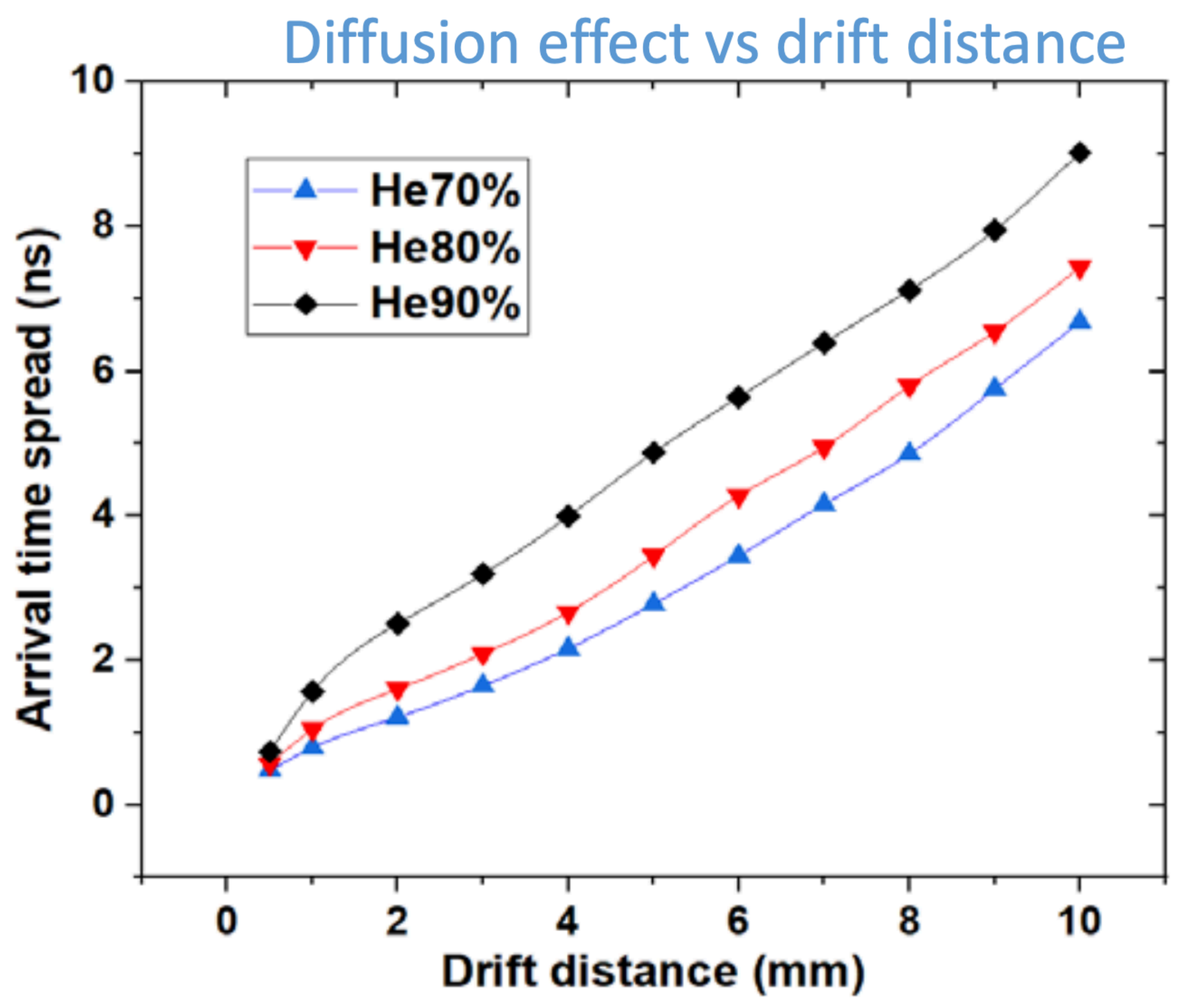
Clustering approach for waveform MC

Shuiting Xin, Guang Zhao, Linghui Wu, Mingyi Dong
xinshuiting@ihep.ac.cn
09/March/2023

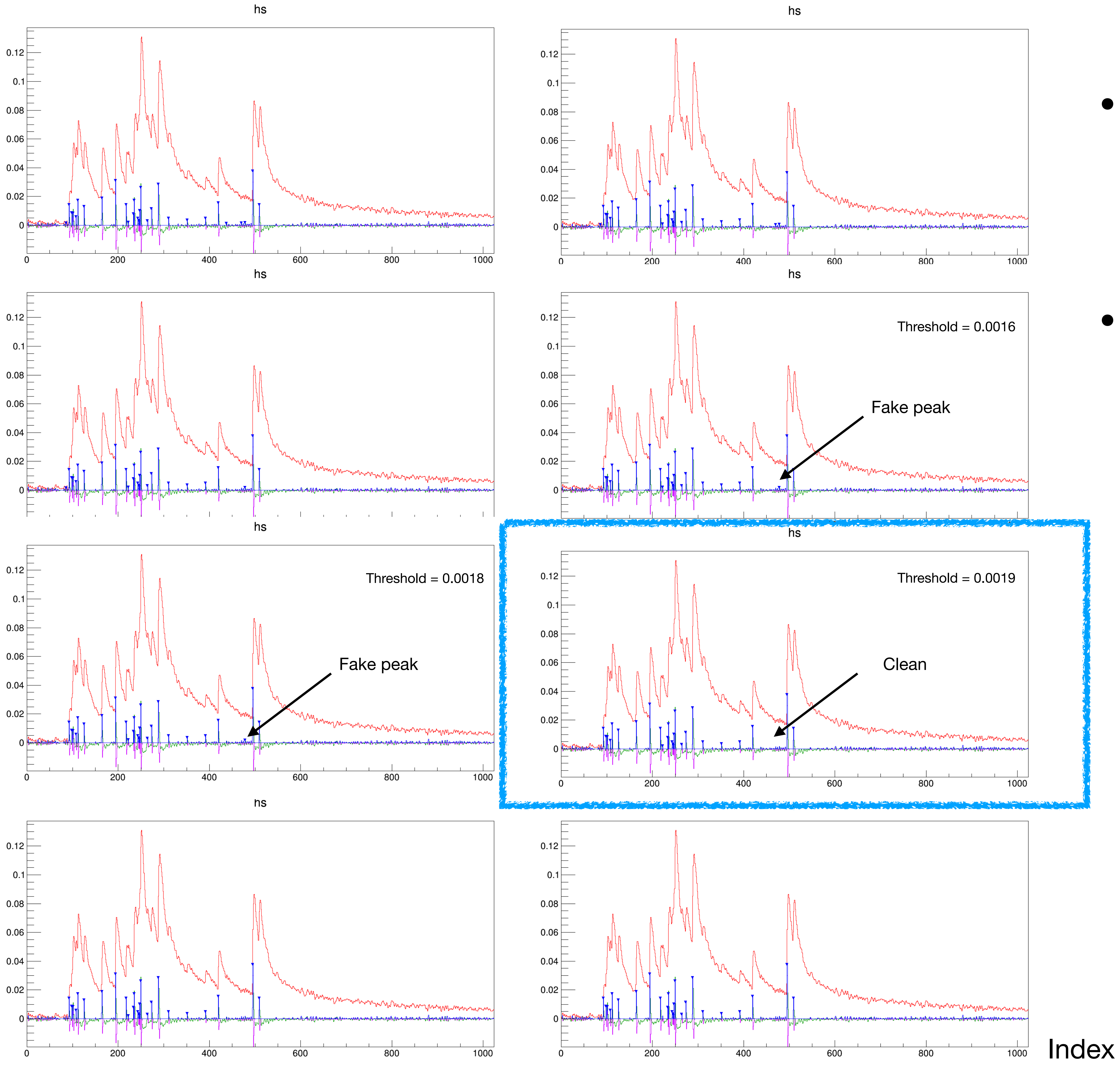
Introduction

- Cluster counting algorithm:
  - The second step of peak finding is to group electron peak into clusters $N_{\text{peaks}} \rightarrow N_{\text{clusters}}$
  - For derivatives, currently the peaks after electronic raise time and sampling smooth are treated as clusters. The time resolution maybe ~2 ns.
  - The single electron diffusion is found to be 2 ~ 4ns, if drift distance is 1.8 mm ~ 4 mm.
  - Clustering strategy: group electron peaks by a time cut, where the cut depends on the drift time.

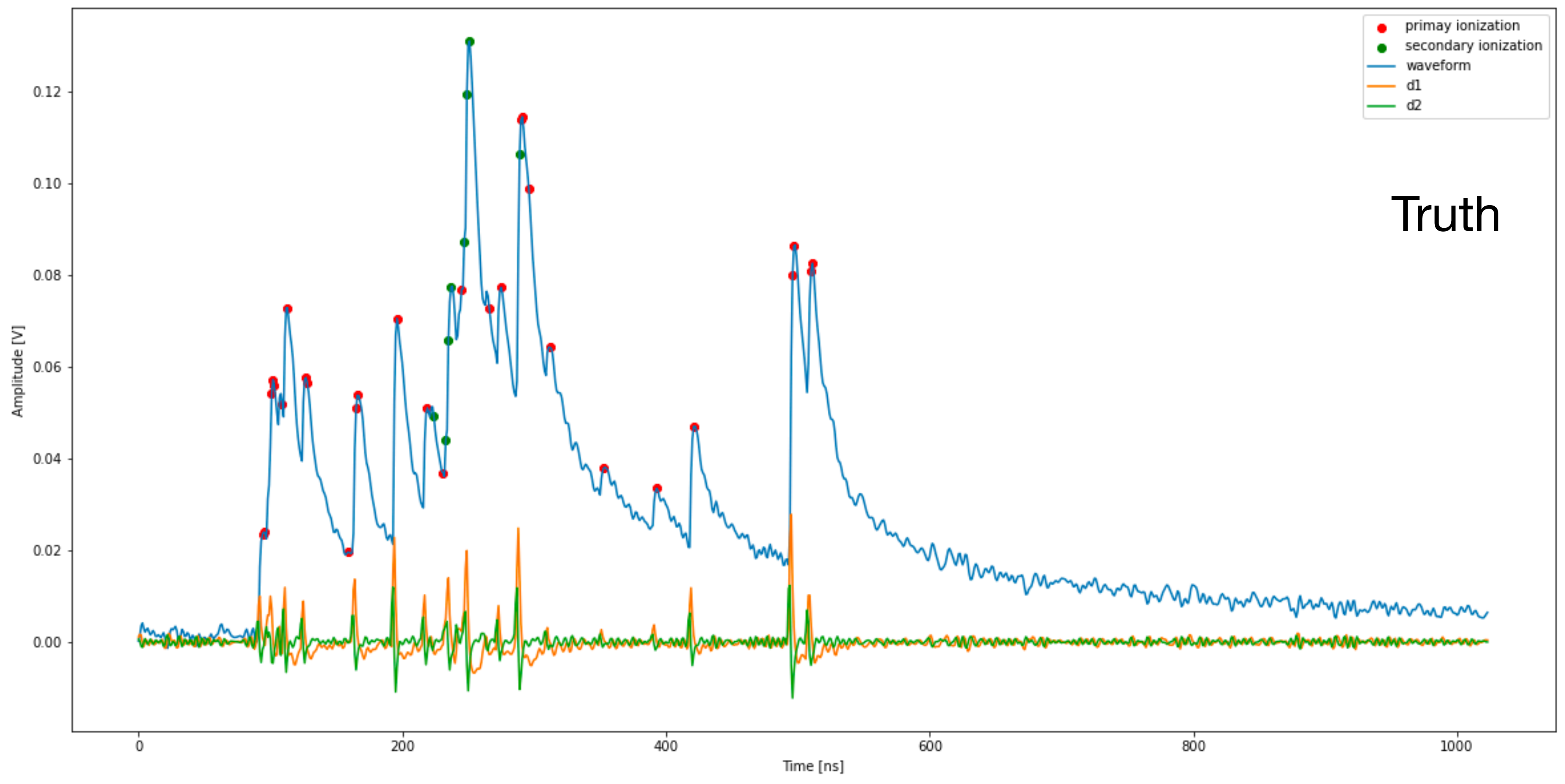




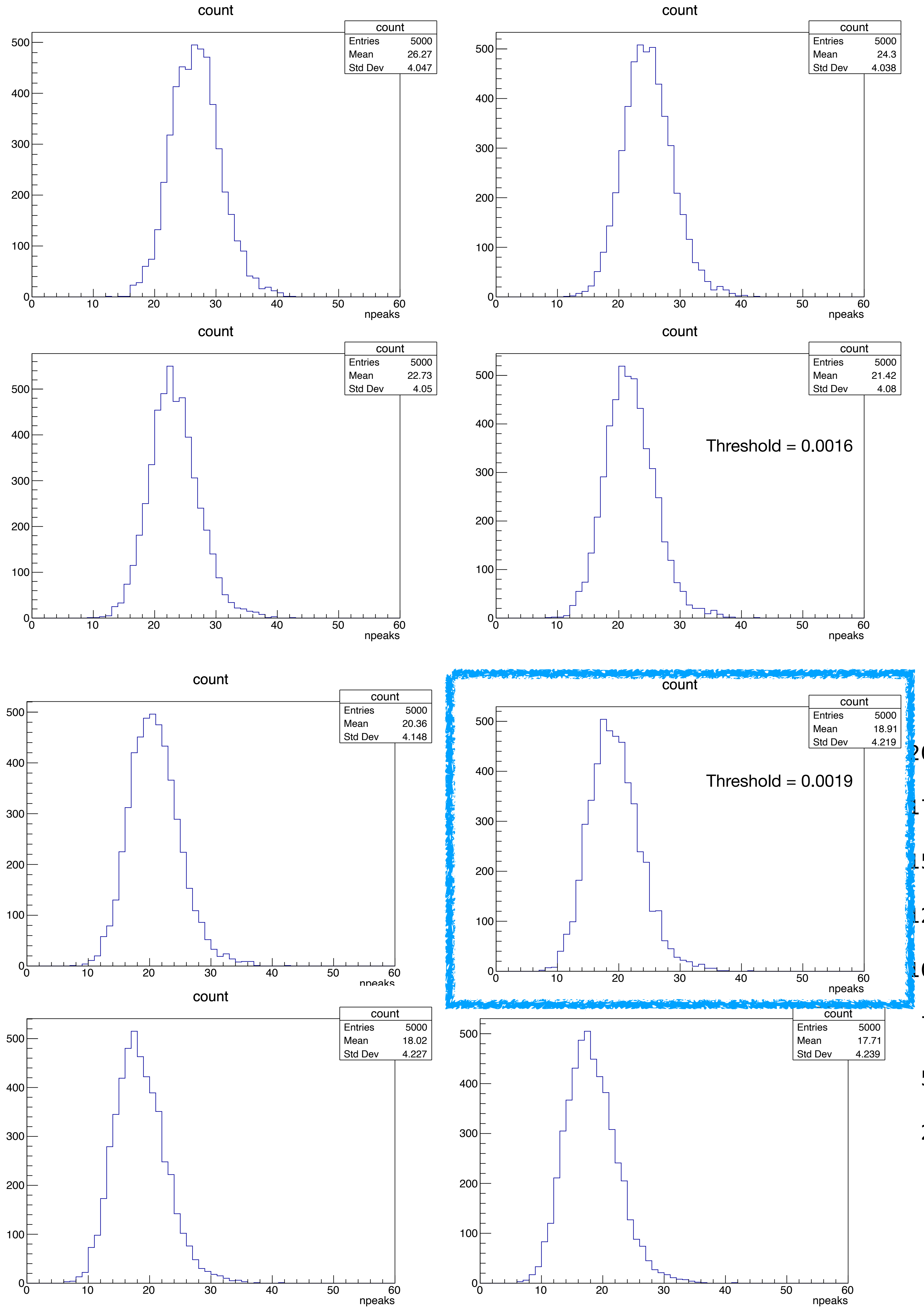
Peak finding with 2th derivative in MC



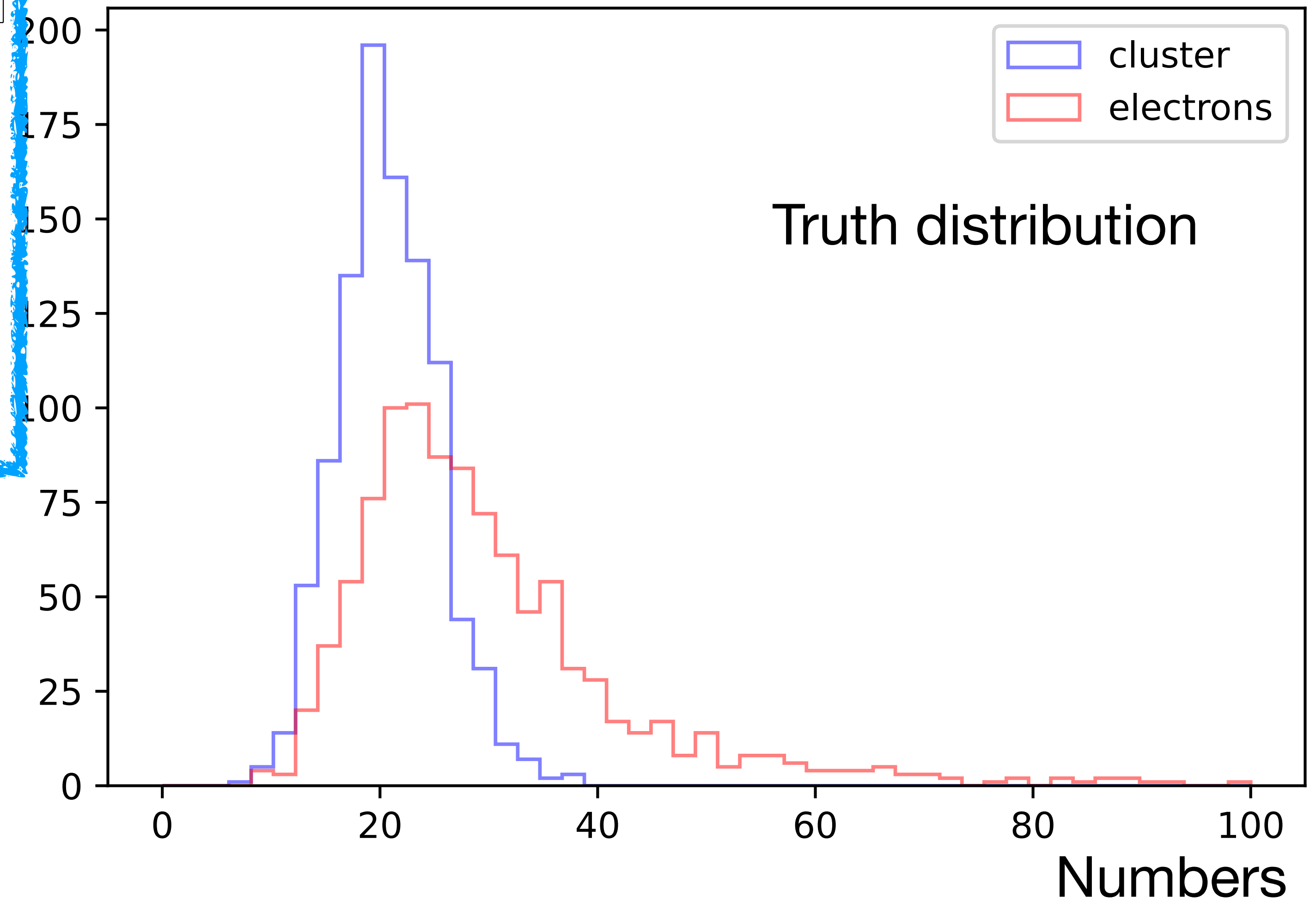
- Our MC :
  - fast simulation based on some electronic assumption, sampling rate = 1.2 GHz, time constant = 2ns
  - Track impact parameter = 4 mm, 1cm cell, 5000 events.
- Peak finding: 2td derivative method.
  - choose a threshold without much fake peaks
  - Cut off t< 50 and t > 541 events

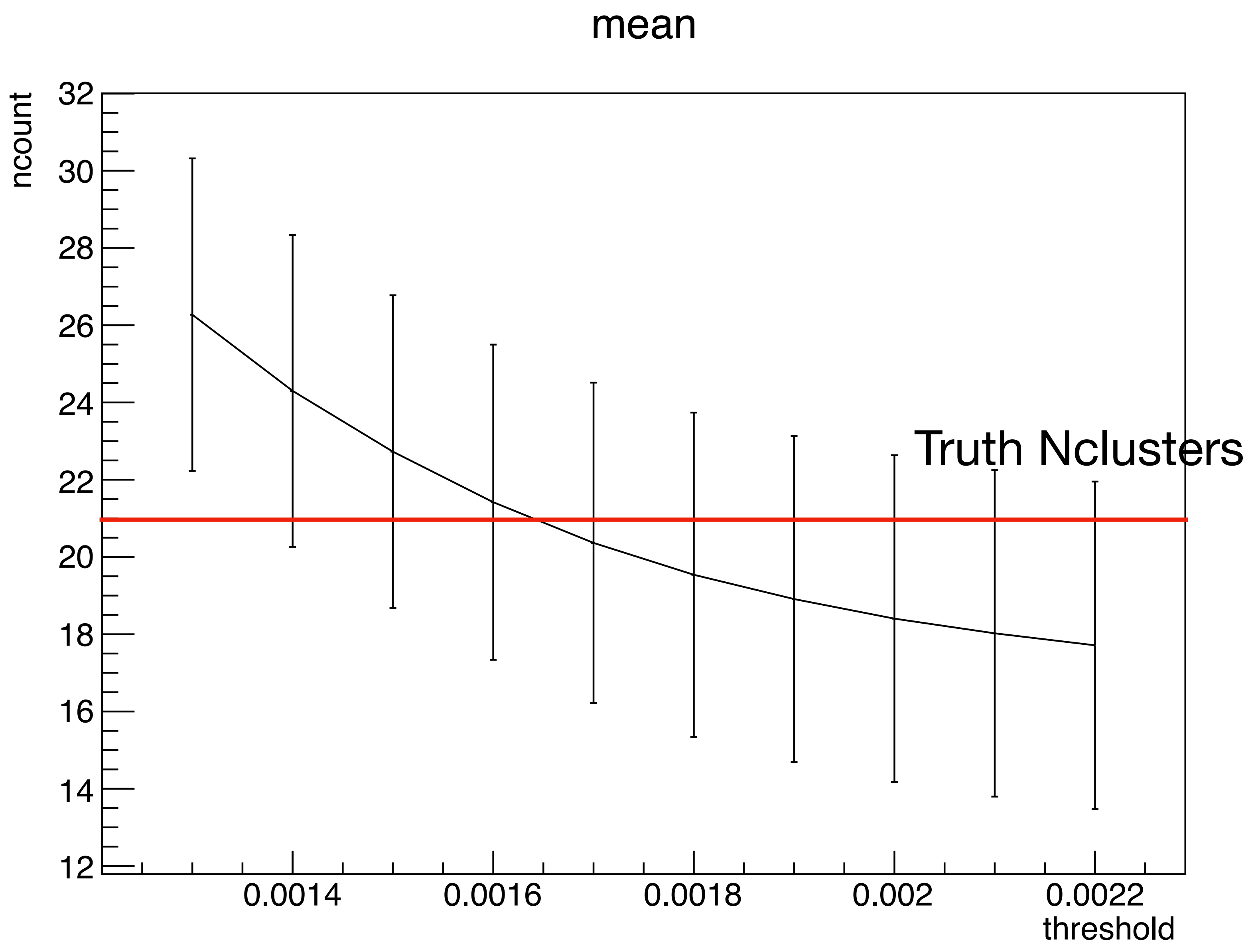


Performance: Npeaks with different threshold



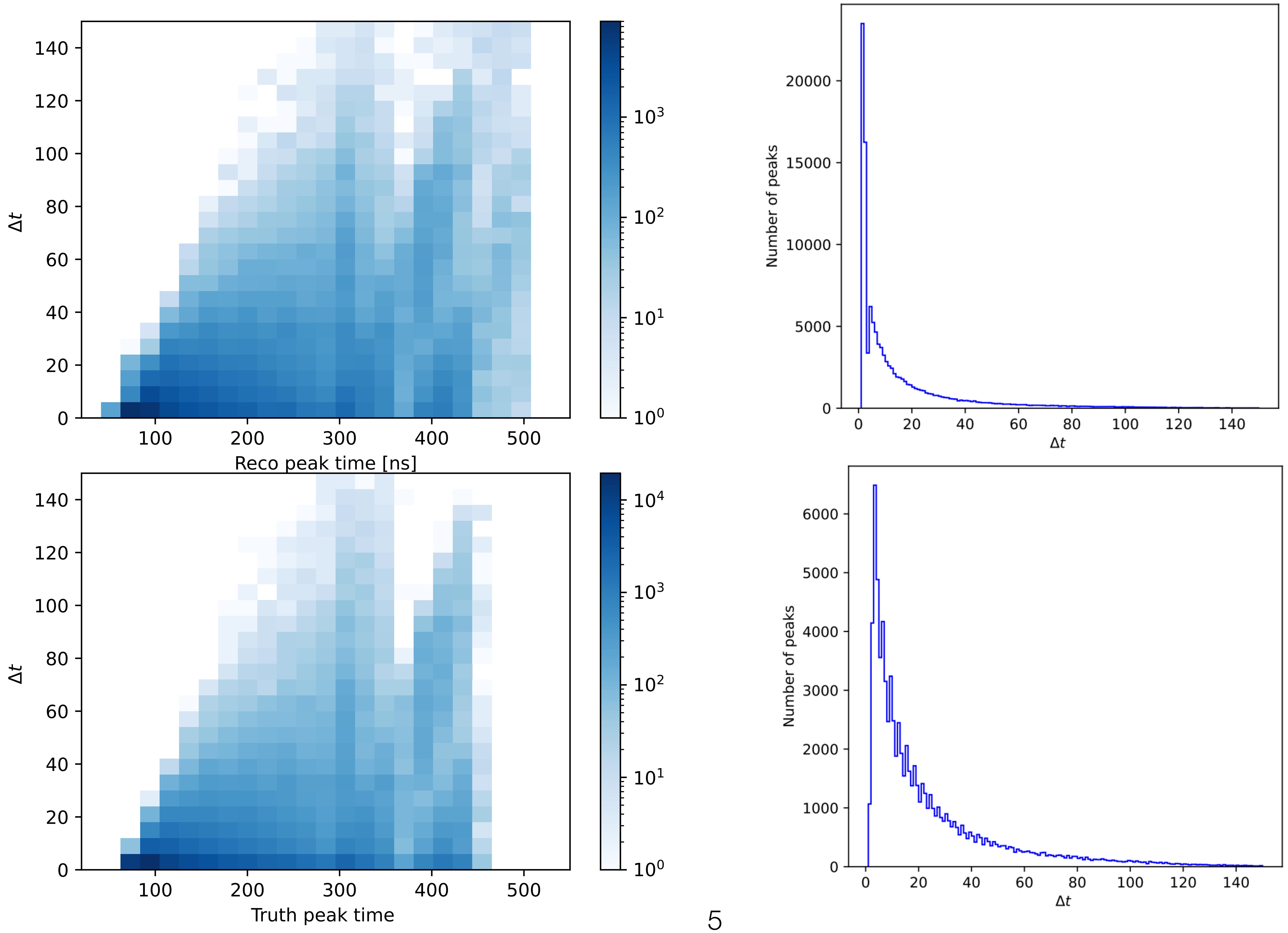
- Found peaks have very small long tail in the right of the distribution
- Truth information: Cluster size ~ 1.47. average #clusters: $\mu_{cls} = 20.9$
- Choose a threshold: 0.0019 (no obvious peaks).





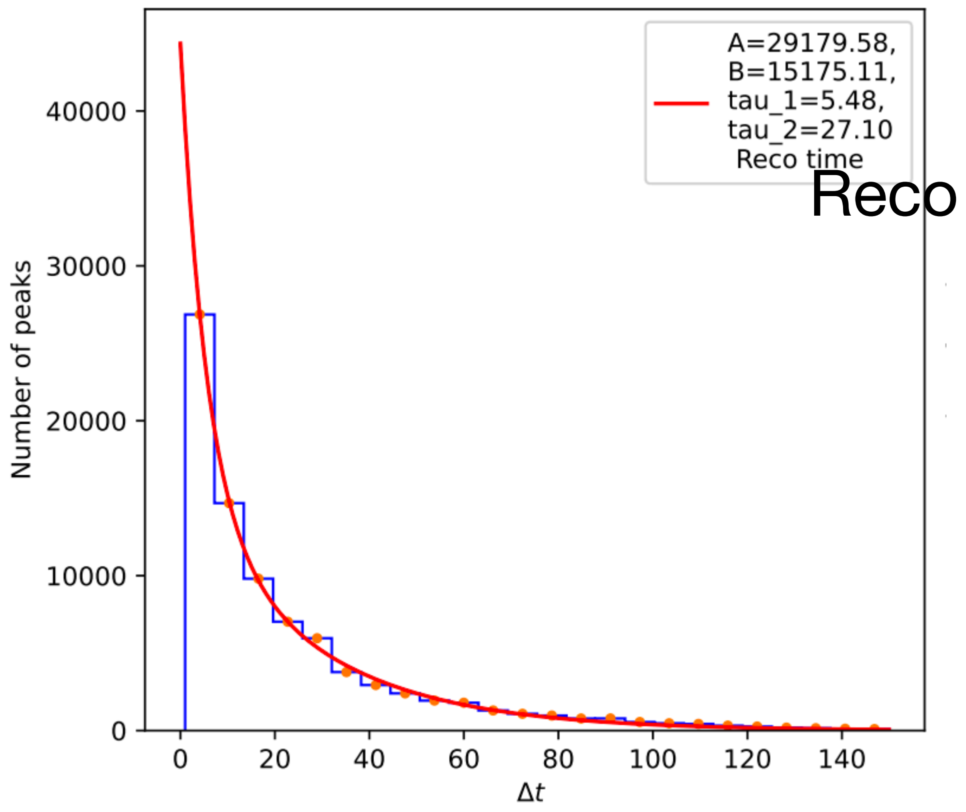
MC Time difference

- Time difference: the time between consecutive peaks
- 80~100 ns region is highly occupied. (assuming $v = 2.5\mathrm{cm}/\mu s$)
- Filled the time difference distribution at the peak time. First line: Truth; Second line: reconstruction.

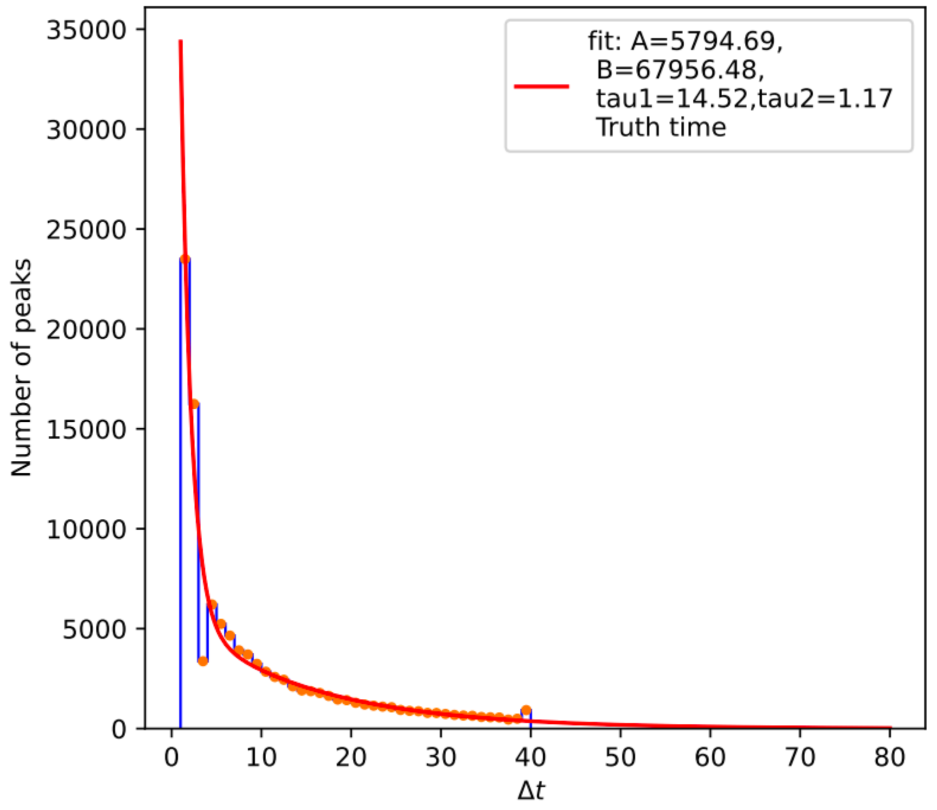


Time difference fit

- Fitted the Δ_t to $f = f1 + f2 = Ae^{-\tau_1/t} + Be^{-\tau_2/t}$ function. (An idea which was presented by Cuna in her previous talk: [Cluster counting algorithm based on first and second derivatives](#)
- Two components means the exponential function of Cluster's/ electron's delta time . Normalized $f1 + f2$ to 1. The intersection point gives us a timing cut.
- The combined fit ($f1 + f2$) is different to individual fit (truth primary and secondary Δ_t), but agreed within same magnitude.
- Not a pure exponential shape, might due to magnetic field, drifting .etc affect.



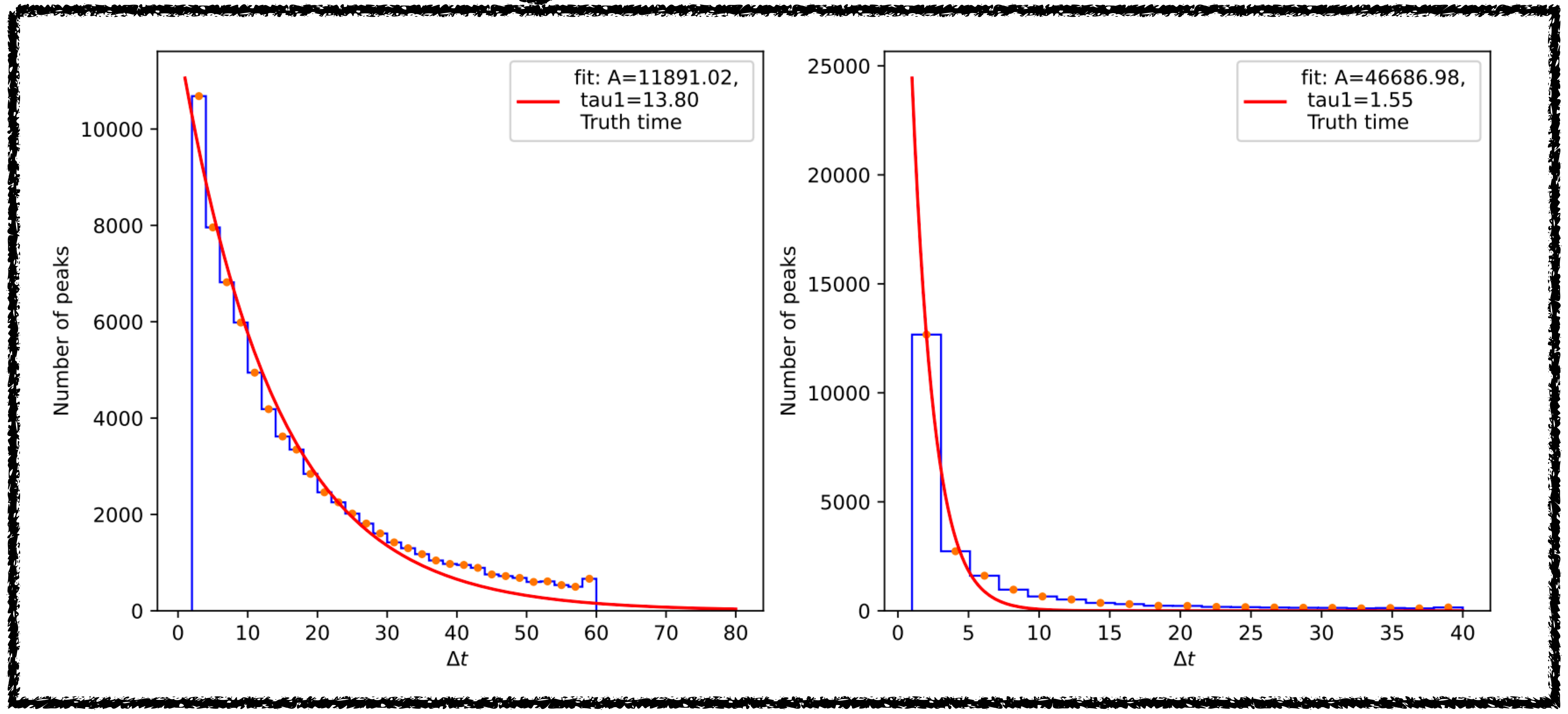
Reco

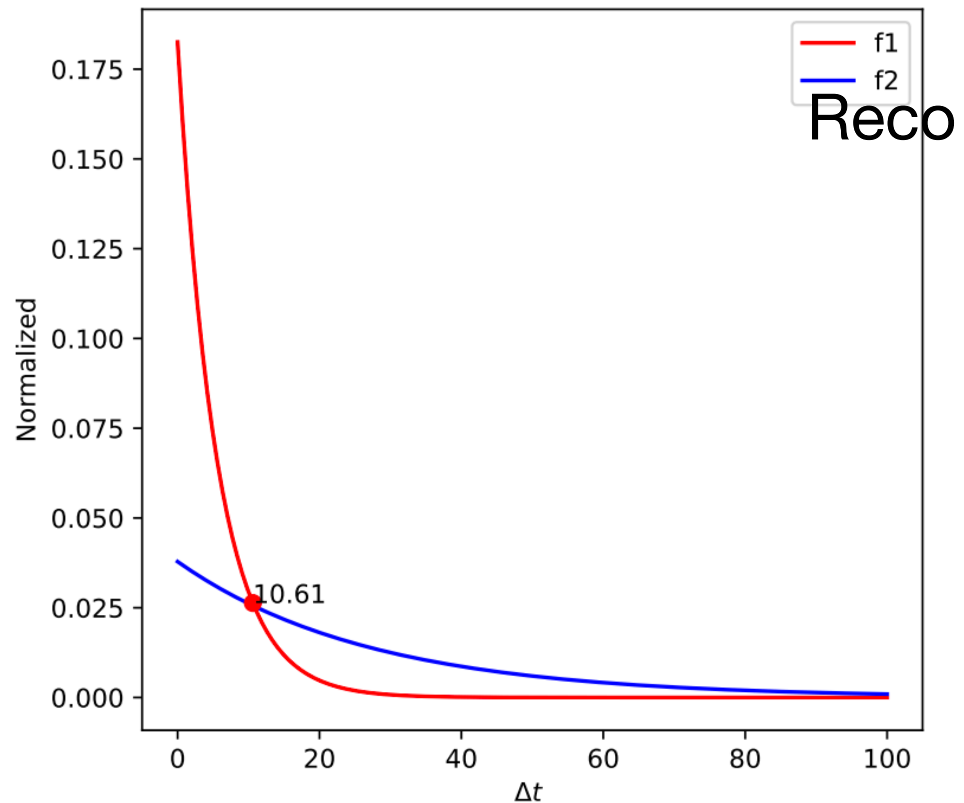


MC

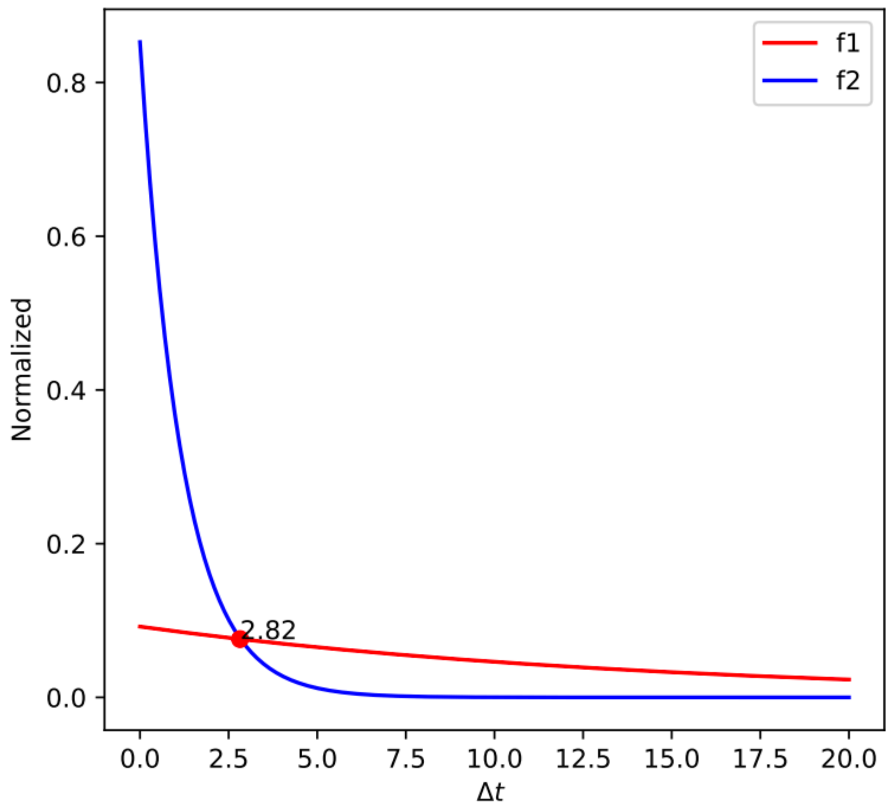
truth primary

Truth secondary





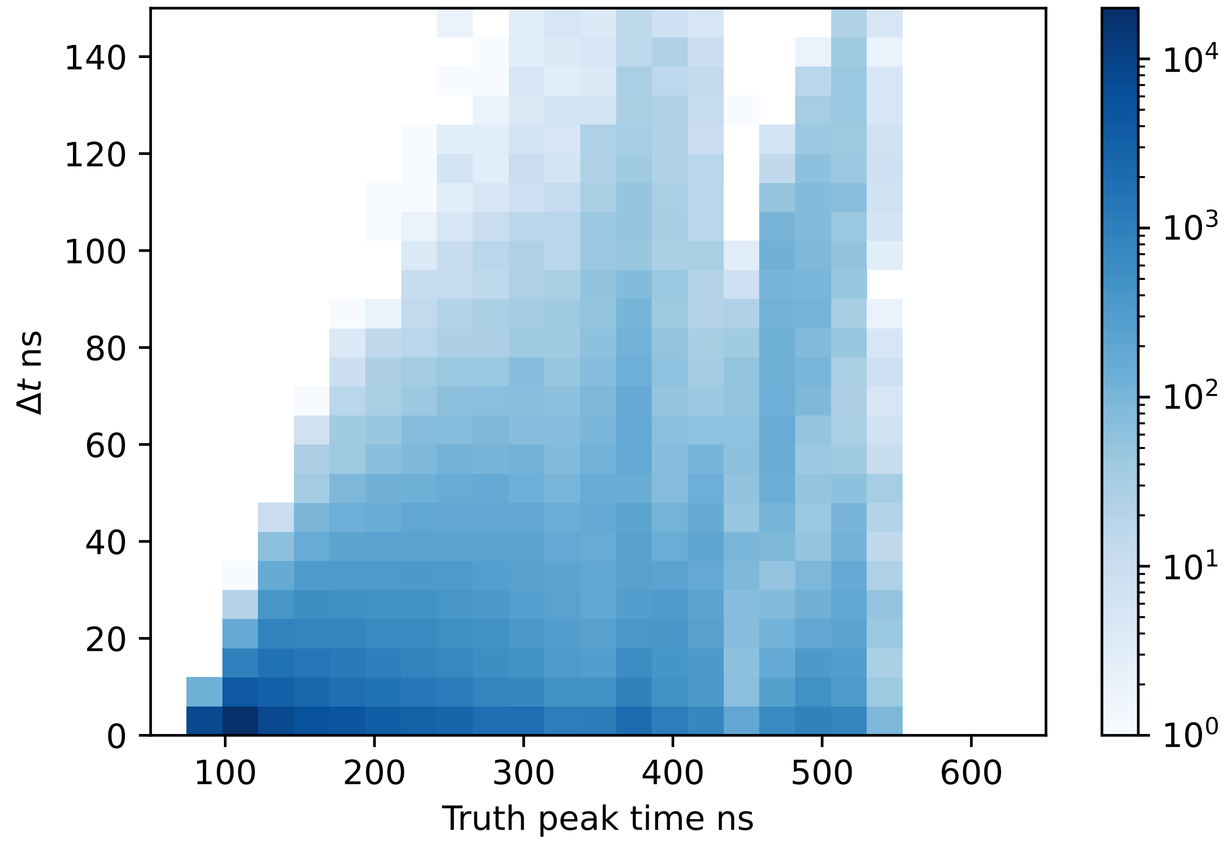
Reco

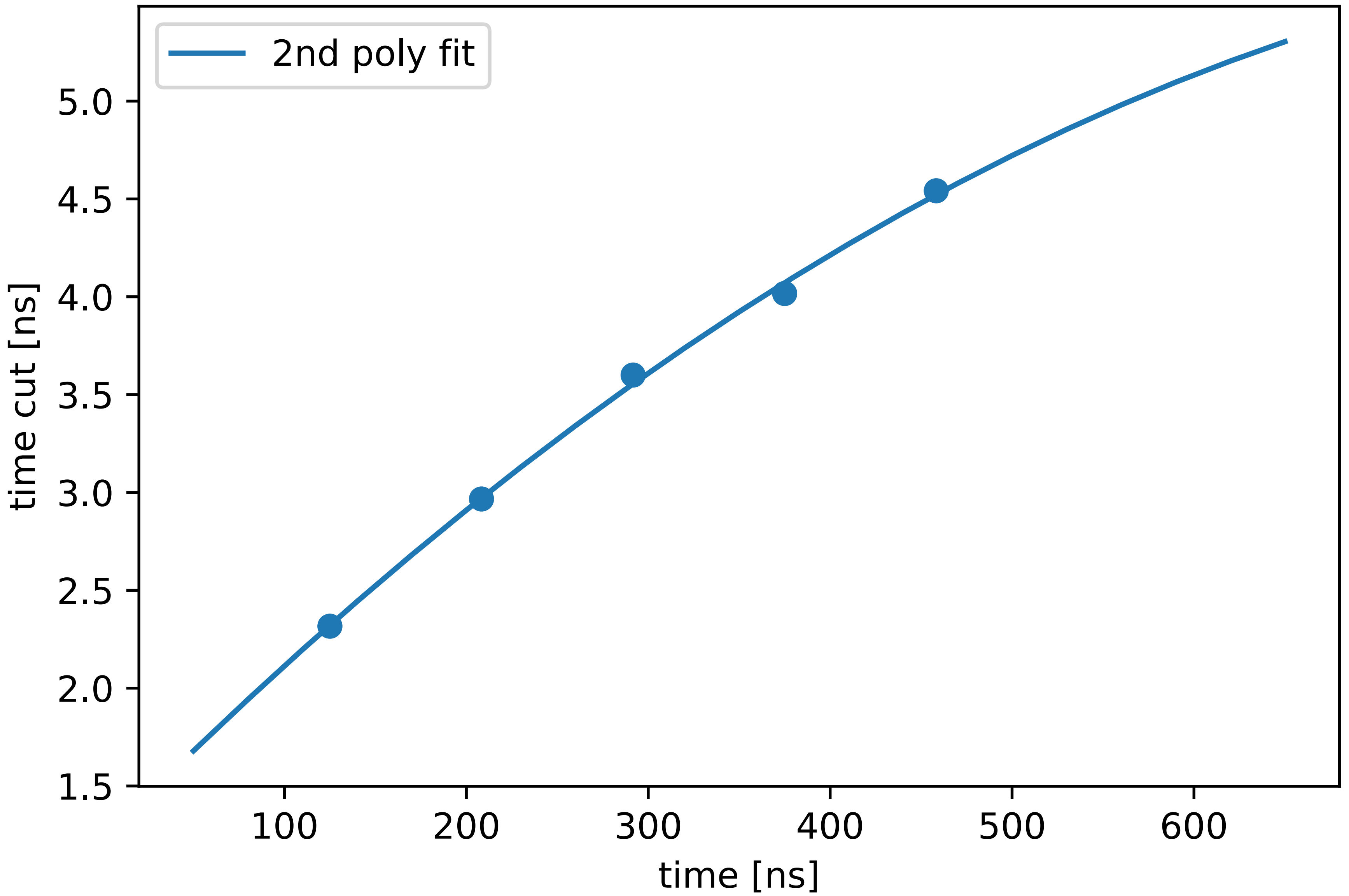


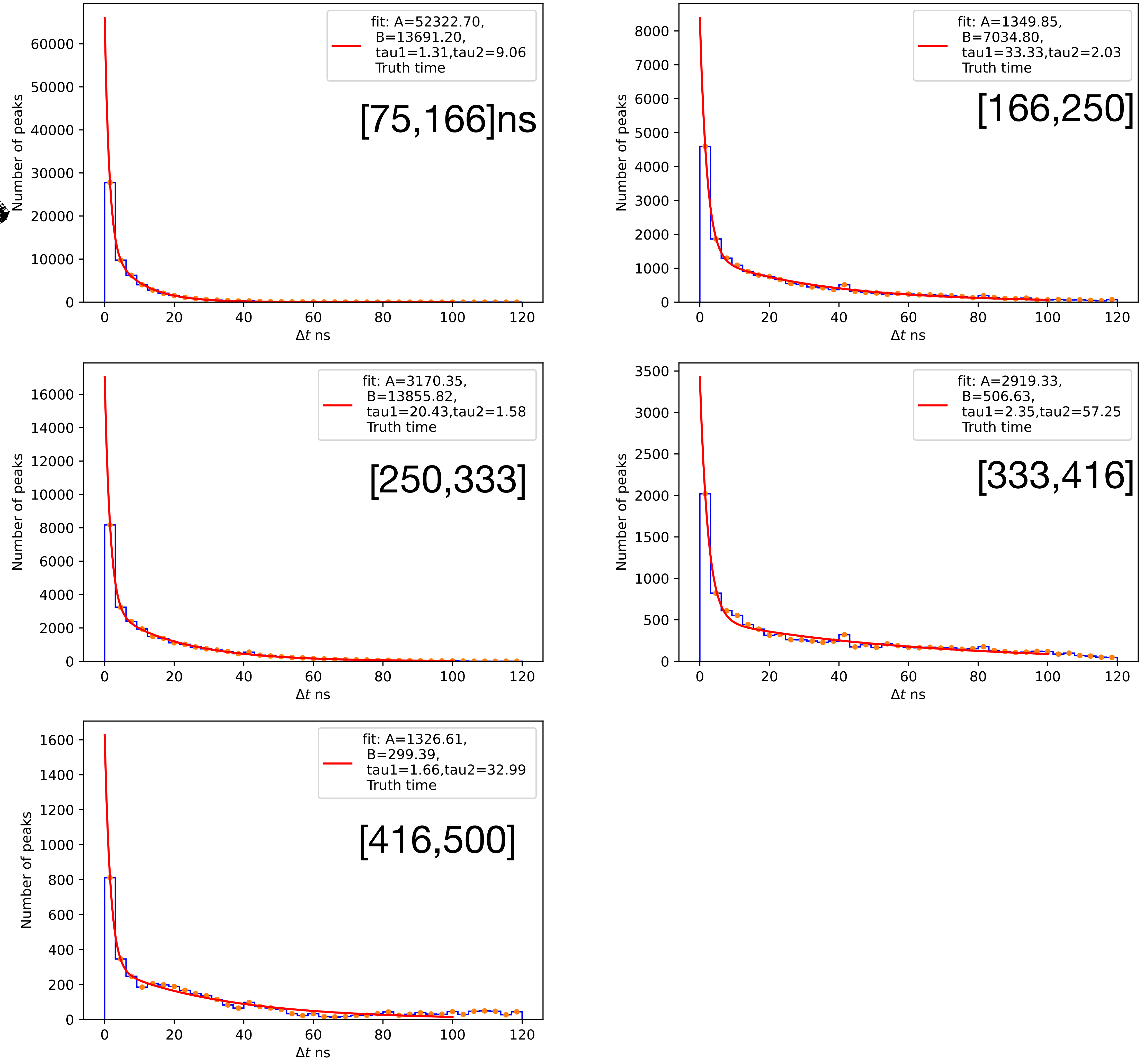
MC

Timing cut estimated from MC truth

- 5 intersection points from several time regions using truth.
- Timing cut determination: fit with a 2nd poly function. The value seems to be small than single electron diffusion (200ns ~ 5 ns)

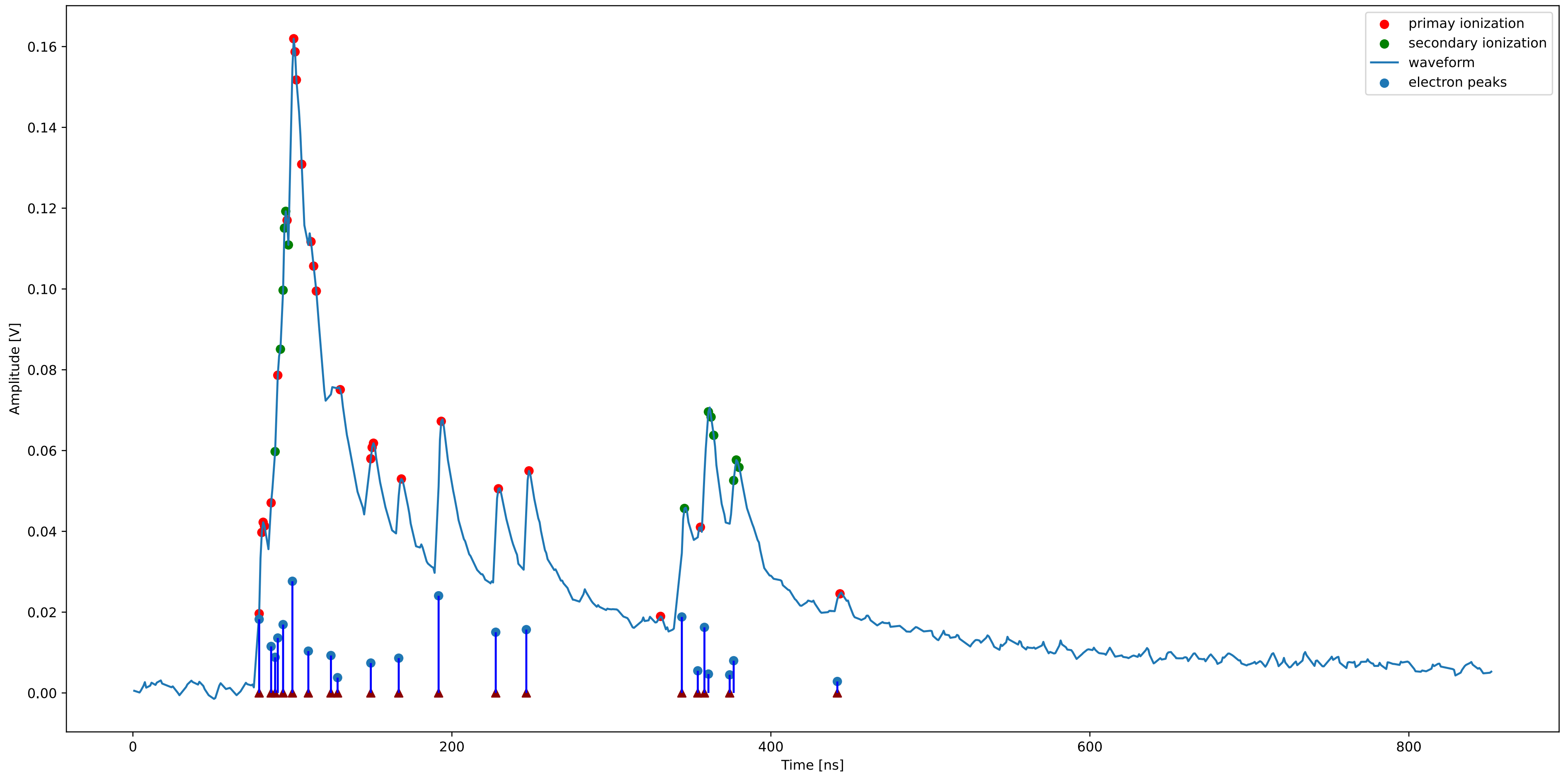


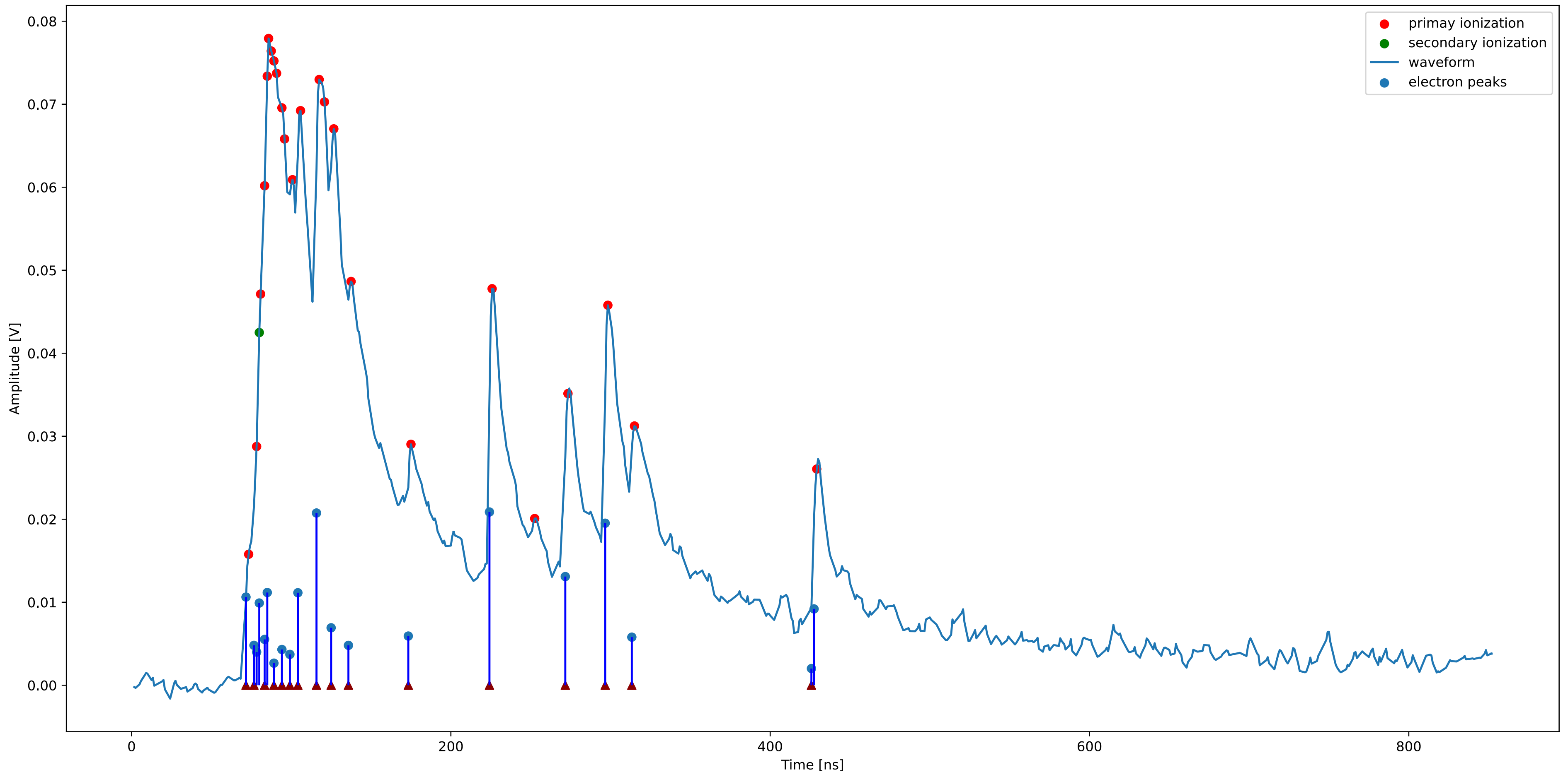


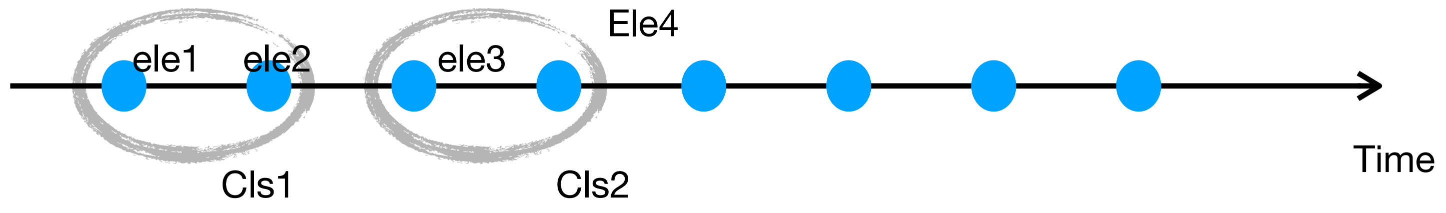


Clustering strategy

- Loop all electron peaks, if the distance of two consecutive peaks is small than the timing cut, merged as a cluster. A cluster is allowed to have a length within 1 timing cut unit.
- Dark red: clusters; blue stick: electrons







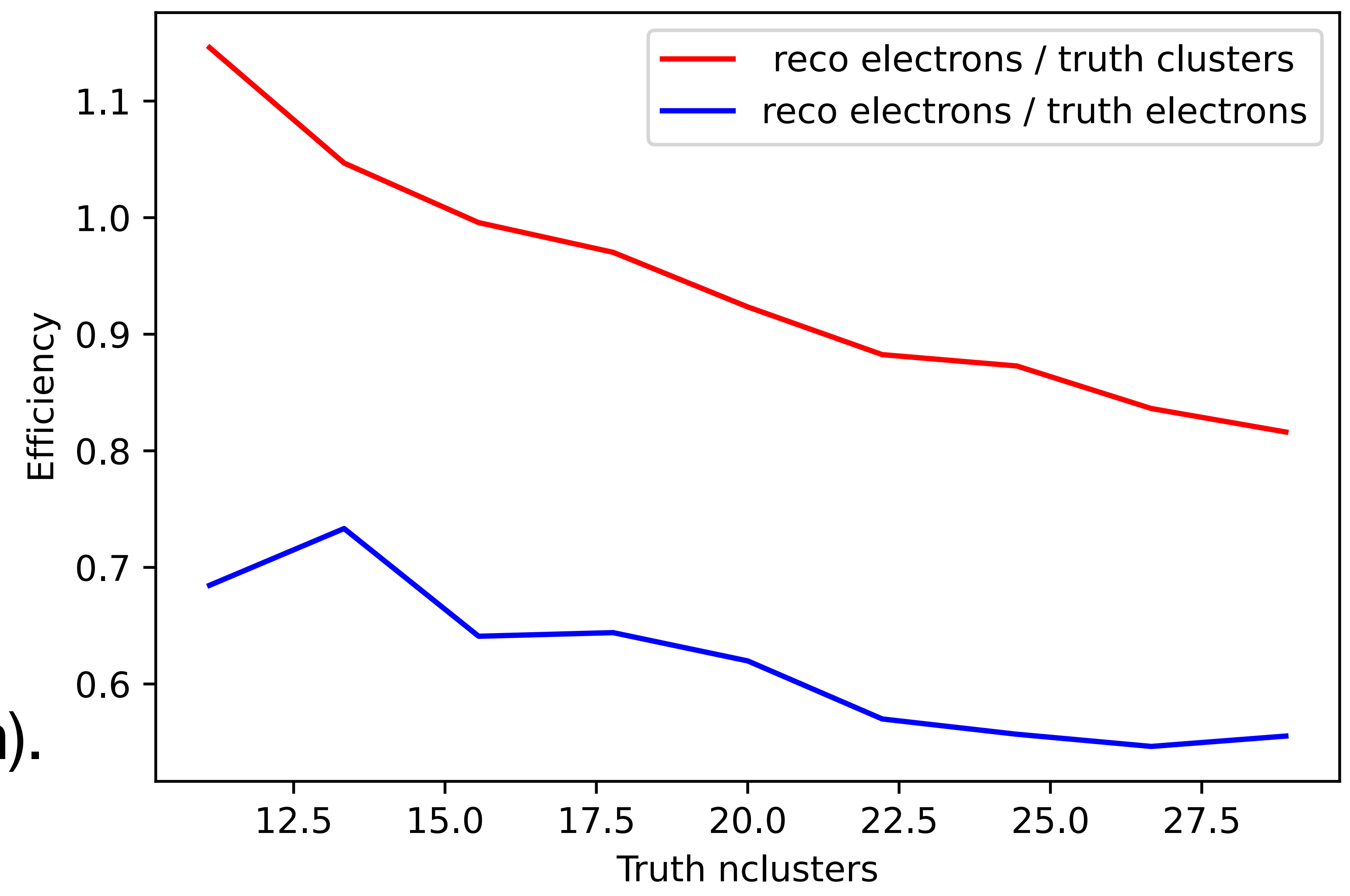
Assuming T[ele1,ele2] = T[ele2,ele3] = timing cut,

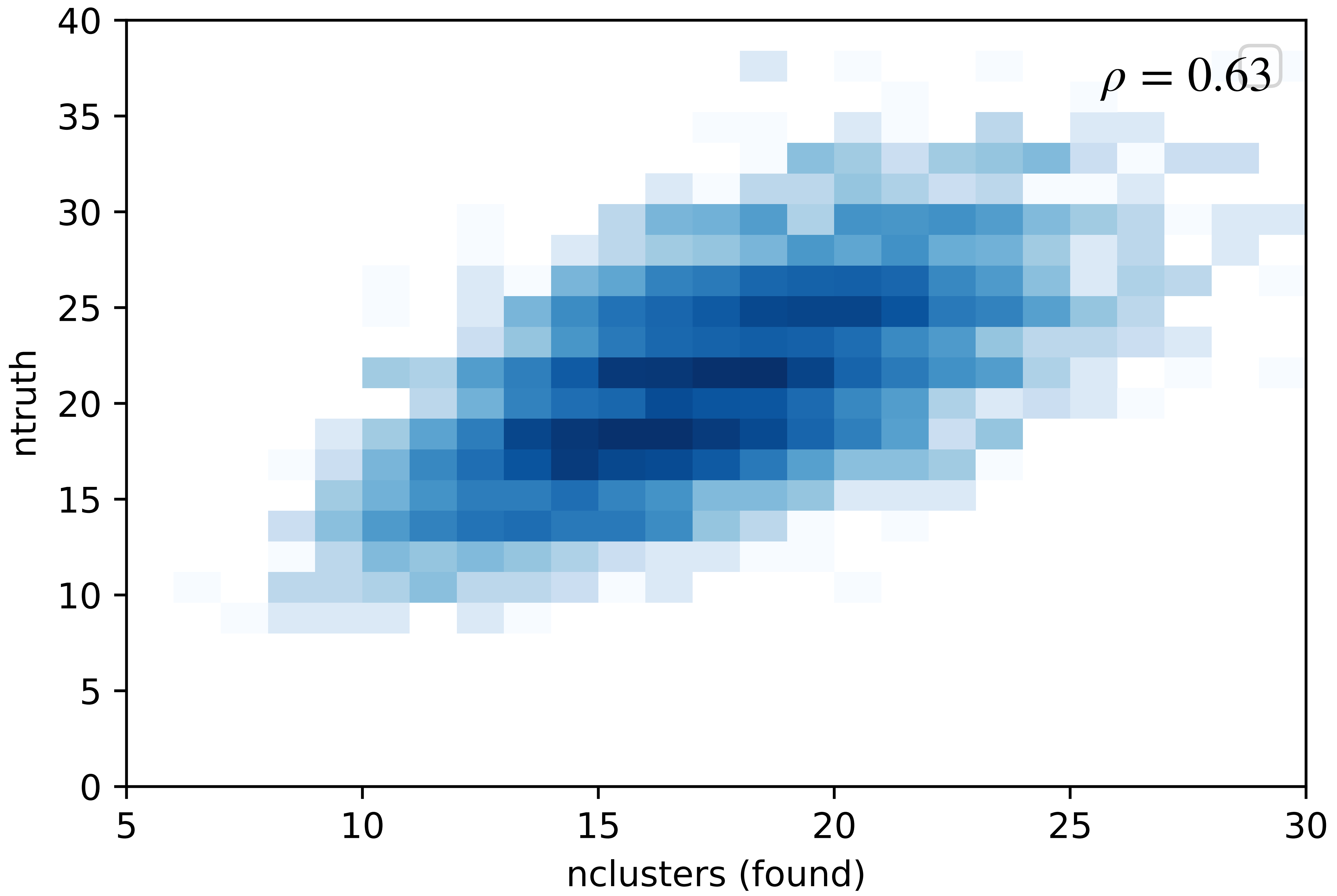
the second electron is grouped to the first electron as one cluster (Cls1),

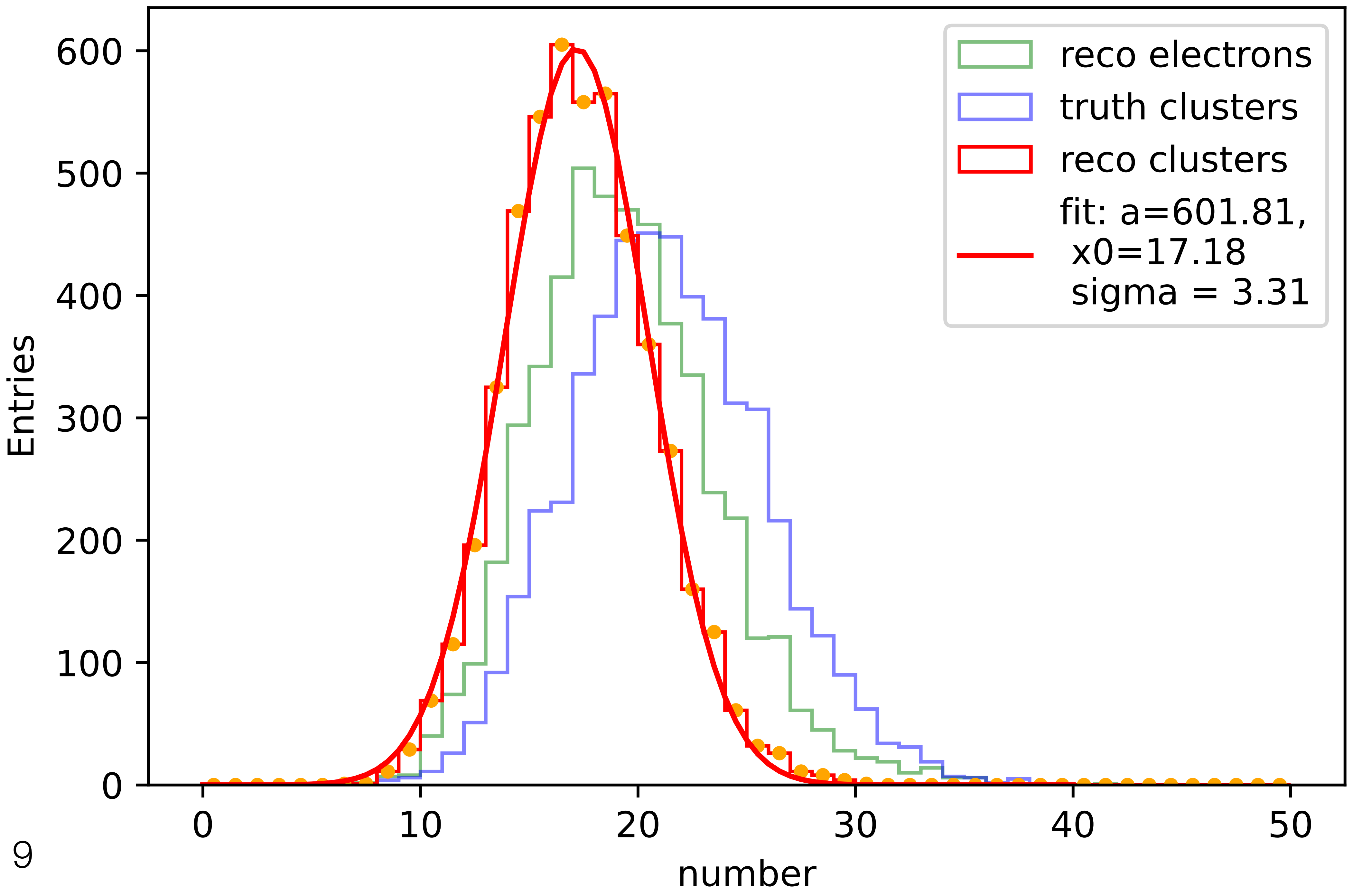
the third electron will not not be merged to Cls1 as T[ele3,ele1] > timing cut, but merged to Cls2 if the 4th electron has T[ele4,ele3] < timing cut.

Clustering performance

- Good linear shape of reconstructed cluster and truth clusters.
- Good gaussian shape after clustering. The resolution is 19.2 % (truth 21%).
- Clustering eff = reco cls / truth cls , slope is about -0.1, which makes the resolution better than the truth one.
- Clustering efficiency (Mean) = 17.18 / 20.9 = 82.2 %
- Results of other threshold has similar feature. Threshold 0.0016 eff = 93.3 %
- Not surprised because there is no much asymmetry in the reconstructed Npeaks (the green).



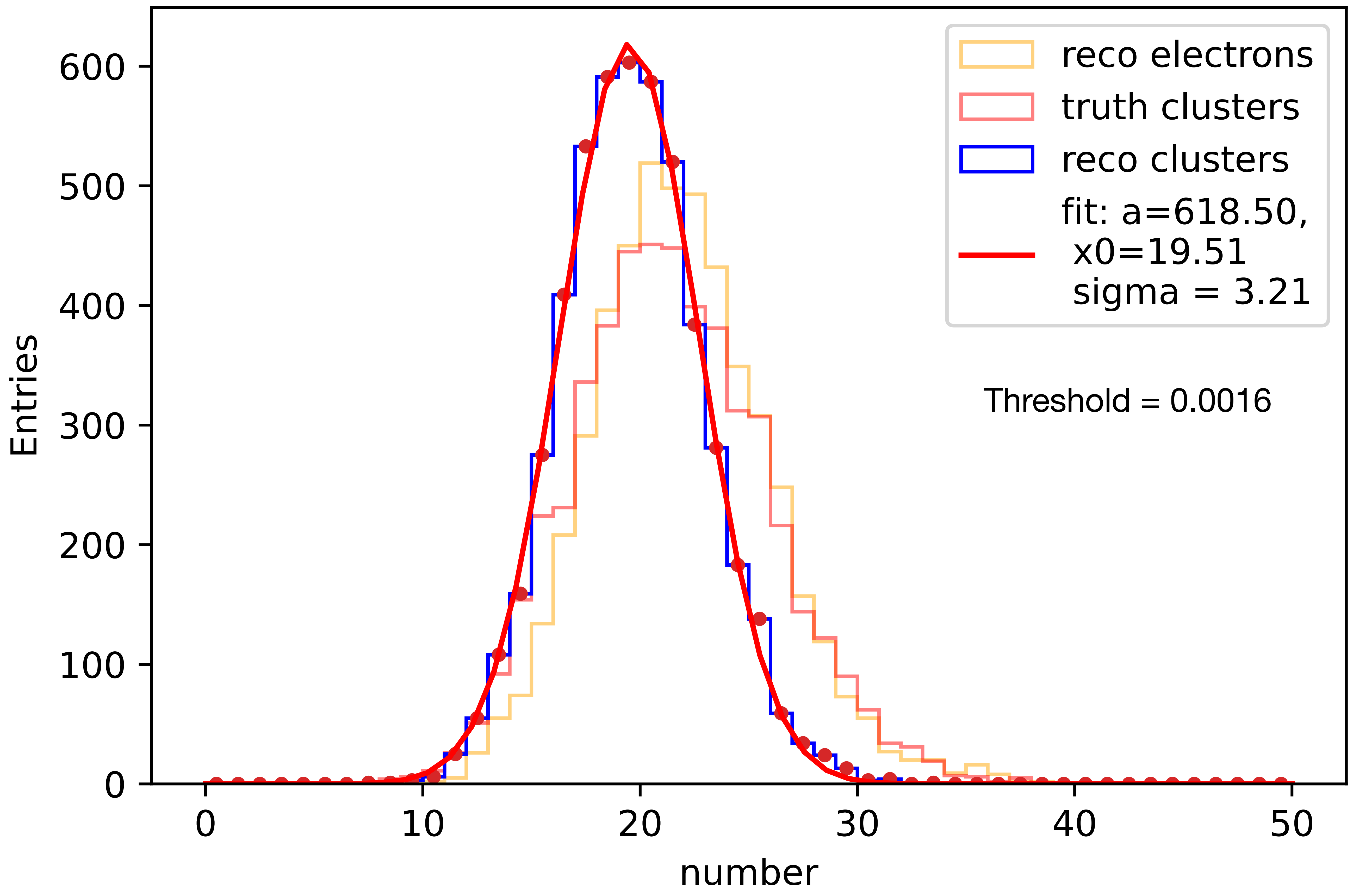




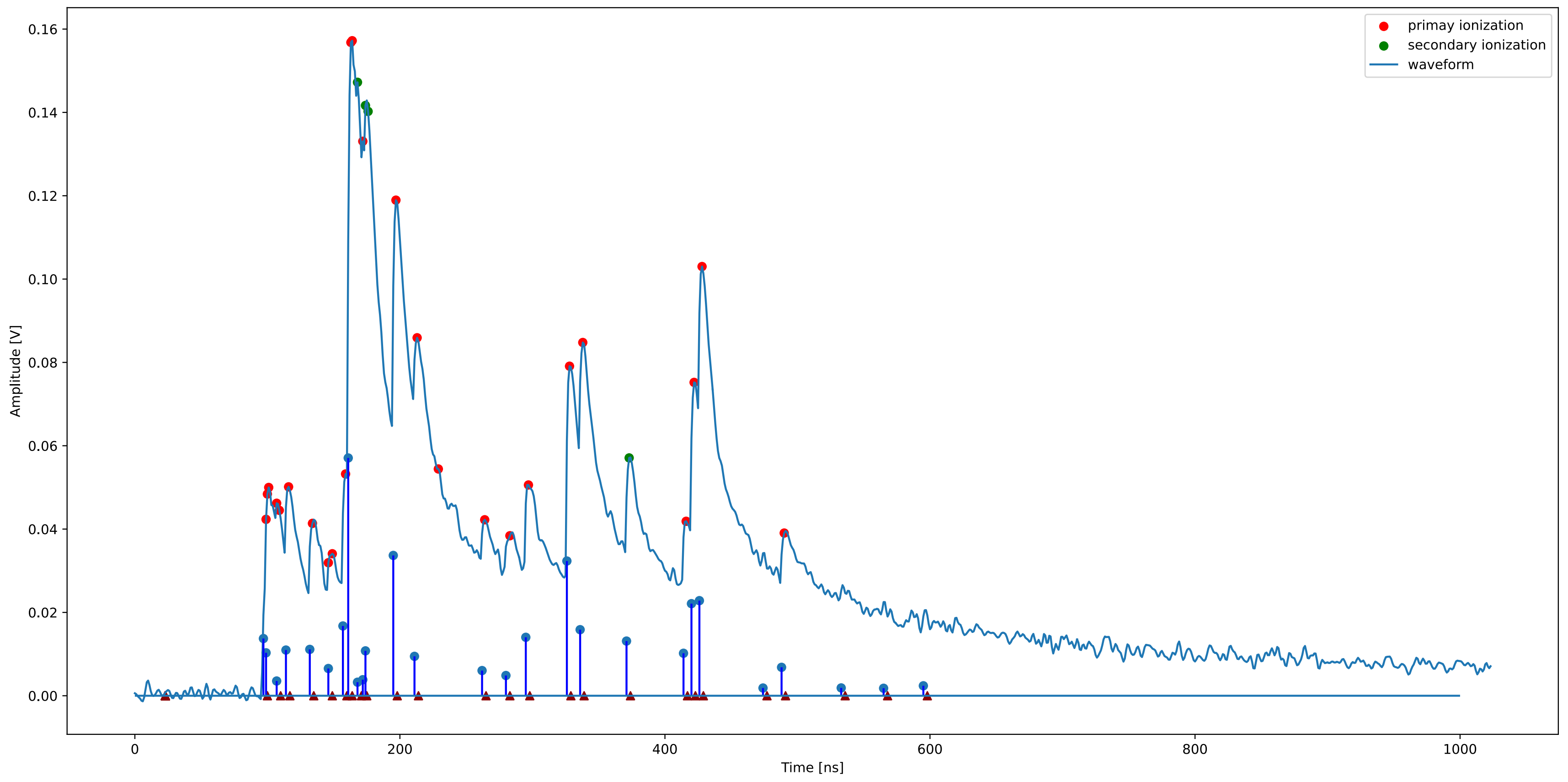
Summary

- The clustering is needed after the peak finding with derivatives
- Encoded a clustering method for MC waveform, employing the time differences idea from Federica Cuna's talk.
- The timing cut is varied from 2 ns to 5 ns, with drift time from 100 ns to 500 ns.
- The cluster coating efficiency is 82.2 % with 19.2 % resolution with 1cm track.
- Plan: Try to apply clustering on test beam data.

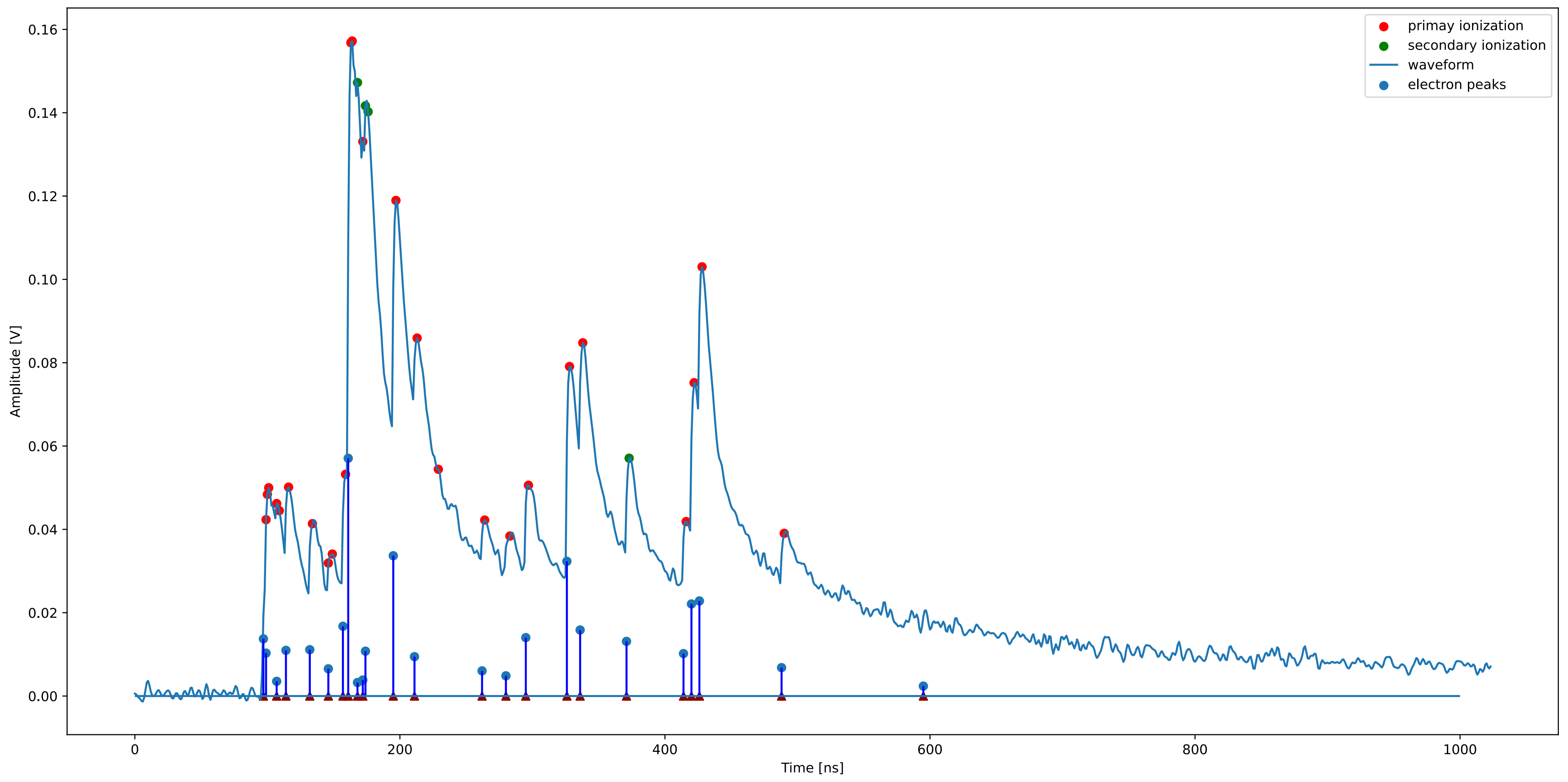
Backup



Threshold = 0.0016



Threshold = 0.0016



Threshold = 0.0019

Backup

- 1000 events threshold = 0.00205