Cluster Counting Results and Progress from Beam Test Data

6th SuperB Collaboration Meeting - INFN - LNF

Jean-François Caron
The University of British Columbia

December 12, 2012

Background

- Single-cell prototypes built at TRIUMF were tested using a ~210 MeV/c beam of electrons, muons, and pions in August, September, and December 2012.
- The single-cell events are used to create “bootstrapped” tracks for PID studies.
- External time-of-flight signal is used to identify particles.
- A broad goal is to help decide between equipment choices, such as amplifiers, cables, wires.
- A more immediate goal is to investigate cluster-counting algorithms, and quantify benefits.

Smoothing Algorithms

Suppose our signal is V(t) where t are discrete steps from 0 to N, and V(t) is a voltage.

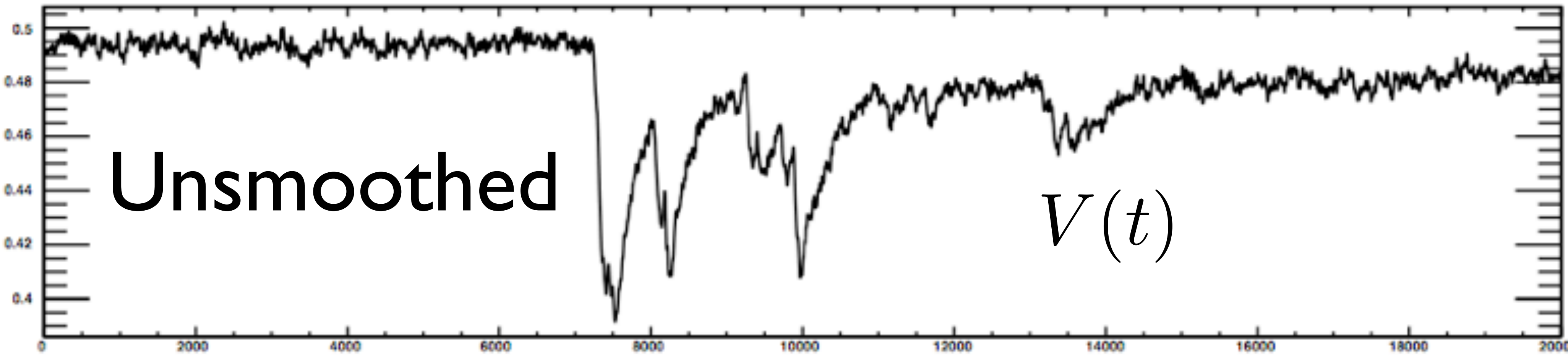
"Boxcar Smoothing" with k frames says (with exceptions for edge-effects):

$$\tilde{V}_k(t) = \frac{1}{k}\sum_{j=0}^{k-1} V(t-j)$$

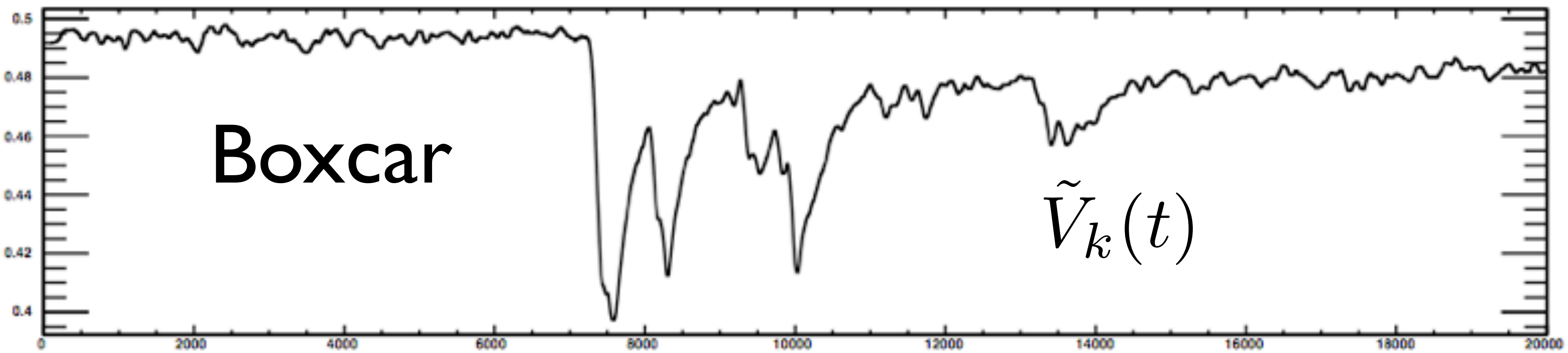
"Averaging" with k frames reduces the total number of samples to M = N div k, with s = N mod k frames skipped.

$$\bar{V}_k(t') = \frac{1}{k}\sum_{j=0}^{k-1} V(s+t'k+j)$$

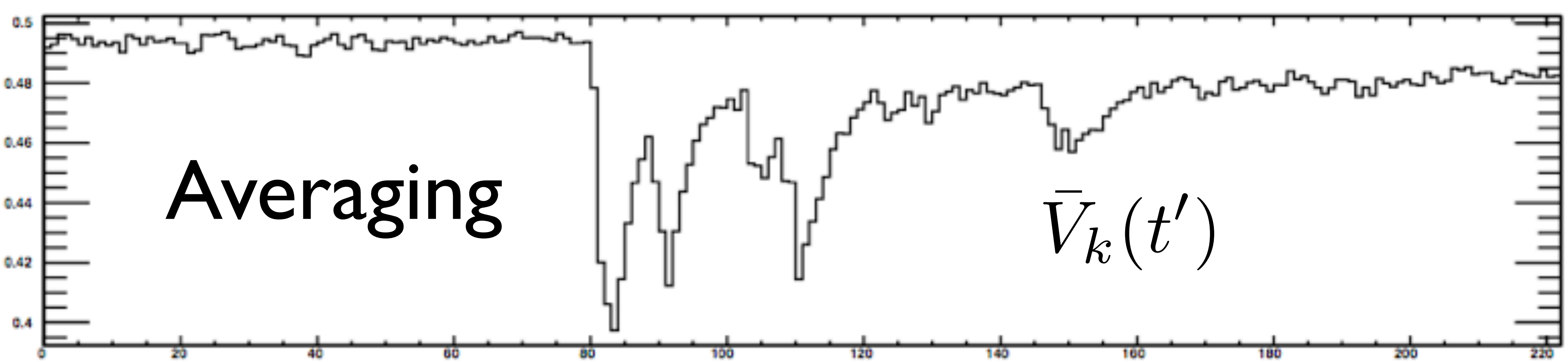
Smoothing Illustrated



N = 20002 samples
50ps width



k = 95 frames
(4.75ns)



k = 95 frames
(4.75ns)
N div k = 210
N mod k = 52

Cluster-Counting Algorithms

- Threshold-over-boxcar average: $Q_0(t) = V(t) - \tilde{V}_k(t-1)$

  Only two parameters (k, threshold)

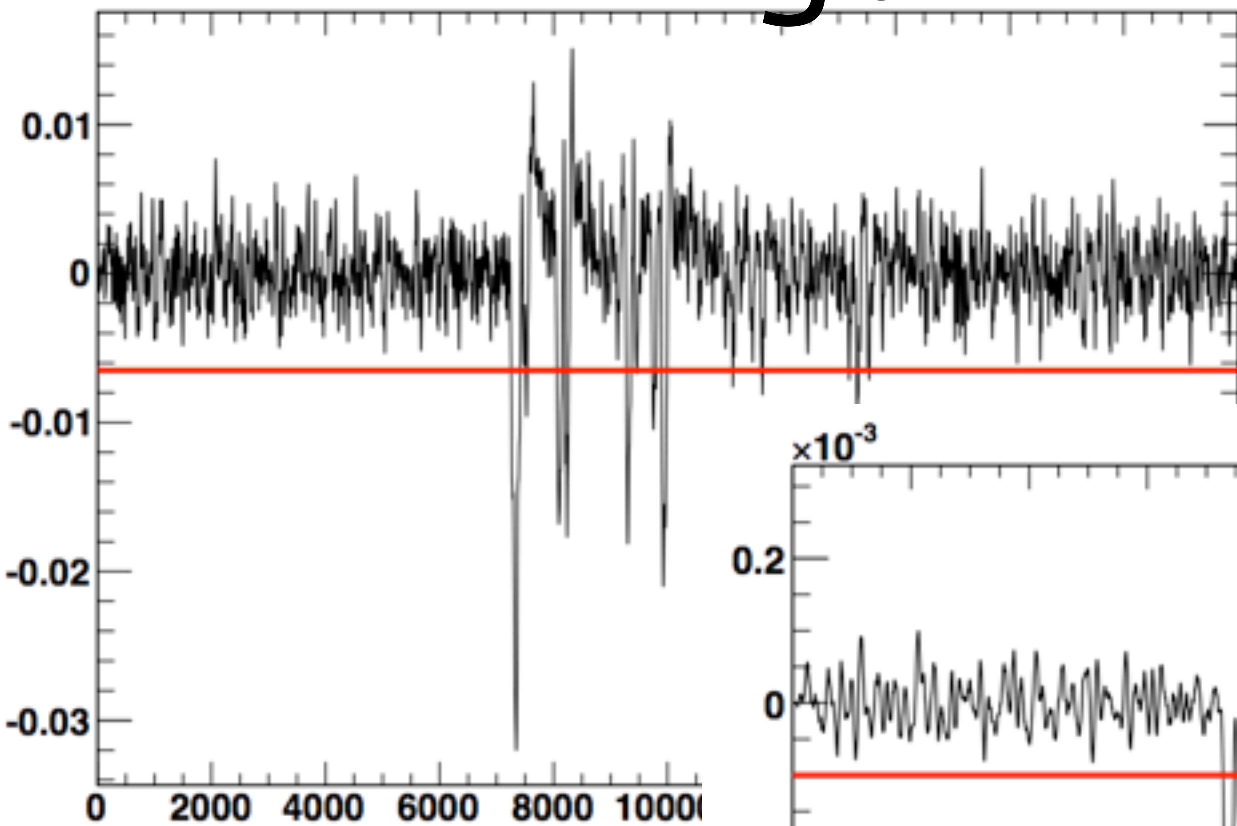
- "Hardware derivative": $$Q_1(t) = \frac{1}{k}\left(\tilde{V}_i(t) - \tilde{V}_j(t-k)\right)$$

  Four parameters (i,j,k,threshold)
  Equivalent to Q_0 for i = 0, j = k_0, k = 1

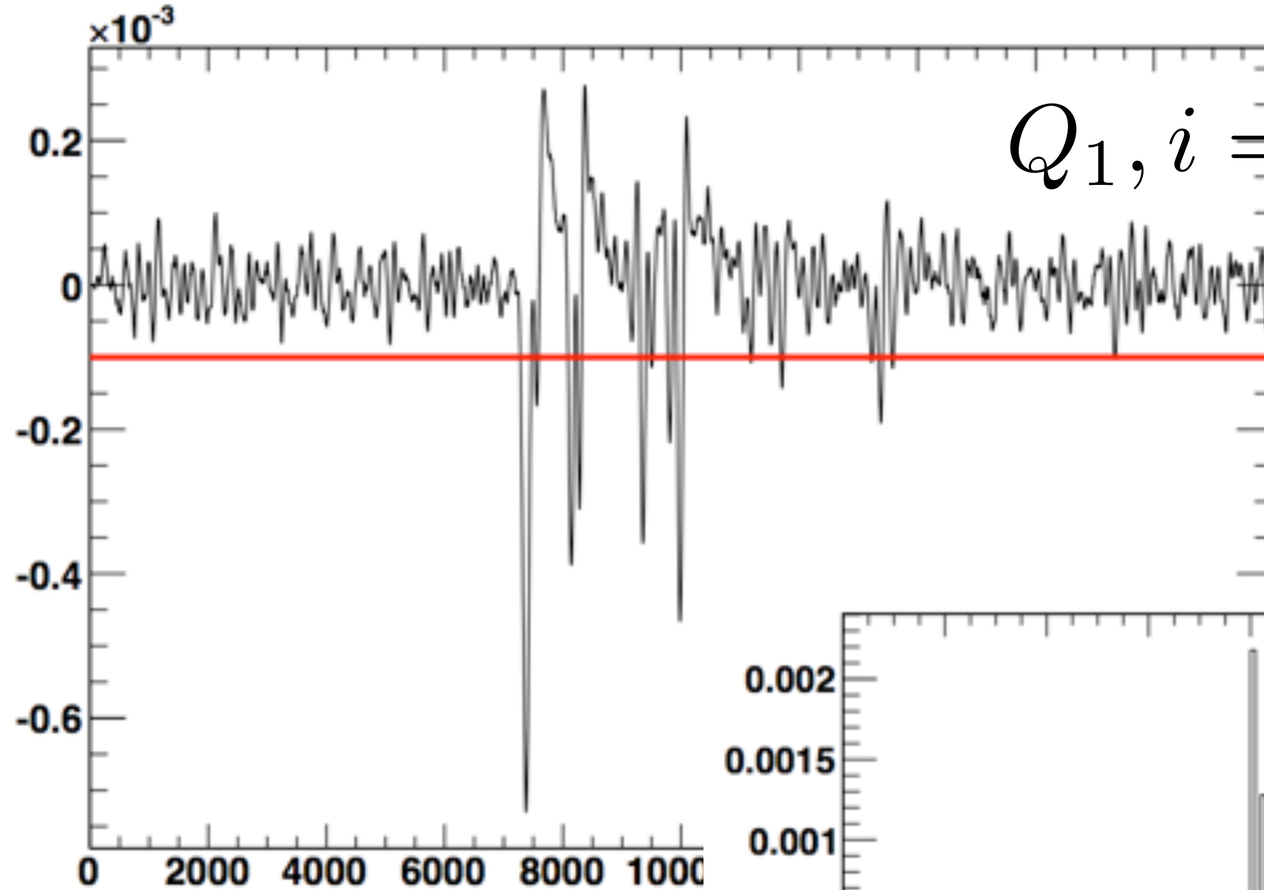
- Second derivative: $$Q_3(t) = \frac{\left(\bar{V}_k(t+2) - \bar{V}_k(t+1)\right) - \left(\bar{V}_k(t+1) - \bar{V}_k(t)\right)}{(\delta t)^2}$$

  Only two parameters (k, threshold)

Algorithms Illustrated

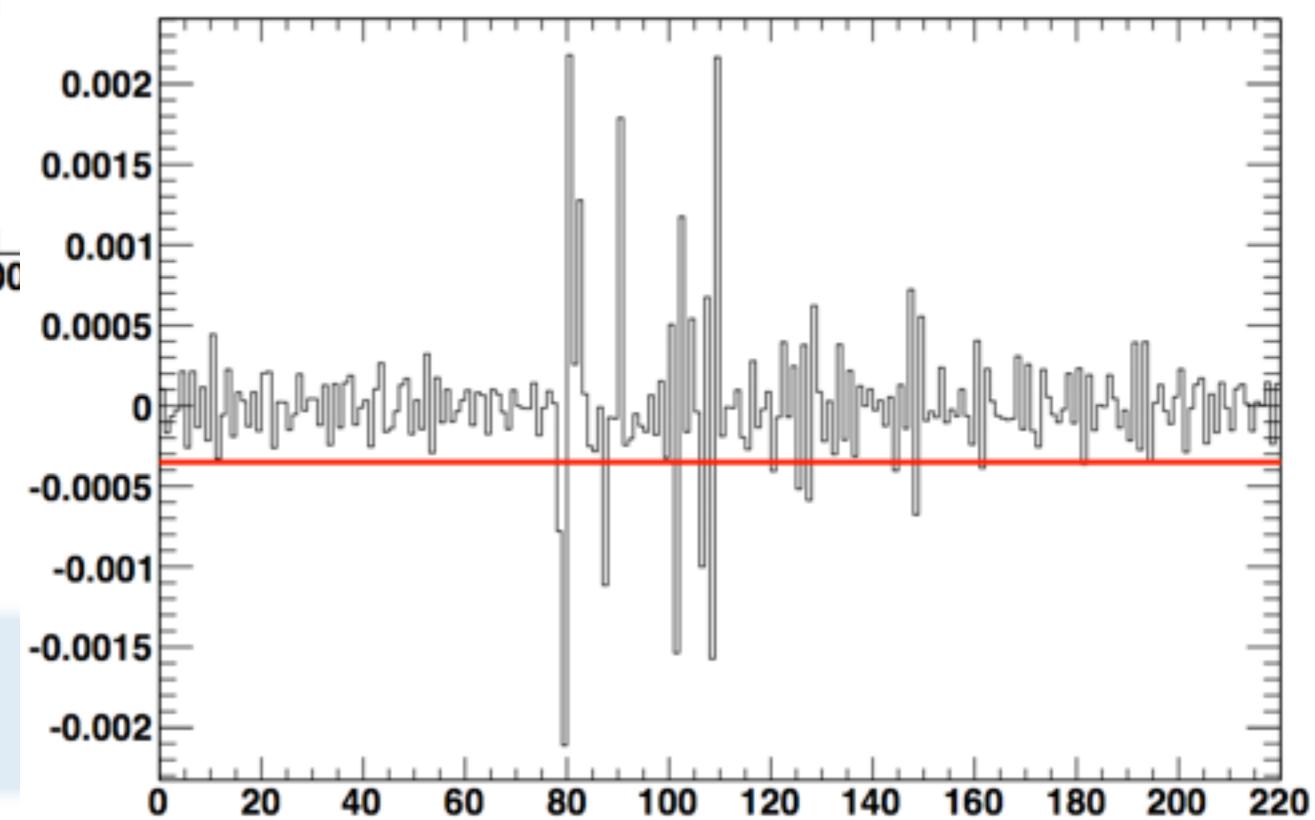


$$Q_0, k = 70, \Delta = 0.0065$$



$$Q_1, i = j = 55, k = 75$$

$$\Delta = 0.0001$$

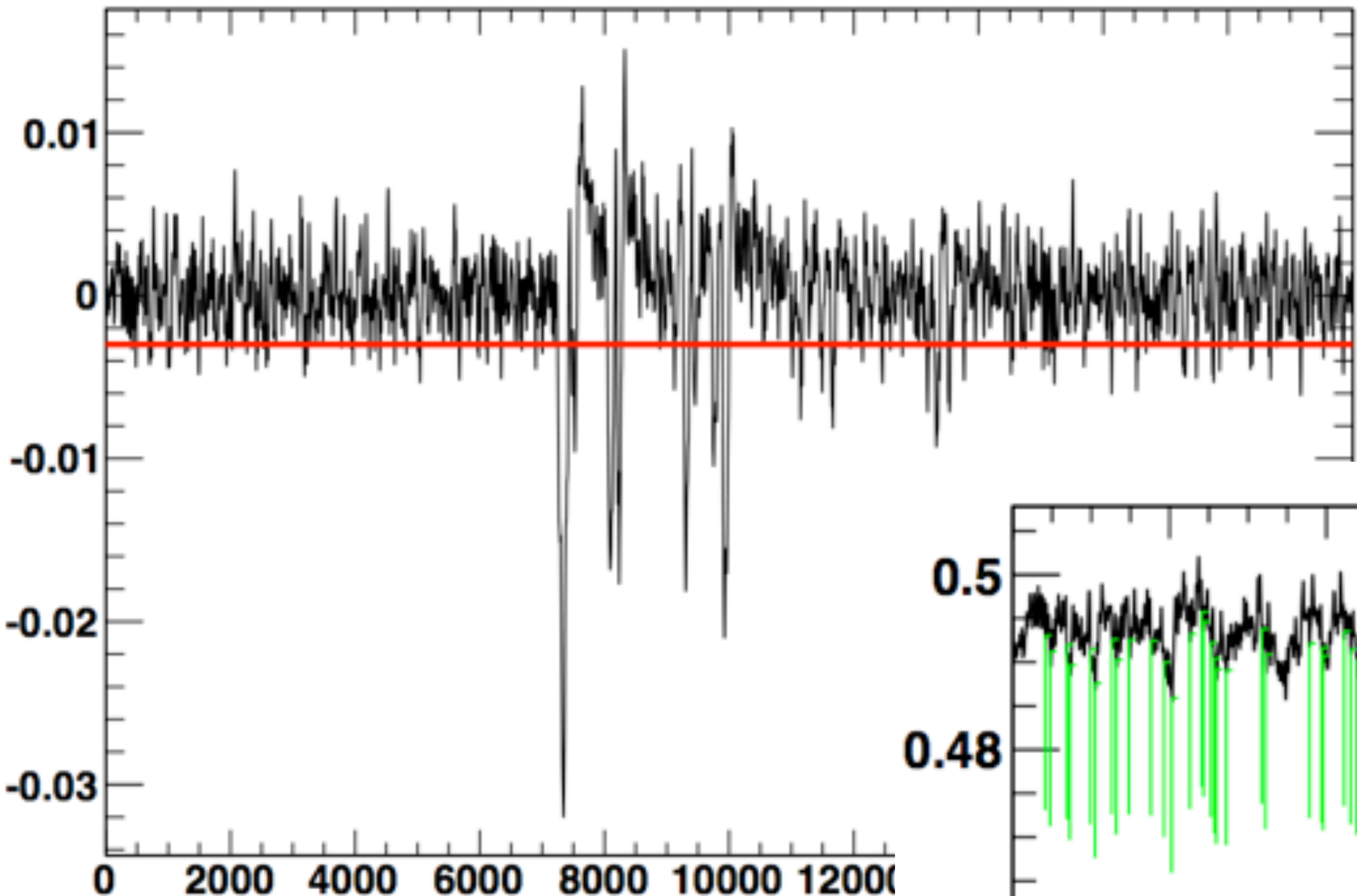


$$Q_3, k = 90, \Delta = 0.00035$$

Timeout Booster

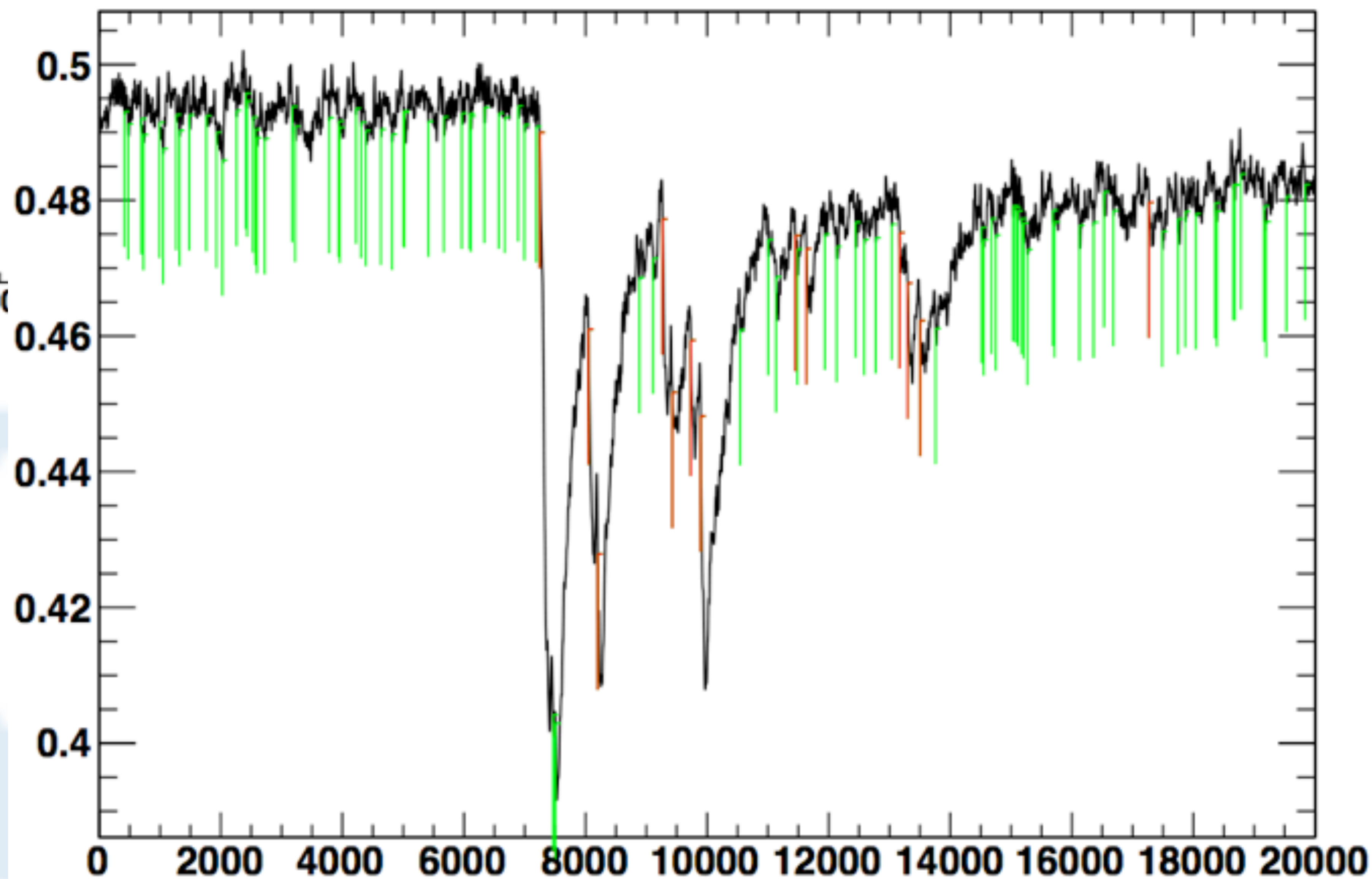
- These threshold-based algorithms only use the “leading-edge” information from each pulse.
- Fake clusters are short-lived pulses, while real clusters have longer tails.
- Can add a “timeout” veto to reject fake clusters.
- Clusters found with algorithm are called “candidates”, and their time is recorded, this time lines up with the leading edge. We check if the voltage recovers to value at the leading edge within a “timeout” window.
- If the voltage recovers faster than the window allows, it is rejected as a fake cluster.
- If the voltage has not yet recovered within the window, it has a longer tail and is kept as a real cluster.
- The real benefit is that it allows lower thresholds to be used without introducing too many fake clusters.

Timeout Illustrated



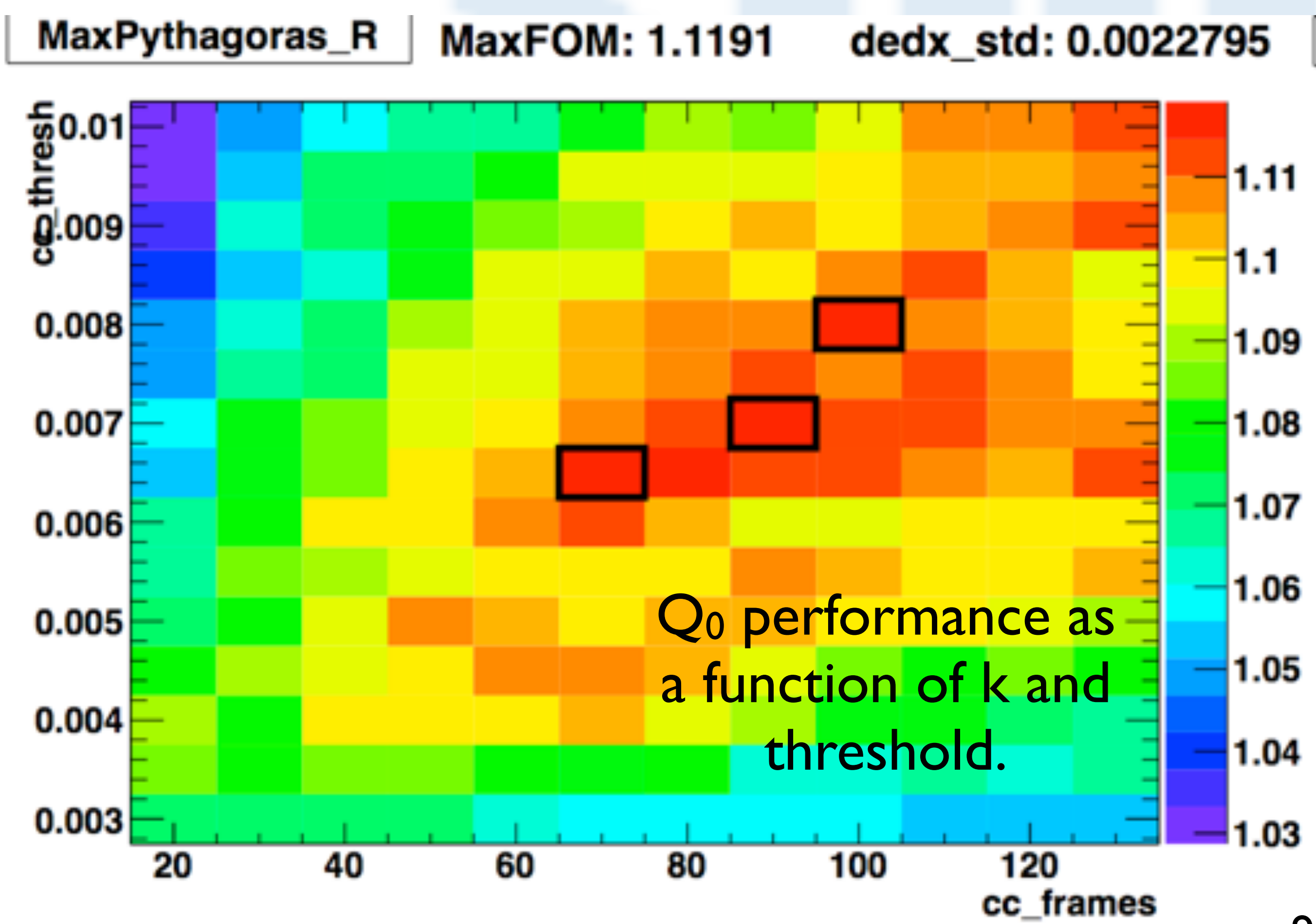
$$Q_0, k = 70, \Delta = 0.0030$$

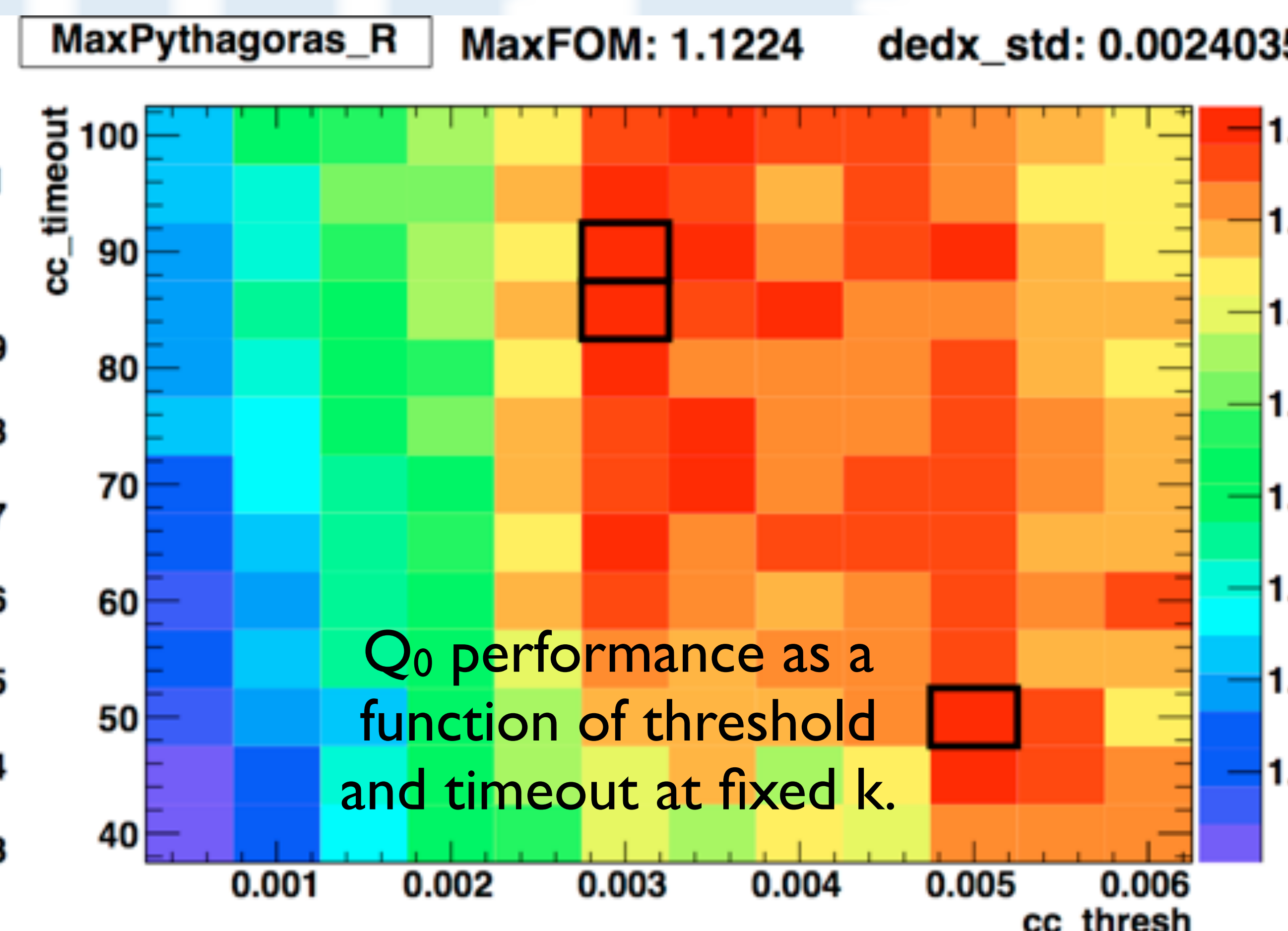
$$T_{\min} = 85$$



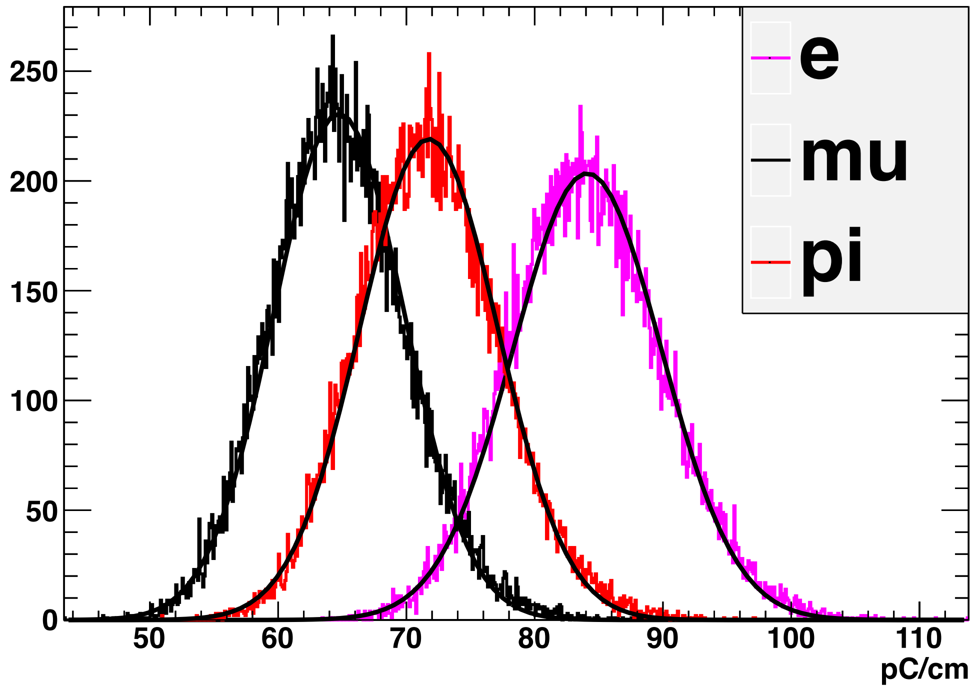
Optimizing Parameters

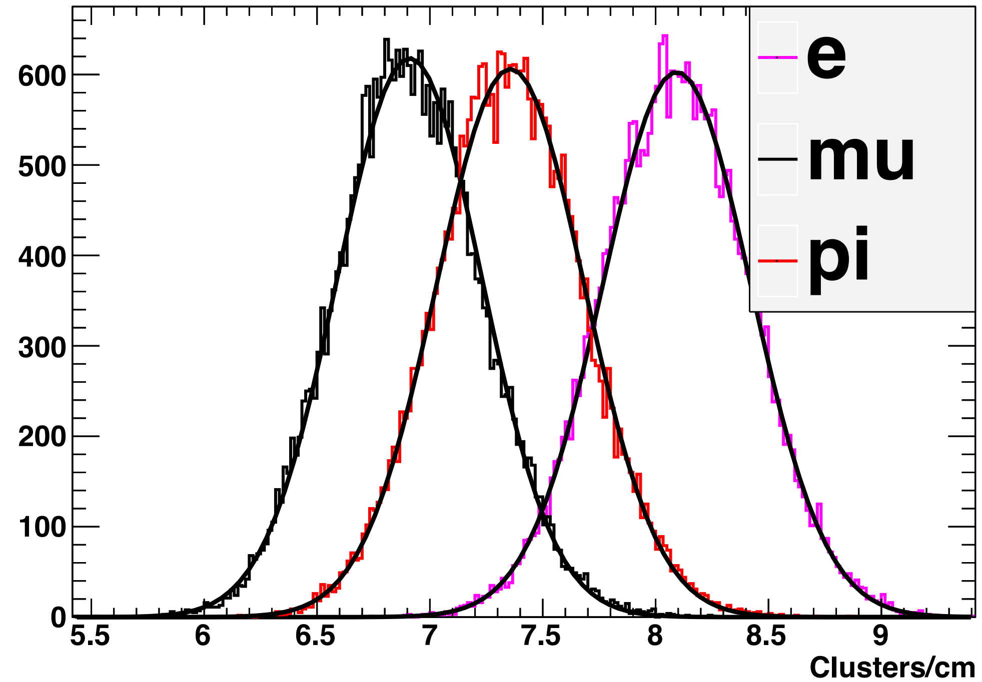
- Optimization is done using figure-of-merit in mu/pi separation plot -> after bootstrapping & combined likelihood fits.
- Do not have a method for automated optimizing -> had to optimize “by hand” by looking at graphs.



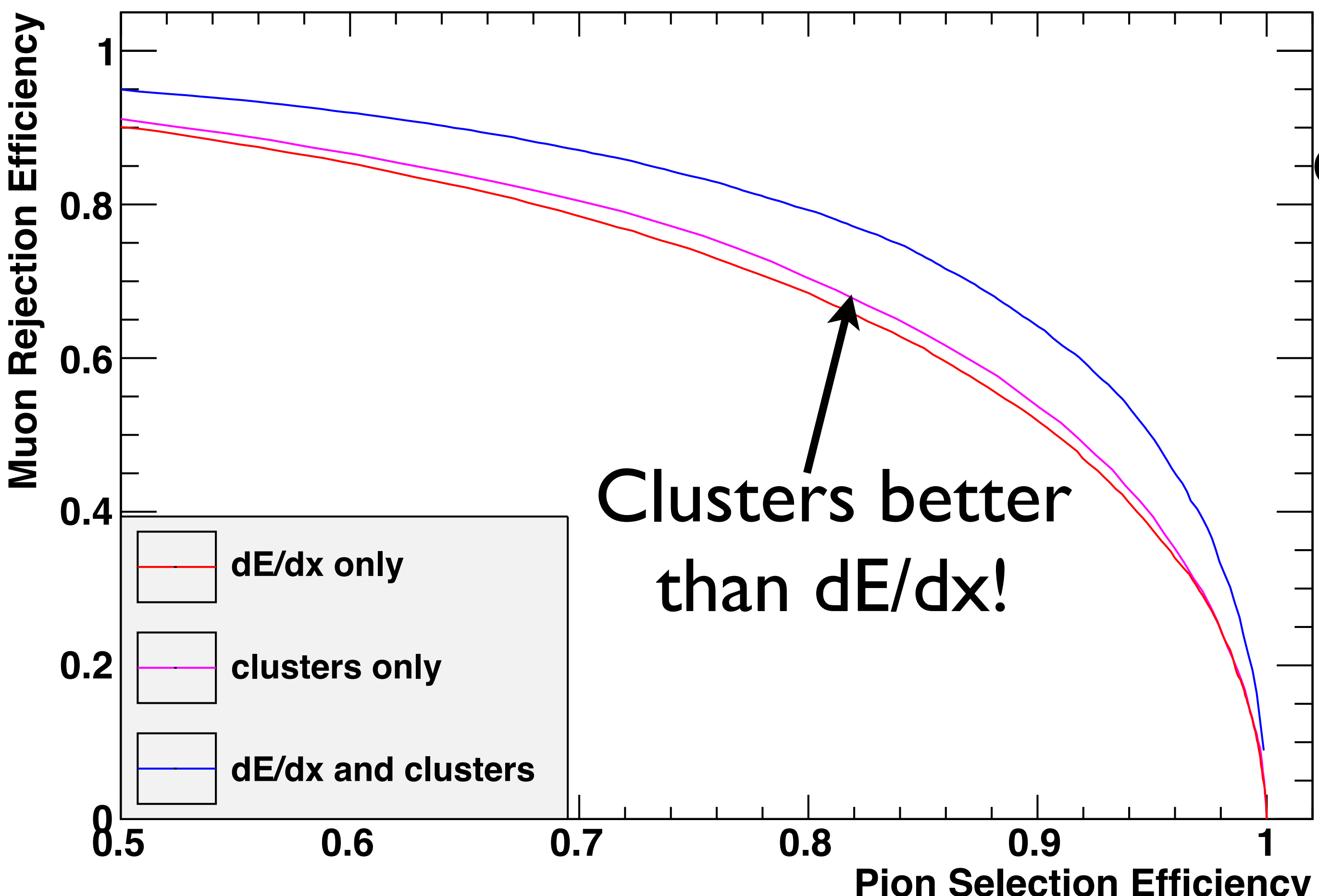


An Optimized Algorithm





Integrated charge and clusters found in bootstrapped tracks.



Run0535 (210 MeV/c)
9200 frame integration time for dE/dx (460ns)
Q_1 with i=j=55 frames (2.75ns), k=75 (3.75ns), threshold 0.0001V/s^2

Comparison of Algorithms

Algorithm	Figure of Merit	Parameters
Q_0	1.119	k=70 threshold 0.0065
Q_1	1.125	i=j=55,k=75 threshold 0.0001
Q_0+time out	1.122	k=70,timeout 85 threshold 0.003
Q_3	1.113	k=90 threshold 0.00035

Variation in figure of merit from dE/dx ~ 0.002

Variation from bootstrapping and intra-run variation still to be studied, but expected to be on the same order.

As expected, a more complex algorithm (more parameters to optimize) offers better results.

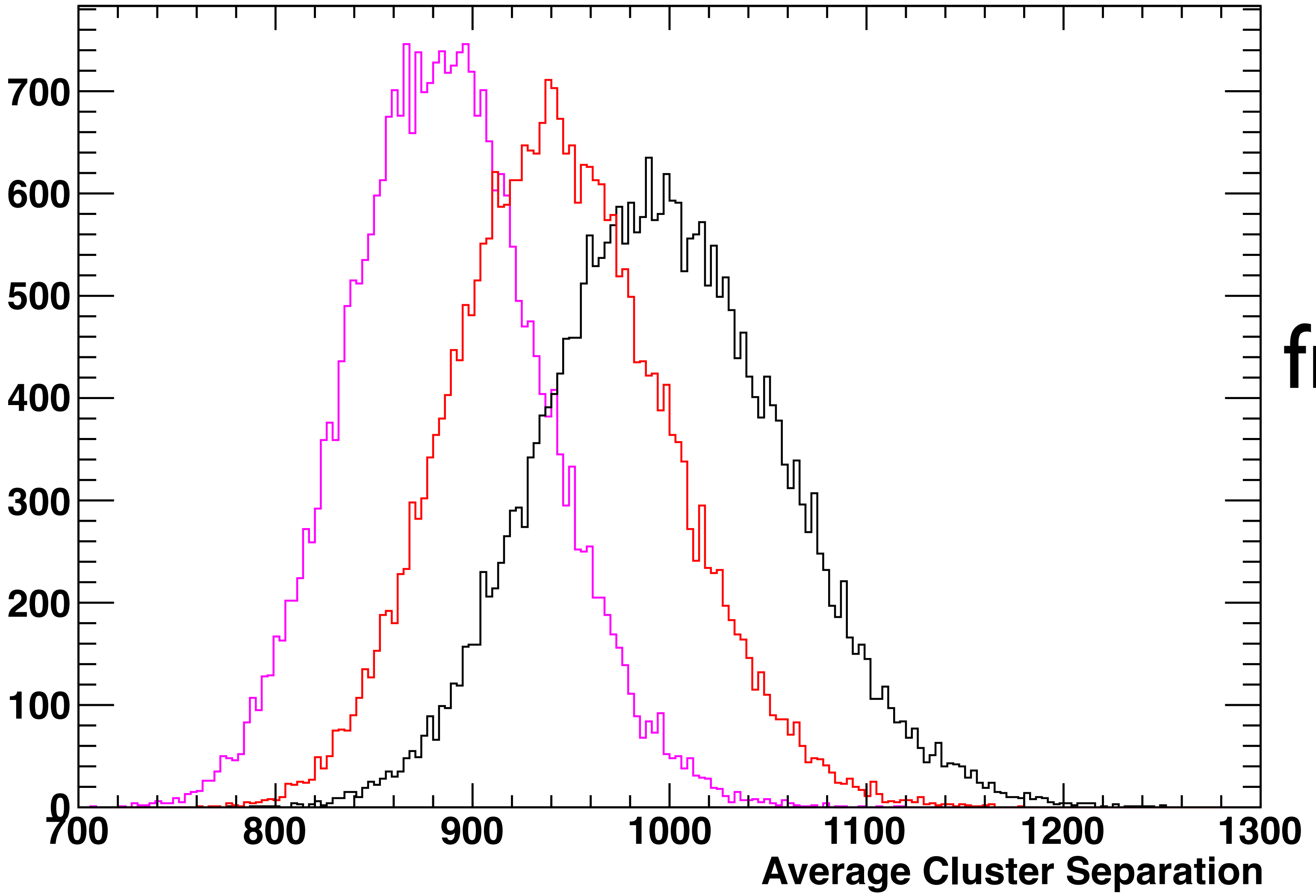
Comparison of Runs

- Variation between runs is consistent across algorithm choices, so we don't need an optimal algorithm to evaluate hardware.
- Gain (HV) variation: lower HV improves performance by ~0.02 FOM.
- Amplifier variation: Non-inverting 370 Ohm amplifier is better than the inverting one.
- The plan is to do a systematic study of all the hardware variations over the next weeks.

Cluster Timing

- Each algorithm can time clusters, in addition to counting them.
- The timing information is also useful for distinguishing different types of particles, so why throw it away?
- We partially reduce the data by storing only the average cluster separation in each cell.
- The track-wise average cluster separation is reconstructed from the cell-wise average separation and the number of clusters in each cell.

Cluster Separation

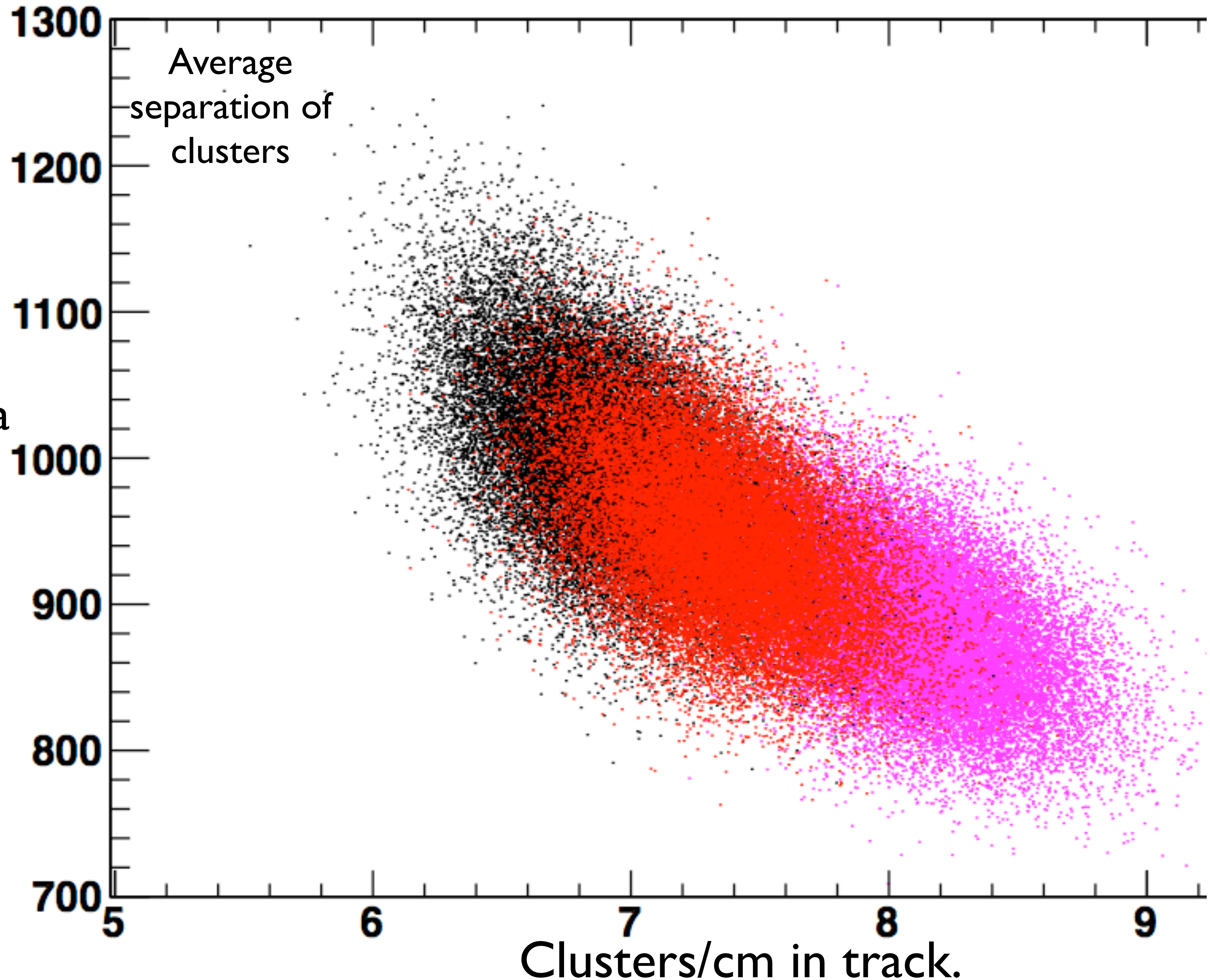


It looks like we get similar separation as from the cluster-counting or dE/dx. Is there new information here?

By a theorem, as the track length goes to infinity, the average separation is just the inverse of the number of clusters per unit length.

Separation and Count

Indications are that the correlation between cluster count and average cluster separation is small enough. Thus a likelihood can be formed from the separation and used for PID.

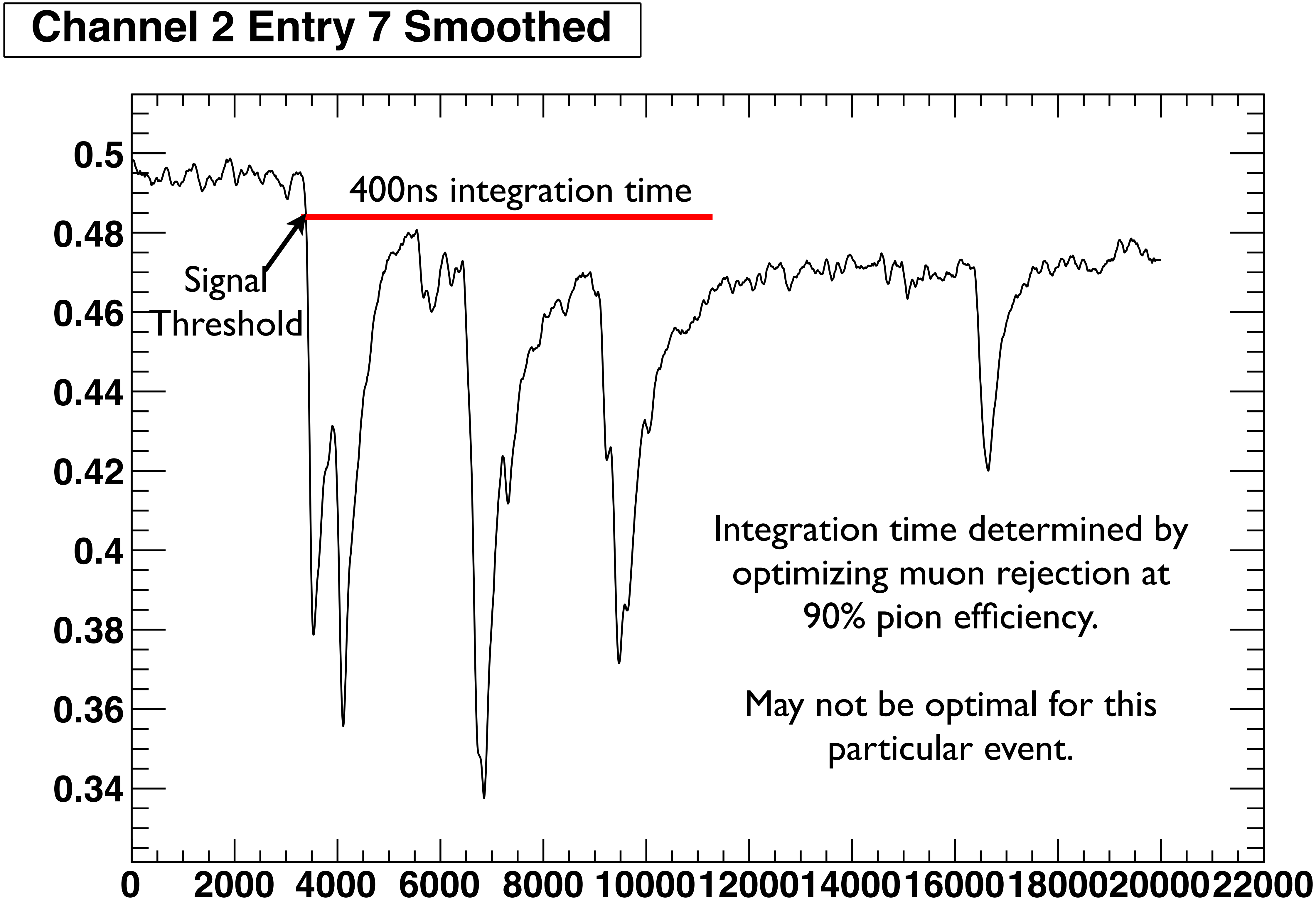


Next Steps

- Systematic study of equipment variations (cables, amplifiers, HV, wire diametre, gas).
- Implementation of cluster timing likelihood.
- Inclusion of data from December runs.
- NIM paper draft & submission by early 2013.

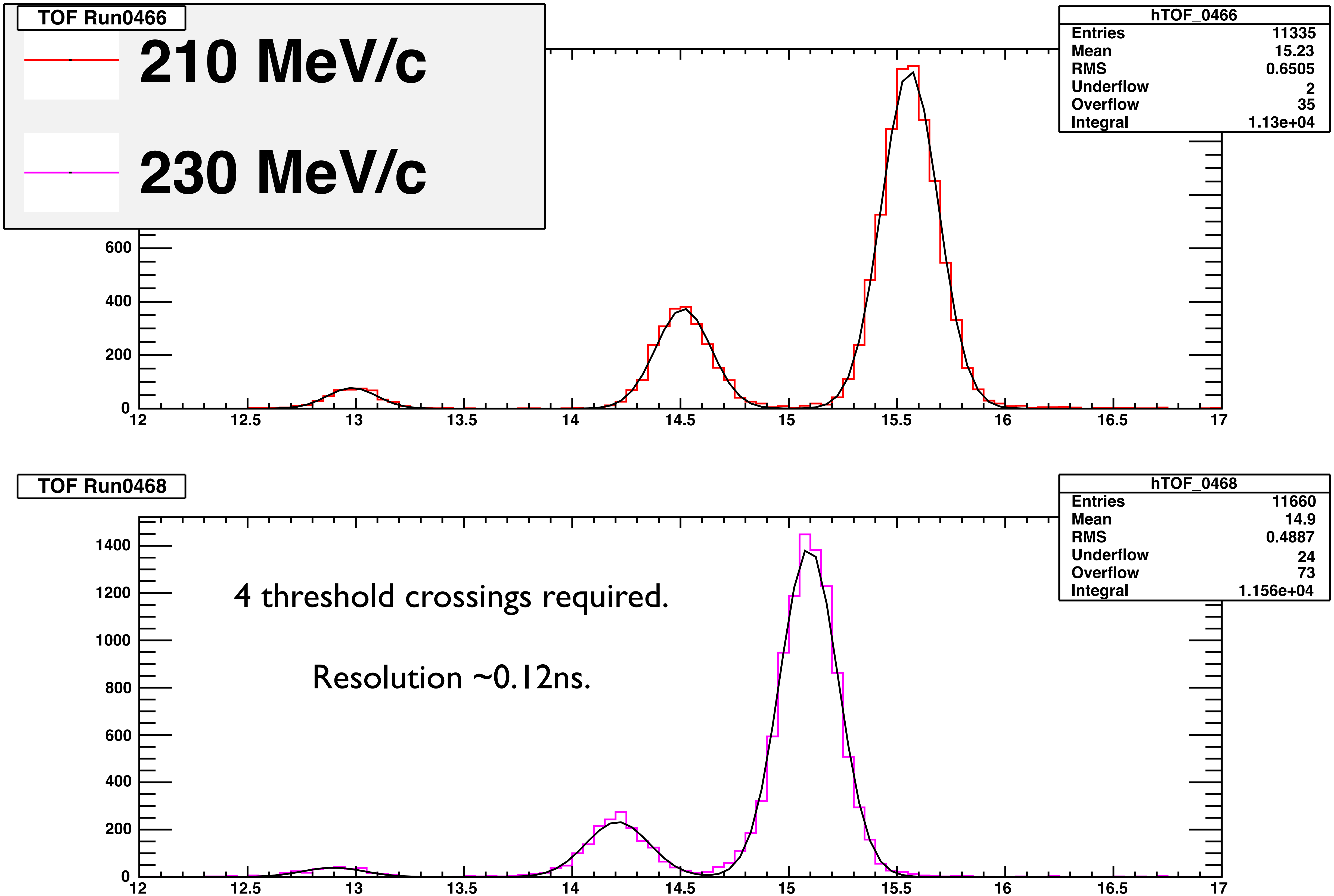
Backup Slides

Signal Determination

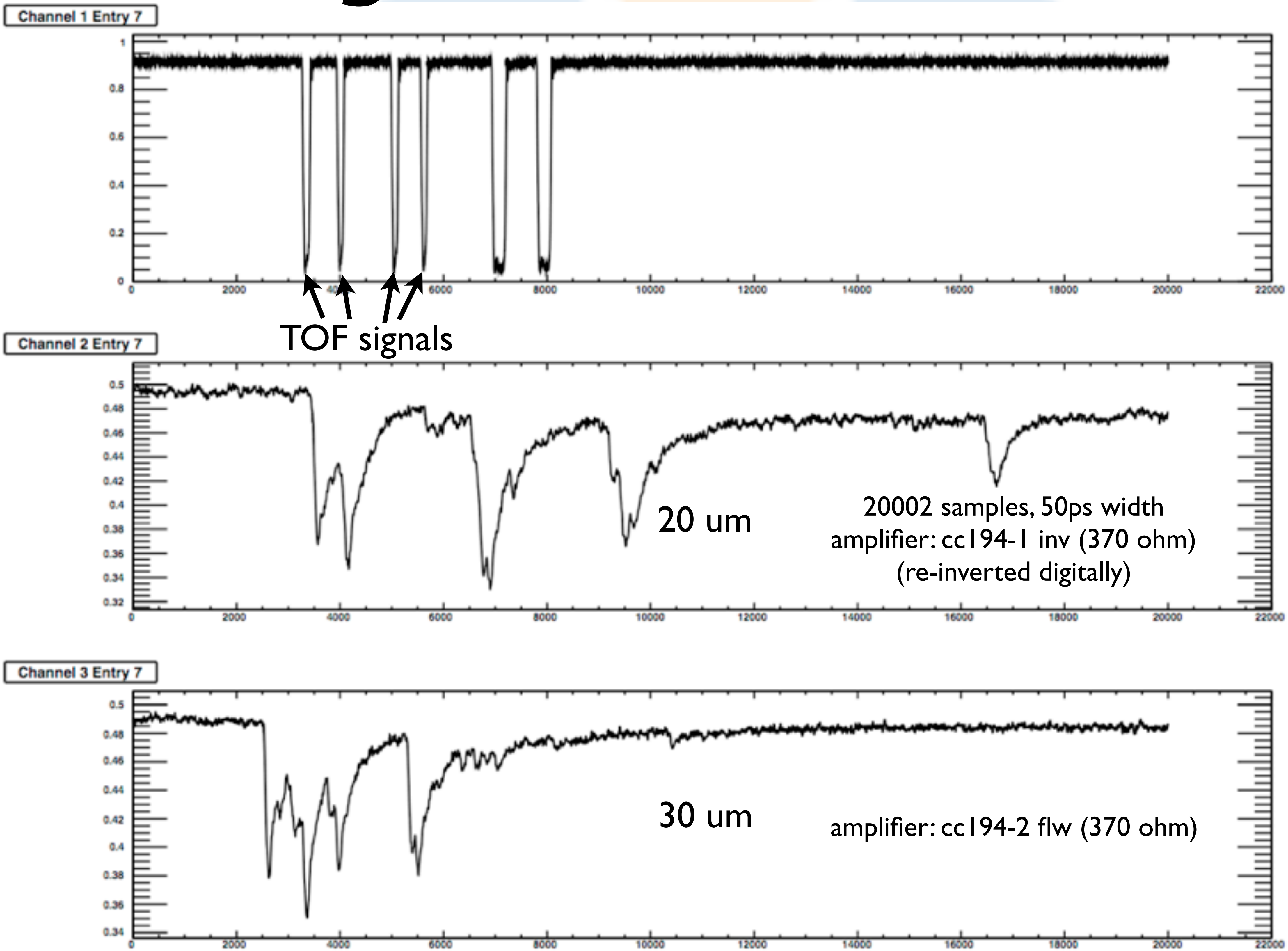


amplifier: cc194-1 inv (370 ohm)

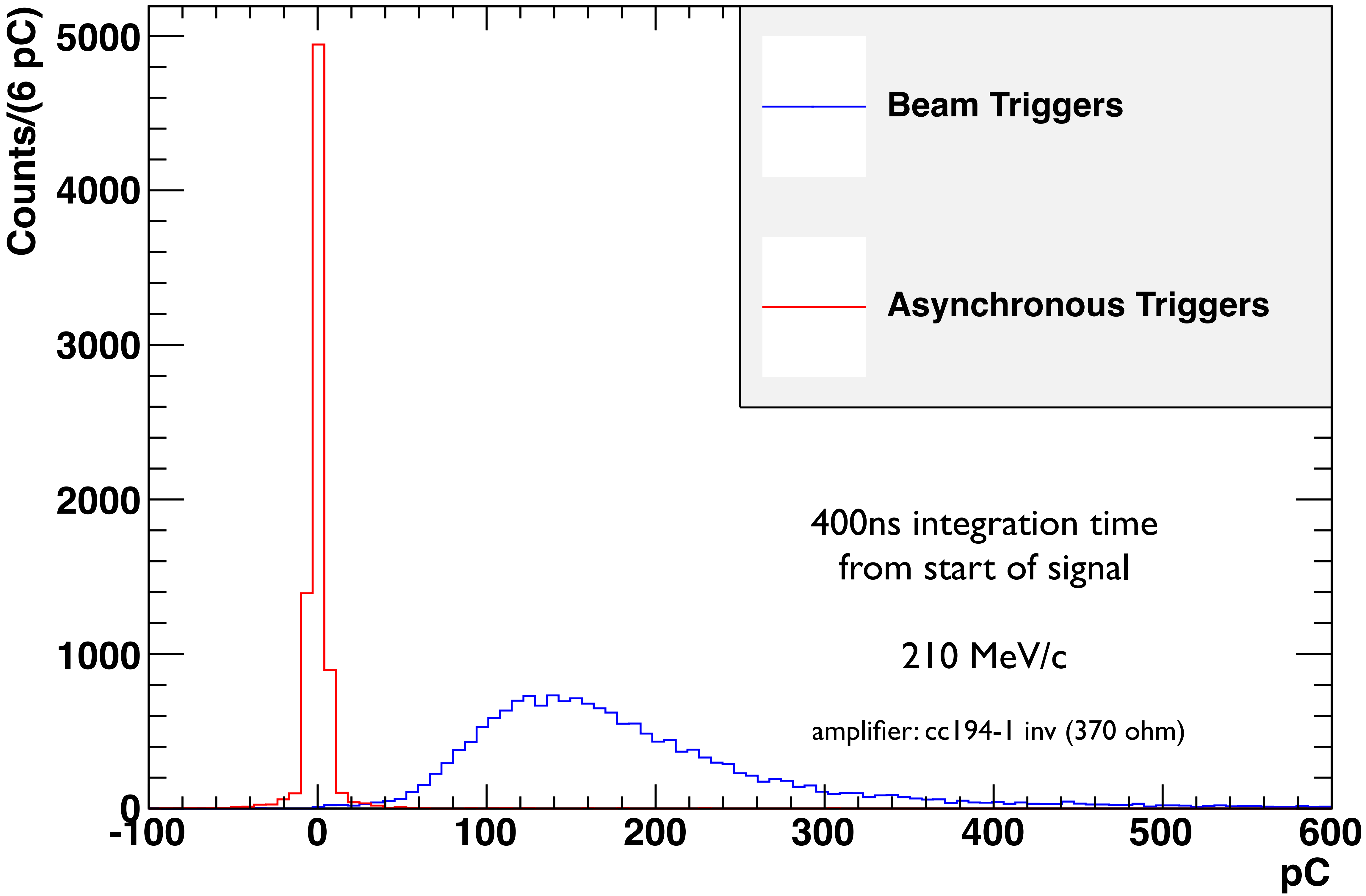
TOF Calculation



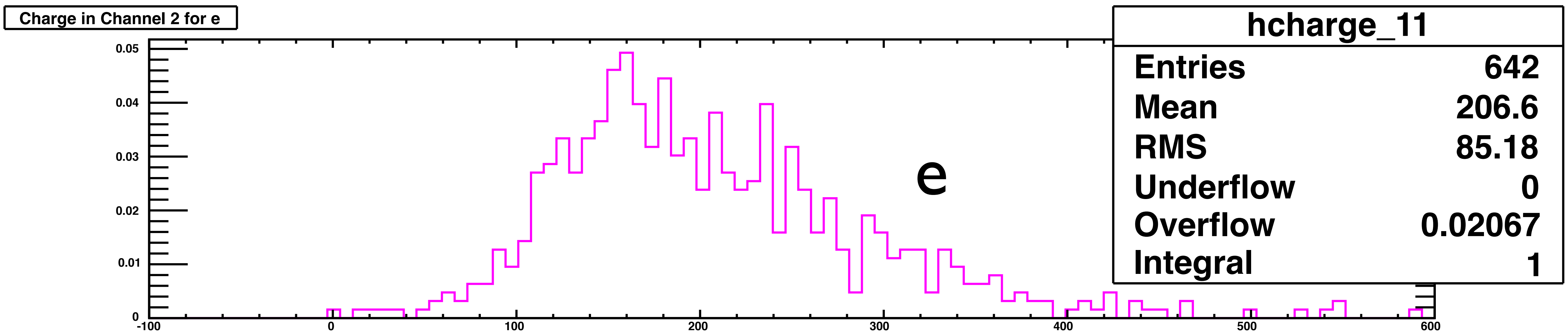
Digitized Waveforms

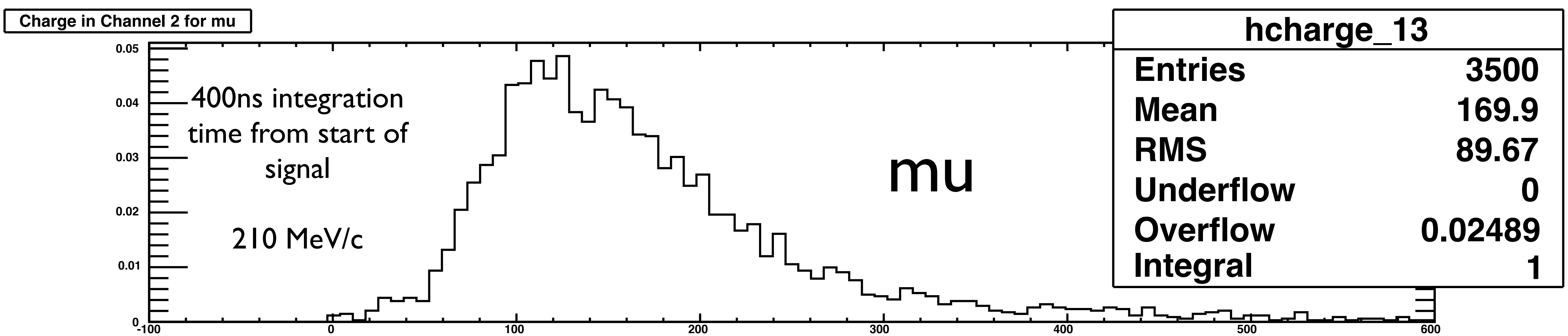


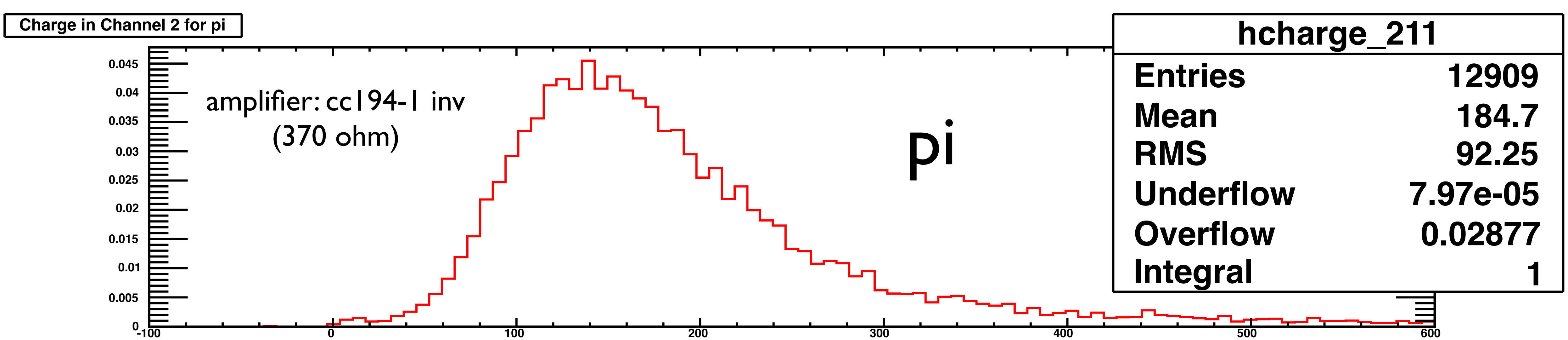
Integrated Charge (Single-Cell)



Integrated Charge (Single-Cell)







TOF Rejections

Events which have zero threshold crossings are usually asynchronous triggers.

Events with 4 TOF threshold crossings are usually beam triggers.

Events with neither 0 or 4 are rejected.

Signal Determination Details

Baseline and threshold is obtained from previous asynchronous trigger.

Baseline = average of smoothed asynchronous event.
RMS = true RMS of smoothed asynchronous event.

(Smoothed signal - Baseline) < 5*RMS is a threshold crossing.

Start time of signal taken 5ns before threshold crossing. Smoothing is an average done over 5ns.

Beam events whose signals did not properly cross the threshold are rejected.

This algorithm is also applied to asynchronous events, if they cross the threshold, a warning is printed but the event is not rejected.

Charge Integration Details

From the signal start, as described earlier, the unsmoothed signal is integrated over a fixed duration.

For asynchronous events, the integration starts at an arbitrary time of 2000ns.

The charge is baseline-corrected by subtracting the integral of the previous uncorrelated trigger.

I determined that an integration time of 400ns to be appropriate for the amplifiers used in the runs shown. The optimal value depends on the specific amplifier.

Bootstrapping Basics

Particles are identified by TOF: must be within 3 sigma of mean of Gaussian fit.

Tracks are composed of randomly-selected same-PID triggers.

Track-wise dE/dx and cluster counts are done.

Bootstrapping

The TOF spectrum is fitted with the sum of three Gaussians, for electrons, muons, and pions. (See slide 3)

Each event is then assigned a PID value depending on its TOF value. If it is within 3 sigma of a fitted Gaussian peak, it is assigned that particle type, otherwise it is unknown.

40-cell tracks are constructed by putting together 40 events with the same PID value. The events are randomly selected with replacement.

For the tracks, the 70% truncated-mean charge and cluster counts are obtained.

The dE/dx and cluster information for each hit in each track are consistent, due to the way the code is structured.

Likelihood Ratio

Fit with Gaussians both dE/dx and cluster distributions for each species. Obtain mean and standard deviation for each.

i Species

Q Track dE/dx

N Track Clusters

P pdf of Gaussian fit

Likelihood (for dE/dx or cluster count) for each track is the probability density of the actual dE/dx or cluster measurement according to the fit. Total likelihood is the product of the dE/dx and cluster probability densities.

$$L_i = P_{i,Q}(Q) * P_{i,N}(N)$$

The likelihood ratio is formed from the total likelihoods of being a pion and muon.

$$R = \frac{L_\mu}{L_\mu + L_\pi}$$