Measuring LAr purity from cosmics tracks

LArIAT Meeting
April 28th, 2015

Roberto Acciarri

Before we start...

- Of the two methods – LAPD like and single track minimization like – described in the last Physics Meeting I concentrated on the latter, given its potentiality to provide a lifetime (τ) estimation using only few tracks

- At the moment, the lifetime measurements obtained are fairly consistent with the simulated ones even for high τ

- On the other side, there are couple of passages in the code that gave a lot to think about to both Michelle and me. I'll present them as clearly as I can

The Method

✔ Select straight tracks crossing the whole TPC

✔ Given a track, compute for each hit of the track the Landau probability l to obtain the observed dQ/dx as a function of the "uncontaminated" value dQ_0/dx and the lifetime τ

✔ Minimize the logarithm of the sum of l over all the hits - max likelihood -ln(L) - respect to τ, dQ_0/dx and σ (width of the Landau)

✔ Eventually, average the values of τ obtained for the selected tracks

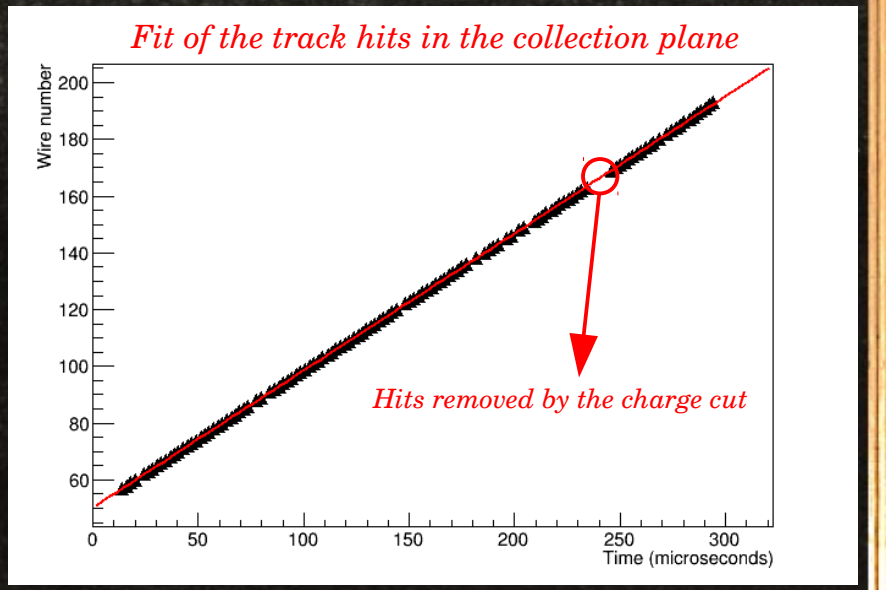
Track Selection

The following cuts are applied to the events:

✔ TrackId==0. Cosmic Tracker is good in finding straight tracks and reconstruct visible delta rays as indipendent tracks.

✔ Track must start and end within 4 cm from wireplanes and cathode and within 10 cm from the upstream and downstream end of the TPC.

✔ Track must contain at least 100 hits. This together with the previous point ensures that the track covers as much of the drift time as possible.



✔ Linear fit of the hit wire vs hit time in the collection plane must have a χ^2 <0.05. Make sure the tracks does not change its picth due to scattering and remove tracks with visible delta rays not reconstructed as indipendent tracks.

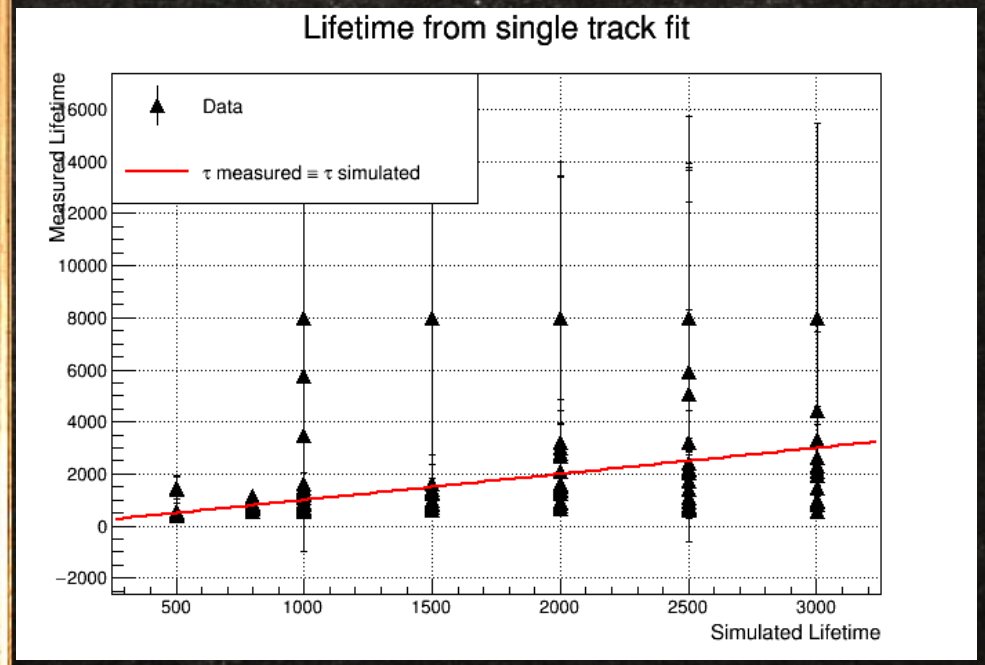
✔ In addition, the minimization is ONLY performed using the hits of the collection plane with charge below a certain value (to remove small delta rays aligned to the track).

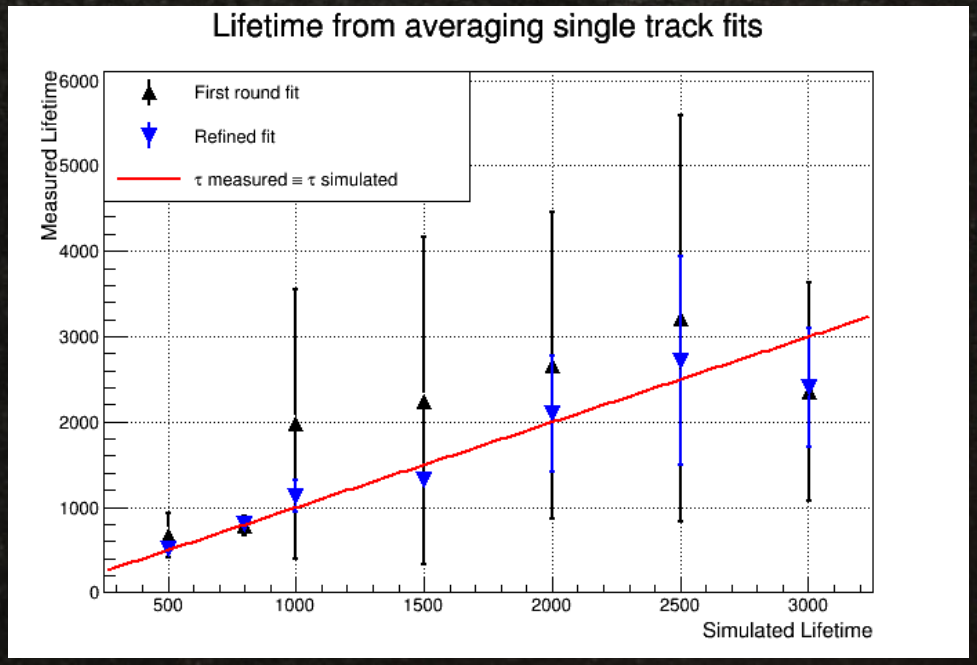
~ 20% of simulated events pass these cuts

Simulation

- 7 τ values considered: 500 μs, 800 μs, 1000 μs, 1500 μs, 2000 μs, 2500 μs, 3000 μs

- A sample of 100 events (~ 2 hours of data taking) have been simulated for each lifetime value





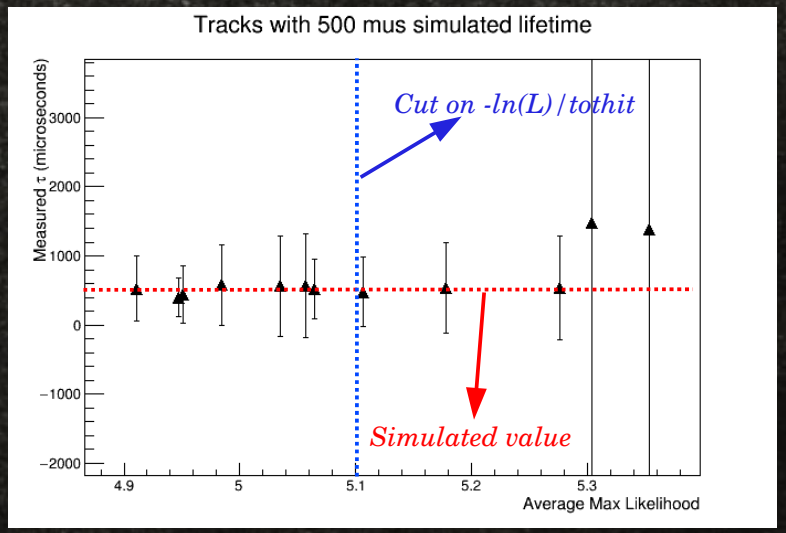
Considering only the black triangles:

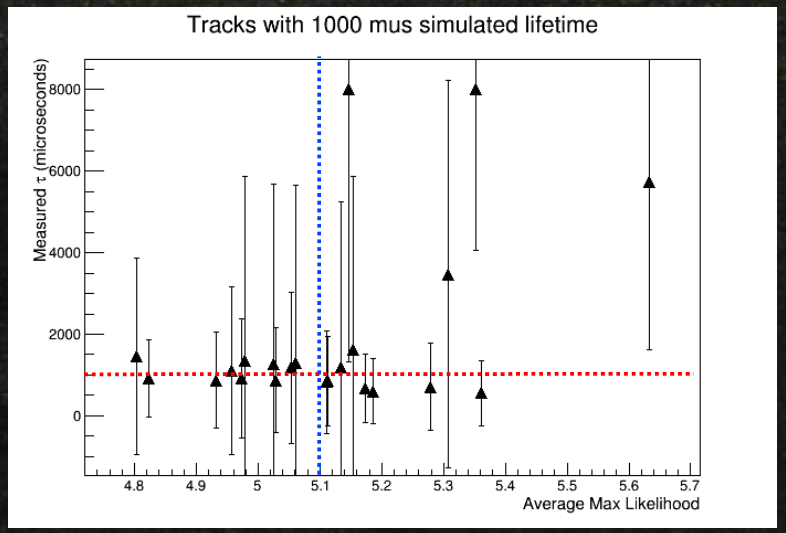
- Average calculated is not weighted with the minimization errors

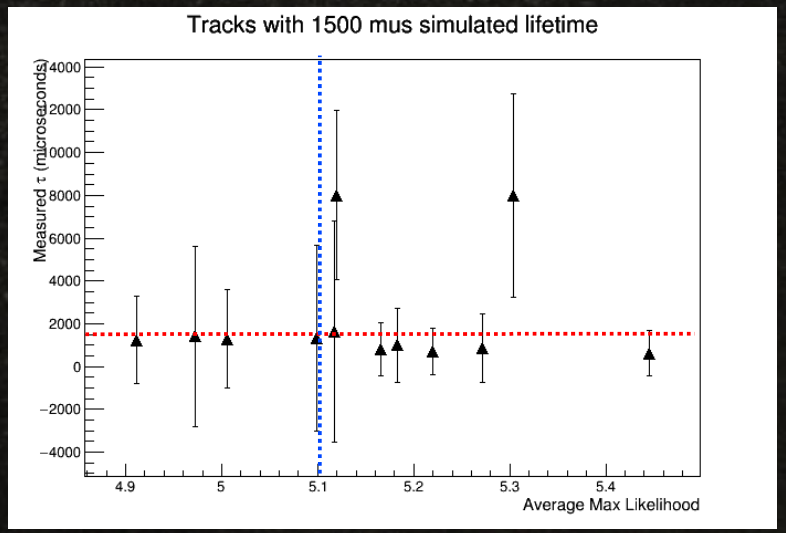
- As τ increases, the measured value progressively deviates from the simulated one, as was happening with the first method I tried

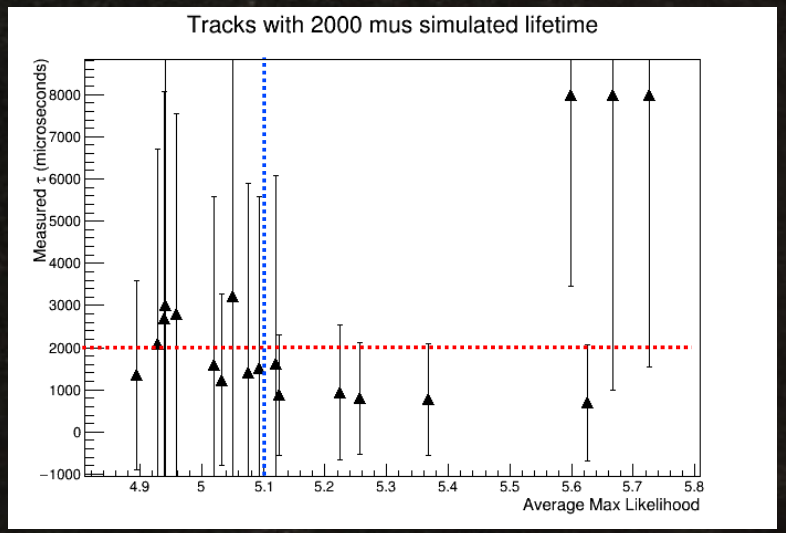
Additional cut – the tricky one

Within each simulated sample, there seems to be a correlation between the value of τ obtained from the minimization and the "average max likelihood" (-ln(L) @ minimum divided by the number of hits in the track)

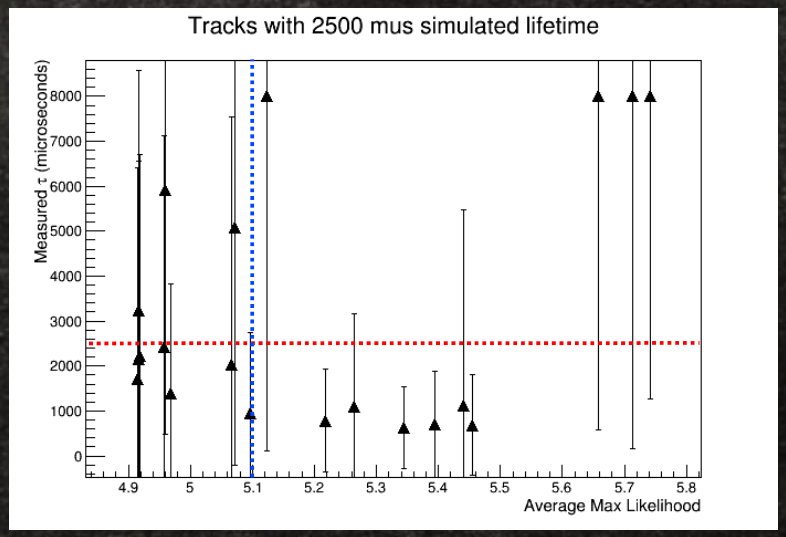


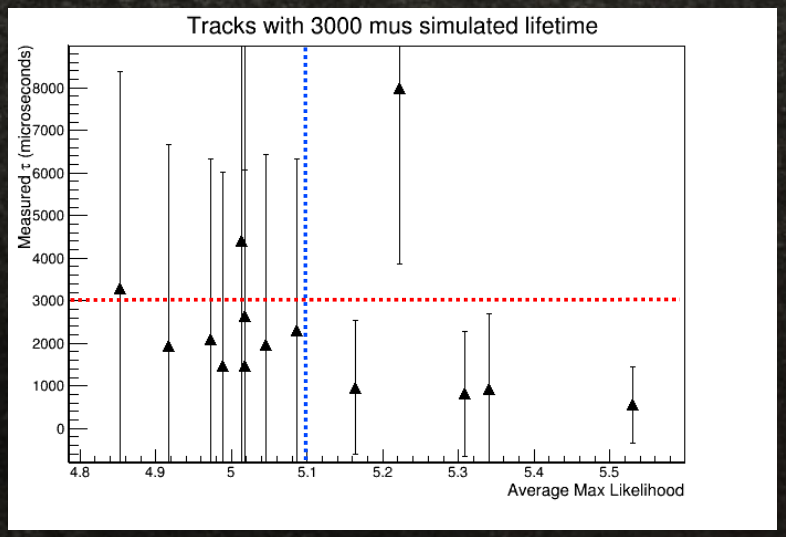






Additional cut – the tricky one





- The measured τ deviates more easily from the simulated one at the highest average max likelihood

- As described in “Pitfalls of Goodness-of-Fit from Likelihood” by J. Heinrich, the max likelihood, even though is sometimes used, is actually not statistically well defined as GOF

- On the other side, there was an initial idea that the errors provided by MINUIT weren't always correct, as the real value of τ doesn't always fall inside the error range of the measurement

- Nonetheless, looking at the minimization curve of -2ln(L), it seems like if error miscalculation may actually not be the problem...

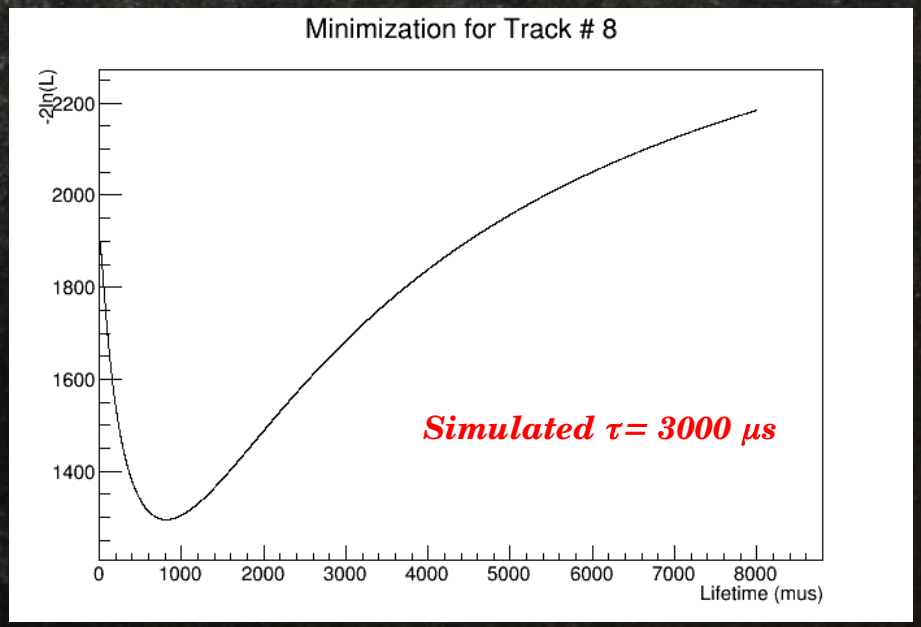
Additional cut – the tricky one

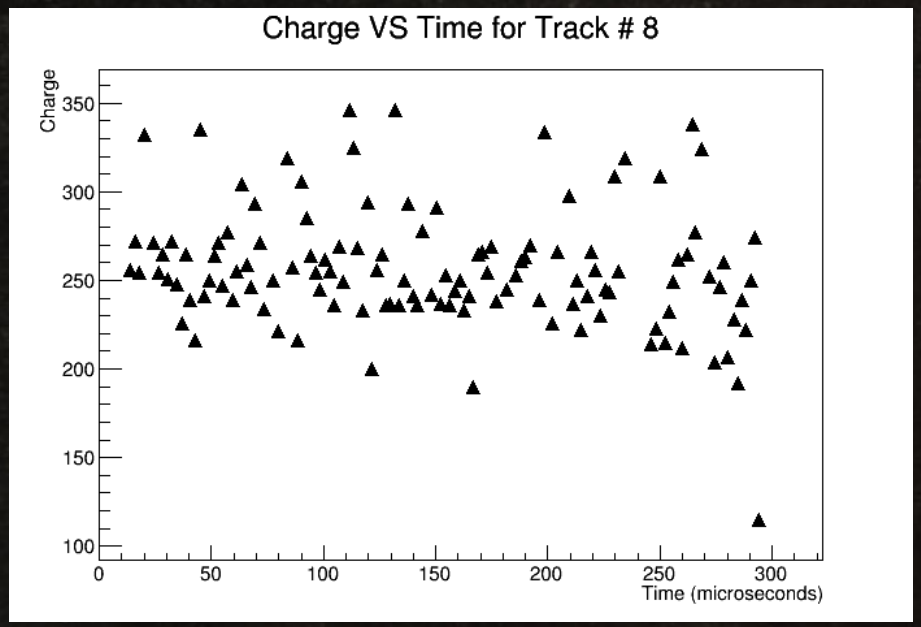
- Top plot is the value of -2*ln(L) as a function of τ. Even applying the correct errors the real value of τ wouldn't be reached

- I suspect the problem being the same found with the first method: given the small size of the TPC, the fluctuation of the charge in the hits is comparable to the charge loss due to contamination

- This may tell us that we are simply not sensitive enough to high purities at our operating electric field

- A "practical temporary solution" - which is <u>theoretically uncorrect</u> but can give an idea of where the real purity value lies - is to average the values of τ obtained for the selected tracks, excluding those values for which -ln(L) is too high.

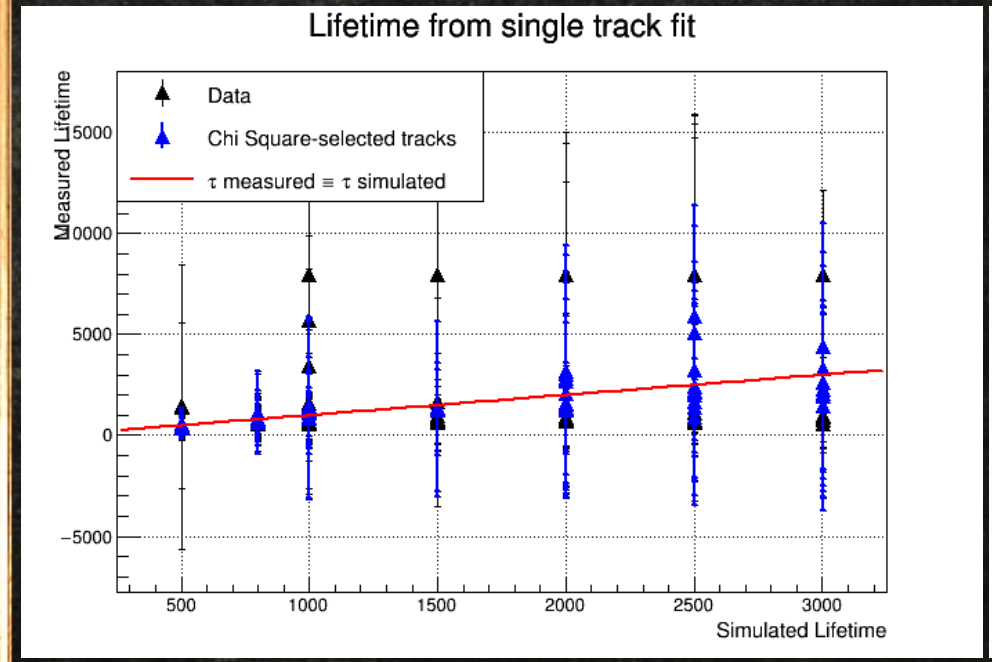


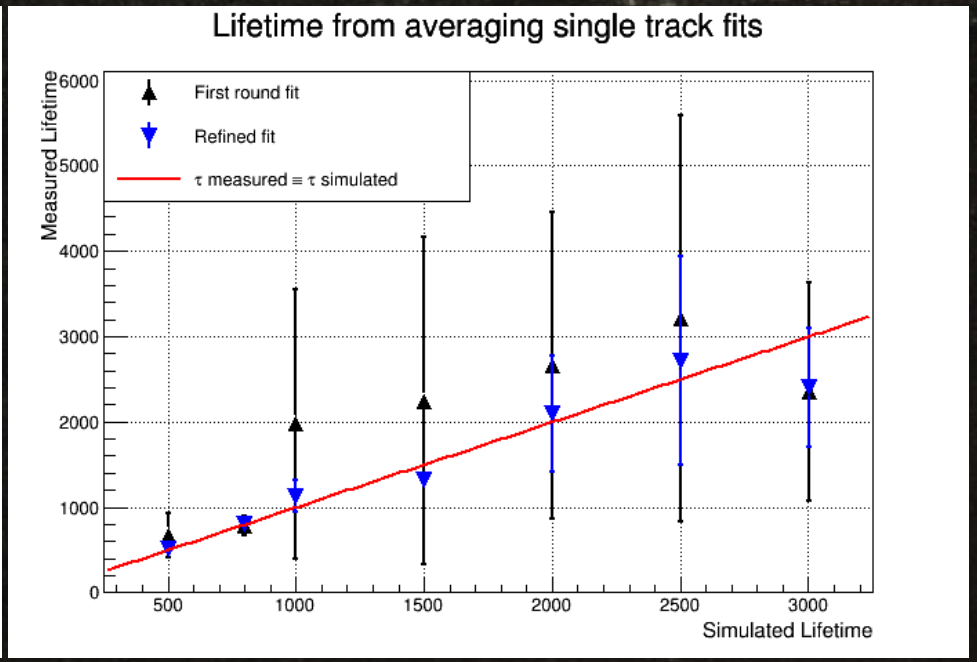


Applying the two sets of cut

- 7 τ values considered: 500 μs, 800 μs, 1000 μs, 1500 μs, 2000 μs, 2500 μs, 3000 μs

- A sample of 100 events (~ 2 hours of data taking) have been simulated for each lifetime value





Black triangles: tracks passing the first set of cuts

Blue triangles: tracks passing the first set of cuts and cut on -ln(L)/tothit

~ 10% of simulated events pass both sets of cuts

Additional test

To make sure of not being biased by a particularly "lucky" simulation, I simulated 600 events with τ = 3000 μs and run the program for 6 subsets of 100 events each, as well as for all the events together

Measured τ (μs)	***Sample (track #)***
2400 +- 701	0-99
3075 +- 1014	100-199
2999 +- 927	200-299
3320 +- 1577	300-399
2882 +- 1802	400-499
3051 +- 1461	500-599
2986 +- 1264	0-599

Once again – this is a temporary solution. I'm working on getting better error estimation without using the cut on the max likelihood

If we have to measure the purity today...

- The code surely needs improvements, a better understanding of the errors, and it'll need to be tuned on real data. But I'm confident it can be used to obtain a first rough estimation of the purity

- My suggestion is to have - once the TPC is full and we made sure we can apply voltage to both cathode and wireplanes – at least 2 hours run at low E-Field (half or even less if possible) to get a good estimation of the purity

- Once we are sure we understand the purity and that the code is correct, we may think to have the code running as cronjob to automatically measure the purity every hour or two. I talked with Johnny and it shouldn't be complicated for him to readout the output of the program and produce an online plot

- The code is now divided in two parts:
  - *boreco.fcl* (fcl file inherited by LAPD), which takes care of track reconstruction using cosmic tracker and creates an AnaTree – a ROOT Tree with all the raw and recosntructed information needed
  - *purity_minuit.C*, a ROOT macro which retrieves information from the Anatree to calculate the lifetime

- ***What I don't have/don't know yet*** is how to translate the events from the structure they're written by the DAQ to a structure readable by a standard fcl file