Global PID

Melissa Uchida

CM46

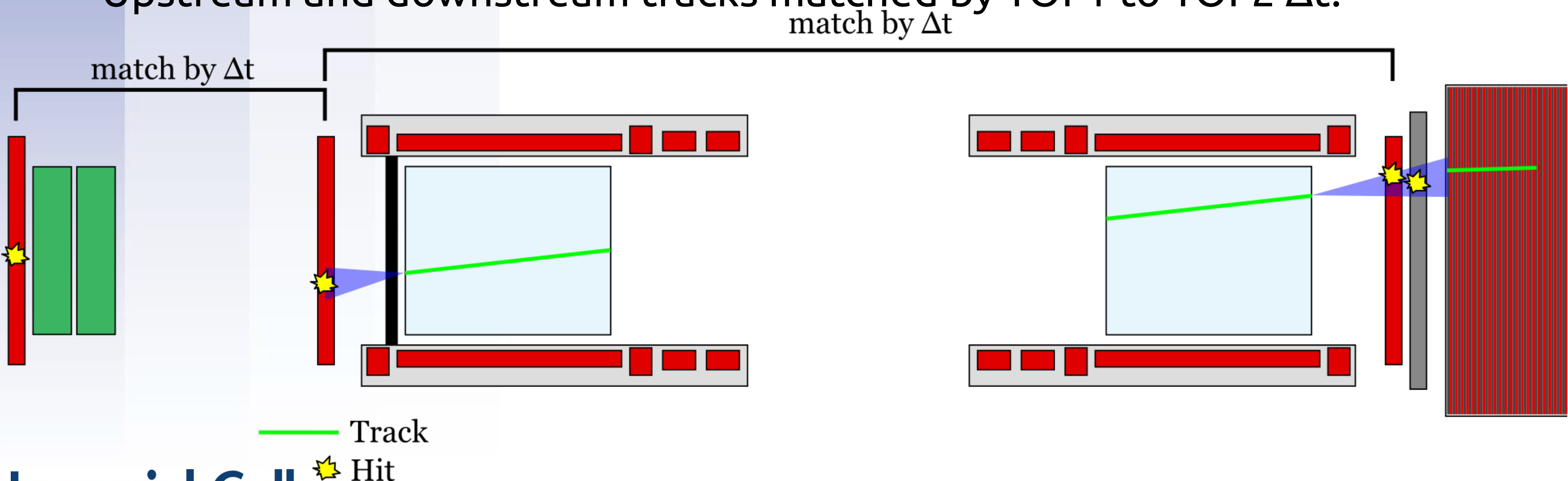
Imperial College London

Summary

- Celeste has now written up and is handing PID over to me.
- PID Overview
- Celeste’s Analysis
- The future

Track matching

- See Jan's talk infrom the software session.
- Track matching occurs outwards from the Tracker tracks
  - Well-reconstructed tracker information crucial.
- Propagation from Tracker track to other detectors with RK4, accept / reject based on agreement.
- TOF0 matched by Δt to TOF1 as quads magnify uncertainties in propagation
- Upstream and downstream tracks matched by TOF1 to TOF2 Δt.



Global PID

- Combines information from detectors into a set of PID variables—quantities that can be used to distinguish between possible PIDs.
- Produces probability density functions (PDFs) of these variables from MC data, for each possible PID.
- Calculates the value of each variable for a track matched data track.
- Compares these values to the corresponding PDFs for each PID, obtaining a likelihood for the track to have been made by a particle with that PID.
- The likelihoods (as log-likelihoods) are combined and the confidence level of the track having each PID is calculated. If one confidence level is sufficiently higher than the others (can be user defined) then the track is assigned that PID.

Celeste Pidcott PID Analysis

- Straight (commissioning) 3-200 muon beams through the LiH absorber
  - **Runs: 07834 07835 07836 07837 07838 07841 07842 07843**
- Matching PID run on MC input from a straight 3-200 g4bl muon beam (actually ~175 MeV/c at the absorber).
  - Mean momentum so distribution overlap, high momentum tails and all variables used in the PID analysis included the TOF (where momentum dependence is somewhat accounted for).
- Efficiency and purity, and PID variable comparison studies were performed.

C.P. PID Analysis Data

	Run #								
	07834	07835	07836	07837	07838	07841	07842	07843	Total
# suitable tracks	1140	10648	9957	9912	10875	3349	1957	9126	56964
# identified	1084	10104	9427	9377	10375	3192	1852	8682	54093
# muons	1013	9591	8952	8879	9816	3023	1736	8241	51251
# pions	64	452	418	448	498	154	102	389	2525
# positrons	7	61	57	50	39	15	14	52	295
# failed	56	544	530	535	500	157	105	444	2871

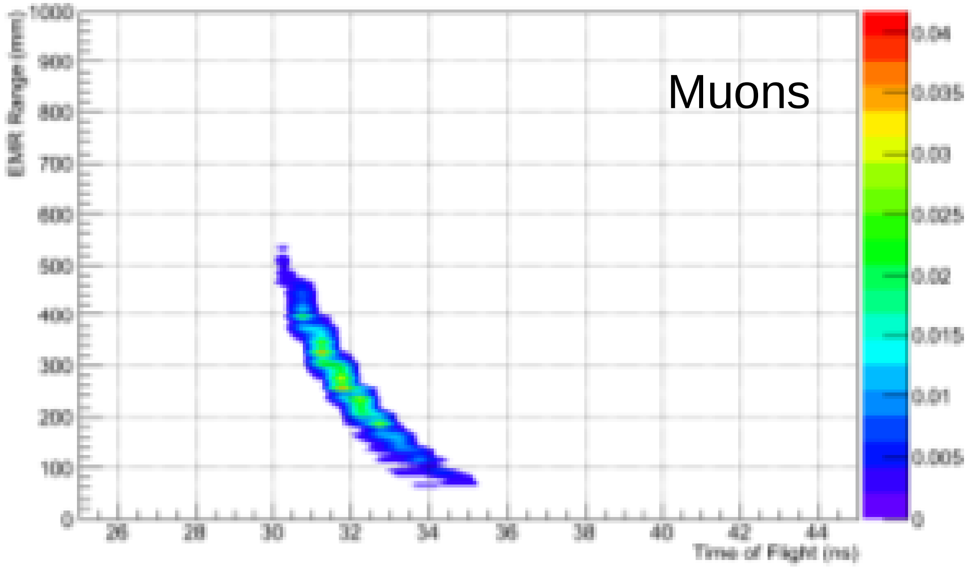
Table 3.10: Results of PID for all 200 MeV/c muon LiH runs.

- Suitable tracks are a track that contains at least enough detector information for at least one PID variable to attempt an ID, but that also doesn't contain extra/spurious/non-physical hits, e.g. doesn't have two TOF1 hits (taken care of by track matching).
- Tracks where PID failed were due to the tracks failing the confidence level cut, ie they sit in a region where the particle species cannot be easily distinguished.

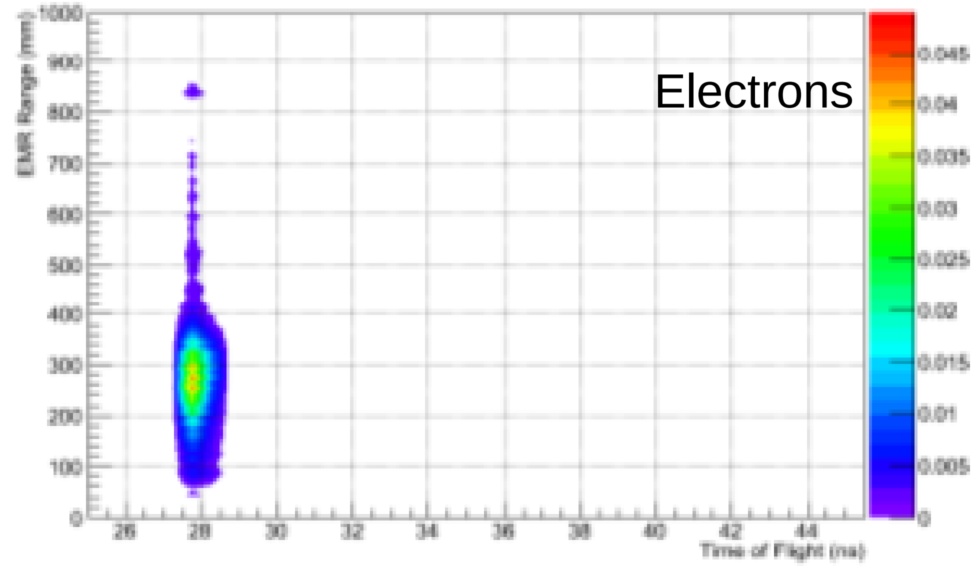
Commissioning PID Variables

Class Name	Variable Name	Description
ComPIDVarA	diffTOF1TOF2	Time of flight between TOF1 and TOF2. Momentum dependent.
ComPIDVarB	KLChargevsDiffTOF1TOF2	KL ADC charge product and time of flight between TOF1 and TOF2. Time of flight reduces dependence of KL measurement on beam momentum.
ComPIDVarC	CommissioningKLADCChargeProduct	KL ADC charge product. Momentum dependent.
ComPIDVarD	CommissioningEMRrange	Range of particle in EMR. Momentum dependent.
ComPIDVarE	CommissioningEMRrangevsDiffTOF1TOF2	Range of particle in EMR and time of flight between TOF1 and TOF2. Time of flight reduces dependence of EMR measurement on beam momentum.
ComPIDVarF	CommissioningEMRdensity	EMR plane density. Momentum dependent.
ComPIDVarG	CommissioningEMRdensityvsDiffTOF1TOF2	EMR plane density and time of flight between TOF1 and TOF2. Time of flight reduces dependence of EMR measurement on beam momentum.
ComPIDVarH	CkovAvsDiffTOF1TOF2	Number of photoelectrons measured in CkovA and time of flight between TOF1 and TOF2.
ComPIDVarI	CkovBvsDiffTOF1TOF2	Number of photoelectrons measured in CkovB and time of flight between TOF1 and TOF2.

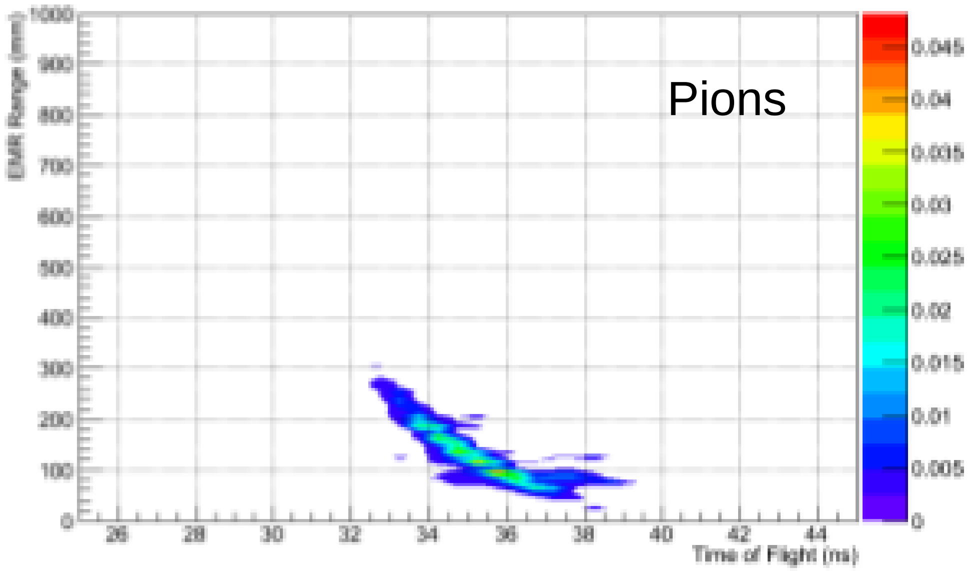
ComPIDvarE PDF



(a)



(c)



(b)

Example PDFs, produced using ReduceCppGlobalPID, of ComPIDVarE for:
a) muons, b) pions and
c) electrons, for a simulated 200 MeV/c muon beam.

PID Performance Validation

- The consistency of the variables with MC, as represented by efficiency/purity plots.
- The consistency between the upstream variables, both on MC and real data, determined by the degree of agreement between the variables on a PID hypothesis.
- The consistency between the downstream variables, both on MC and real data, determined by the degree of agreement between the variables on a PID hypothesis.
- The consistency between upstream and downstream PID hypotheses.

PID Efficiency and Purity

- Efficiency:

$$\epsilon_{PID} = \frac{N_C}{N_{ID} + N_F} = \frac{N_C}{N_C + N_W + N_F}$$

with Binomial error

- Purity:

$$\rho = \frac{N_C}{N_{ID}} = \frac{N_C}{N_C + N_W}$$

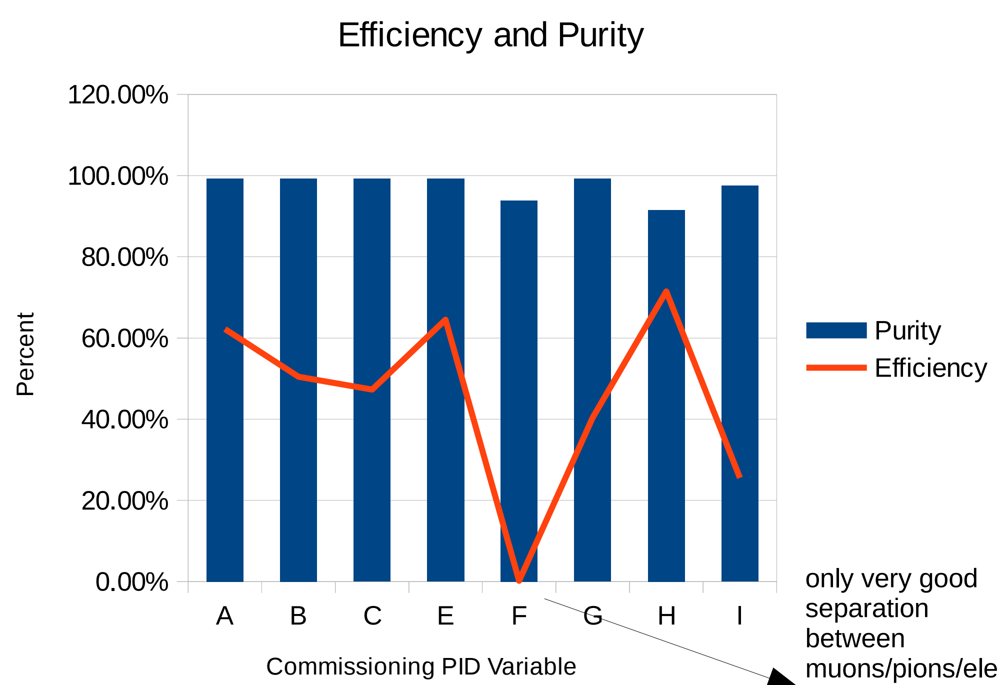
Purity error:

$$\sigma_\rho^2 = \left(\frac{N_W}{(N_C - N_W)^2}\right)^2 N_C + \left(\frac{-N_C}{(N_C - N_W)^2}\right)^2 N_W$$

where N_C is the number of correctly identified tracks, N_{ID} is the total number of identified tracks, N_W is the number of incorrectly identified tracks, and N_F is the number of tracks for which the PID failed to identify the track, despite it being suitable.

PID Analysis
Efficiency and Purity (MC)

Variable	Purity	Efficiency
A	99.3%	62.2%
B	99.3%	50.4%
C	99.2%	47.3%
E	99.2%	64.5%
F	93.9%	0.189%
G	99.2%	40.4%
H	91.5%	71.5%
I	97.5%	25.5%
Overall	99.9%	96%



only very good separation between muons/pions/electrons at plane density > 0.95

Variables are used in combination and tuned for purity on individual particle species therefore efficiency takes a hit (by variable) but purity is maintained and overall PID (combined) efficiency is good.

C.P. PID Variable Consistency

A	B	C	D	E	F	G	H	I
A	99.5%	99.1%	26.9%	99.6%	100%	99.2%	100%	99.3%
	B	99.2%	27.1%	99.1%	100%	98.4%	100%	99.2%
		C	30.4%	98.9%	100%	98.7%	100%	99.4%
			D	27.2%	81.2%	45.7%	57.5%	25.2%
				E	100%	98.9%	95.1%	98.3%
					F	93.3%	100%	100%
						G	90.6%	95.2%
							H	83.3%

Consistency between 2 PID variables determined by measuring what proportion of particles are identified as muons/not muons by both variables (ComPIDVars D, E, F and G inc EMR).

Table 3.4: Consistency between ComPIDVars on MC.

A	B	C	D	E	F	G	H	I
A	99.0%	98.6%	34.0%	78.2%	92.6%	98.2%	86.1%	81.9%
	B	97.8%	31.5%	78.8%	88.9%	97.0%	85.5%	81.3%
		C	34.2%	71.9%	100%	95.6%	89.5%	82.0%
			D	32.8%	93.6%	34.2%	36.5%	37.0%
				E	30.8%	71.5%	64.4%	64.0%
					F	65.8%	98.1%	83.7%
						G	81.0%	82.6%
							H	74.1%

Calculated as the proportion of matched tracks where both PID variables determined the particle species as muon or not muon. (see Celeste's CM44 talk/thesis for details).

Table 3.5: Consistency between ComPIDVars on data.

Helical PID Variables

Class Name	Variable Name	Description
PIDVarA	diffTOF1TOF0	Time of flight between TOF1 and TOF0. Momentum dependent.
PIDVarB	diffTOF1TOF0vsUSTrackerMom	Time of flight between TOF1 and TOF0 and momentum measured in upstream tracker.
PIDVarC	KLChargeProdvsDSTrackerMom	KL ADC charge product and momentum measured in downstream tracker.
PIDVarD	KLADCChargeProduct	Same as ComPIDVarC.
PIDVarE	EMRrange	Same as ComPIDVarD
PIDVarF	EMRrangevsDSTrackerMom	Range of particle in EMR and momentum measured in downstream tracker.
PIDVarG	EMRdensity	Same as ComPIDVarF.
PIDVarH	EMRdensityvsDSTrackerMom	EMR plane density and momentum measured in downstream tracker.
PIDVarI	CkovAvsUSTrackerMom	Number of photoelectrons measured in CkovA and momentum measured in upstream tracker.
PIDVarJ	CkovBvsUSTrackerMom	Number of photoelectrons measured in CkovB and momentum measured in upstream tracker.

The Future

- Improved user documentation.
  - **Existing documentation update in process on workbook.**
  - **Full super easy to follow documentation coming soon!**
- PDFs produced for a good sample of run settings by M. Uchida.
- Collaboration wide use of globals.
- Helical analysis.
- Identify and implement performance improvements.
  - Inc speeding up track matching.
- Rinse and repeat...

Thank you

The Method

- **Existing documentation update in process on workbook.**
- **Full super easy to follow documentation coming soon!**
- Either pick your chosen pre-created PDF (available in a couple of weeks) or create your own.
  - Go to bin/Globals and edit process_datacard.py (set particle hypothesis to kmuplus) and pdf_example_datacard.py for your MC file.
  - Run pid_pdf_generator.py
  - This will create a set of directories (one for each particle hypothesis, 3 total). Combine PDFs.
- Run Track matching using process_global.py but setting particle hypothesis to all and ensuring the mappers MapCppGlobalReconImport and MapCppGlobalTrackMatching are included.
- Run PID:
  - Take your sample of MC/data that you want to run PID on, that has already been processed through track matching.
  - Copy and edit pid_example_datacard.py as above (pid_config, pid_beam_setting and pid_beamline_polarity must be the same as they were when the PDFs were made)
- Run GlobalPID.py --configuration_file edited_pid_example_datacard.py
- This will produce your output root file with PID'd global tracks in it, that can be analysed with a C++/python script or a reducer in MAUS.

Example PDFs

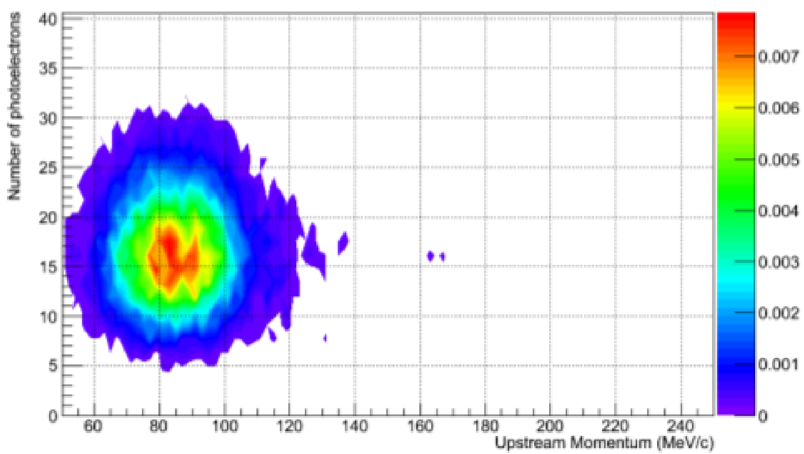
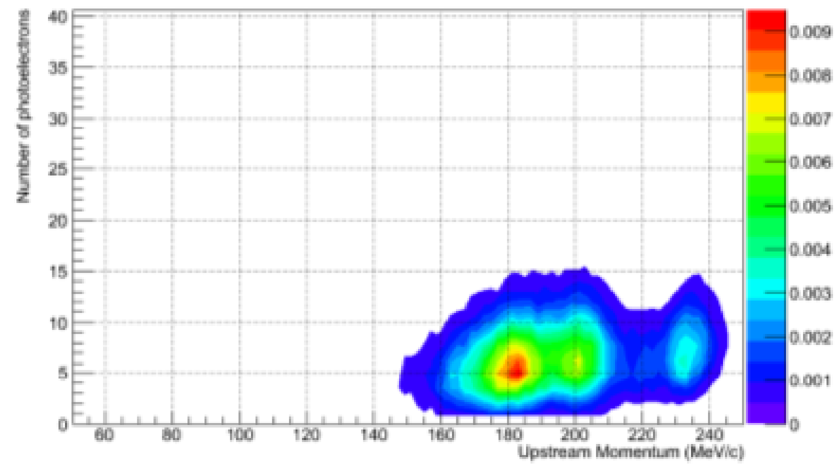
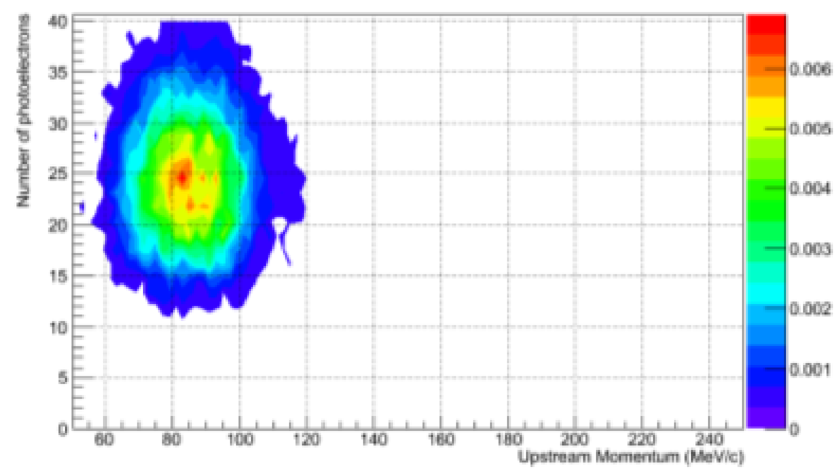


Figure 3.19: Example PDFs, produced using ReduceCppGlobalPID, of PIDVarI for electrons, for a simulated 200 MeV/c pion beam. For the momentum of this simulated beam muons and pions did not produce a signal in the detector.



(a)



(b)

Figure 3.20: Example PDFs, produced using ReduceCppGlobalPID, of PIDVarJ for muons (a) and electrons (b), for a simulated 200 MeV/c pion beam. For the momentum of this simulated beam pions did not produce a signal in the detector.

PID Variable Consistency

- Only tracks identified by both variables are considered.
- And (assuming the particle was identified as a muon:

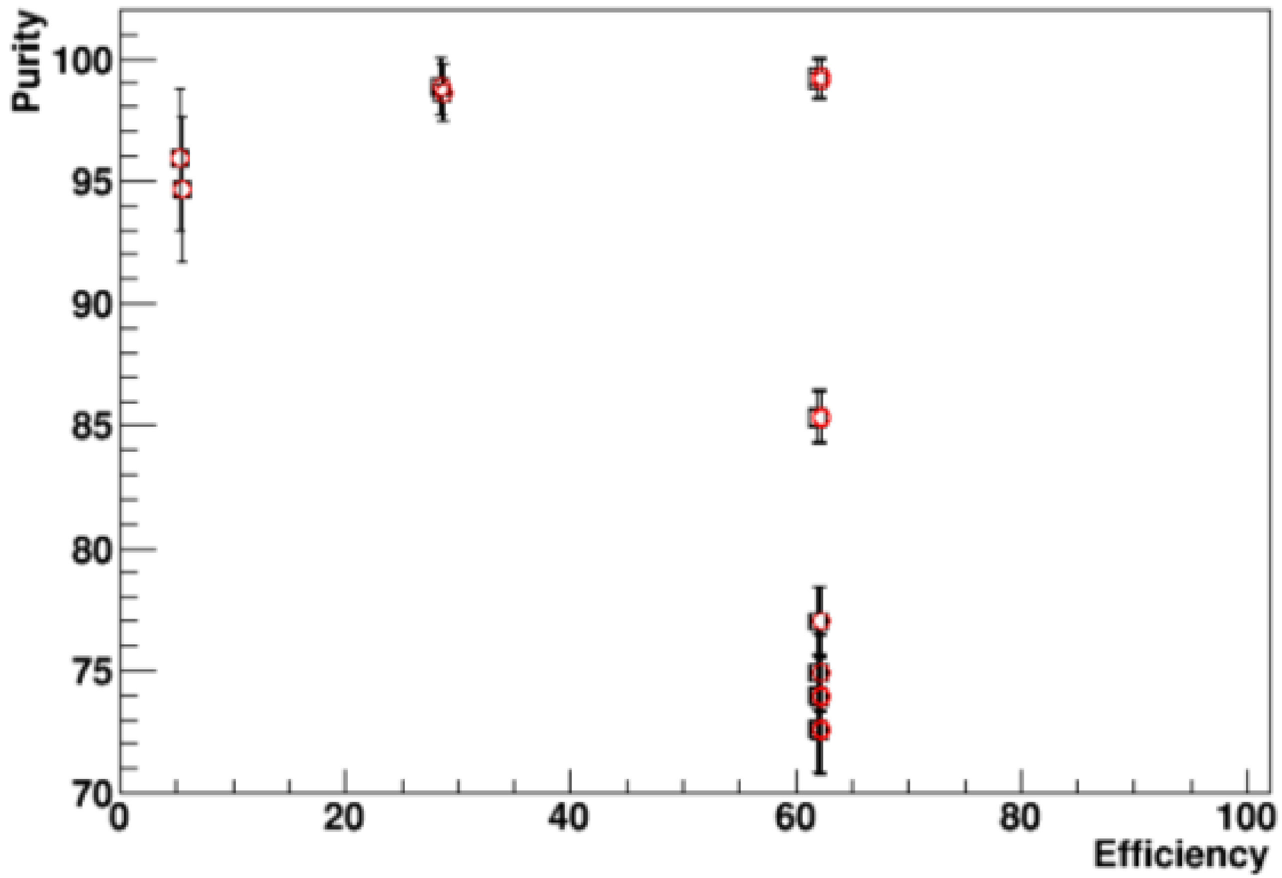
$$P(\mu) + P(\text{other}) = 100\%$$

- These tracks must therefore have been identified with a probability that is greater than the probability of any other particle by at least the value of the confidence level cut:

$$P(\mu) > P(\text{other}) + CL$$

- CL = 10% used for these studies
- Therefore P(μ) minimum > 45%.

PID Var A Eff/Purity



Each point corresponds to a different cut setting. black squares = overall particle identification efficiency red circles = muon identification efficiency