Data-driven Background Measurement

- Outline
  - Reco efficiency study
  - Background measurement

Heng-Ye Liao,
Hadron analysis meeting
Jan. 13, 2022

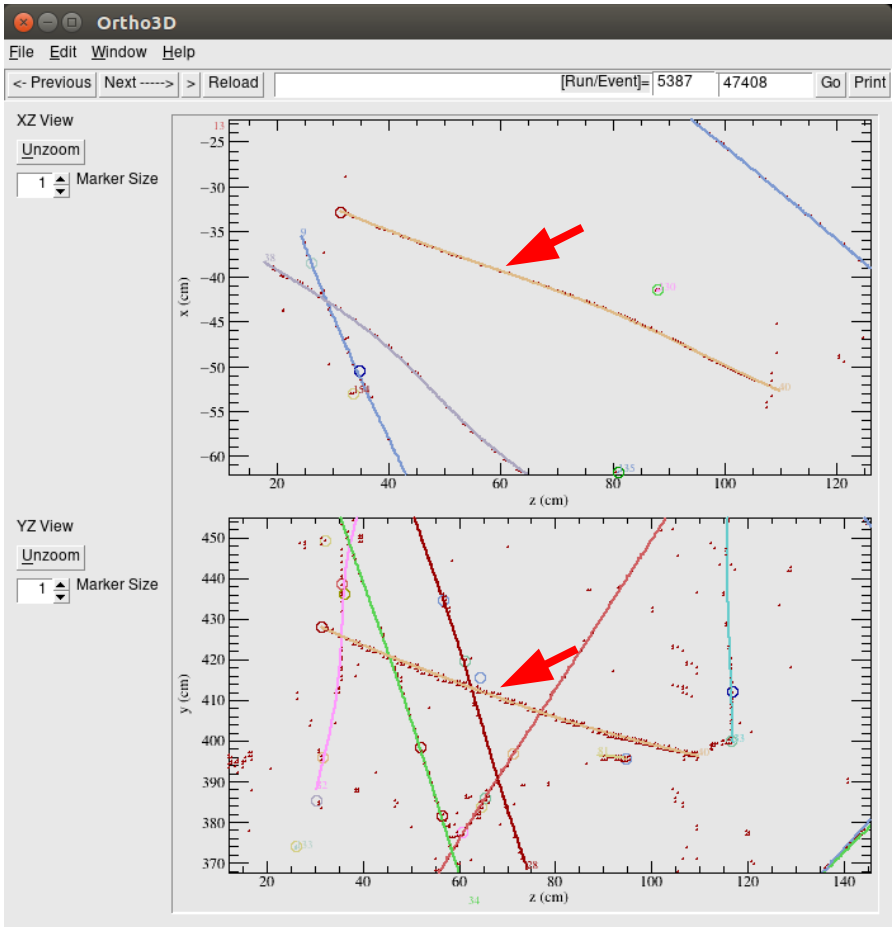
Reco Efficiency Study

- Truncation method:
  - Method developed by Yinrui & Tingjun
  - Select long tracks, truncate them to given lengths
  - $$\textbf{Reco efficiency} = \frac{\textbf{\# of reconstructed tracks}}{\textbf{\# of selected long tracks}}$$

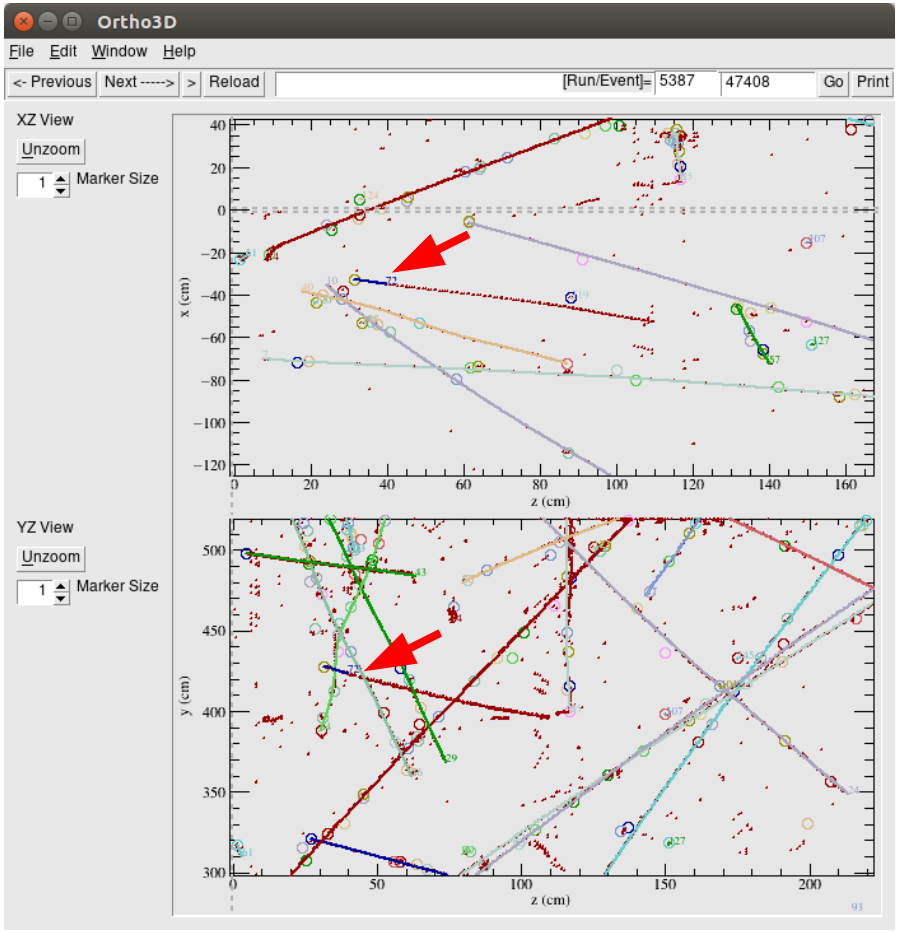
- Selection cuts for long tracks:
  [1] Beam quality cuts:
  Beam instrumentation cut (data)/PDG cut (MC) to select protons
  + Pandora slice cut + CaloSize cut + Position cut + CosΘ cut
  [2] Track length cut to select long tracks: **75-140 cm**

- Configurations:
  - Data: protodune-sp_runset_5387_reco2_v09_30_00_v0
  - MC: PDSPProd4a_MC_1GeV_reco1_sce_datadriven_v1_08
  - Statistics of selected long track events:
    + Data: 2696
    + MC: 3743

Event Display (Data)

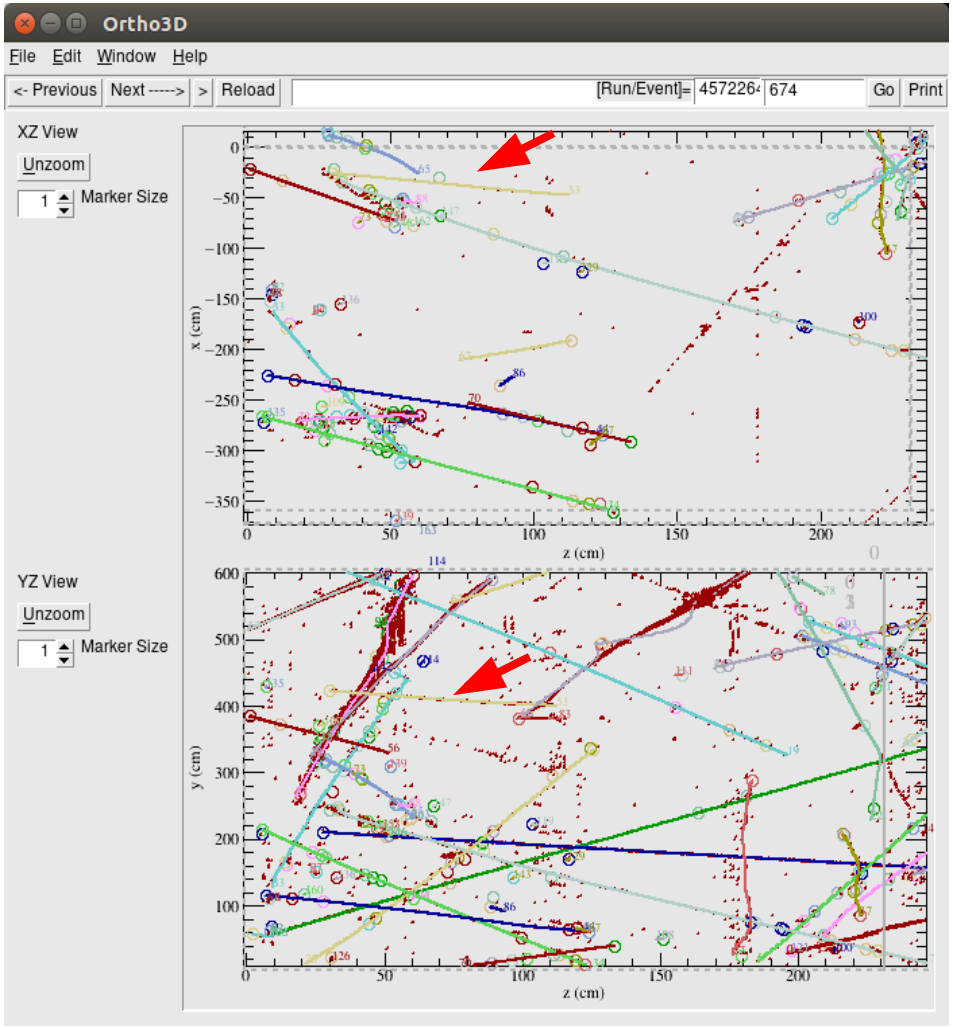


Long proton track before truncation

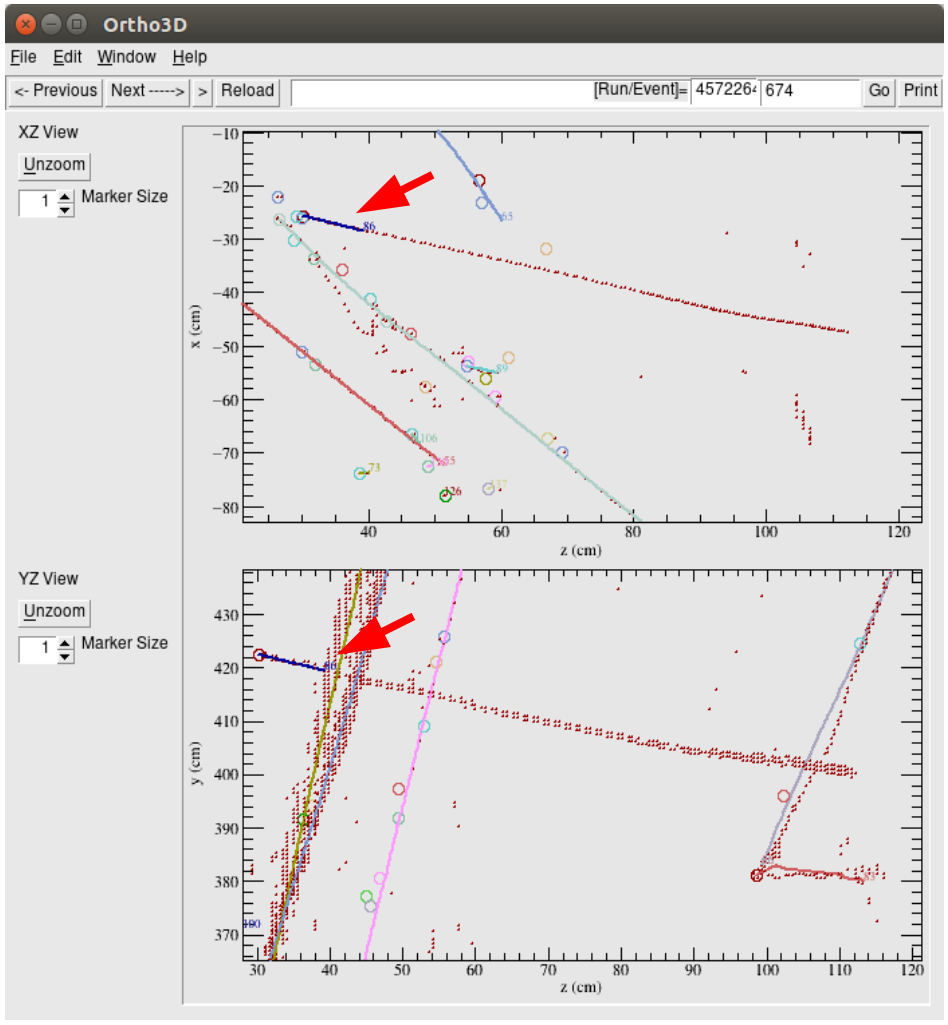


Long proton track after truncation

Event Display (MC)

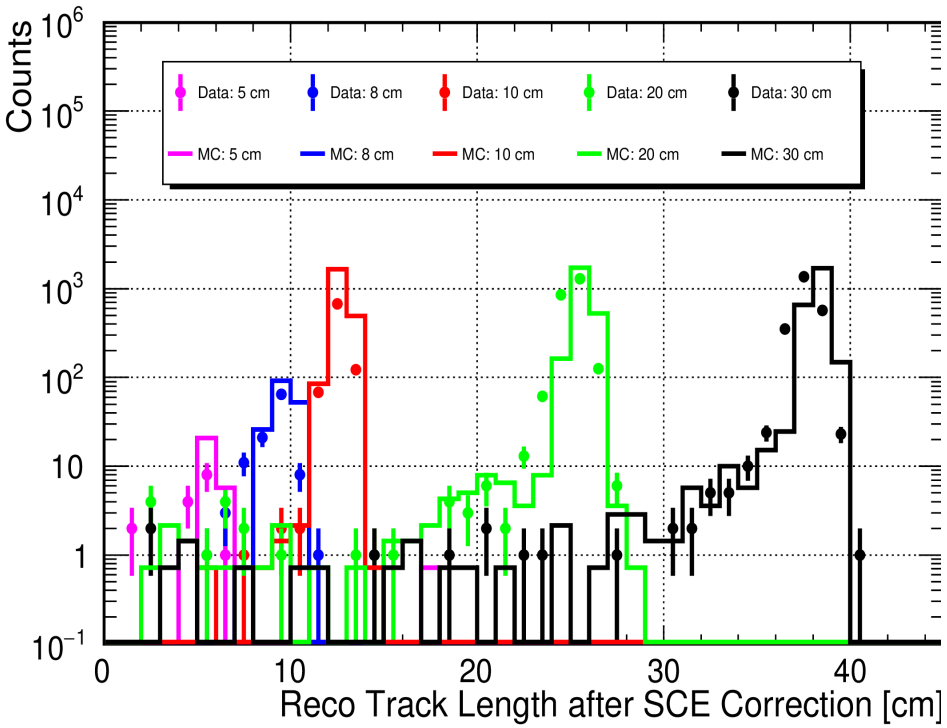


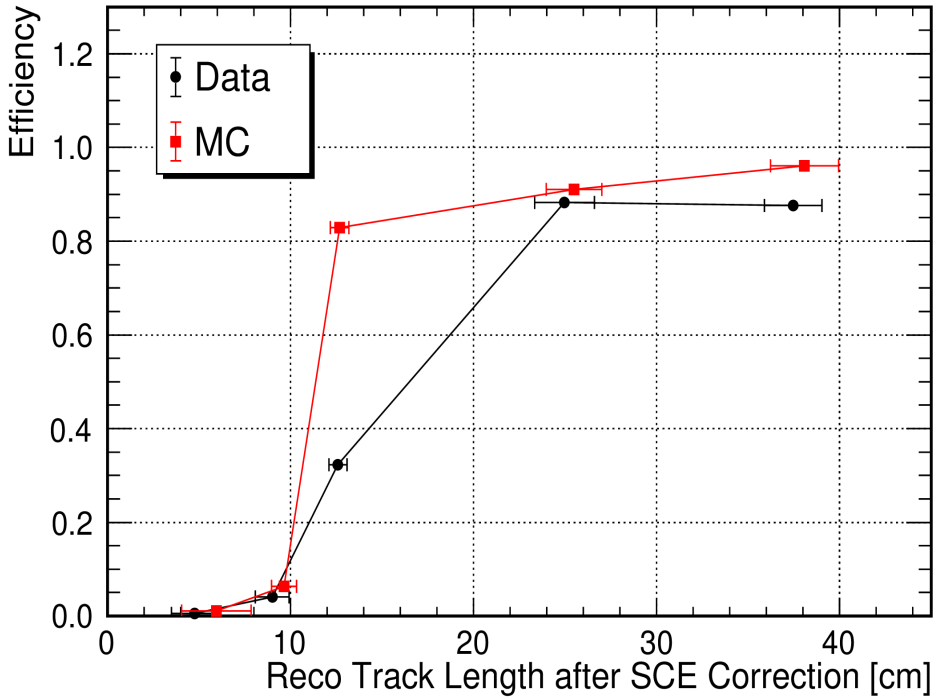
Before truncation



After truncation

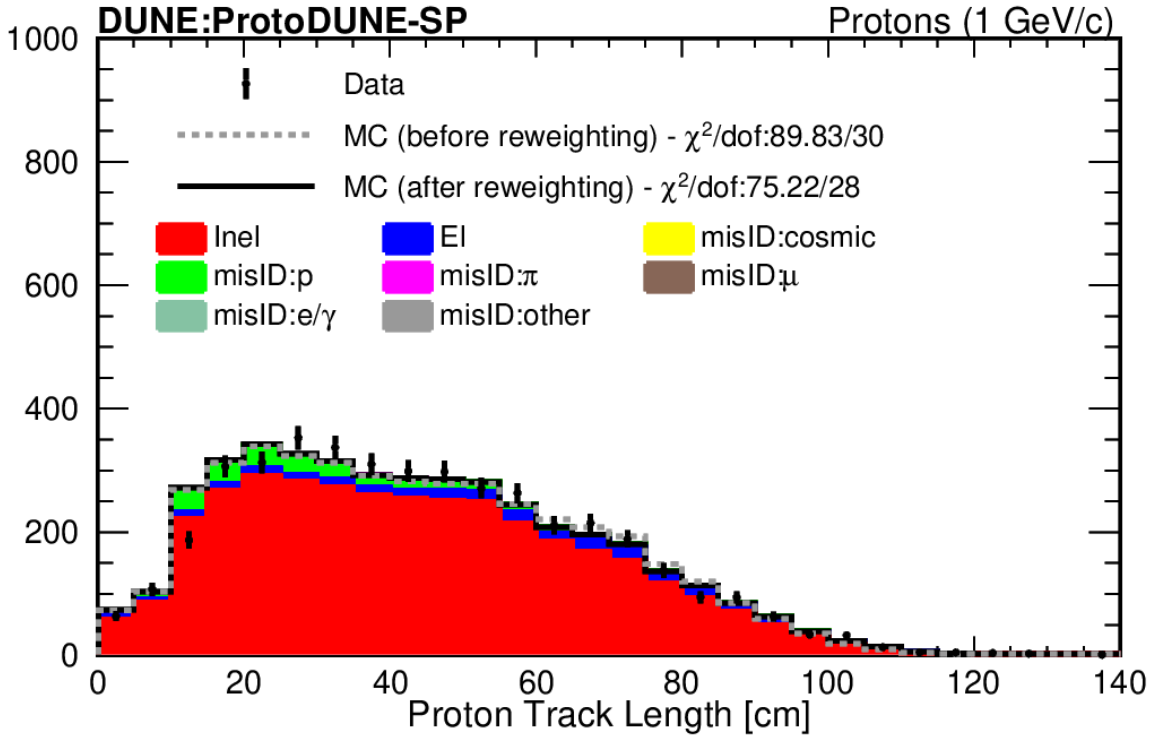
Reco Efficiency: Protons





- Truncated lengths: 5 cm, 8 cm, 10 cm, 20 cm, 30 cm
- Error definitions in reco efficiency:
  - δx: Standard devitation
  - δy: 1-σ boundaries for the Clopper-Pearson interval
- Reco efficiency: MC> Data
  - Larger discrepancy between 10 to 25 cm (plan to have more points)
  - Less than 5 cm, data/MC agree better

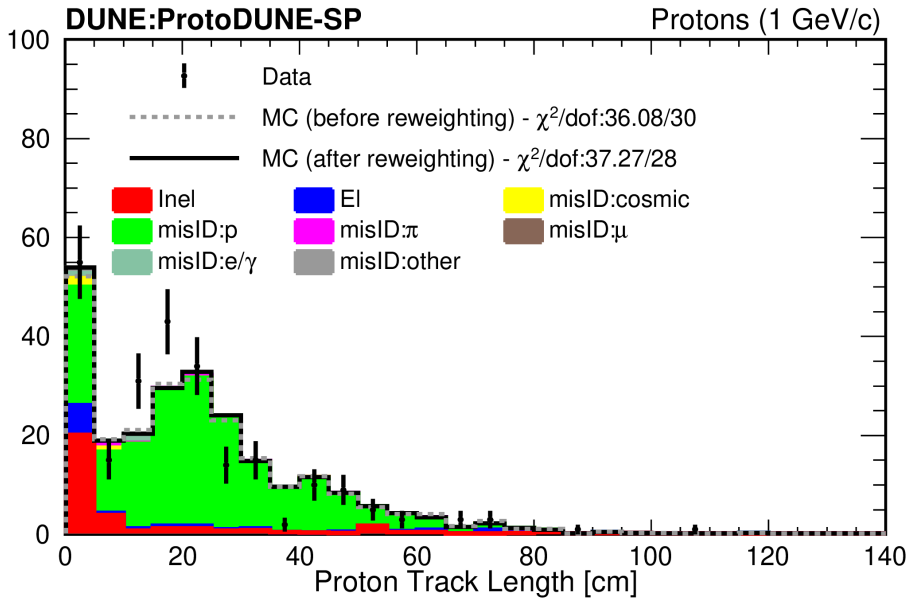
Background Measurement

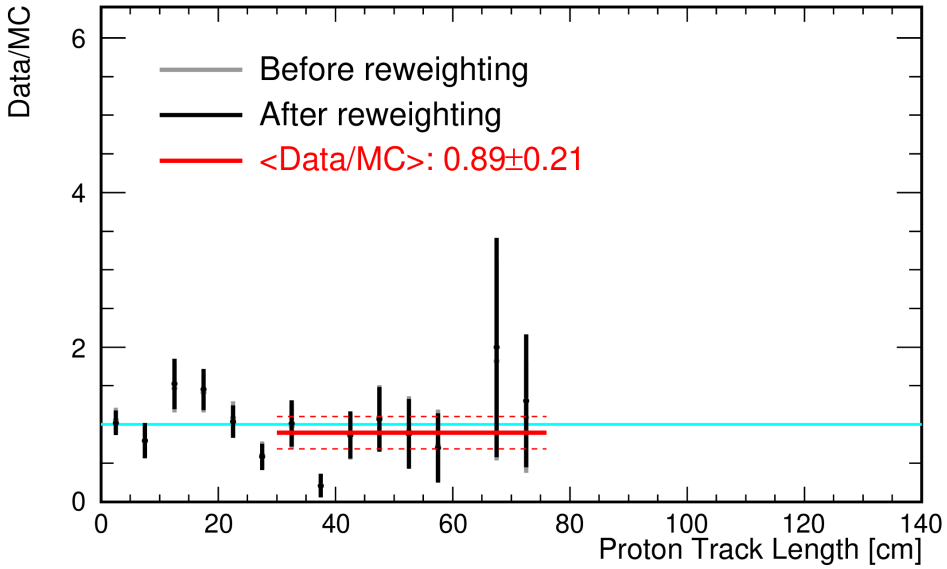


- Main backgrounds in the proton XS analysis:
  - Elastic scattering protons (El)
  - Secondary protons (MisID:p)

- Use background-rich samples to constrain backgrounds

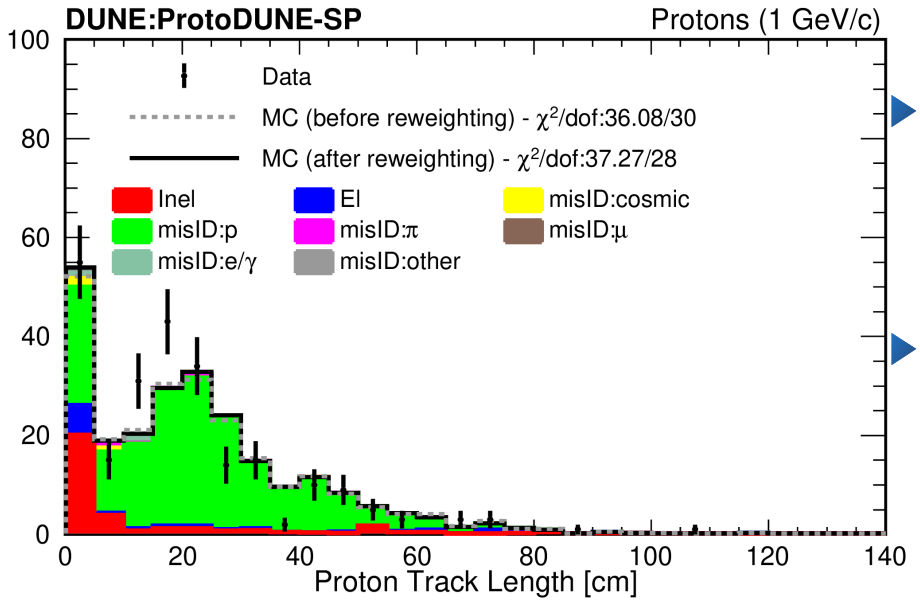
MisID:P-Rich Sample





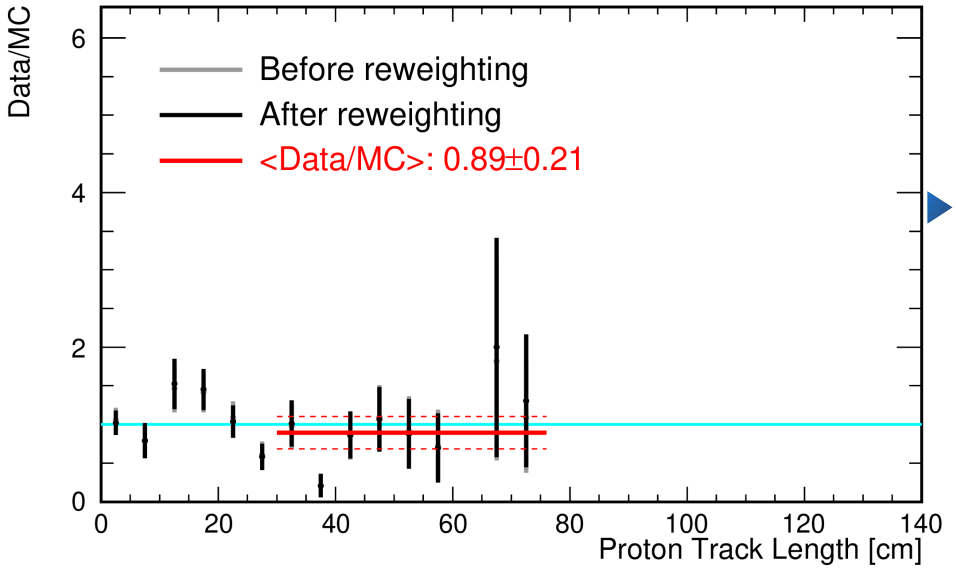
- CosΘ<0.9 to select MisID:P-rich events
- Nearly elastic-free after 10 cm
  → Good to constrain other non-El backgrounds
- Before any correction,
  <Data/MC> = 0.89±0.21
  [over-estimation of MC]
- MisID:P correction factor
  - α:=Best fit for Data-h1-α*h2
    h1: All MC except misID:p
    h2: MC misID:p
  - **α=0.68±0.19 (best-fit)**

MisID:P-Rich Sample



- Before any correction,
  <Data/MC> = 0.89±0.21

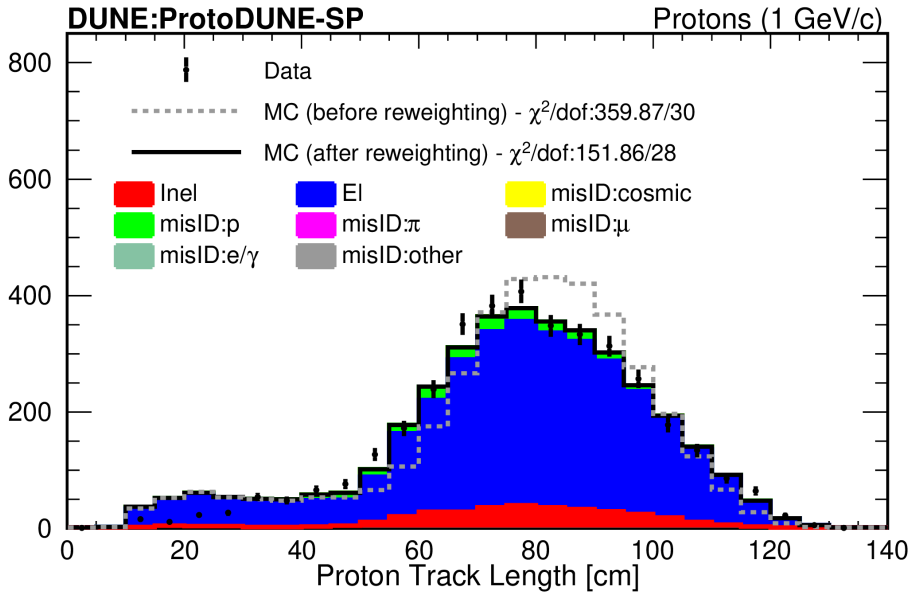
- El. correction factor
  - β, γ:=Best fit for Data-h1-α*h2-γ*h3
    - h1: All MC except MisID:P&Inel
    - h2: MC MisID:P
    - h3: MC Inel

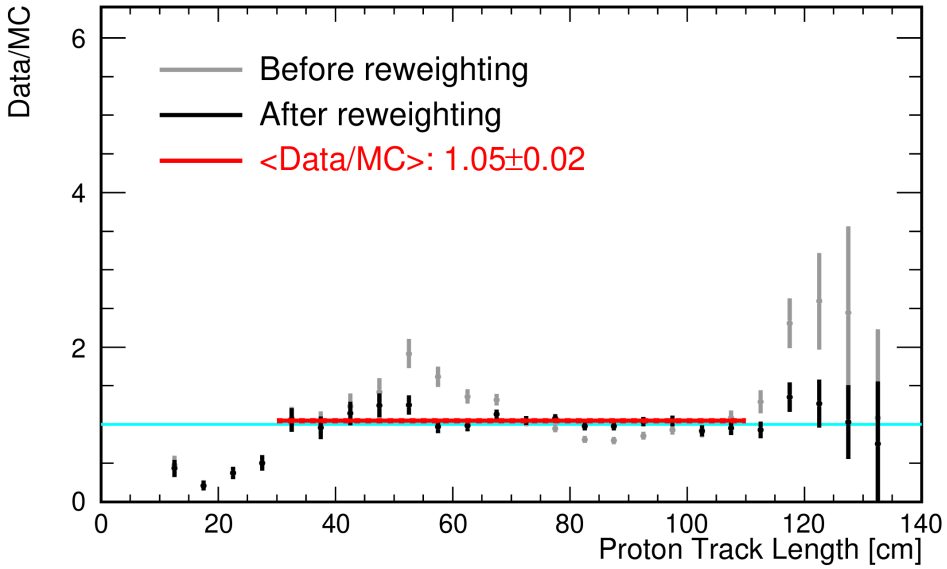


α=0.77±0.36
γ=0.14±9.24

- Big uncertainty on Inel component using simultaneous fit
  → Safe only using α correction, not (α,γ) correction

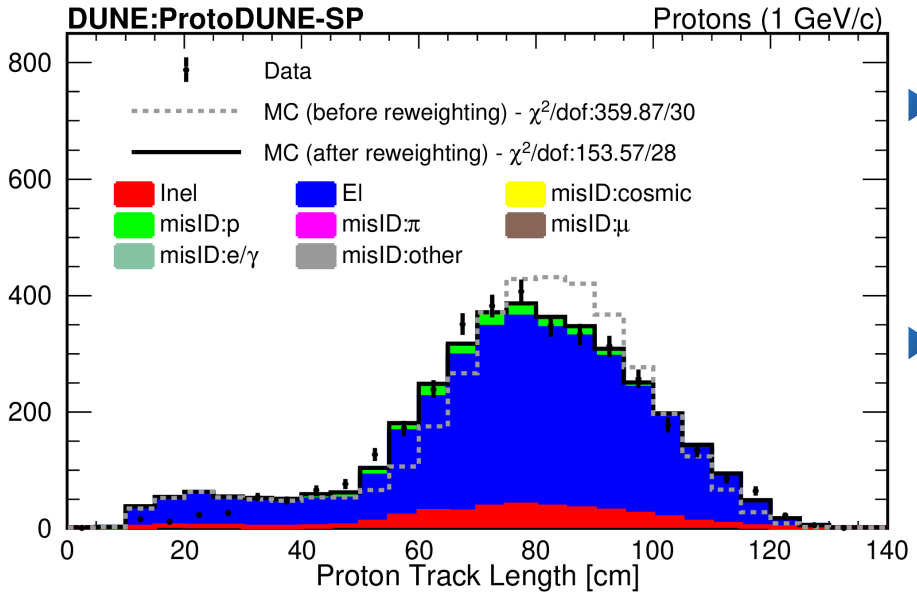
El-Rich Sample

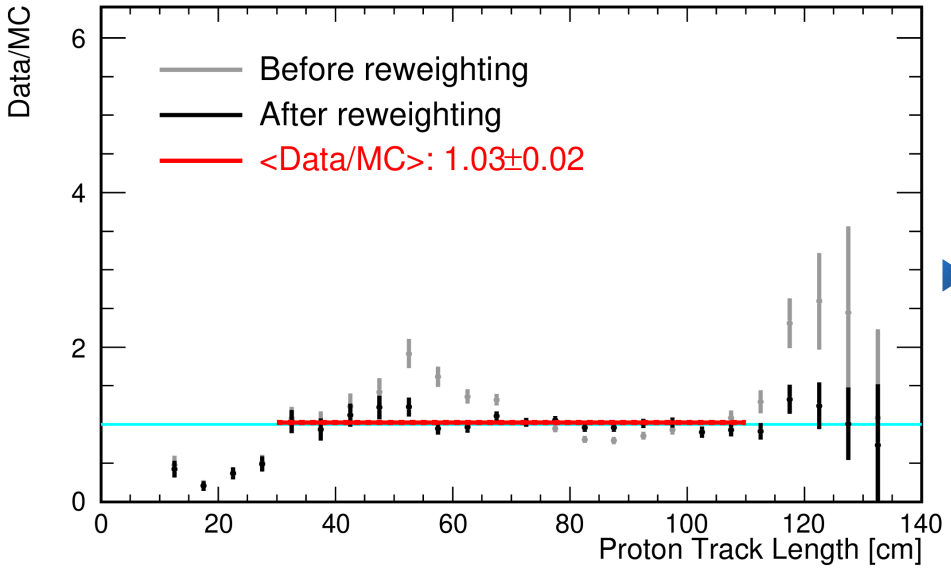




- χ^2-PID cut to select elastic-rich events
- Before any correction, **<Data/MC> = 1.05±0.02**
- El. correction factor
  - β:=Best fit for Data-h1-β*h2
    h1: All MC except El
    h2: MC El
  - **β=1.03±0.02 (best-fit)**

El-Rich Sample (After α Correction)



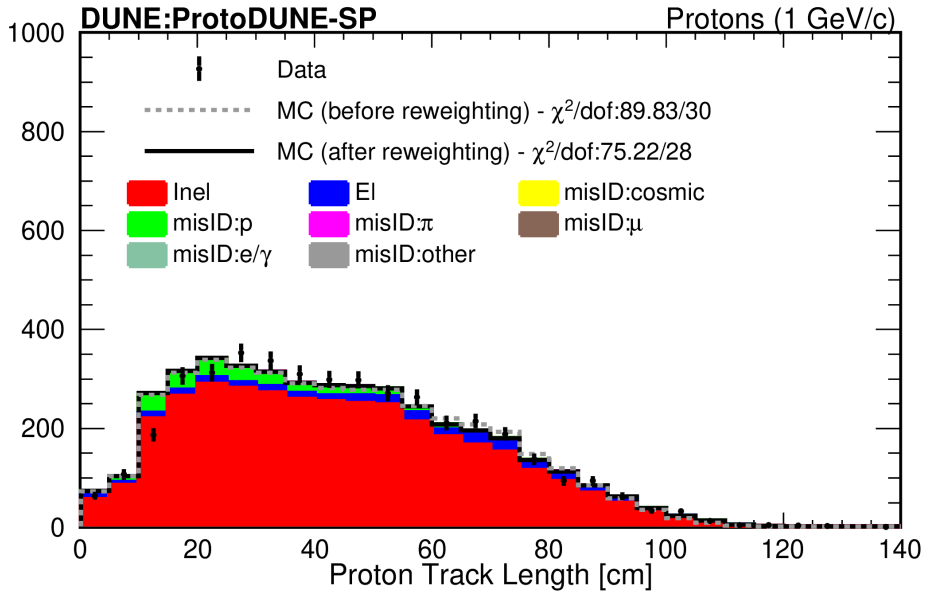


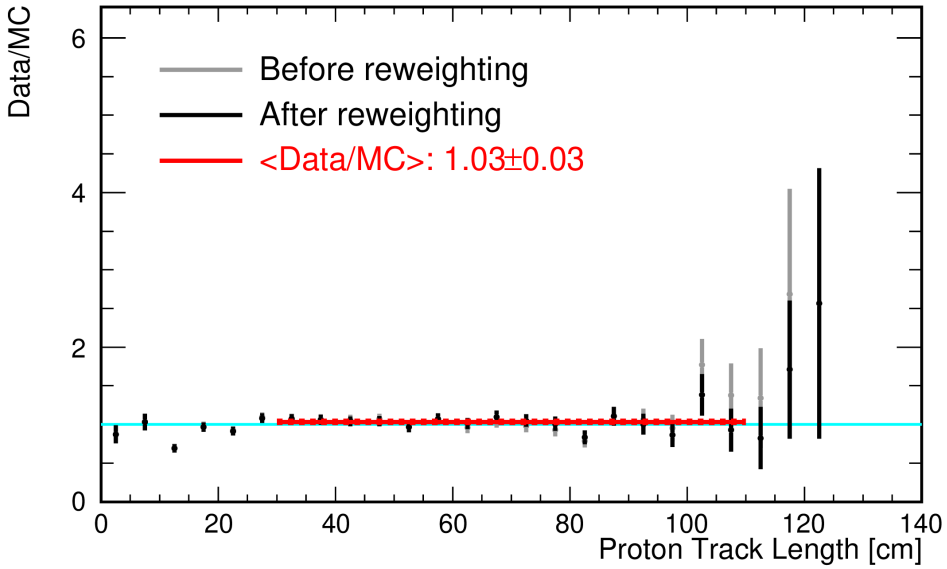
- Apply α correction,
  <Data/MC> = 1.03±0.02
  (Improvement after α corr.)
- El. correction factor
  - β:=Best fit for Data-h1-α*h2-β*h3
    h1: All MC except MisID:P&El
    h2: MC MisID:P
    h3: MC El

α=0.68±0.19 (derived from MisID:P-rich sample)
β=1.00±0.02 (best-fit)

- It seems that after α correction, we do not have to make correction on El. background

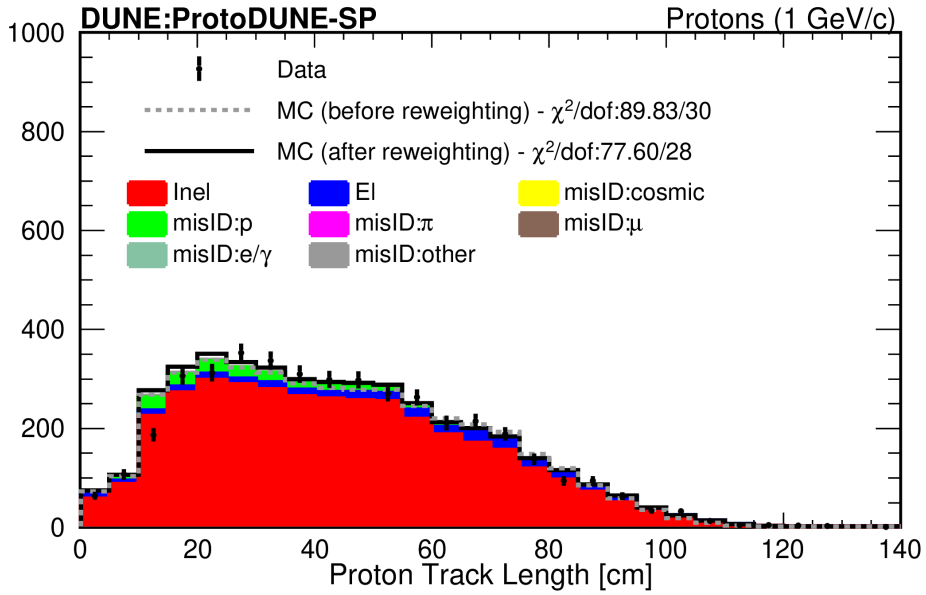
Inel-Rich Sample (Signal Channel)

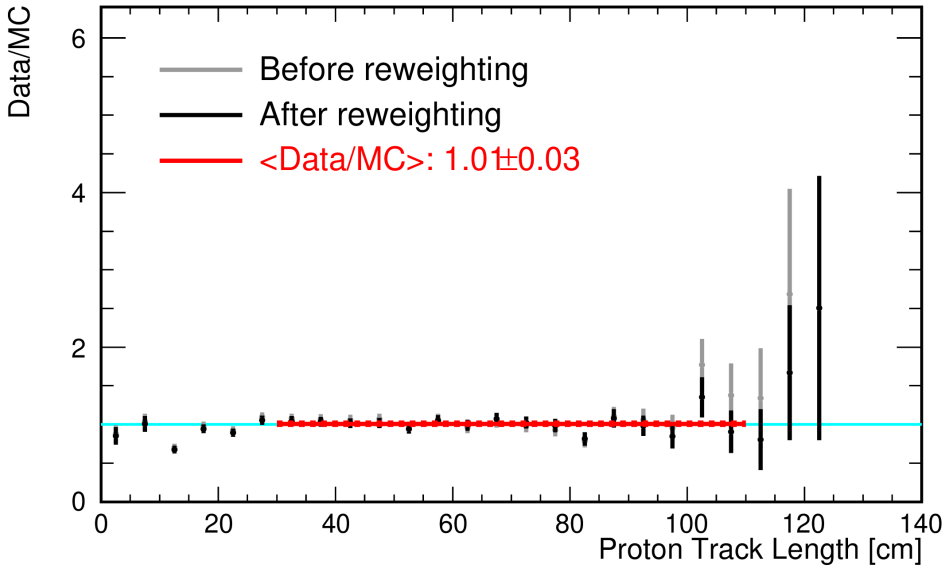




- Before proton corrections, **<Data/MC> = 1.03±0.03**

Inel-Rich Sample (Signal Channel)

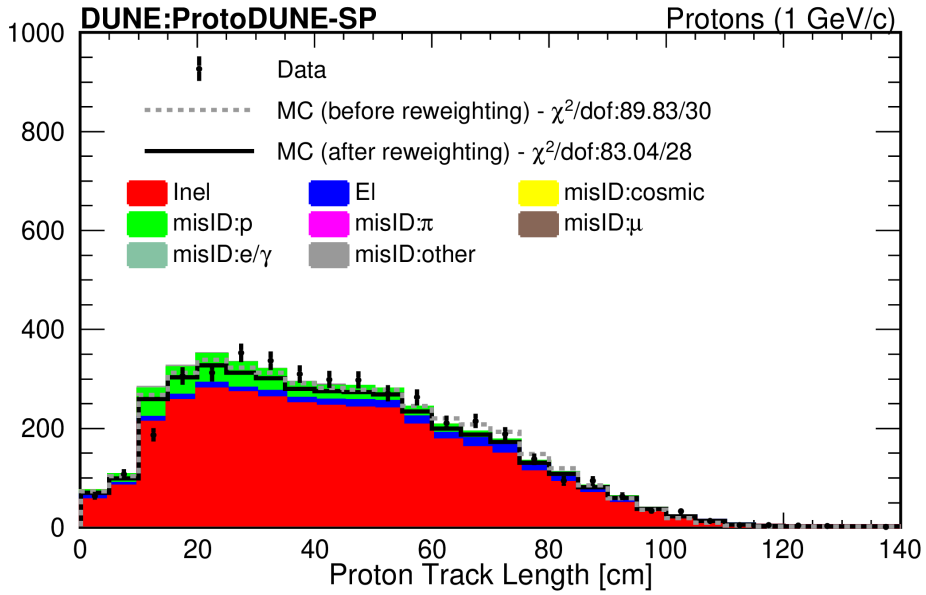


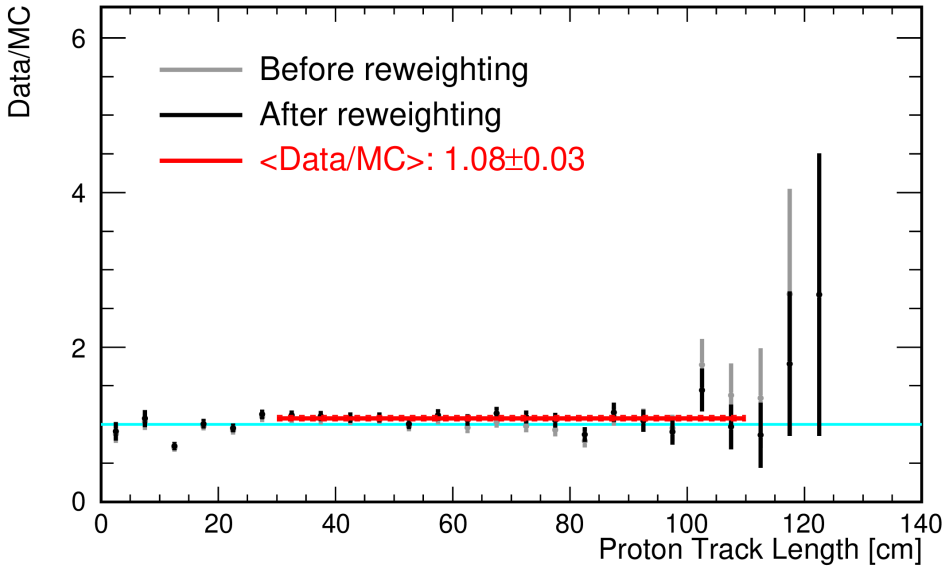


- Apply proton corrections on the signal channel (α, β)=(0.68±0.19, 1)
- After proton corrections, **<Data/MC> = 1.01±0.03**

  (Improvement after α corr.)

Inel-Rich Sample (Signal Channel)





- Before proton corrections,
  <Data/MC> = 1.03±0.03
- Best fit for Data-h1-α*h2-β*h3
  h1: All MC except MisID:P&El
  h2: MC MisID:P
  h3: MC El

  α=1.65±0.50 (best-fit)
  β=0.96±0.31 (best-fit)
- After proton corrections,
  <Data/MC> = 1.08±0.03
  (bad idea to use signal channel to constrain backgrounds!)

Summary on Data-driven Proton Correction

Channel	Data/MC	Proton Correction Factor		
		α (MisID:P)	β (El.)	γ (Inel.)
MisID:P-rich	0.89±0.21	0.68±0.19	-	-
	0.89±0.21	*0.77±0.36*	-	*0.14±9.24*
El-rich	1.05±0.02	-	1.03±0.02	-
El-rich with α corr.	1.03±0.02	**0.68±0.19**	1.00±0.02	-
Inel-rich (signal channel)	1.03±0.03	-	-	-
Inel-rich with α corr.	1.01±0.03	0.68±0.19	--	--
Inel-rich	1.08±0.03	*1.65±0.50*	*0.96±0.31*	--

Summary & Outlook

- Reco efficiency:
  Larger discrepancy between data and MC for track length between 10 cm and 25 cm

- Background measurement:
  - Data/MC has good agreement for track length>30 cm
  - Proton correction:
    Scaling factor of MisID:P: 0.68±0.19

- Next:
  - Include background subtraction in the XS analysis
  - Proton KE