

# Zcounting Meeting

---

14/02/2018

HARRY LYONS

# Histogram Suppression

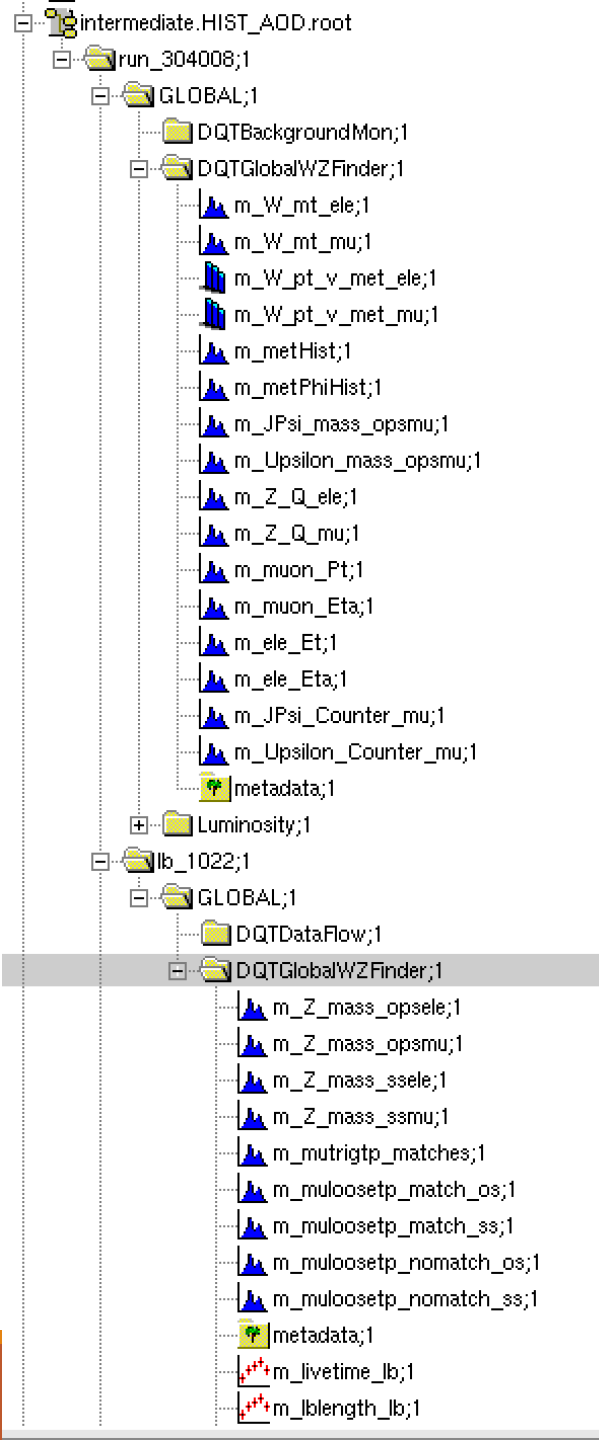
---

- Used this for the Simulation flag:  
`thisEventInfo->eventType(xAOD::EventInfo::IS_SIMULATION)`
- The fix as it stands: Completely remove any LB dependent Histograms when running over MC. Leaves only GLOBAL Histograms
- More work could be done on neatness, currently rather ugly in places. Possibly better create a variable like `m_isSimulation` for compactness.
- Output when running over data is Identical to before. Introduces the 2D Histograms only when running over MC.

# Histogram Suppression

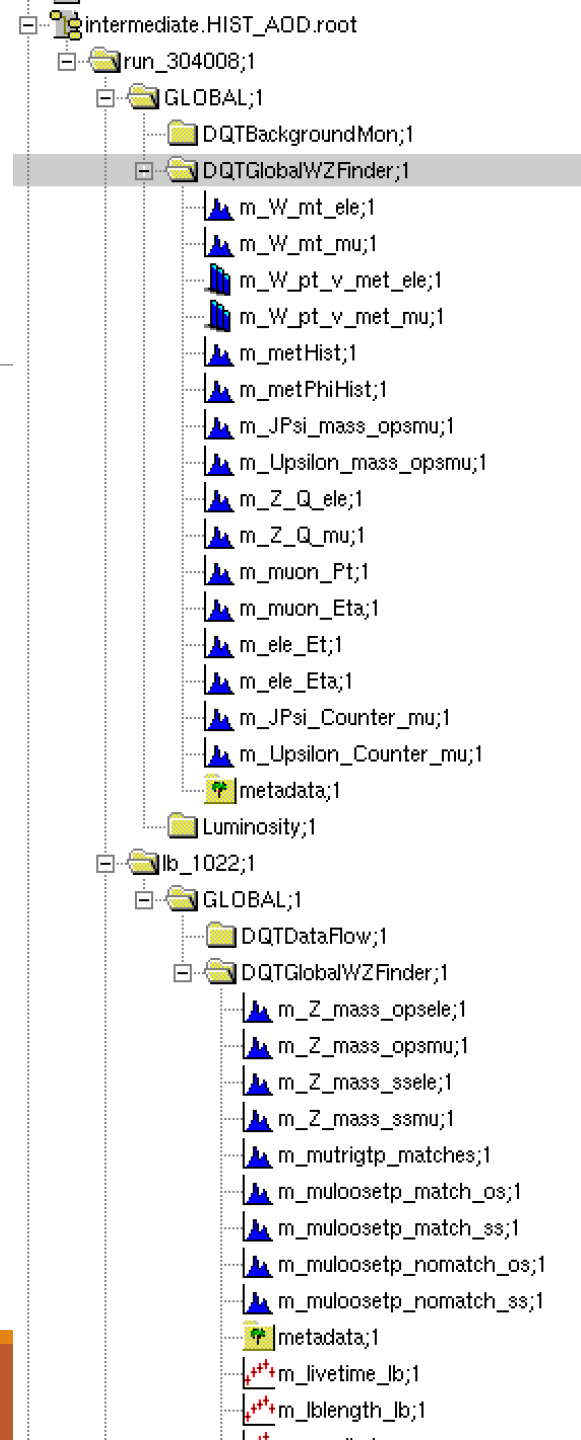
---

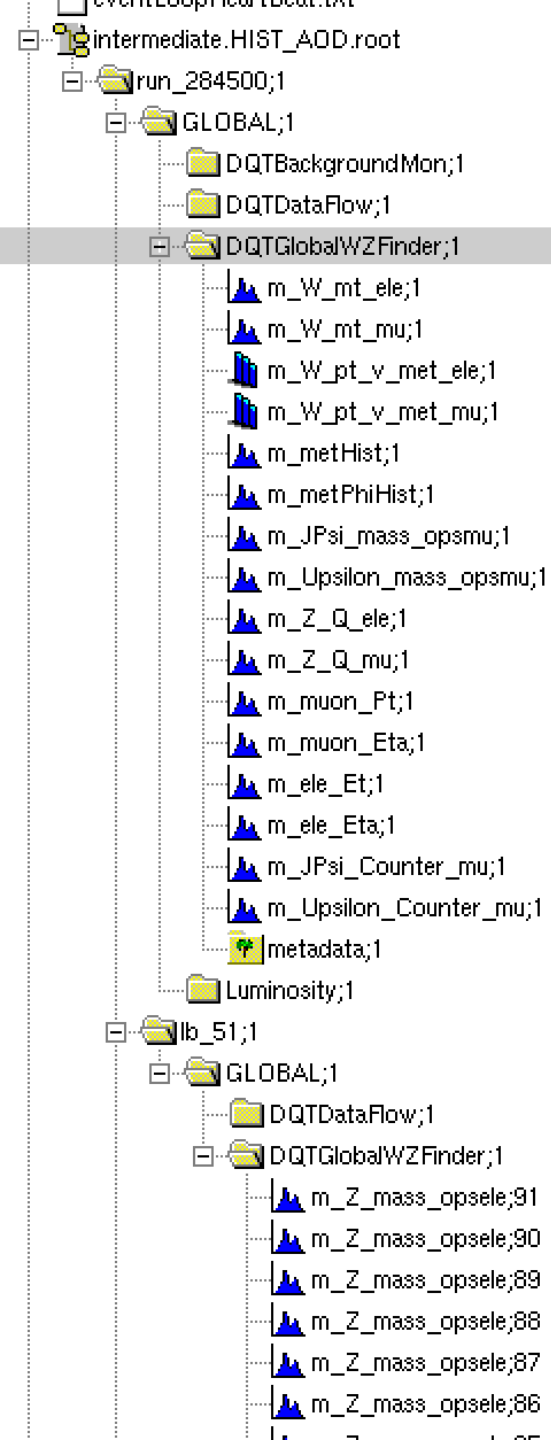
- Had to flag some histograms within DQTDataFlow, including the lowstats Histograms.
- Currently not filling m\_mcmatch Histograms (from the doMuonTruthEff function). These should only fill when it is a simulation, but they also suffer from the many histograms problem. How should this be dealt with? Though not sure if they are used anywhere?



DATA  
Before

DATA  
After

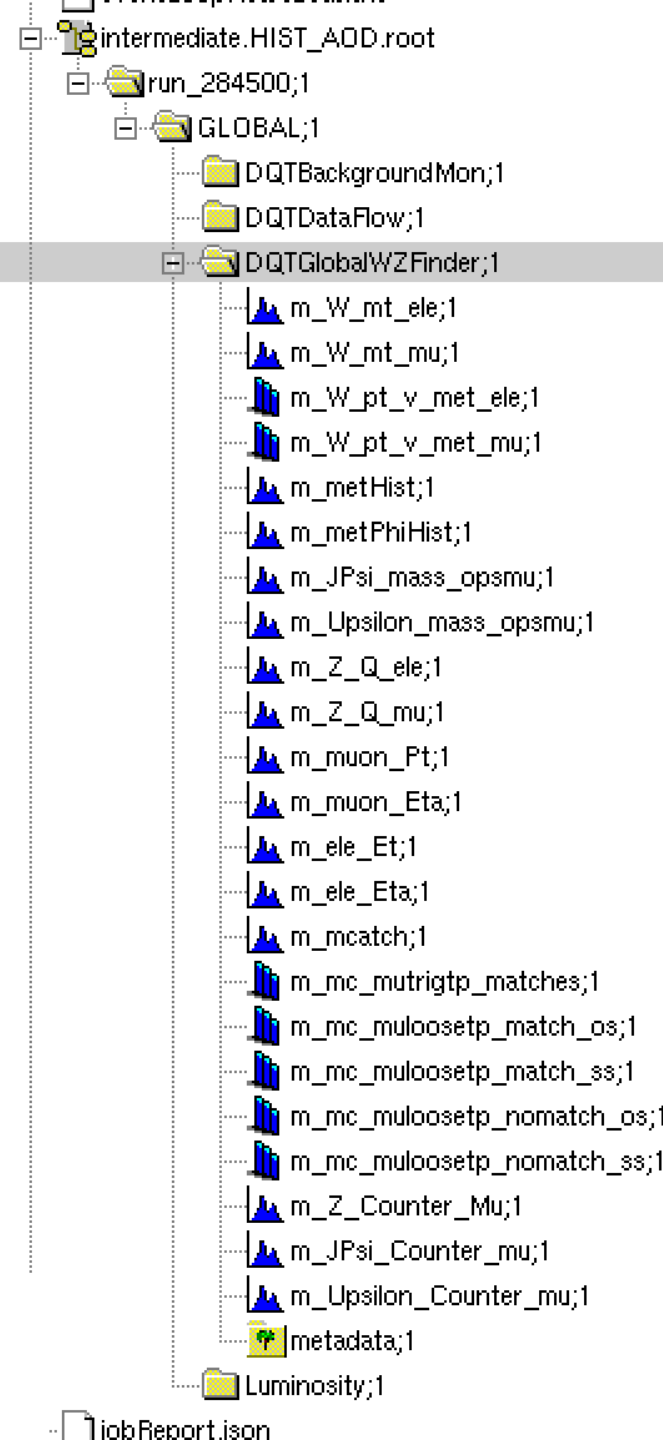




MC  
Before

etc...

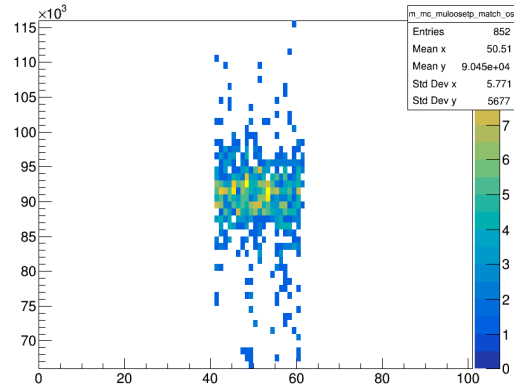
MC  
After



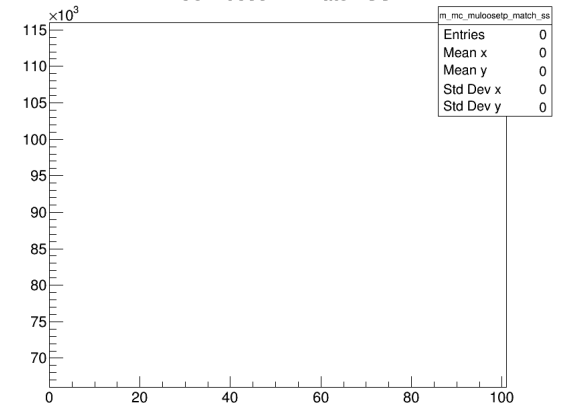
Histograms needed for  
"alleff" python script.

LB on x-axis

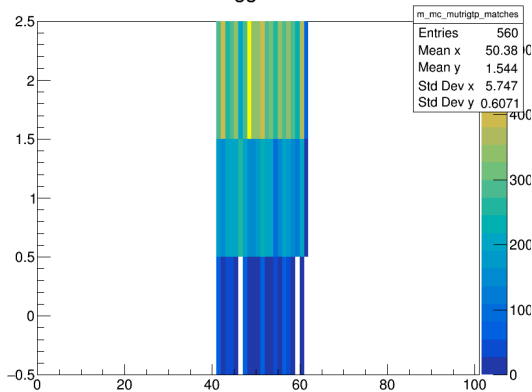
Muon loose TP match OS



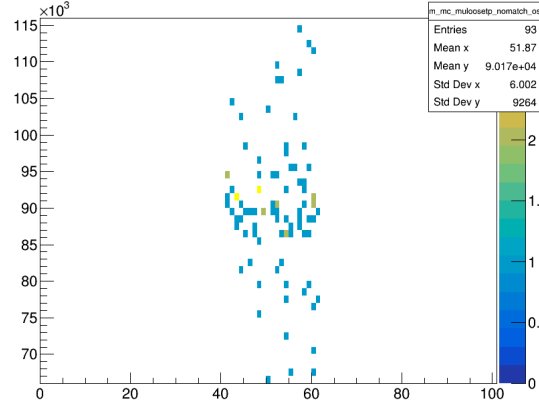
Muon loose TP match SS



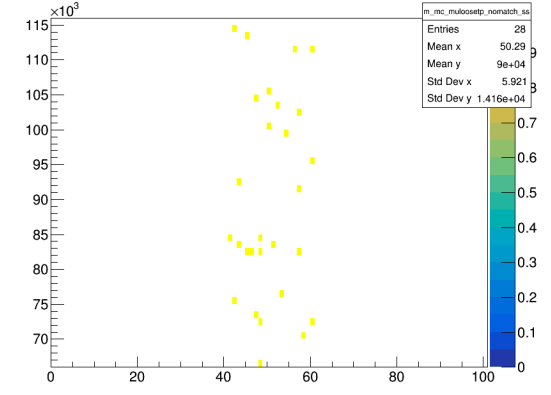
Muon trigger TP stats



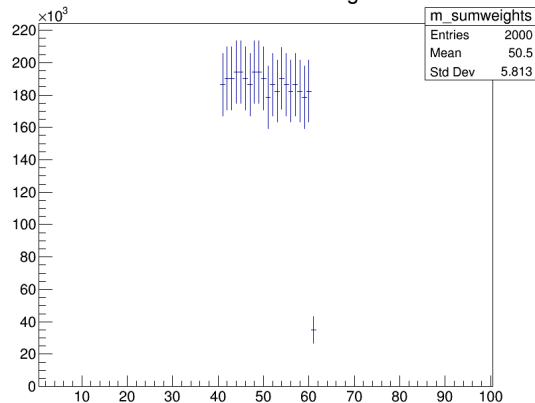
Muon loose TP nomatch OS



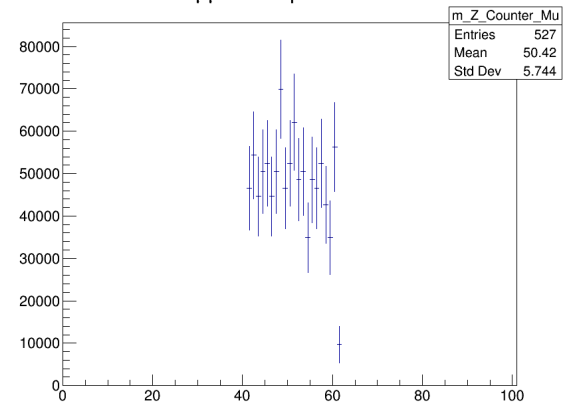
Muon loose TP nomatch SS



Sum of MC event weights



Z → μμ Count per Lumi Block



# How much space is saved?

---

- Have been testing on a 2.2GB file from the total 54TB:  
mc16\_13TeV.361107.PowhegPythia8EvtGen\_AZNLOCTEQ6L1\_Zmumu.merge.AOD.e3601\_s3126\_r9425\_r9315  
  
AOD.11693591.\_012413.pool.root.1
- Output file from the 2.2GB test file originally was ~14MB, scaling (naively) to the full 54TB this could potentially be an output of ~320GB
- Output from the modified code is ~90KB (~0.6% of original)
- Scaled (naively) to the full 54TB is an output of ~2GB.. Much more manageable.

# Open Questions

---

- Is the goal to produce the 2D Histograms just for the simulated data, or to create them for data also?
- Have any changes I've made caused memory issues? Memory leak? I haven't tested this yet.
- How well will (if at all) this modified output interact with DQHistogramMerge.py? I haven't tested this yet.
- Currently the MC output root file has no LB folders, but these are used for the LB iteration in the python script. How exactly should modifying the python script be tackled?



# Zcounting (Updated)

---

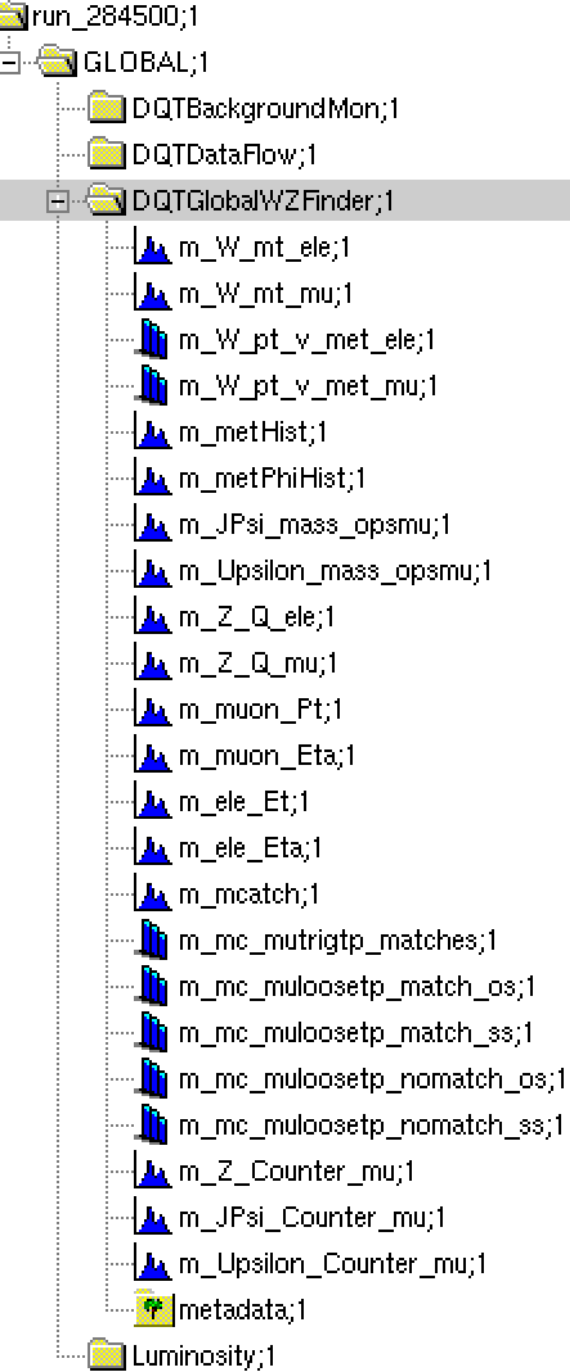
19/02/2018

HARRY LYONS

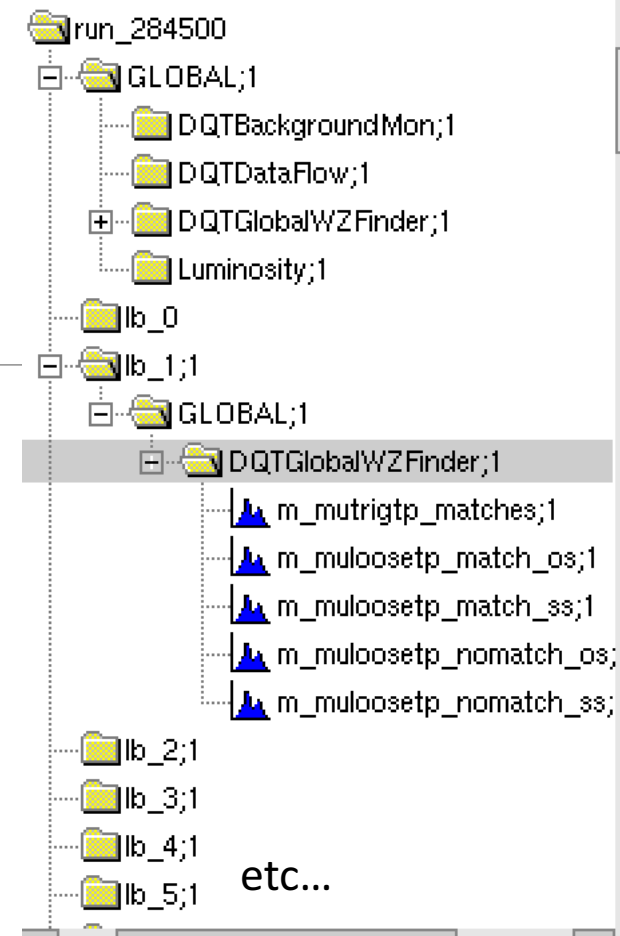
# Updates

---

- For flagging if simulation or not, I created a boolean `m_isSimulation` variable, which makes things a little neater. Only applied to the `DQTGlobalWZFinder.cxx`, which leaves `DQTDataFlowMonTool.cxx` still a little messy, though less lines were modified in this file anyway.
- To use this modified output with the “alleff” python script the file needed to be reformatted, i.e. bring back the LumiBlock directories and stick the 1D slices of the 2D Histograms in there. This was accomplished with a basic python script.



Python  
Script



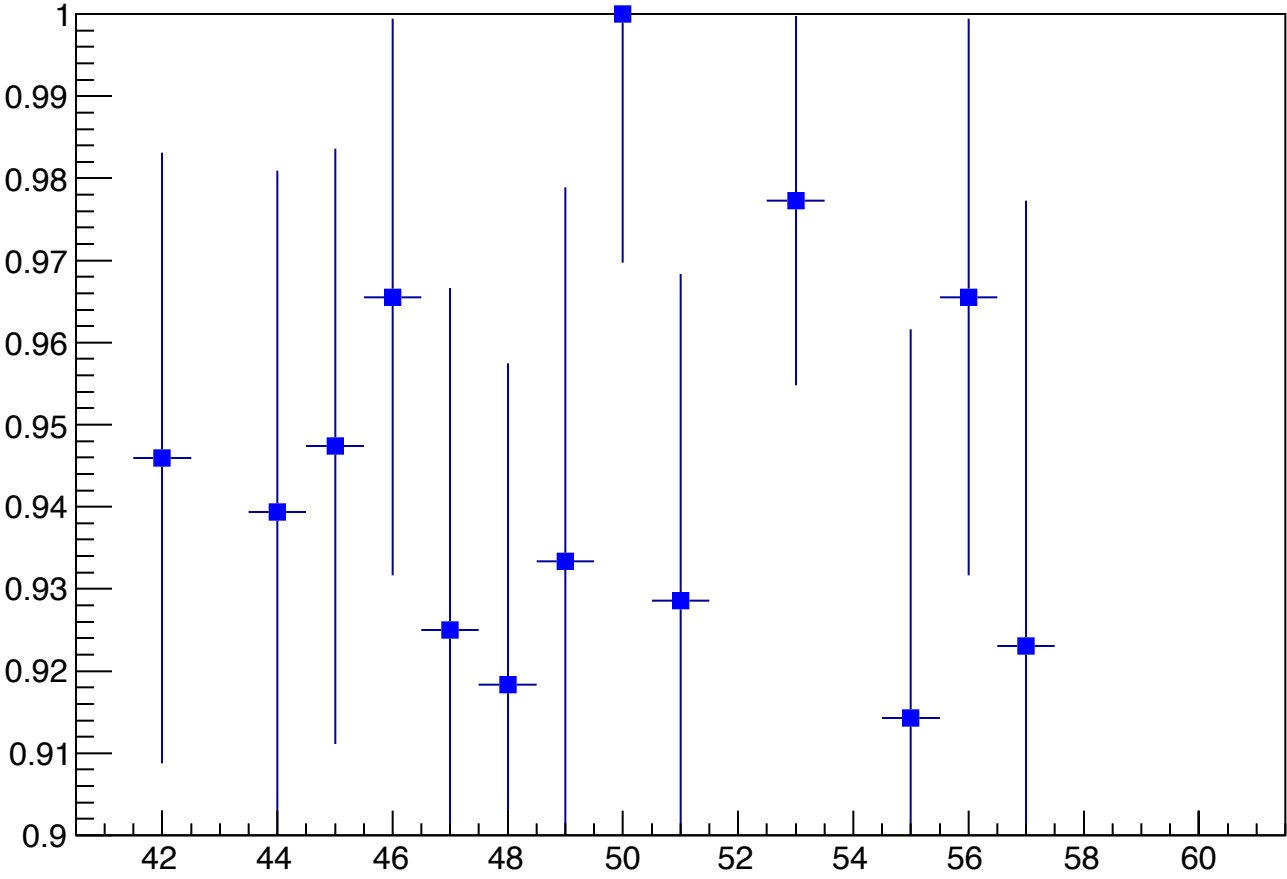
etc...

# Comparing Old and New

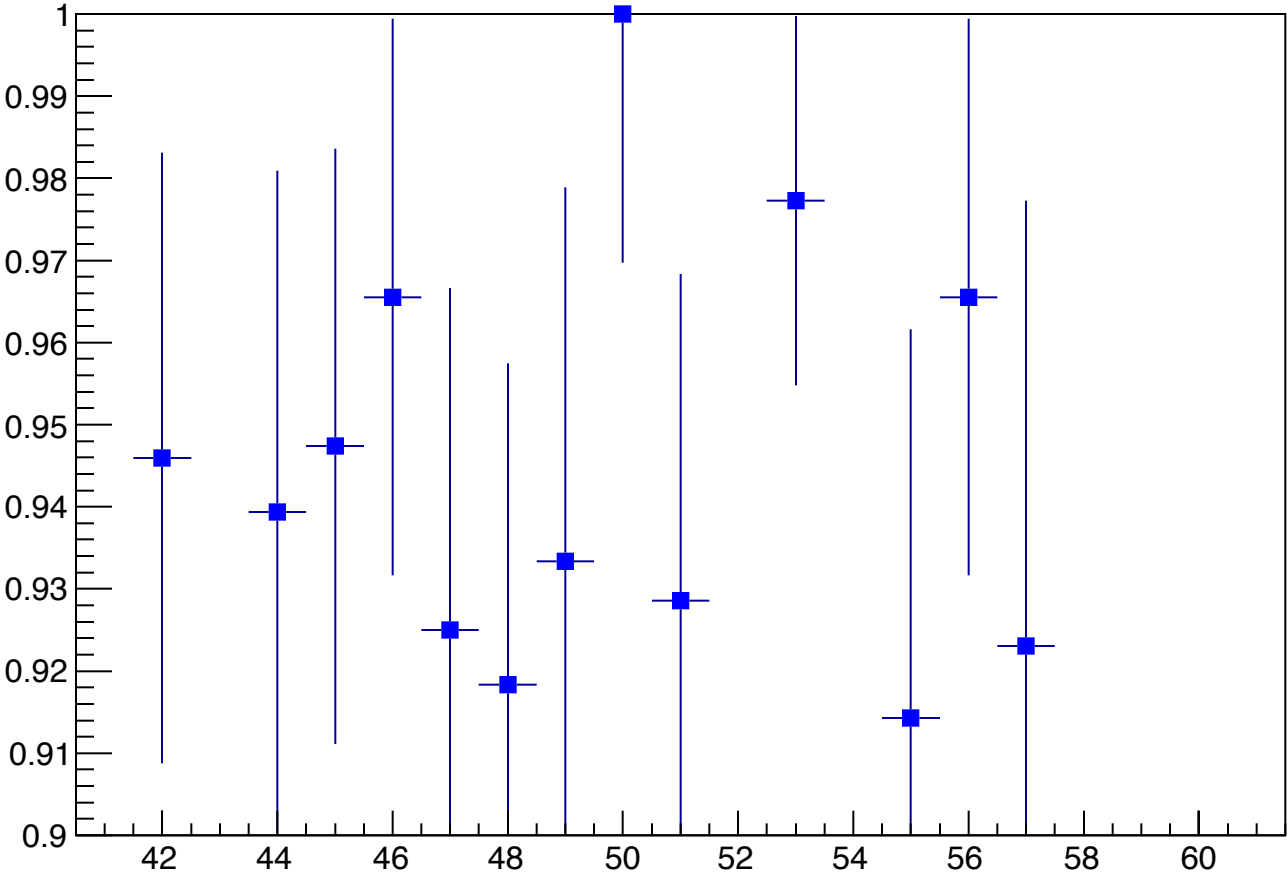
---

- Feeding the reformatted output into alleff should yield the exact same plots and fit result as what was originally being used.

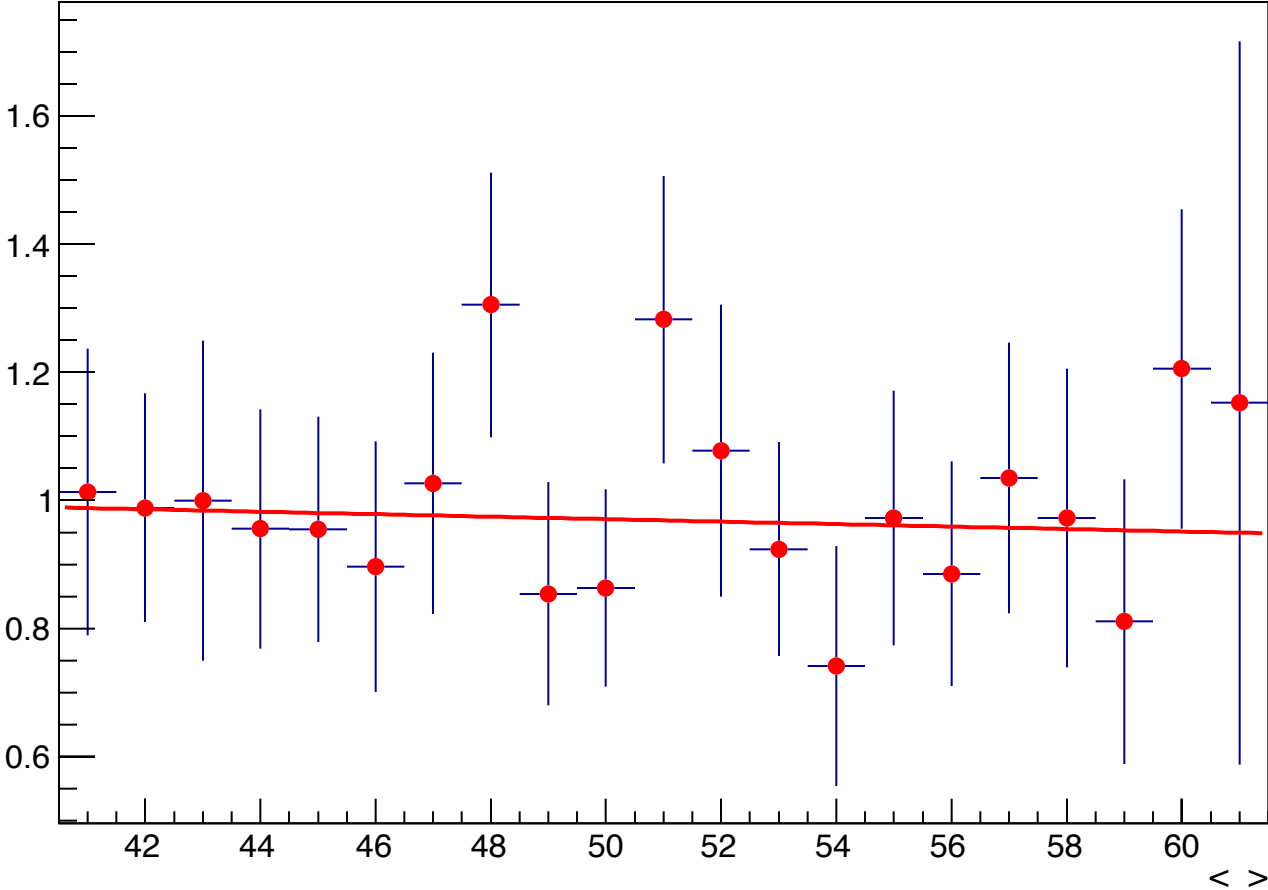
### Loose muon reco efficiency



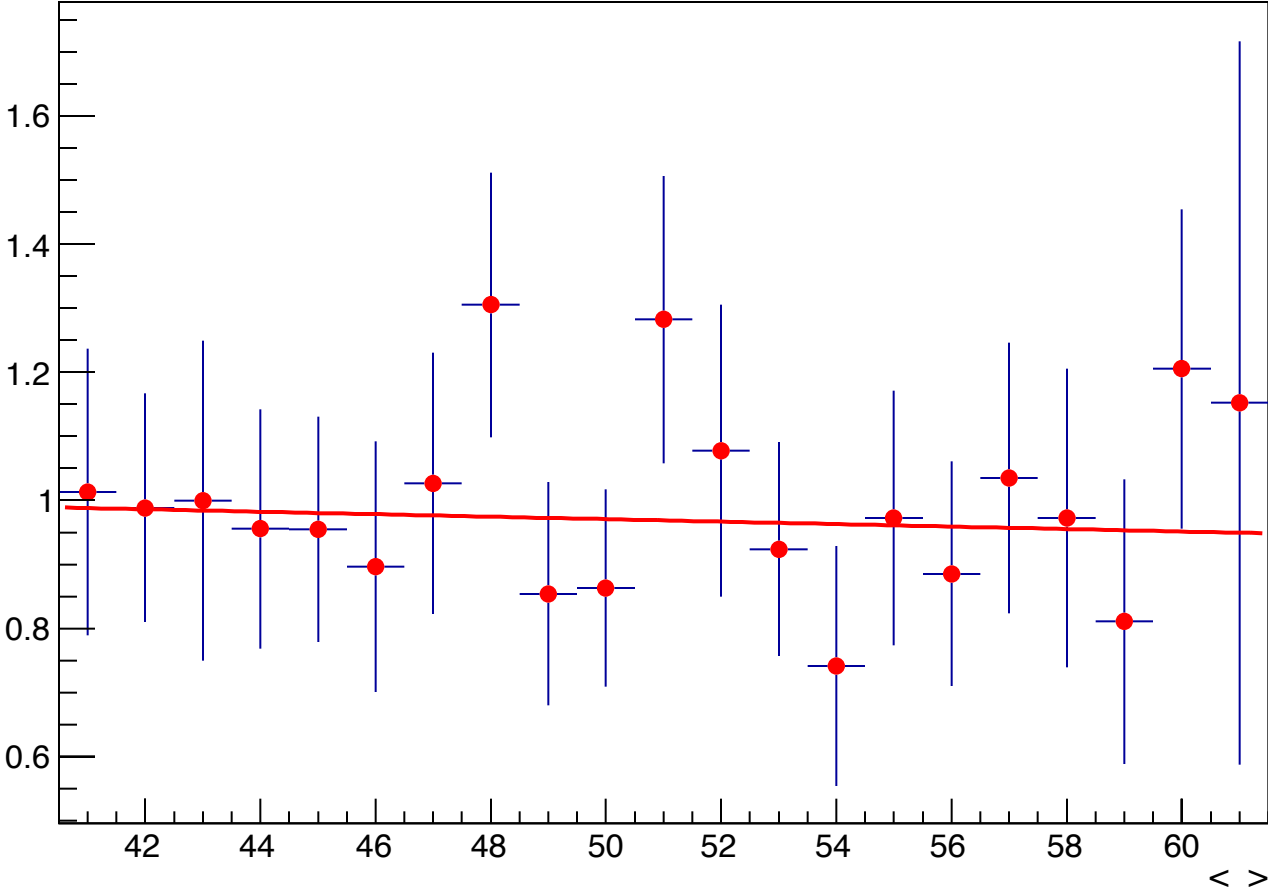
### Loose muon reco efficiency



### MC Correction Factor



### MC Correction Factor





# Mu Correction Fit Results

---

OLD

```
*****
Minimizer is Linear
Chi2          =      9.33542
Ndf           =      19
p0            =      1.06698 +/- 0.405557
p1            =     -0.0019285 +/- 0.00801704
Info in <TCanvas::Print>: eps file plots/284500_tp_correction.eps has been created
[shelver@lxplus026 ~]$ cd
```

NEW  
(USING 2D)

```
*****
Minimizer is Linear
Chi2          =      9.33542
Ndf           =      19
p0            =      1.06698 +/- 0.405557
p1            =     -0.0019285 +/- 0.00801704
Info in <TCanvas::Print>: eps file plots/284500_tp_correction.eps has been created
```

# Open Questions

---

- Unsure exactly how this will interact with multiple input files (and eventually when running over the full data set on the grid). Does it all collect in one file or will it produce multiple files?
- The `m_mcmatch` histogram is produced ONLY for simulation, however it suffers from the “many histograms” problem. Currently I have set it to not fill, but how important is it? Can it be changed to a 2D Histogram?