Jet validation and other stories
Qualification task report

Christian Kirfel

31st March 2021

UNIVERSITÄT BONN

Introduction

Technical work on validation input

- Adding variables and collections to the NTUP_PHYSVAL production
- Examples of collections and variables

Validation histograms

- Production of overlaid histosgrams for validation
- Typical procedure and an example

A tutorial to jet validation

- The steps to do a jet validation
- A step by step example

Work on the framework

Technical work on the JetValidation

- Merge Request

JetSubstructureHistos

- Class to manually calculate variables
- Tau21, Tau32, Tau21_wta, Tau32_wta, C1, C2, D2

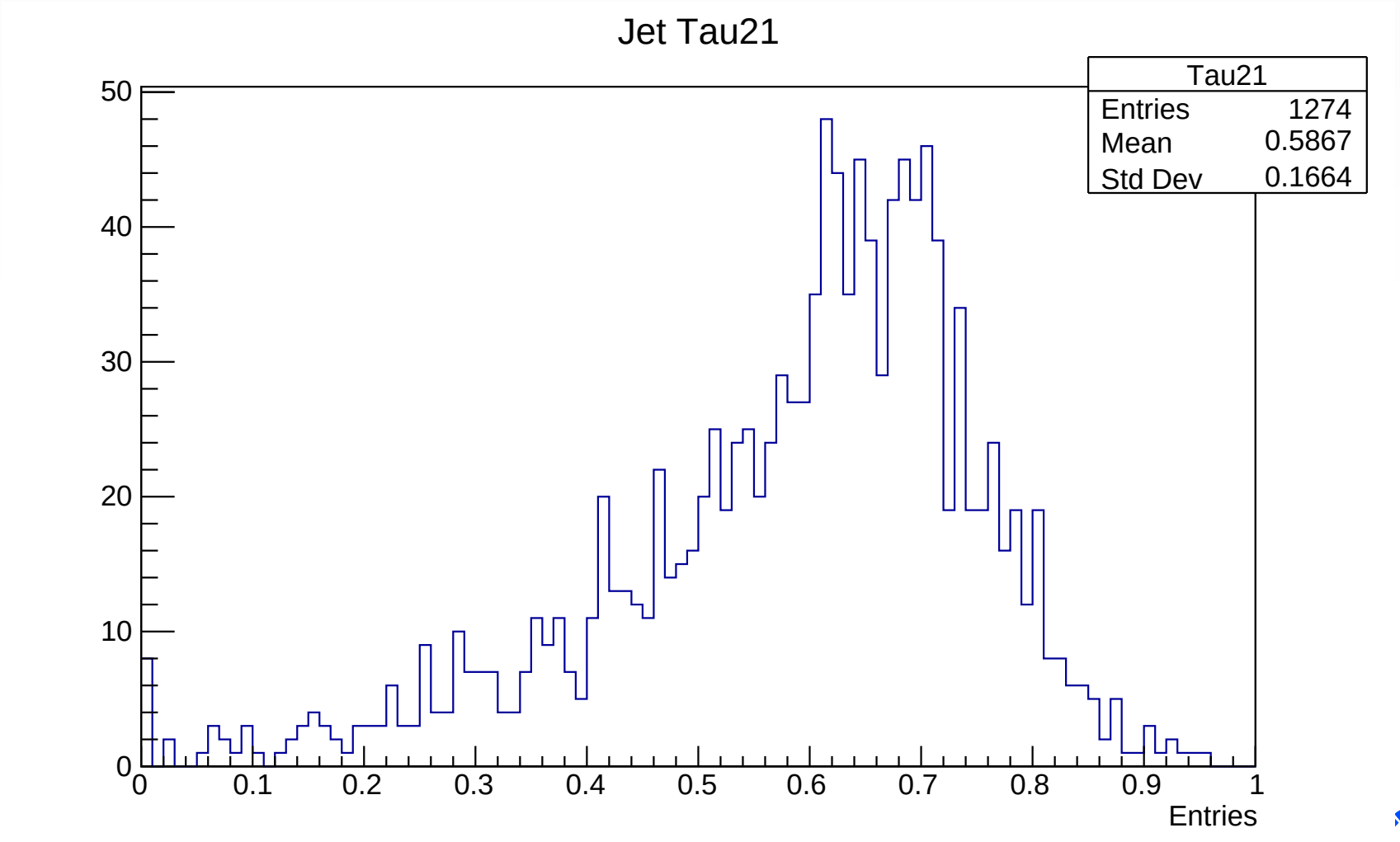
Clearance

- Removed empty histograms
- Rearranged according to new collections

AntiKt4EMPFlow

- Added: DFCommonJets_QGTagger_NTracks, DFCommonJets_QGTagger_TracksWidth, DFCommonJets_QGTagger_TracksC1, DFCommonJets_fJvt

Example of an added variable



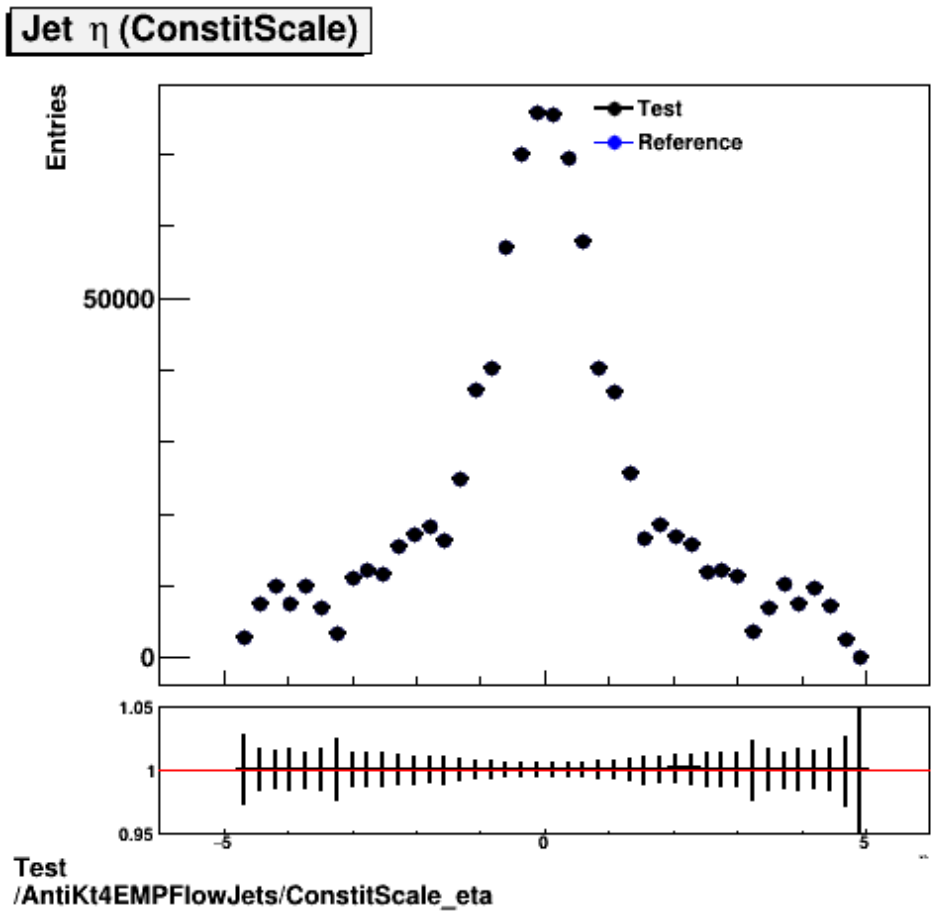
Jet validation
An introduction

What is validation

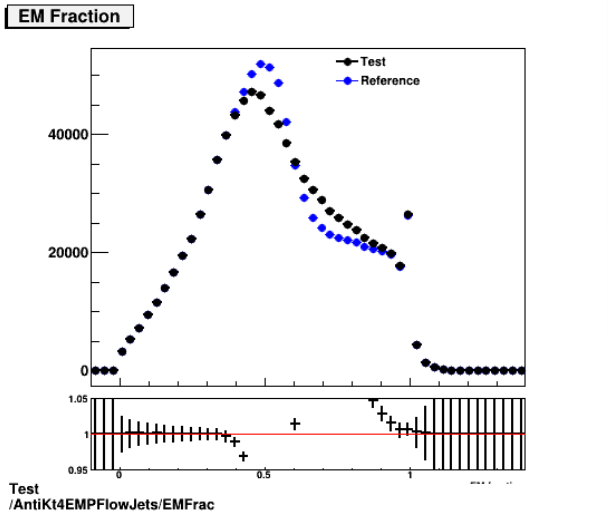
- Necessary to make sure that the switch to r22 is correctly done
- Starting point were representative r21 MC samples for each group
- Format: Events to Hits to AOD to DAOD_PHYSVAL to NTUP_PHYSVAL
- NTUP_PHYSVAL is just a histogram format useful for comparison
- Downloaded and overlaid plots are being created, evaluated and investigated

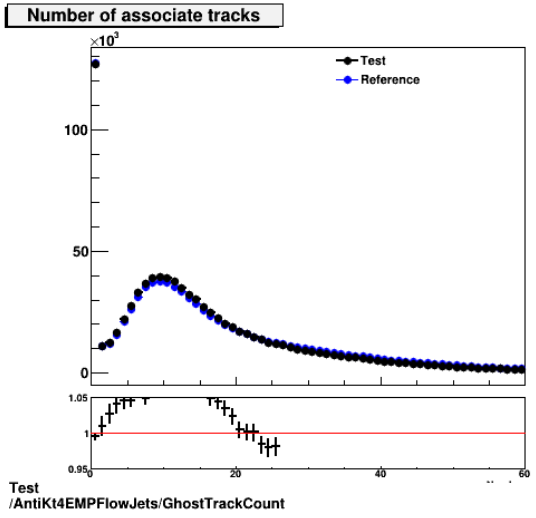
An example

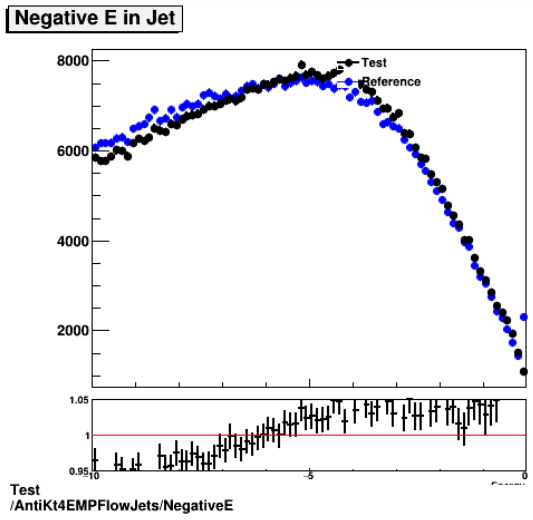
- Validation of jet reconstruction using FlowElements instead of PFO for release 22
- Sample: JZ7
- Collections: AntiKt4EMPFlowJets, AntiKt4EMPFlowFEJets



Red plots







- Electromagnetic fraction, negative energy and number of associated tracks show bad agreement.
- The reason now has to be investigated

Explanation of the disagreements

Negative energy

Explained by the compression of topocluster moments in the AOD.
Turning this off leads to perfect agreement.

Number of associated tracks

Thinning applied at AOD level leads to TRT hits not being stored.
Therefor PFlow had more tracks.

Electromagnetic fraction

For very high energy deposits in EMB2 a default value was chosen for PFO.

A tutorial to validation

Producing validation plots

Why would you need any of this?

- You want to make plots for validation
- You want a nice tool to easily make comparison plots from ntuples
- It is actually quite easy!

What are the requirements

- A grid certificate (if you want to download files via rucio)
- Access to athena
- A gitlab account

An example task

- The details of the task are unimportant
- It shows good behaviour for an example

Details of the task

- Description: Validation of G4fix quasi-stable particle simulation for Run2 re-processing
- JIRA: https://its.cern.ch/jira/browse/ATLPHYSVAL-701
- Reference: valid1:valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nc
- Test: valid1:valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nc
- Webpage

Setup

```
git clone ssh://git@gitlab.cern.ch:7999/ckirfel/jetphysvalidation.git
cd jetphysvalidation
setupATLAS
lsetup 'rucio -w'
voms-proxy-init -voms atlas
```

Set up rucio as wrapper to avoid polluting environment This works only for CLI and no rucio APIs available

Getting the files

```
./JetVal.sh
```

```
TASK: Validation of G4fix quasi-stable particle simulation for Run2 re-processing,test vs test3 JIRA: https://its.cern.ch/jira/browse/ATLPHYSVAL-701
Fetching -1 files from the grid (0 means about 100, -1 means ALL)
Reference sample: valid1:valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.NTUP_PHYSVAL.e4993_s3634_r12287_p4360_p3821 has 1 files
Sample exists!
Test sample:      valid1:valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.NTUP_PHYSVAL.e4993_s3227_s3653_r12320_p4360_p3821 has 1 files
Sample exists!
Downloading to exampleValidation_CKIRFEL/data_ref
rucio download valid1:valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.NTUP_PHYSVAL.e4993_s3634_r12287_p4360_p3821
The reference size is 1
Download successful!
Downloading to exampleValidation_CKIRFEL/data_new
rucio download valid1:valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.NTUP_PHYSVAL.e4993_s3227_s3653_r12320_p4360_p3821
The test size is 1
Download successful!
Copy the files..
Number of ref files is: 1
Number of test files is: 1
exampleValidation_CKIRFEL/data_ref/valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.NTUP_PHYSVAL.e4993_s3634_r12287_p4360_p3821/NTUP_PHYSVAL.23611053._000001.pool.root.1
/afs/cern.ch/user/c/ckirfel/JetPhysValidation/jetphysvalidation
exampleValidation_CKIRFEL/data_new/valid1.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.NTUP_PHYSVAL.e4993_s3227_s3653_r12320_p4360_p3821/NTUP_PHYSVAL.24355197._000001.pool.root.1
/afs/cern.ch/user/c/ckirfel/JetPhysValidation/jetphysvalidation
Compare downloaded directory sizes..
1 (ref) 1 (test) in 100MB units
The next step is to run JetValHan.sh
If you did not setup rucio in a wrapper, you should do this in a separate session to avoid conflicts!
```

Making the comparison plots

```
./JetValHan.sh
```

```
Set status..
Green=3
Red=4
Found Status (R/Y/G): 4/0/3
Created exampleValidation_CKIRFEL Move it where it belongs.
cp -r exampleValidation_CKIRFEL/exampleValidation_CKIRFEL*  /afs/cern.ch/atlas/groups/validation/JetEtMiss/
Are you sure to start to copy to CERN? [y/N] y
```

Checking the plots

flowElements_ttbar	MissET	2021/03/21	FlowElements in r22	2/0/0	20.7.3.5 20.7.3.5	C.Kirfel
flowElements_JZ7W	MissET	2021/03/21	FlowElements in r22	2/0/0	20.7.3.5 20.7.3.5	C.Kirfel
flowElements_JZ3W	MissET	2021/03/21	FlowElements in r22	2/0/0	20.7.3.5 20.7.3.5	C.Kirfel
exampleValidation_CKIRFEL	MissET	2021/03/29	Validation of G4fix quasi-stable particle simulation for Run2 re-processing,test vs test3 JIRA: https://its.cern.ch/jira/browse/ATLPHYSVAL-701	4/0/3		C.Kirfel
comparison_ttbar_log	MissET	2018/02/12	None	-/-/-		
comparison_ttbar	MissET	2018/02/12	None	-/-/-		
comparison_signal_log	MissET	2018/02/12	None	-/-/-		

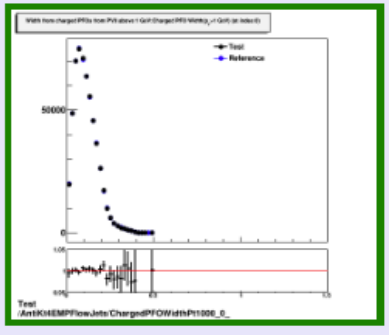
Validation webpage

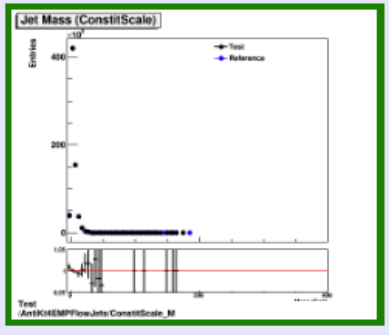
Good agreement

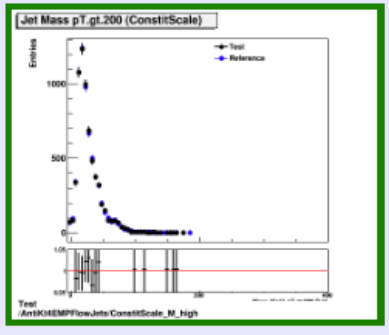
Click on images for details and full size.

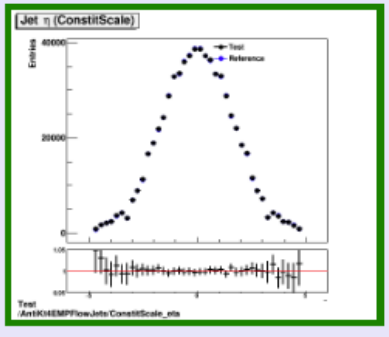
Test AntiKt4EMPFlowJets: Green only

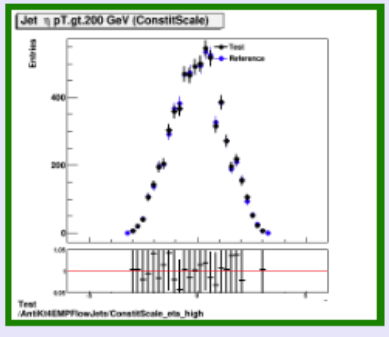
[Only Red] [Only Yellow] [Only Green]

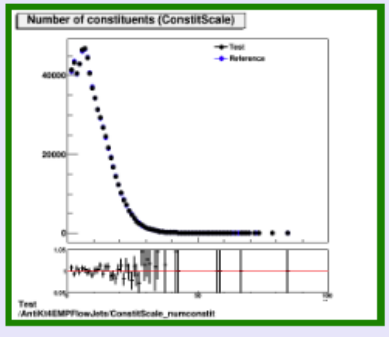


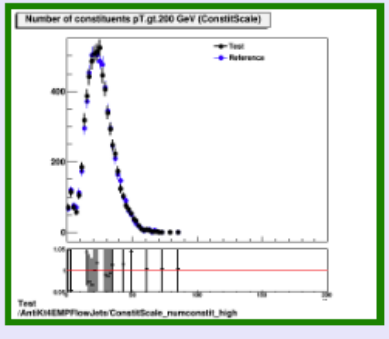


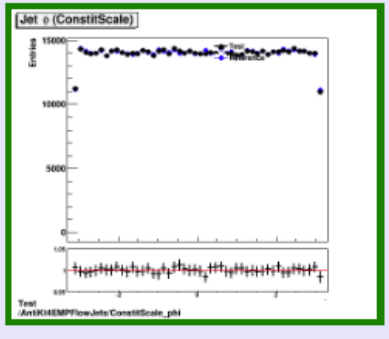


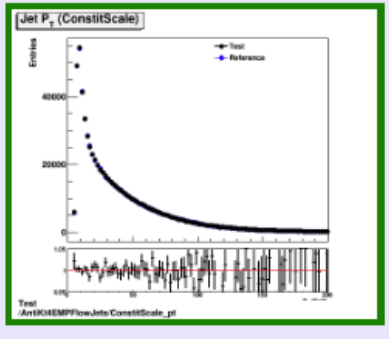










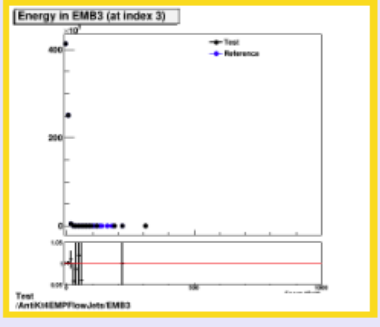


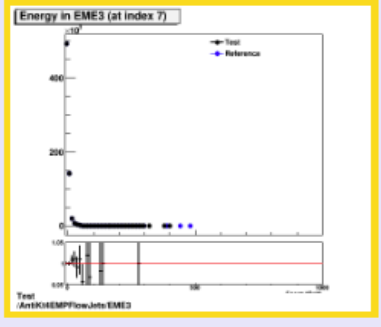
Medium agreement

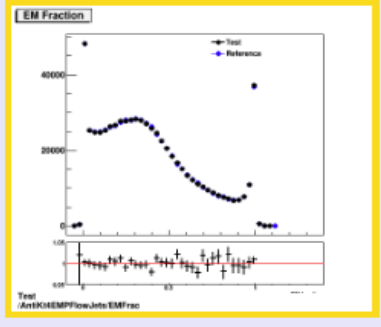
Click on images for details and full size.

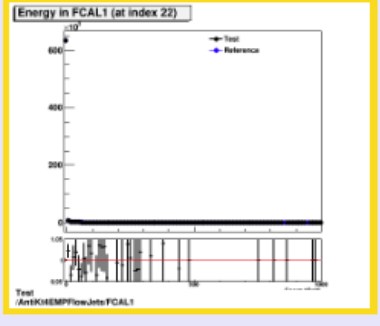
Test AntiKt4EMPFlowJets: Yellow only

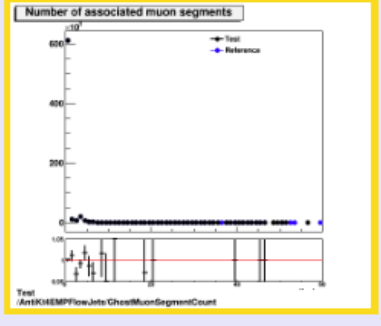
[Only Red] [Only Yellow] [Only Green]

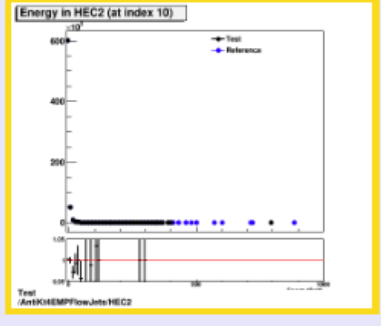


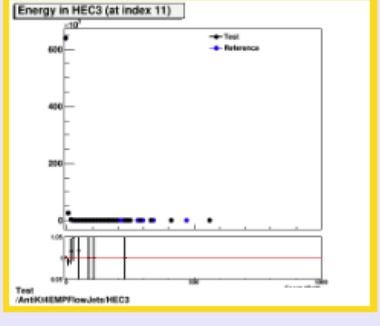


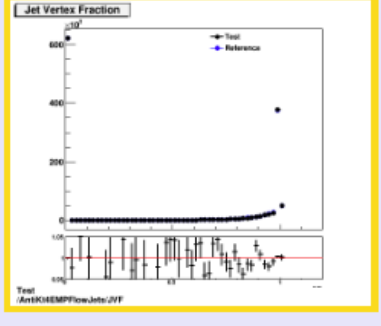


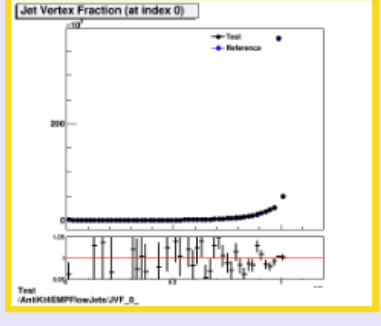








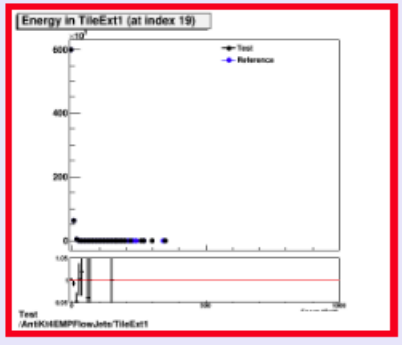




Bad agreement

Test AntiKt4EMPFlowJets: Red only

[Only Red] [Only Yellow] [Only Green]



A plot in detail

[Back]

Test AntiKt4EMPFlowJets/ptleadingjet

Location in ROOT file: Jets/AntiKt4EMPFlowJets/ptleadingjet

Assessment Details:

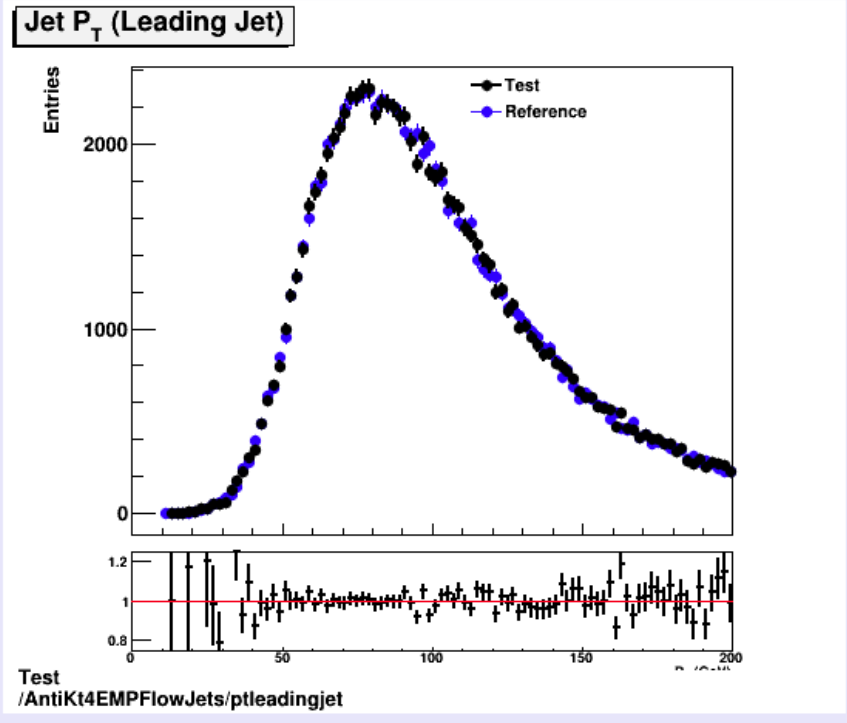
Name: ptleadingjet
Status: Yellow
Algorithm: RepeatAlgorithm
Num. of Entries: 99999
Underflow: 0
Overflow: 7555

Configuration Parameters:

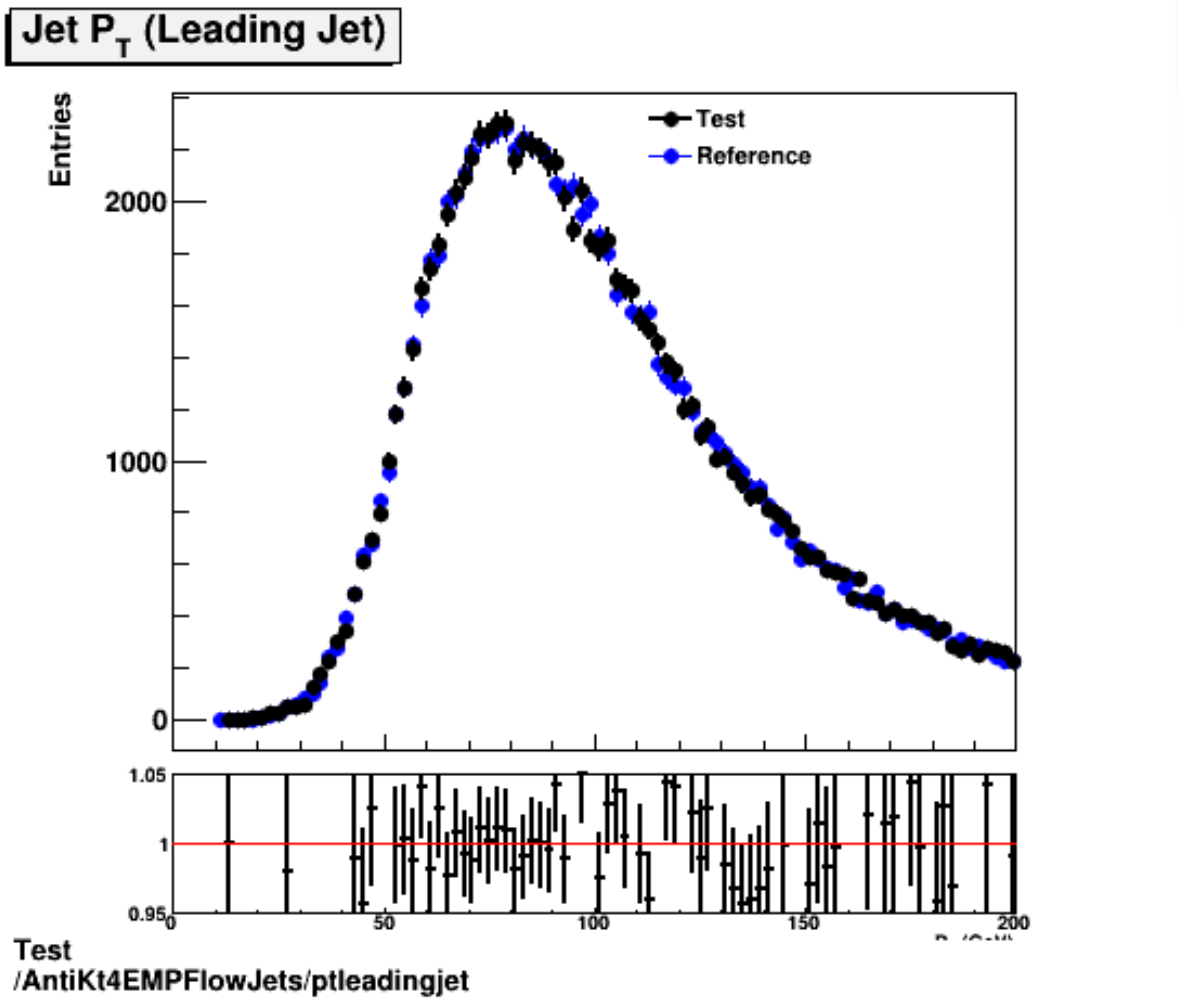
AuxAlgName--Chi2Test_Chi2_per_NDF: 1
RepeatAlgorithm--ResultsNEntries: 1

Chi2_per_NDF
XXXXXXXI XXXXXXXI XXXXXXX
1 1.5

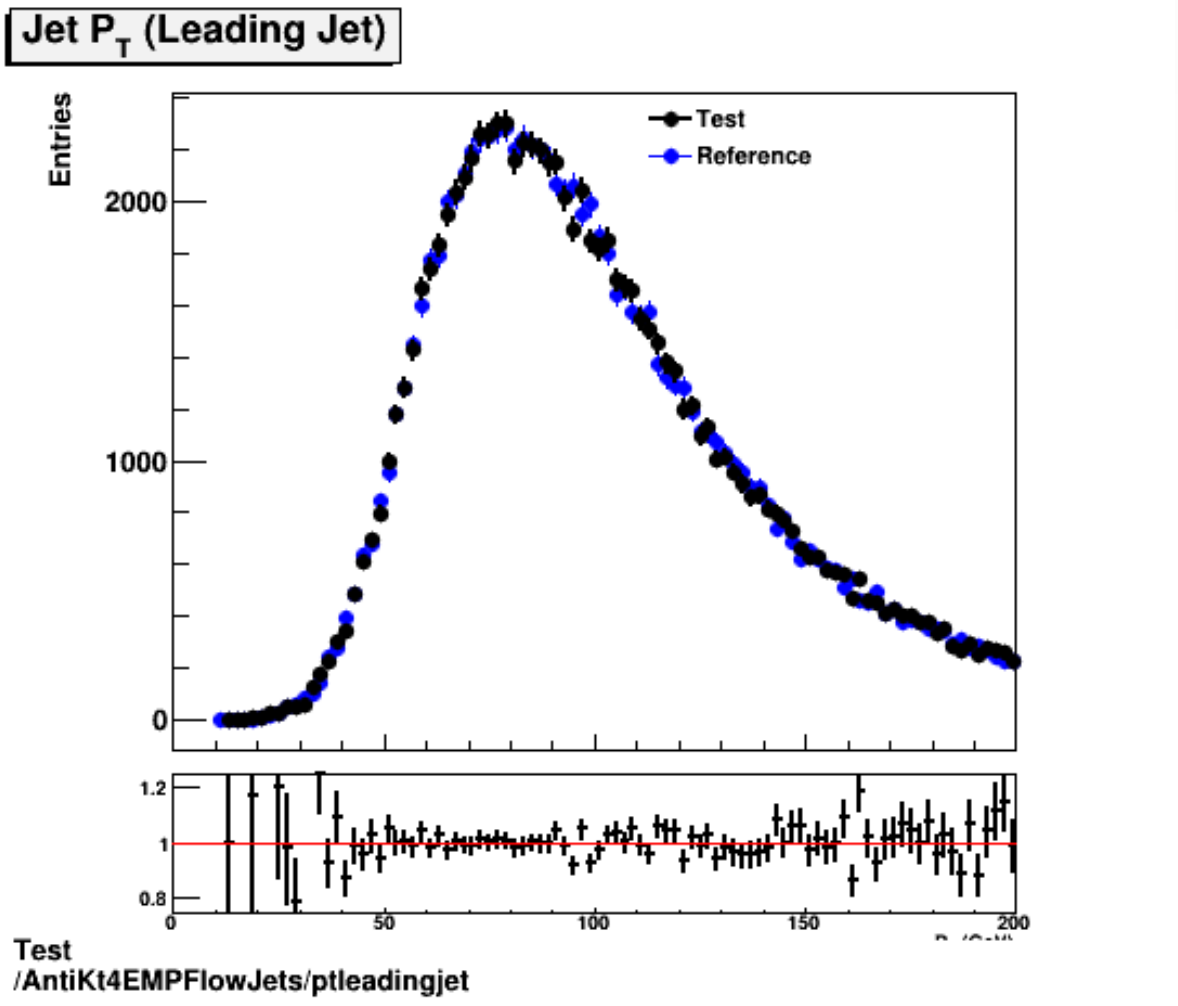
refSource: Multiple



Checking the range



Checking the range



Utilized scripts 1: The web display

- Tool to adjust the webdisplay
- Commonly used adjustments are arguments
- Some details have to be changed in the code
- TWiki

```
physval_make_web_display.py
--reffile Reference:val_ref/output_ref.root
--outdir=$PROJECT --title Test val_new/output_new.root
--startpath=Jets
--ratio --ratiorange=0.25
--normalize (or better scaleref)
```

Utilized scripts 2: The merger

- Code used to merge multiple input files

```
NTUPMerge_tf.py
--inputNTUP_PHYSVALFile=$MERGE_REF
--AMITag="p3821"
--outputNTUP_PHYSVAL_MRGFile="output_ref.root"
```

Troubleshooting. Do not make my mistakes

- You are using athena scripts, if you switched to a newer version that might have broken the code
- You might have used a local version of a script and forgot to switch back to the athena version
- Check that you have rights to push to the webpage
- You can always manually check that all files are in place and filled
- Check that you are not using the same file as test and ref by accident...

Conclusion

- Producing validation plots is a fast and engaging processing.
- Mix of coding and investigation of plots.
- With good tutorials and easily understanable code anybody can do it.
- Write tutorials and keep them updated!

Backup

Selection of validations

JSV

- r22 Feb vs March
- New tracking cuts
- Digitization

Validation meetings

- Single r-tag
- Quasi-stable particles
- New postprocessing
- Final AF3 validarion