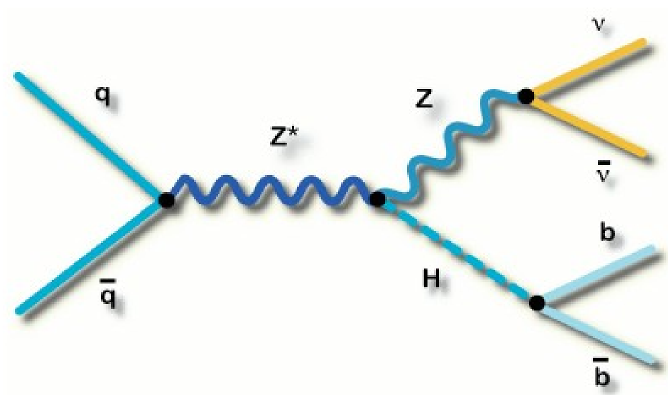
$ZH \to \nu \bar{\nu} b \bar{b}$

PROSPECTS FOR SUMMER

Murilo Rangel for the nunubb team

OUTLINE:

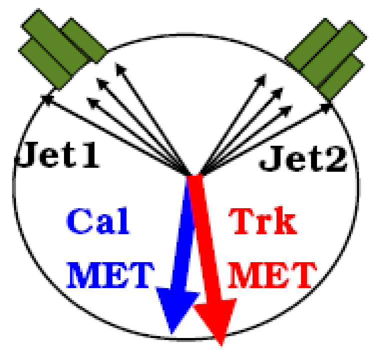
- Introduction
- Analysis Updates
- Summary



Introduction

Search for Missing transverse energy (MET) + 2 b-jets.

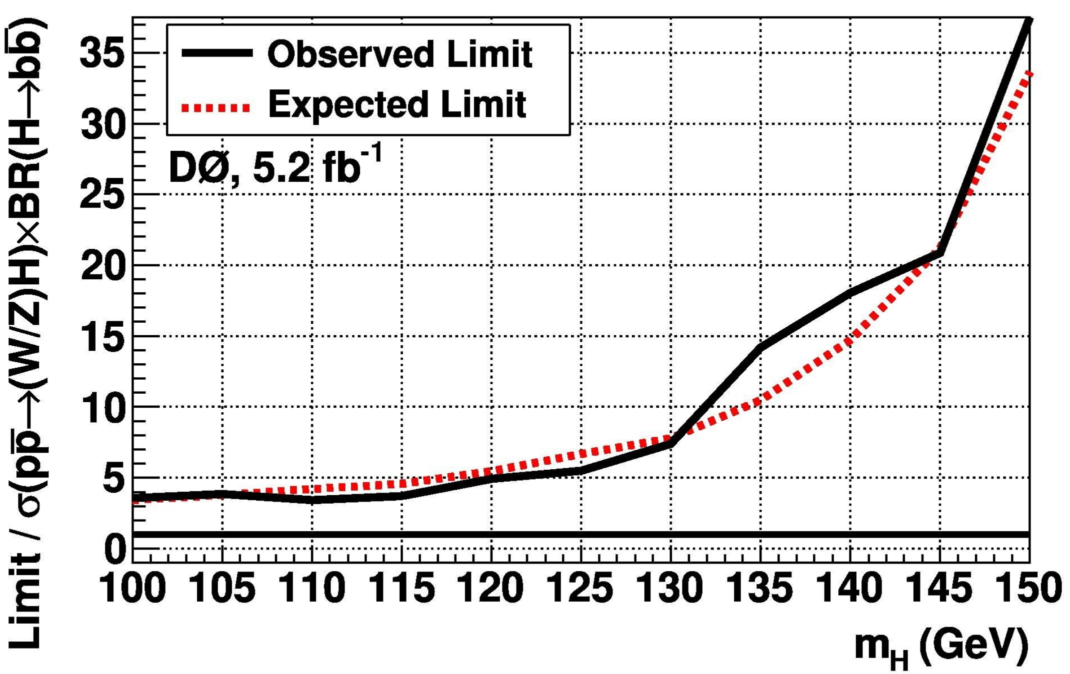
- Data is trigger is based MET+jets.
- Trigger simulation applied in MC.
- QCD background estimated from data.



- Selection:
  cuts for MET+jets topology;
  QCD veto using a dedicated multivariate discriminant;
  Light jets veto using b-tagging;

- Final limits are set with a multivariate discriminant against SM background.

- Published PRL 104, 071801.



Analysis Updates

- Add new Data (in progress):

Summer data set → 21.5% increase in the data used for publication.

- Multivariate discriminants (done):

Code updated to TMVA v4.0.3. Better discrimination achieved.

- B-taggigng (done):

Using new bottom-light jet discriminator (MVAbl) → improved discrimination between signal and light jets background.

- Jet Energy Resolution (JER) Improvements (in progress):

Additional corrections used to improve di-jet invariant mass resolution.

- New ideas:

Merged Jet Taggability and Jet Vertex Confirmation SF → reduce number of corrections and systematics (done).

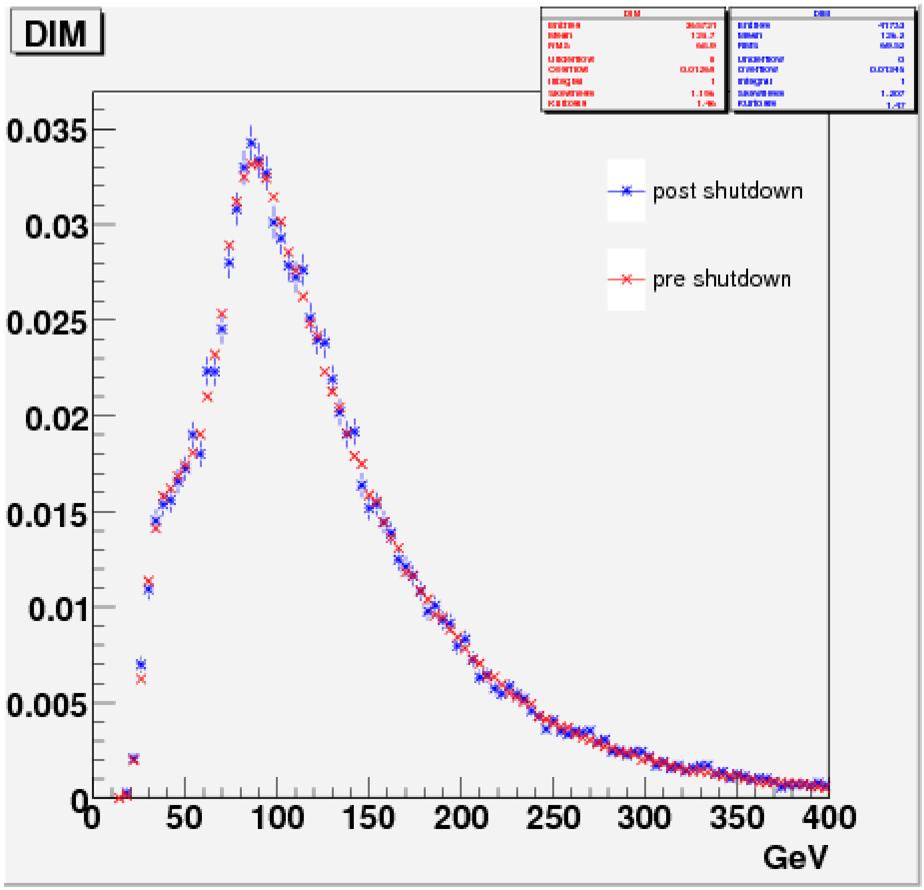
1 taggable jet → recover 30% of signal in the single-tag channel. 2-3% improvement in the final limit (done).

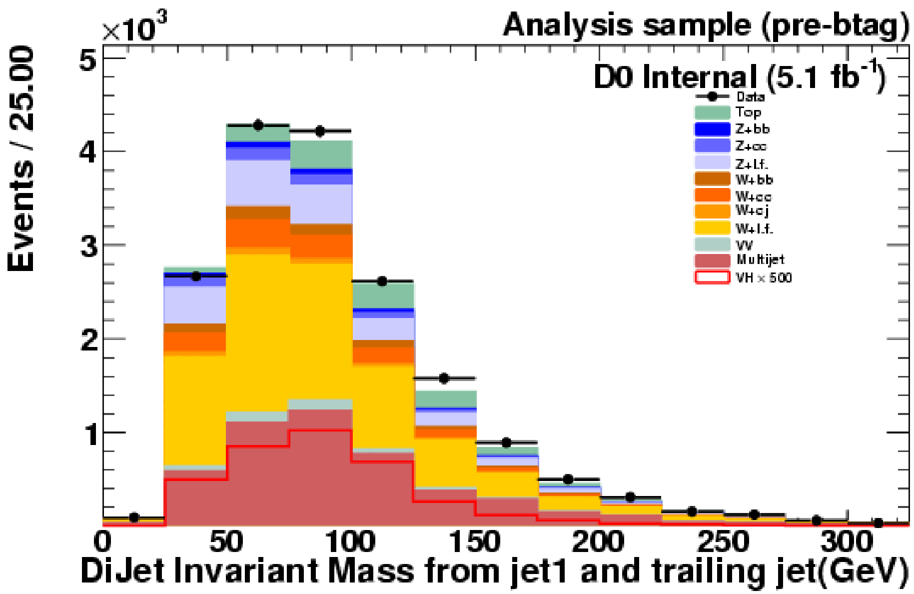
Bottom-charm jet discriminator (MVAbc) → improve final discrimination against charm jets (in progress).

Use MVAbl as input of final discriminators → use full information / correlation for bottom-light discrimination (in progress).

New data

We already looked over the new data.
Comparisons between previous data and new data show good agreement.

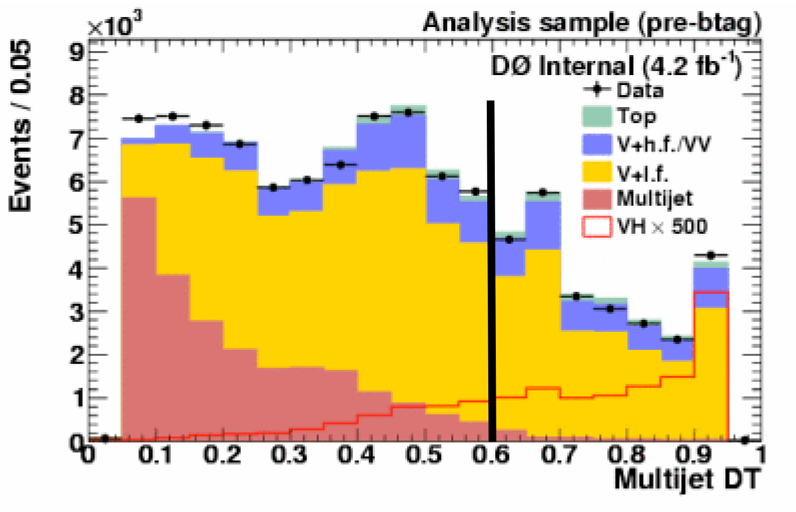


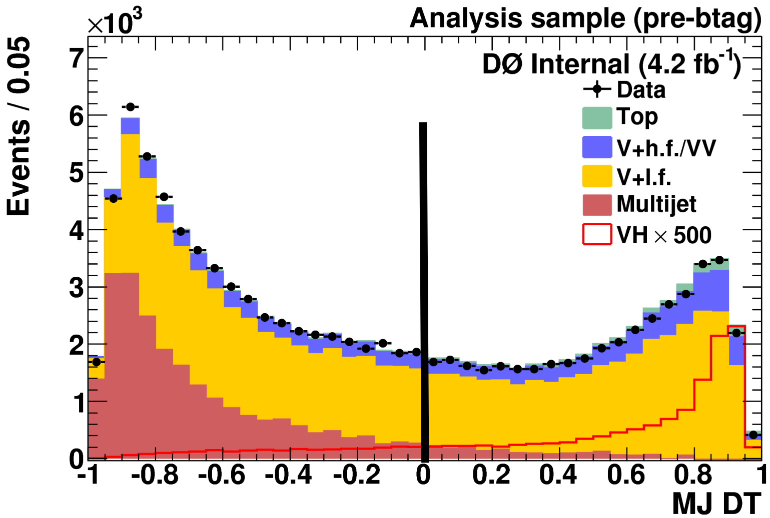


New data will be added without many problems.

Multivariate Discriminants

We have updated training code to TMVA 4.0.3





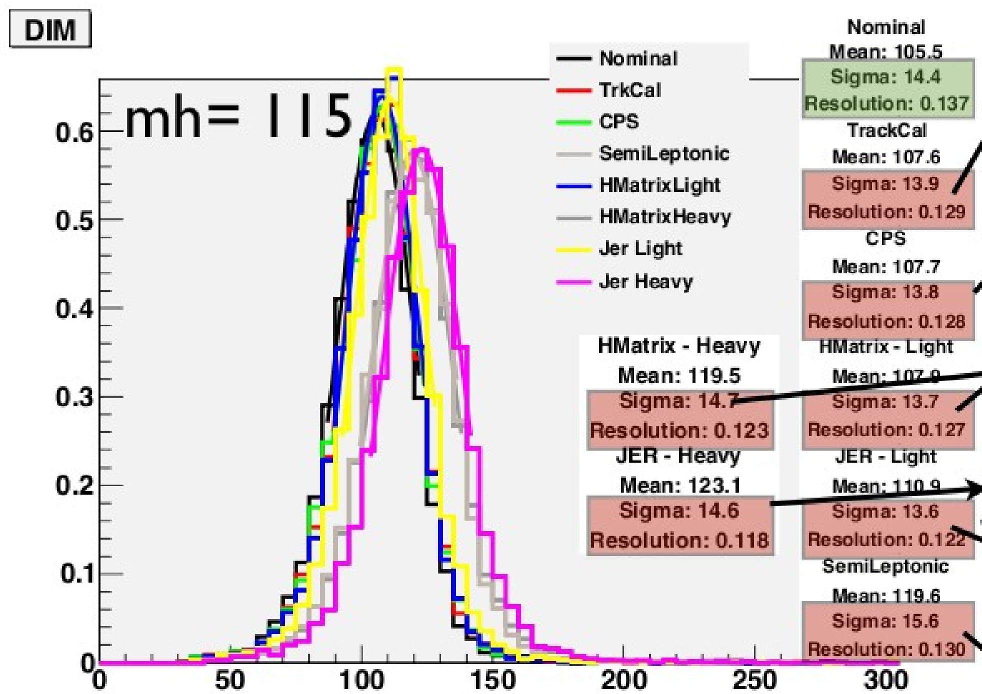
Signal efficiency increases ~10% for a given MJ rejection. For a first pass, we chose MJDT>0.

Publication	TMVA
MJ Rejection Eff **96%**	MJ Rejection Eff **93%**
Signal Selection Eff **71%**	Signal Selection Eff **82%**

We have also updated the Physics training. We get **3%** improvement in CLFast.

JER (1)

The Jet Energy Resolution improvements are only valid in different kinematic spaces



sigma/mean Improvements:

TrackCal: 6%

CPS: 7%

HMatrix (light): 7%

HMatrix (heavy): 10%

JER Combo (heavy): 13%

JER Combo (light): 11%

Semi-leptonic: 5%

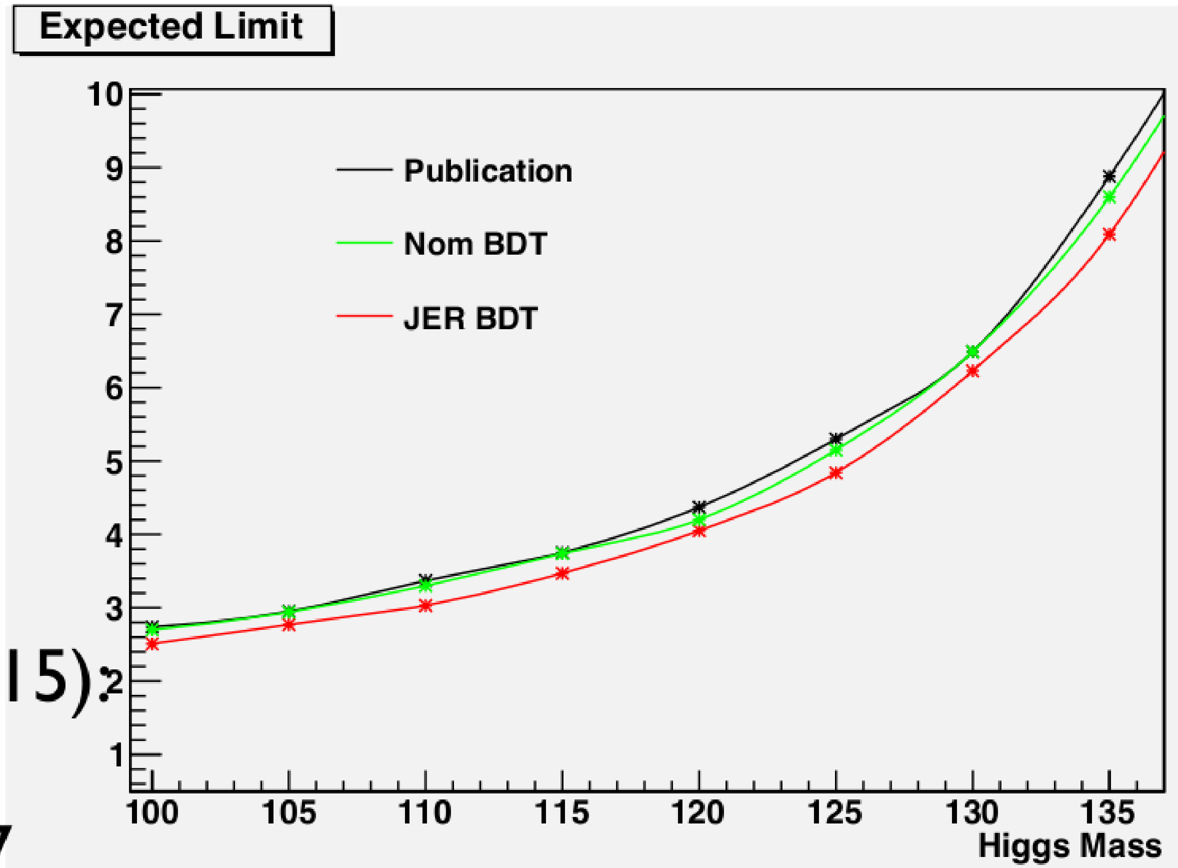
If we confine ZH->nunubb to the central calorimeter (|eta| <0.8), we recover the resolution gain in the individual improvements

Improvements in the CC

JER (2)

Using all of the JER Dijet Invariant Masses in the training of all BDT's we see an improved limit at the level of 5-10% over the entire Higgs mass range

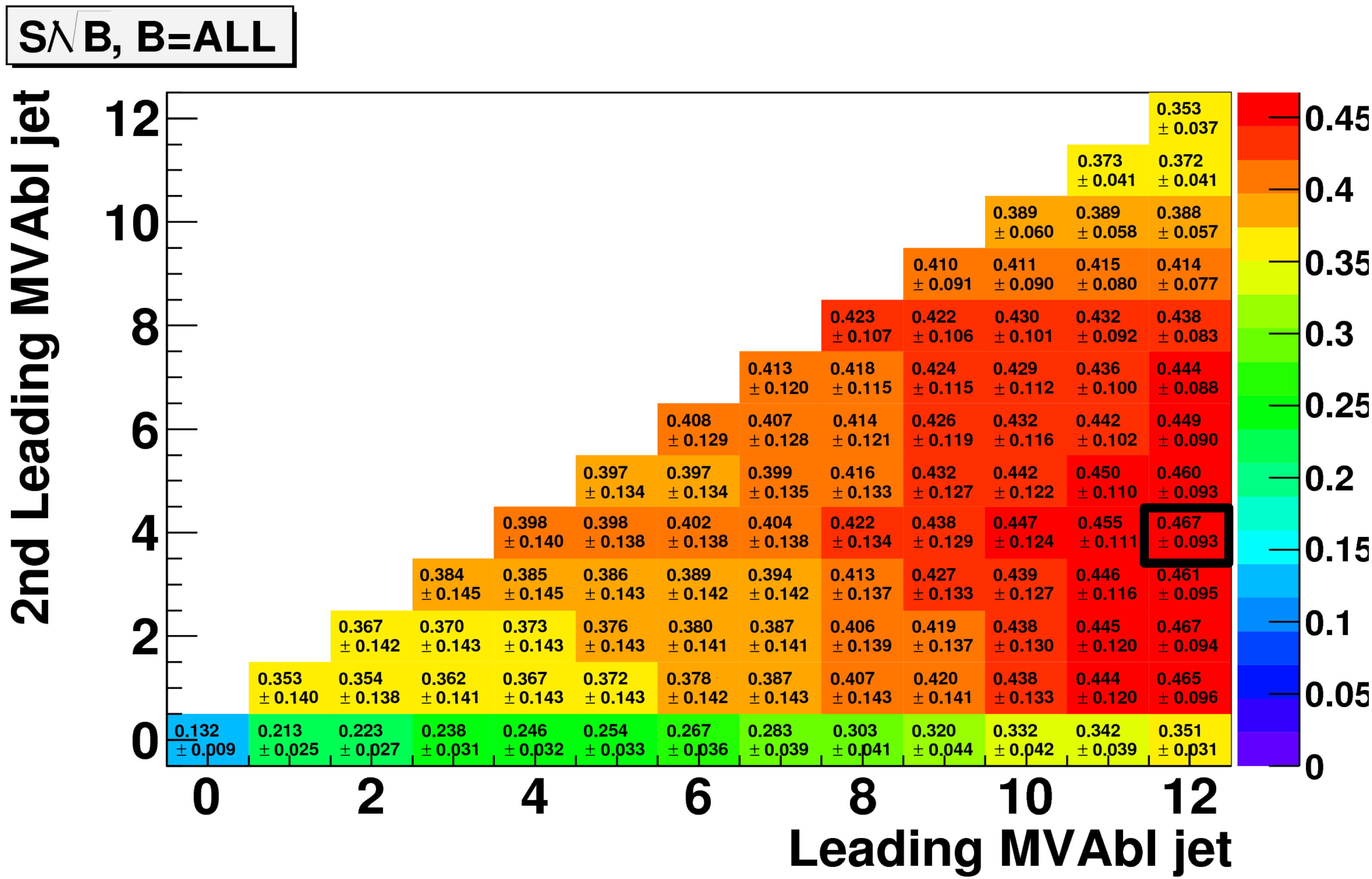
Expected limits (mh=115):
Publication: 3.75
w/JER variables:3.47



These results are for the old training code.
We are updating them with the latest framework.

Finding the best MVA bl OP (1)

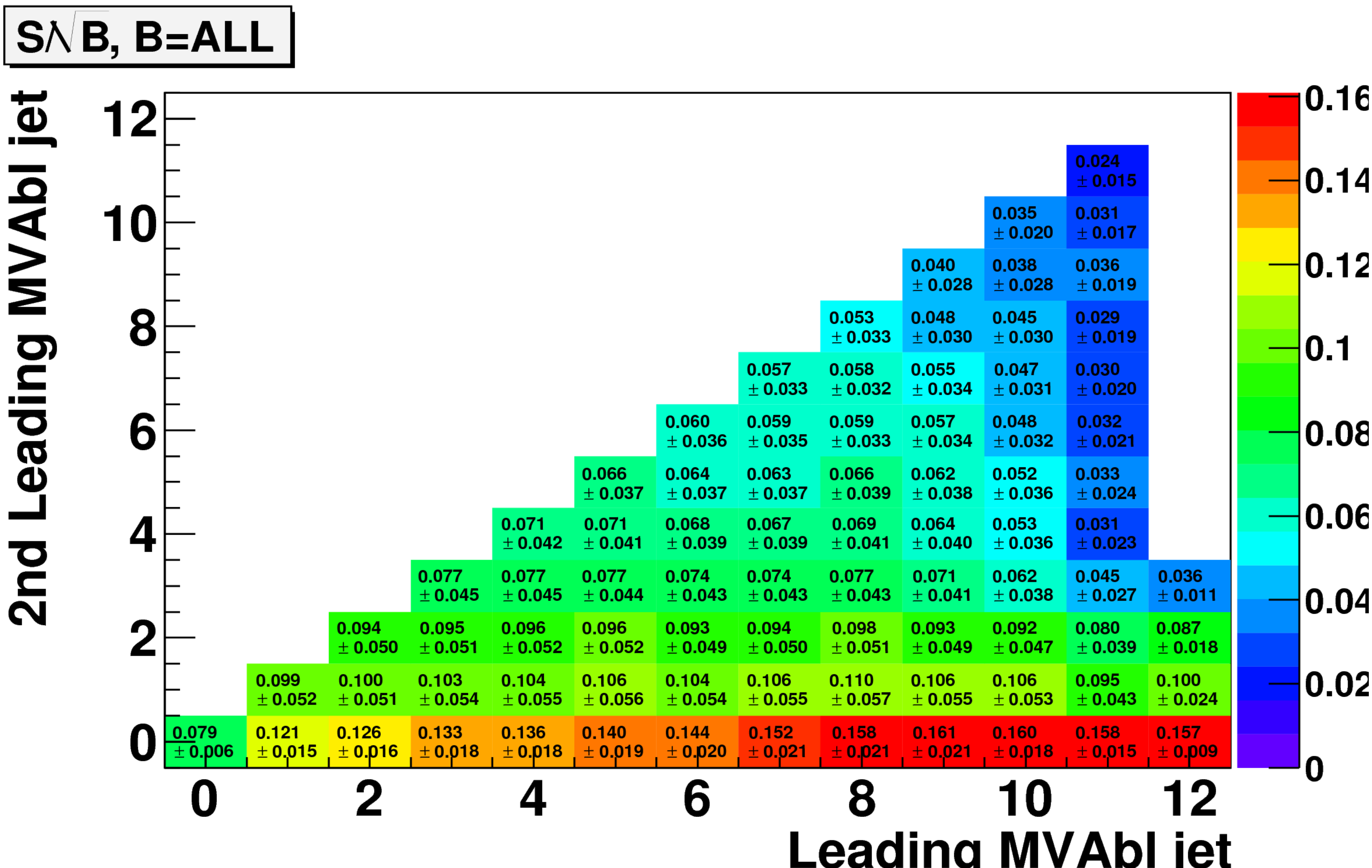
To select best Operating points (OP) for b-tagging, we selected events after Multi-jet veto and in the high S/sqrt(B) of the Final discriminant and compared S/sqrt(B) for all possible double-tags.



Chose MegaTight and L3.

Finding the best MVA bl OP (2)

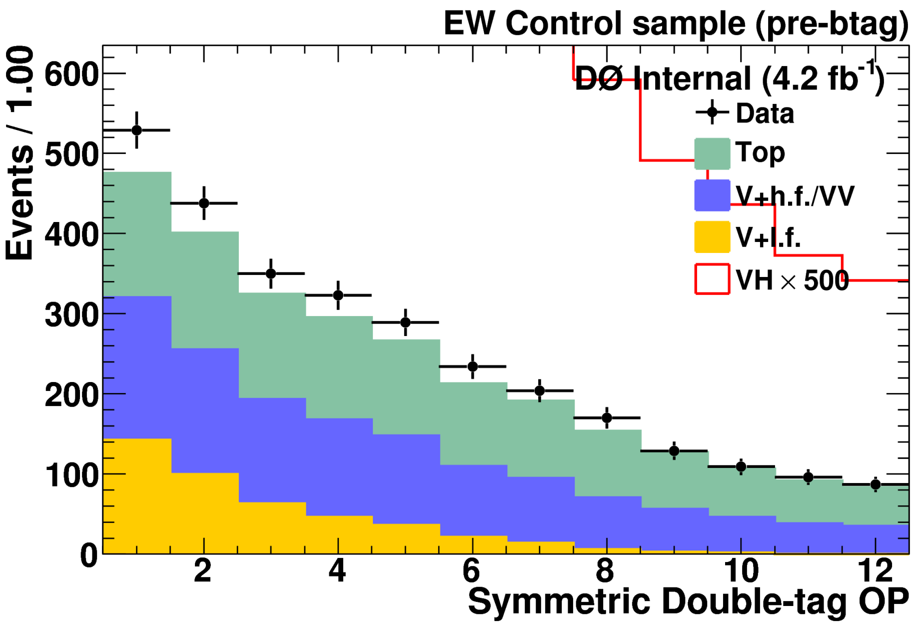
Removing the double-tag events, we can search for the best single or other double tag.

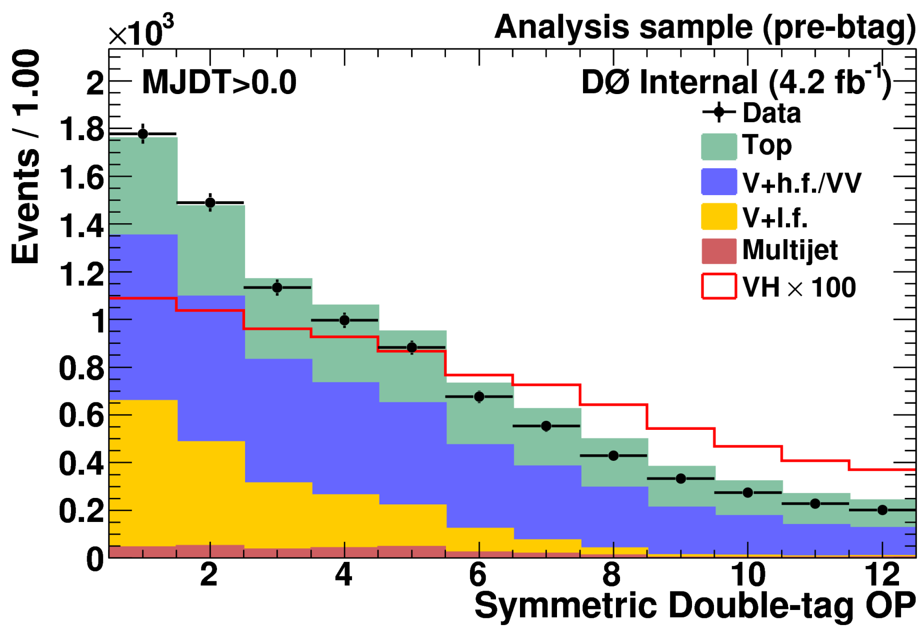


We could choose the Tight (OP=9) single tag, but since the improvement is very low, for simplicity we decided to use MT!L3.

MVAbl in the Final Discriminant

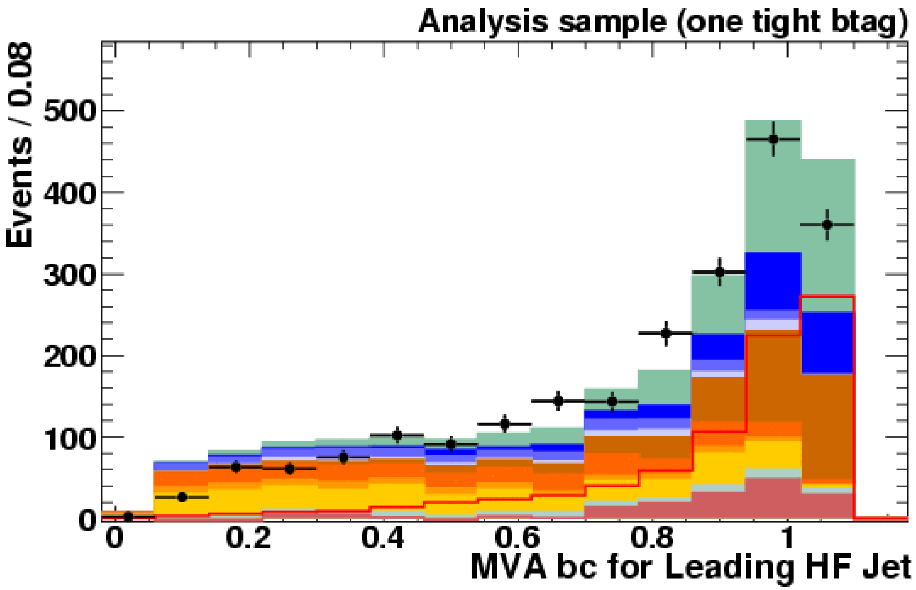
To proceed with the idea of using binned continuous MVA bl, we studied data x MC agreement in a 2D plane (Highest MVA bl vs. 2[rd] Highest MVA bl).
We found that the tighter we cut on MVAbl the lower the data / MC ratio.

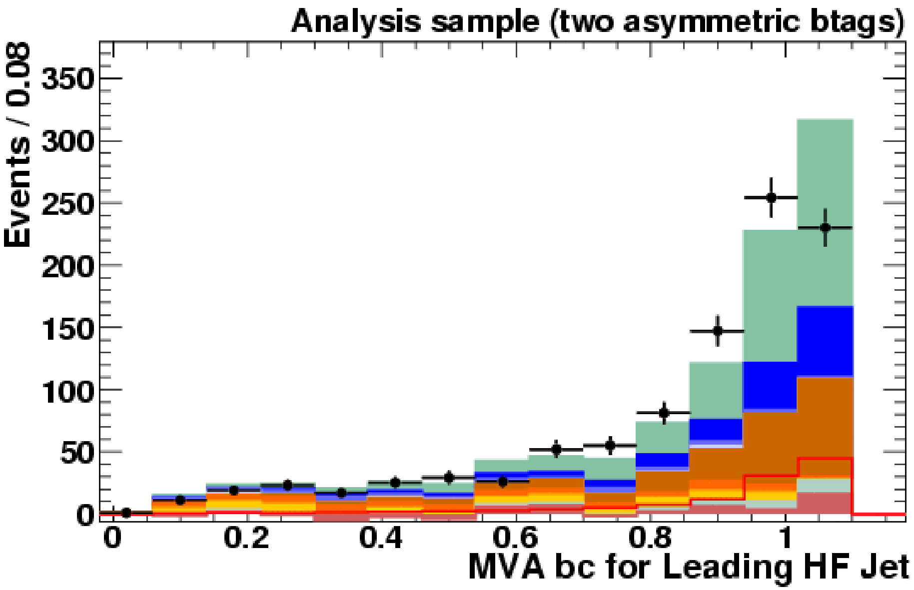




We could not achieve an overall agreement using all OP.
Using them in MVA discriminant, also gives a strange MC x data agreement (backup).
There are also studies in the WH channel. See these talks for more information:
http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a10703&id=a10703s1t46/transparencies
http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a10679&id=a10679s1t66/transparencies
http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a10644&id=a10644s1t33/transparencies

MVAbc / MVAbb





Bottom-charm discriminators (MVAbc) show good discrimination.

Data-MC agreement is not perfect. We have to decide if it is good enough to be used in the final discriminator.

MVAbc scale factors (to be provided) are expected to help.

Summary

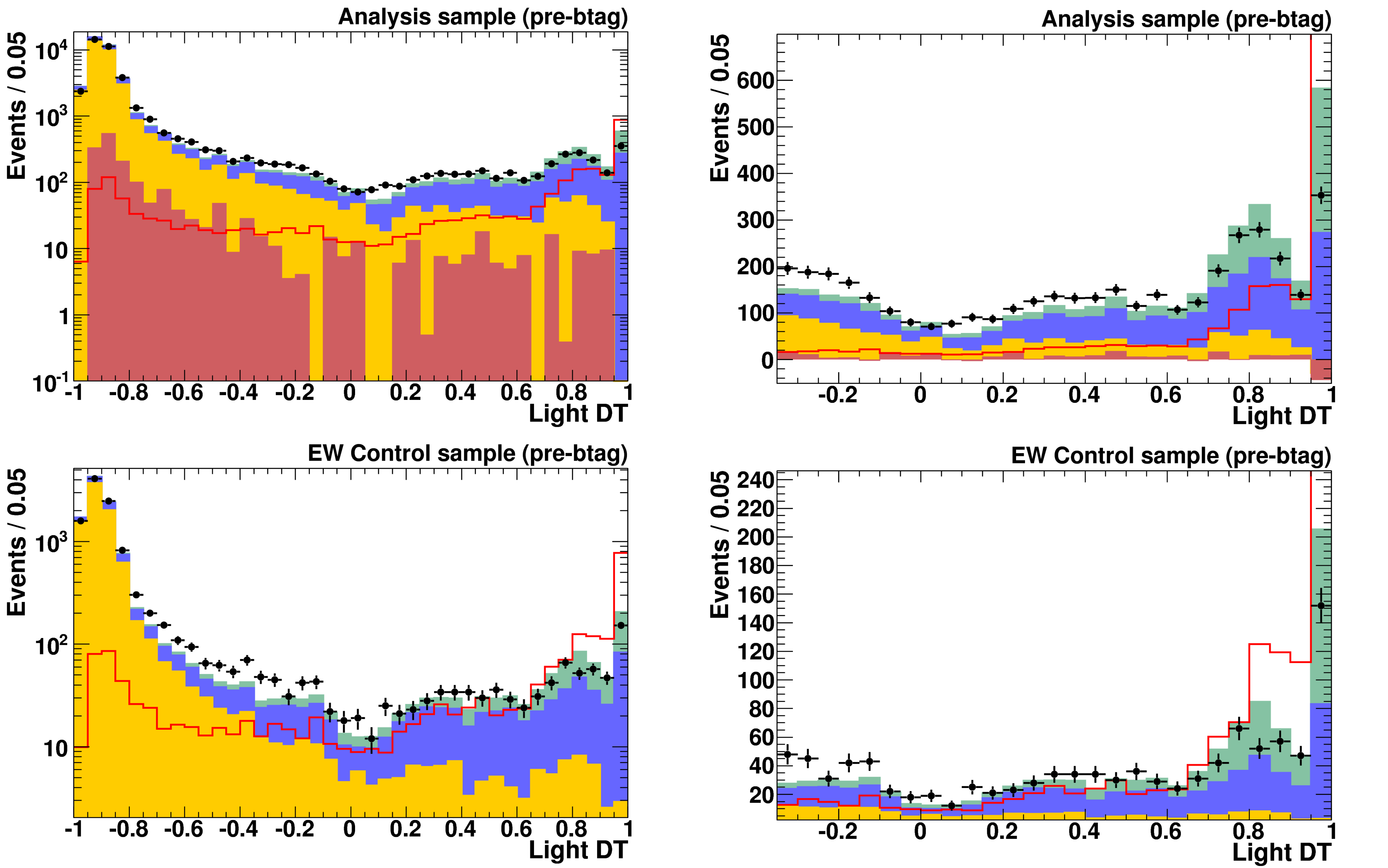
- Updates for summer look very promising,
  but there are still many challenges ahead.

- With 21.5% more data * (
  7% from TMVA and MVAbl + 2% from one taggable channel +
  + X% from JER and MVAbc)
  We can have ~20-X% better limits than PRL.

BACKUP

MVAbl in a MVA

Using in the training the 2 jets MVAbl and their pT. Only using the light MC as background.



Scale Factors

In the process of understanding the MVAbl behavior in our samples, we concluded that the Scale Factors (Vjets, SHF) used for plotting should not be necessarily the same between EW and Analysis samples.

Very recently, we started to think to have different set of SFs.
A top cross section SF should also be derived.

Using the agreed cross section uncertainties in a chi2 minimization, we can fit SFs for each sample (selection).

For **EW sample**, we use as inputs the 2 and 3 jet bins for 0-tag, 1-tag and 2-tag.

For **Analysis sample before MJDT cut**, we use the pre-tag MJDT distribution and the 2 and 3 jet bins for 0-tag, 1-tag and 2-tag. This also gives the QCD normalization.

For **Analysis sample after MJDT cut**, we use the 2 and 3 jet bins for 0-tag, 1-tag and 2-tag.

In a sense, we will have a simplistic version of post-COLLIE fit distributions.

Scale Factors (2)

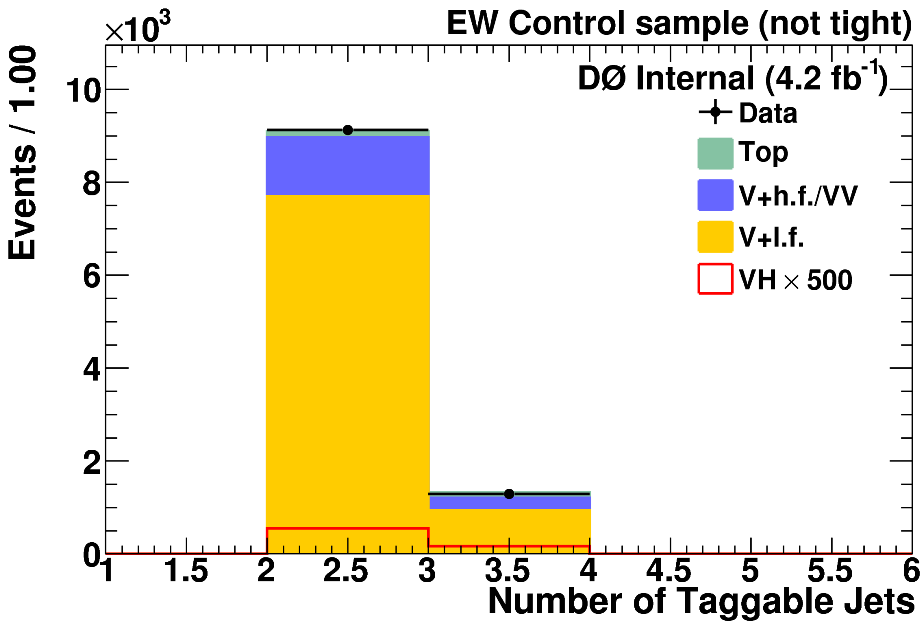
SFs for the EW sample:
Svjets = 1.013 +- 0.017
SHF = 0.960 +- 0.099
Stop = 0.999 +- 0.078

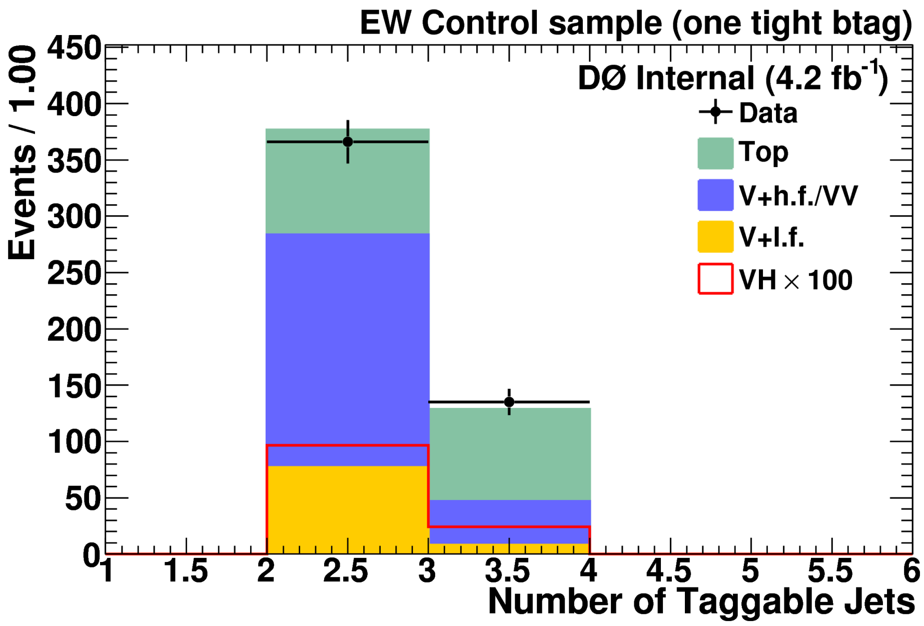
SFs for the Analysis sample before MJDT cut:
Svjets = 0.989 +- 0.010
SHF = 0.962 +- 0.065
Stop = 0.892 +- 0.077
SQCD_pretag = 1.923

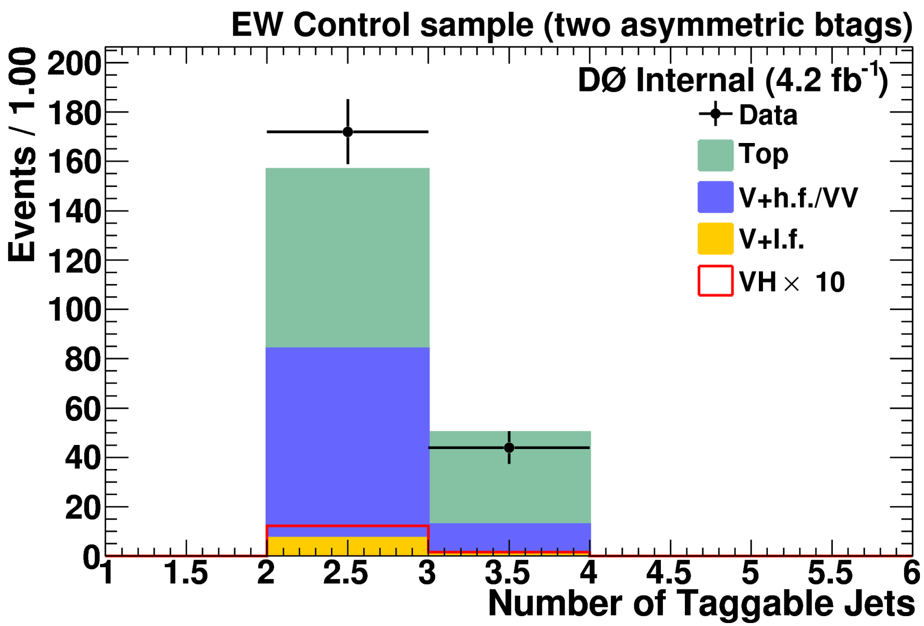
SFs for the Analysis sample after MJDT cut:
Svjets = 0.989 +- 0.012
SHF = 0.864 +- 0.072
Stop = 0.874 +- 0.075
SQCD_pretag = 1.923

Remarks:
1) The Svjets are all consistent within errors
2) The SHF are well consistent between EW and Analysis sample before MJDT cut
   The Stop are “consistent” within errors (these are independent samples)
3) The Stop are consistent without and with MJDT cuts (but the constraint is 10%)
   There is a substantial reduction of SHF (within the 20% constraint).
   Is this related to the pT dependence of data/background vs. operating point
   seen by Sebastien ?

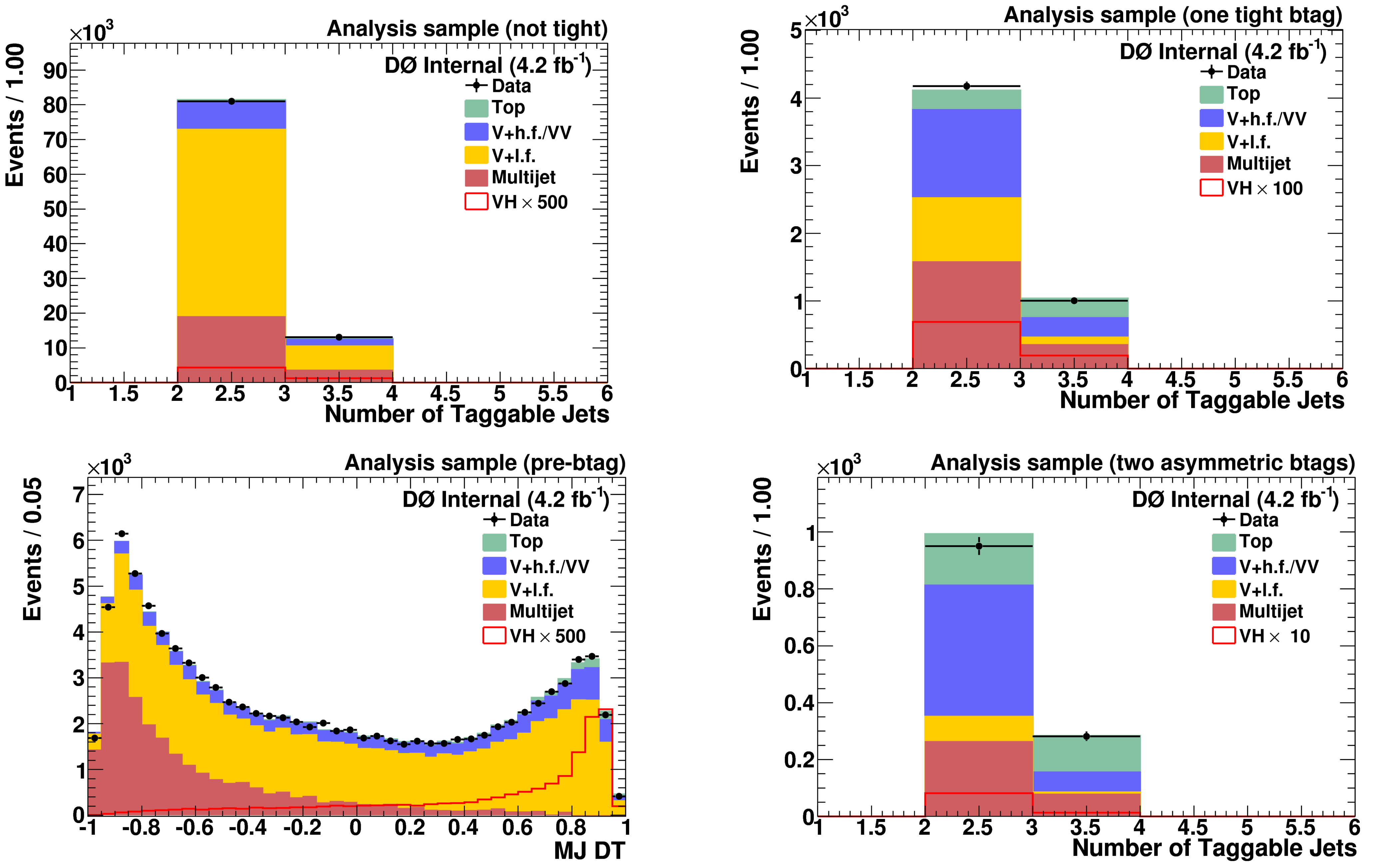
EW Sample


EW Control sample (not tight)
×10^3
DØ Internal (4.2 fb^-1)
Data
Top
V+h.f./VV
V+l.f.
VH × 500
Events / 1.00
Number of Taggable Jets

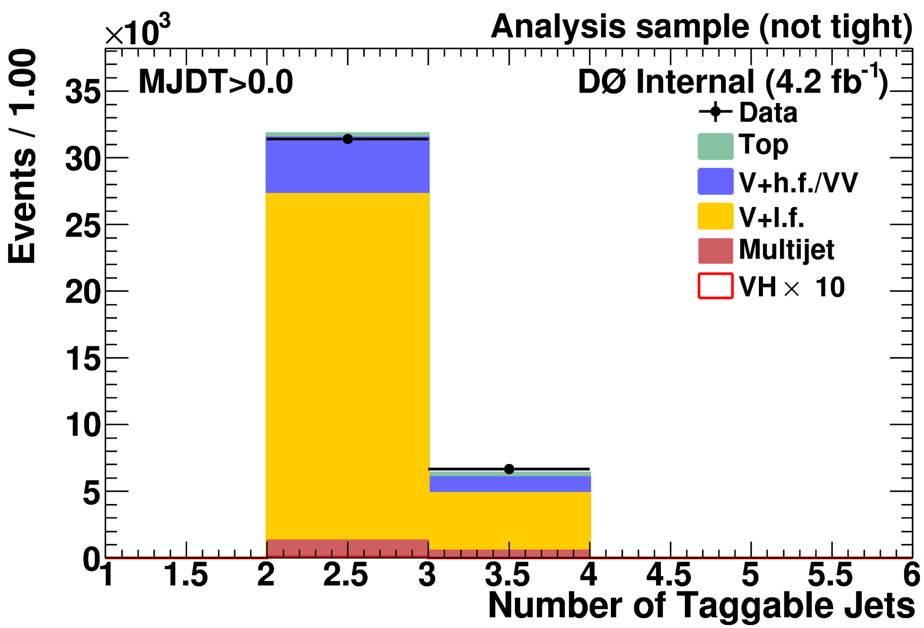

EW Control sample (one tight btag)
DØ Internal (4.2 fb^-1)
Data
Top
V+h.f./VV
V+l.f.
VH × 100
Events / 1.00
Number of Taggable Jets

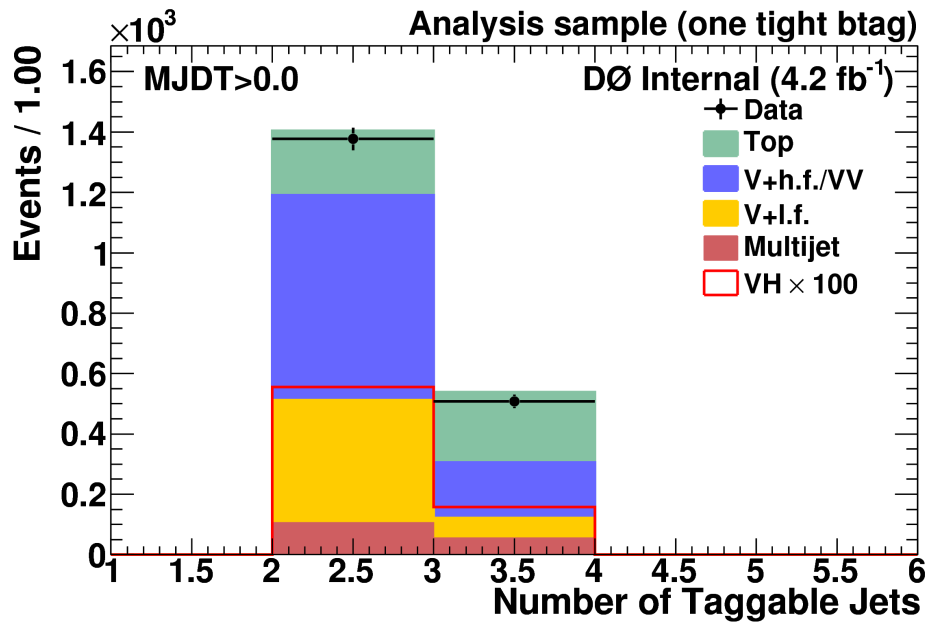

EW Control sample (two asymmetric btags)
DØ Internal (4.2 fb^-1)
Data
Top
V+h.f./VV
V+l.f.
VH × 10
Events / 1.00
Number of Taggable Jets

Analysis Sample before MJDT cut

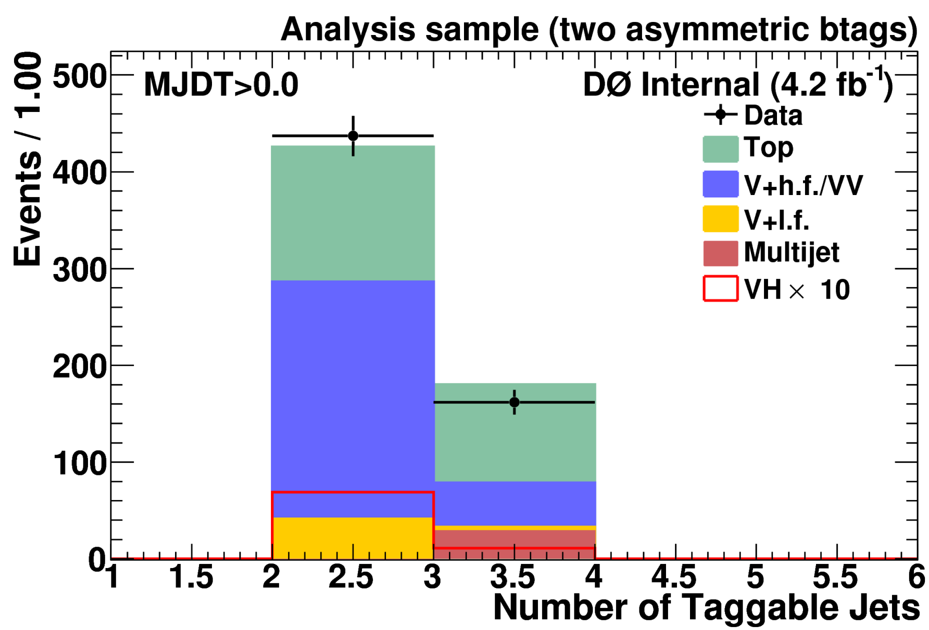


Analysis Sample after MJDT cut

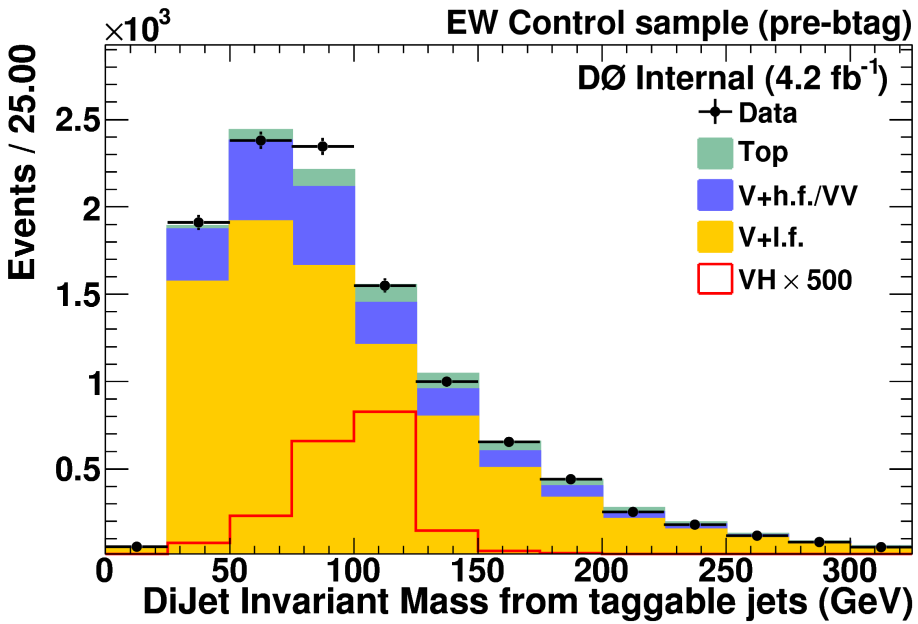




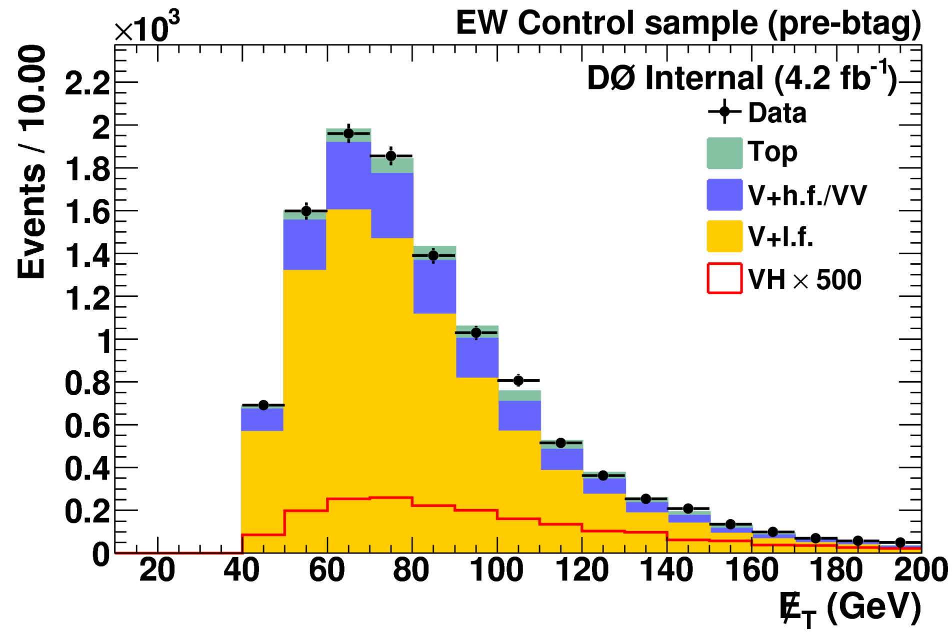
The agreement is fair.



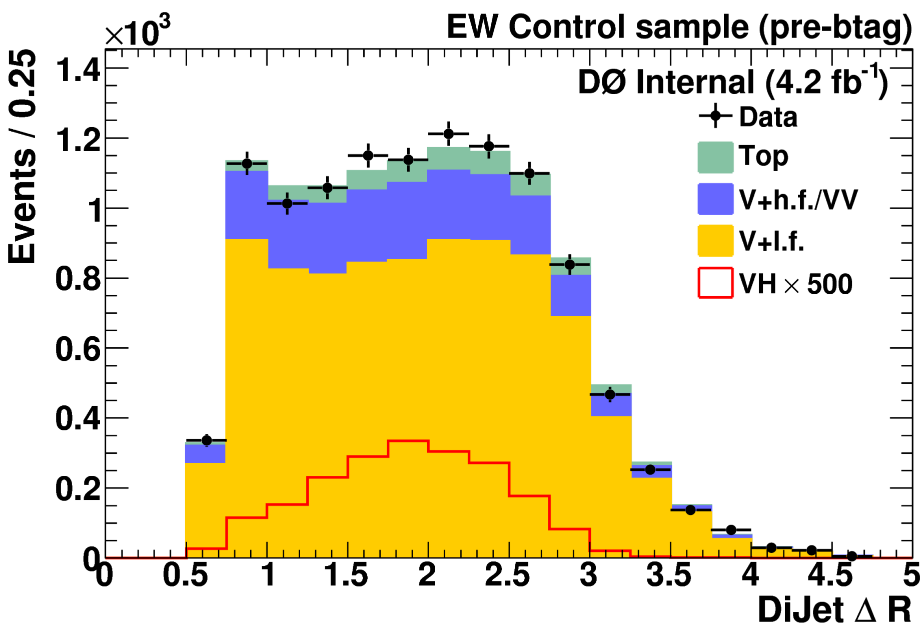
EW Sample



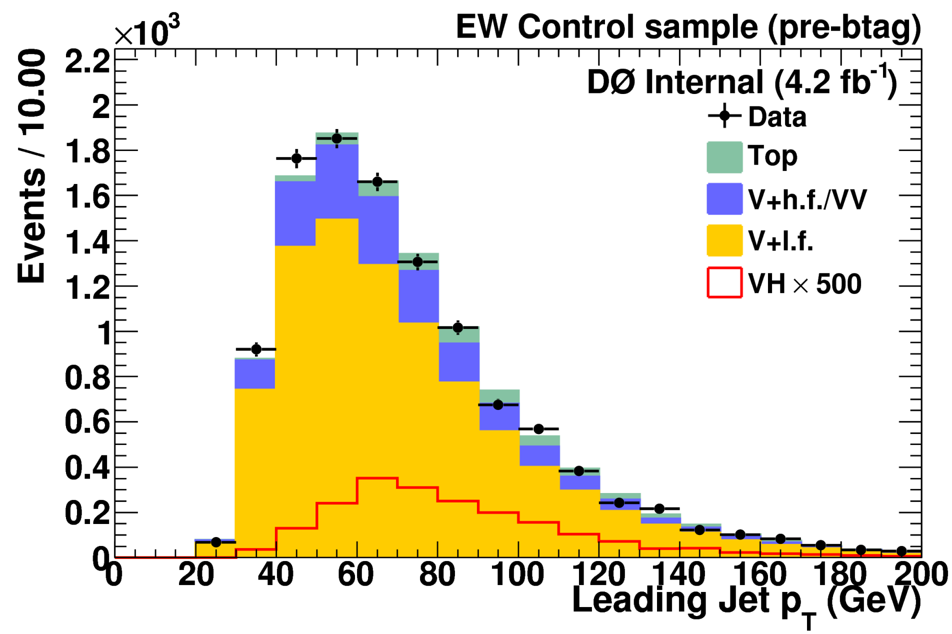
EW Control sample (pre-btag)
DØ Internal (4.2 fb^{-1})
Data
Top
V+h.f./VV
V+l.f.
VH × 500
Events / 25.00
×10^{3}
DiJet Invariant Mass from taggable jets (GeV)



EW Control sample (pre-btag)
DØ Internal (4.2 fb^{-1})
Data
Top
V+h.f./VV
V+l.f.
VH × 500
Events / 10.00
×10^{3}
$\not{E}_T$ (GeV)

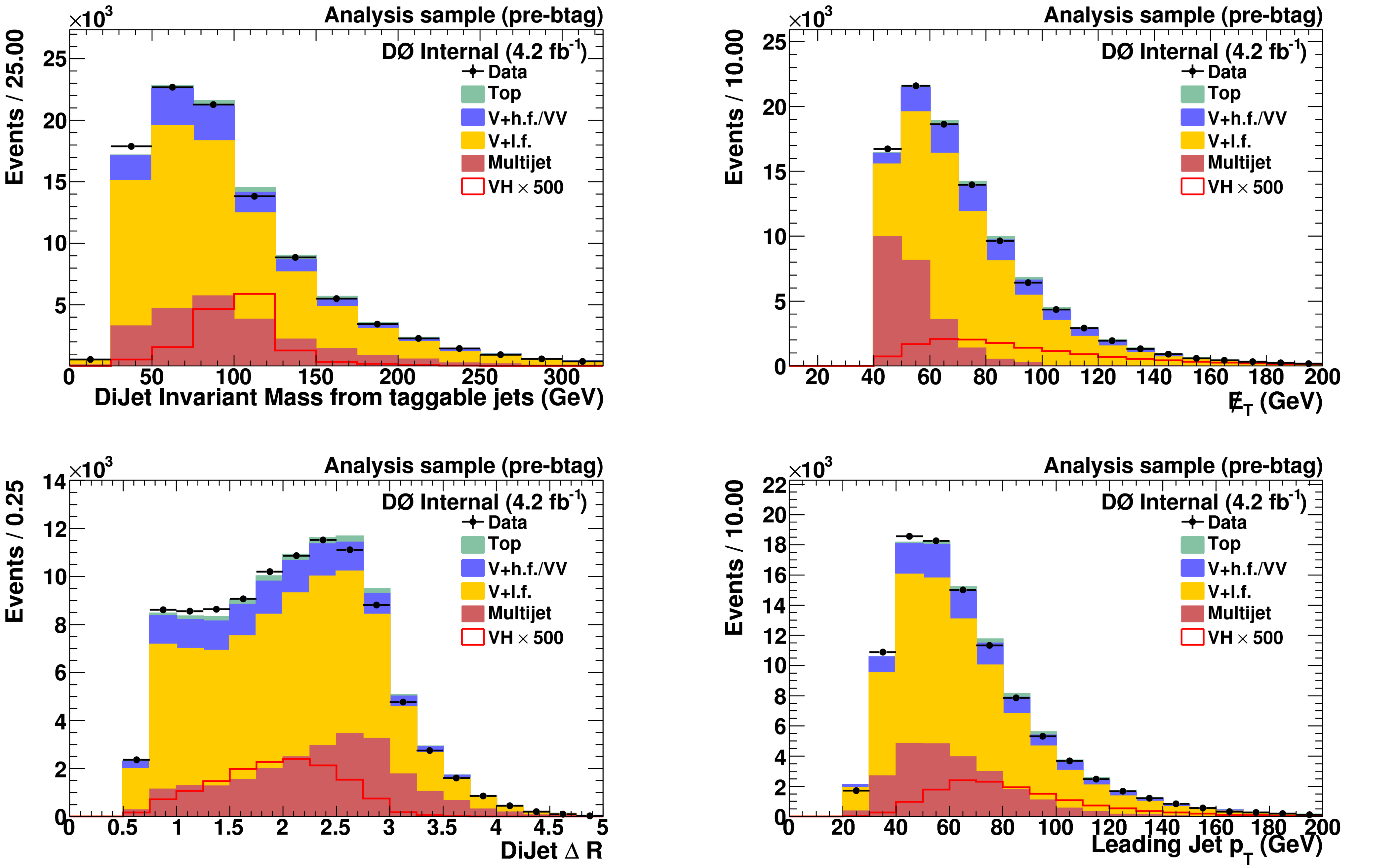


EW Control sample (pre-btag)
DØ Internal (4.2 fb^{-1})
Data
Top
V+h.f./VV
V+l.f.
VH × 500
Events / 0.25
×10^{3}
DiJet Δ R

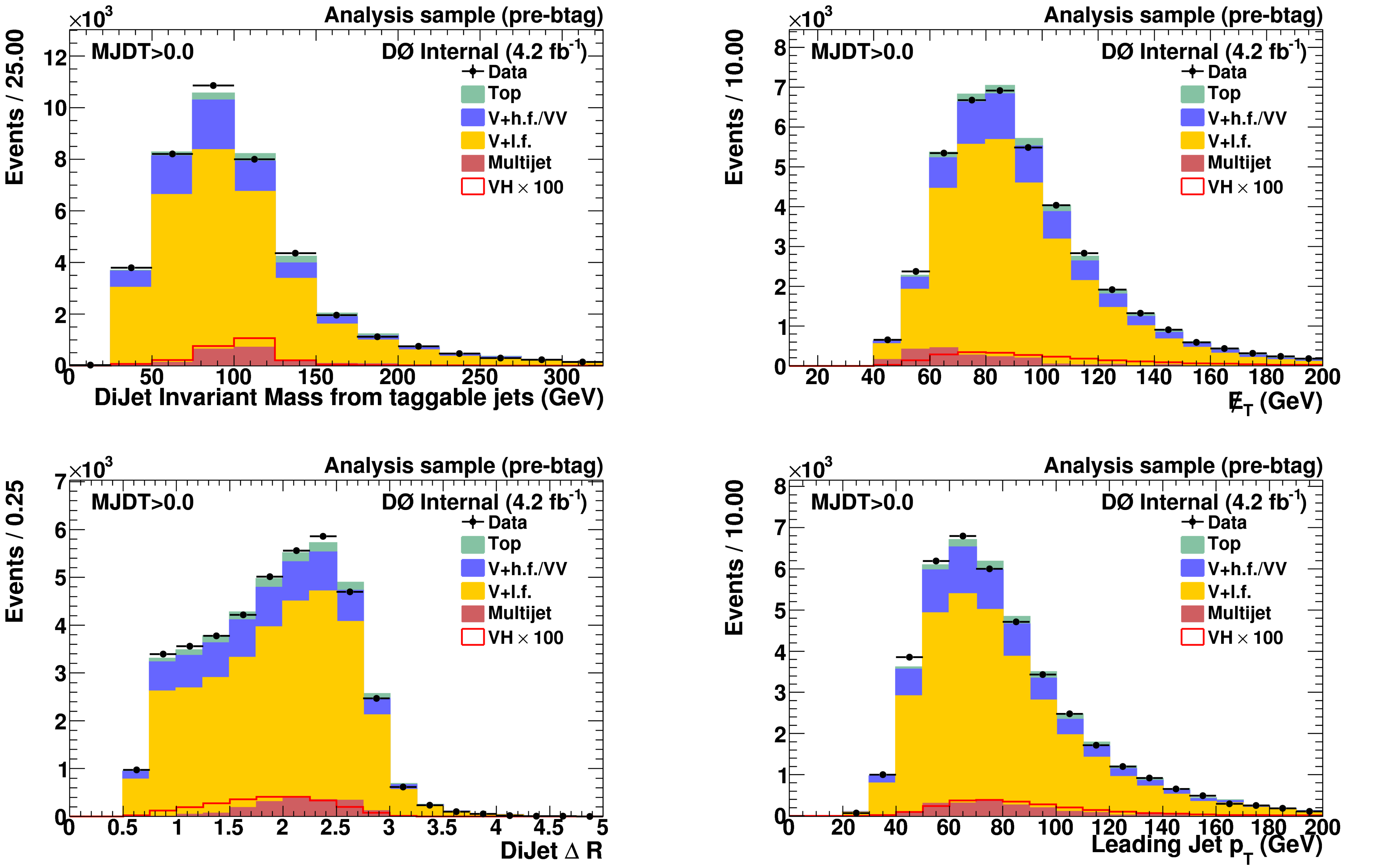


EW Control sample (pre-btag)
DØ Internal (4.2 fb^{-1})
Data
Top
V+h.f./VV
V+l.f.
VH × 500
Events / 10.00
×10^{3}
Leading Jet p_T (GeV)

Analysis Sample before MJDT cut



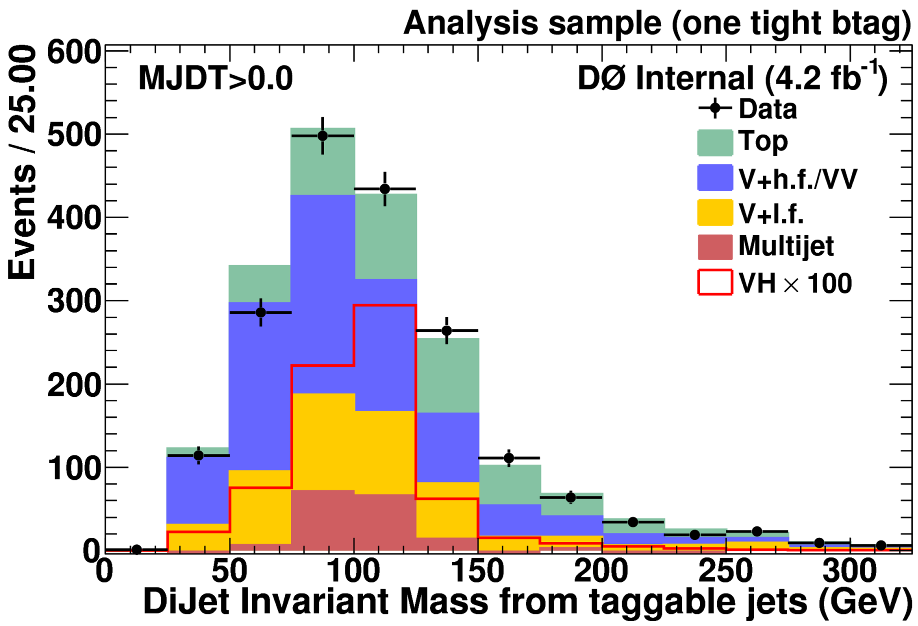
Analysis Sample after MJDT cut

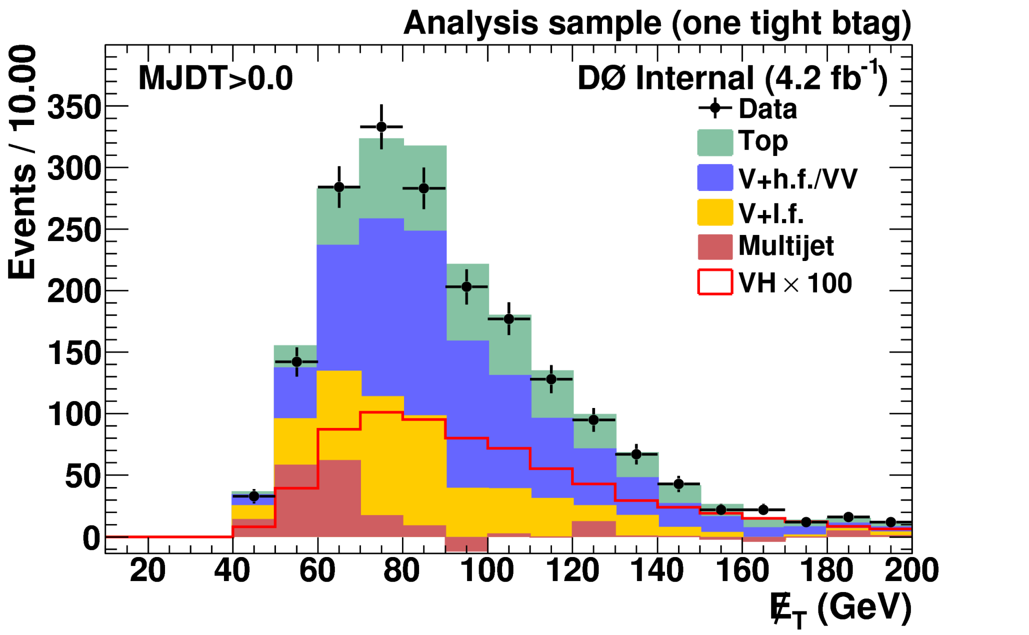


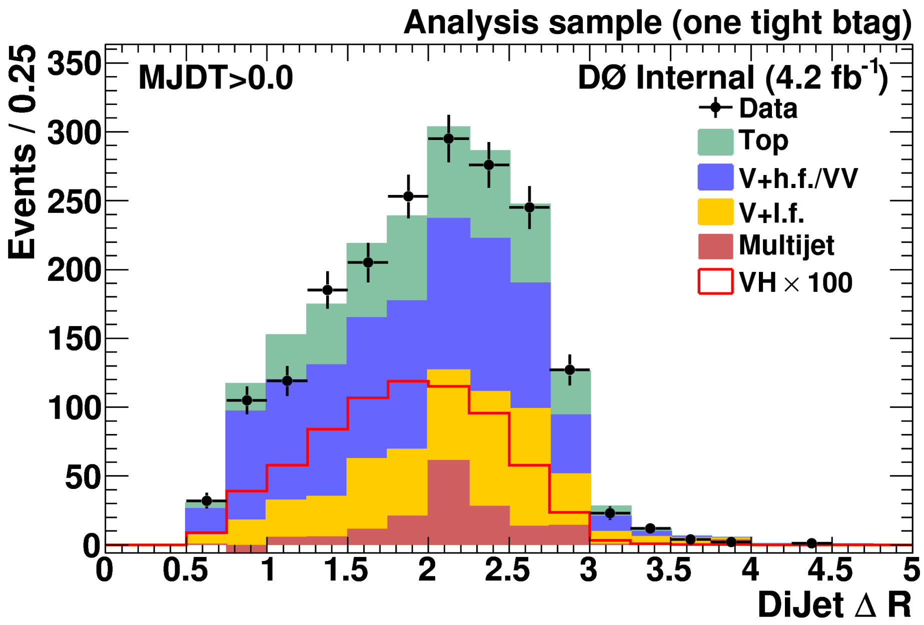
D0 France 2010-May-03

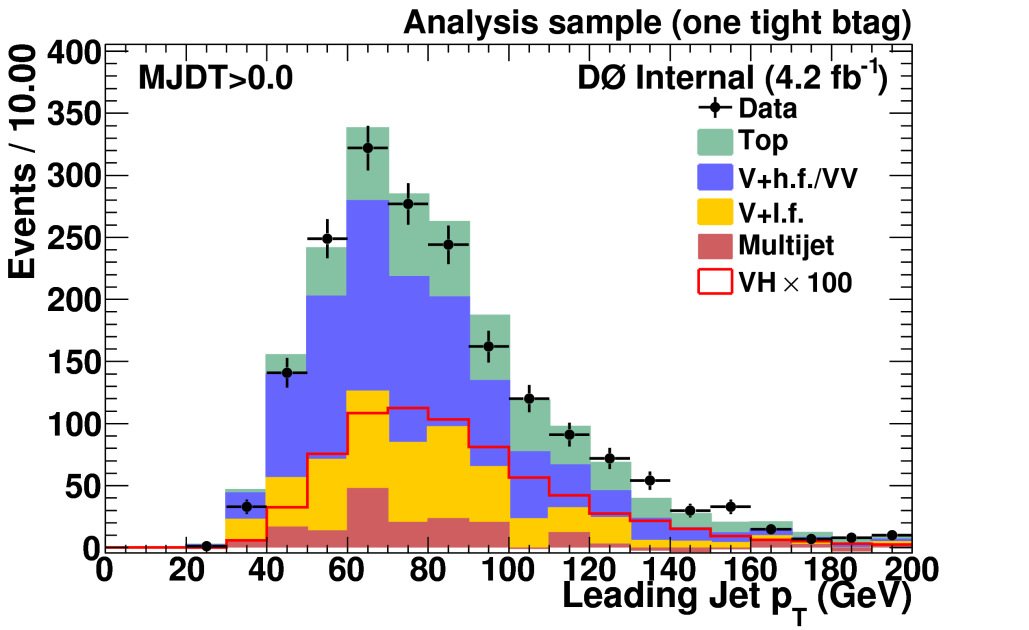
M. Rangel

Analysis Sample after MJDT cut

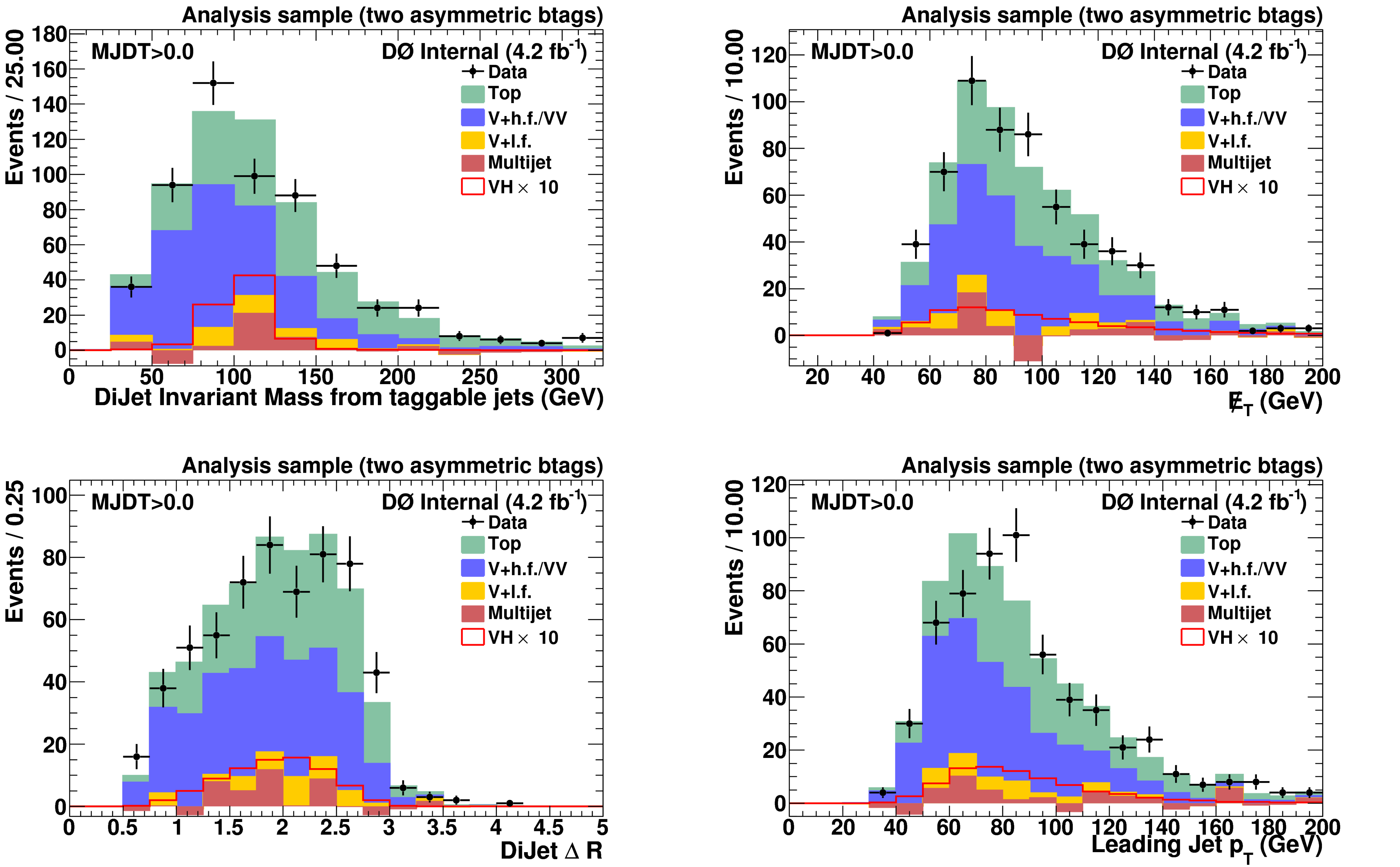




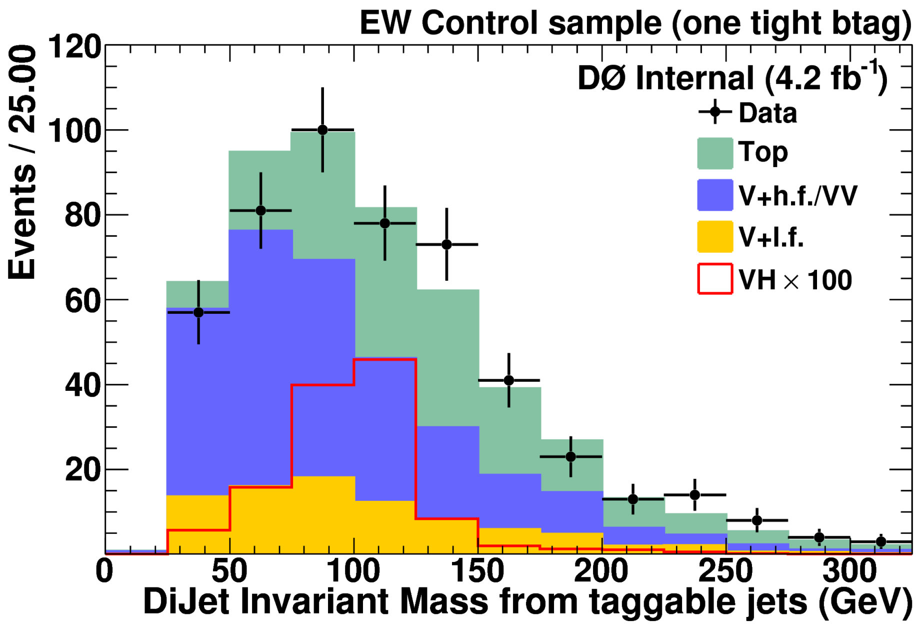


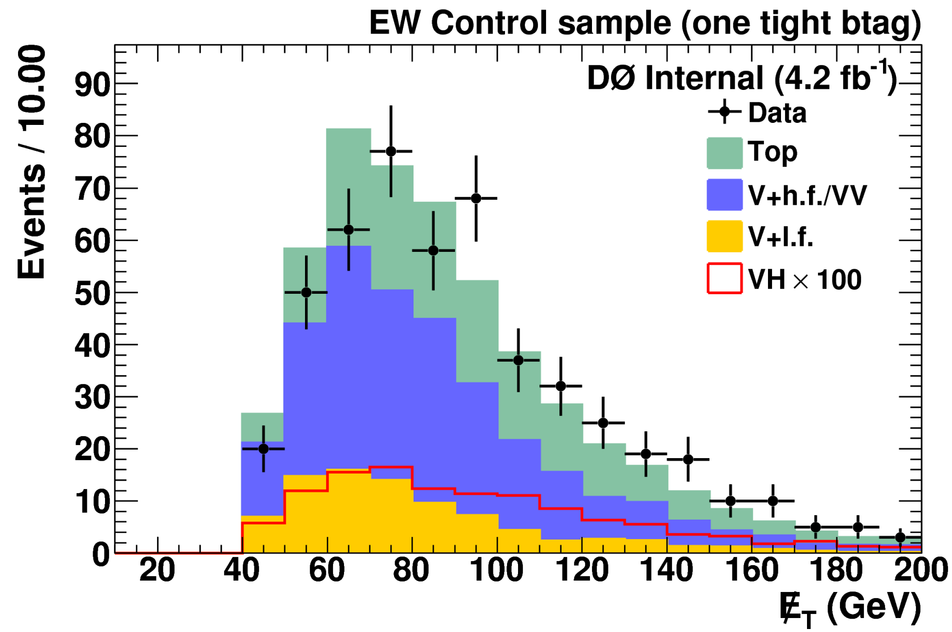


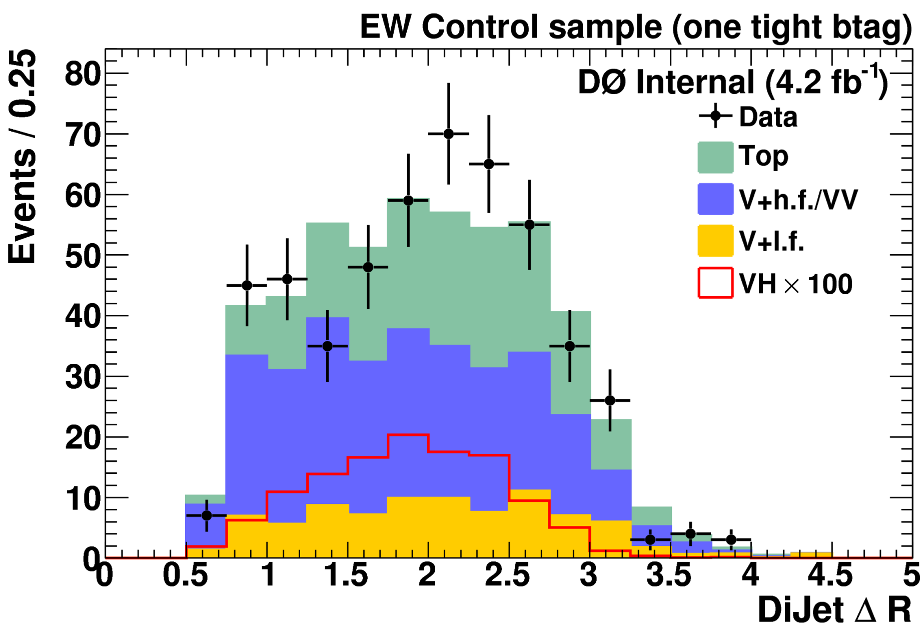
Analysis Sample after MJDT cut

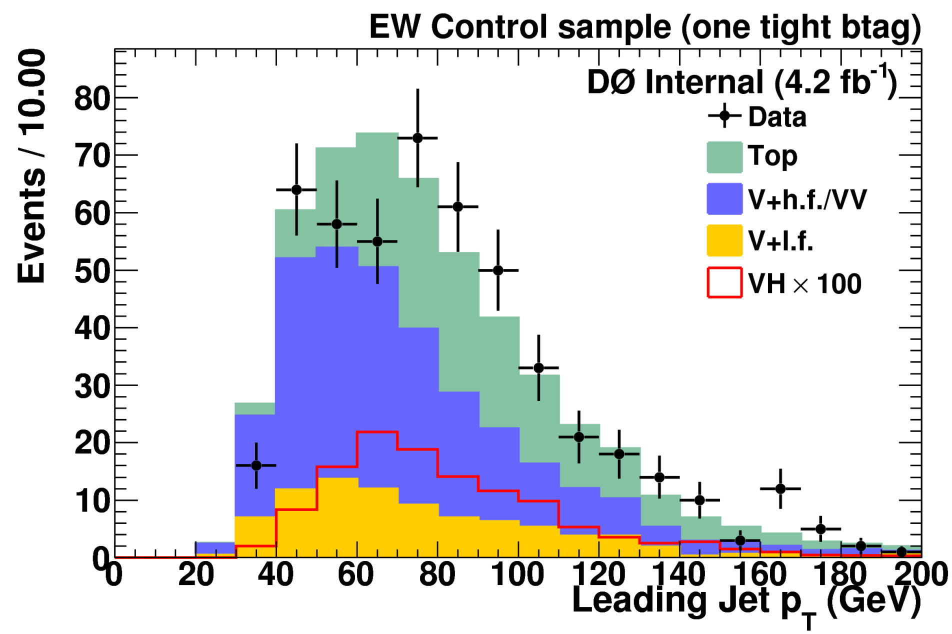


EW Sample with SFs

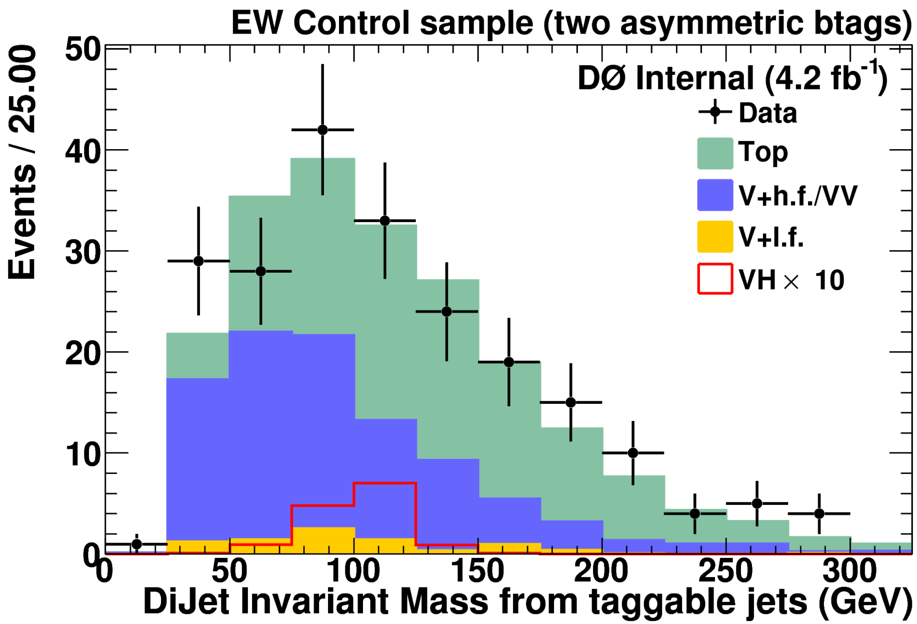


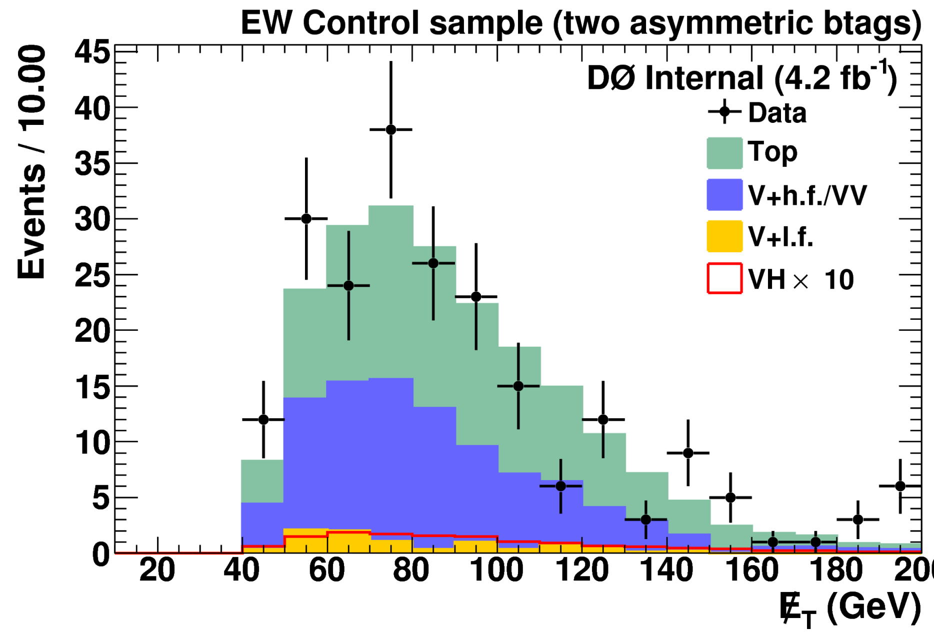


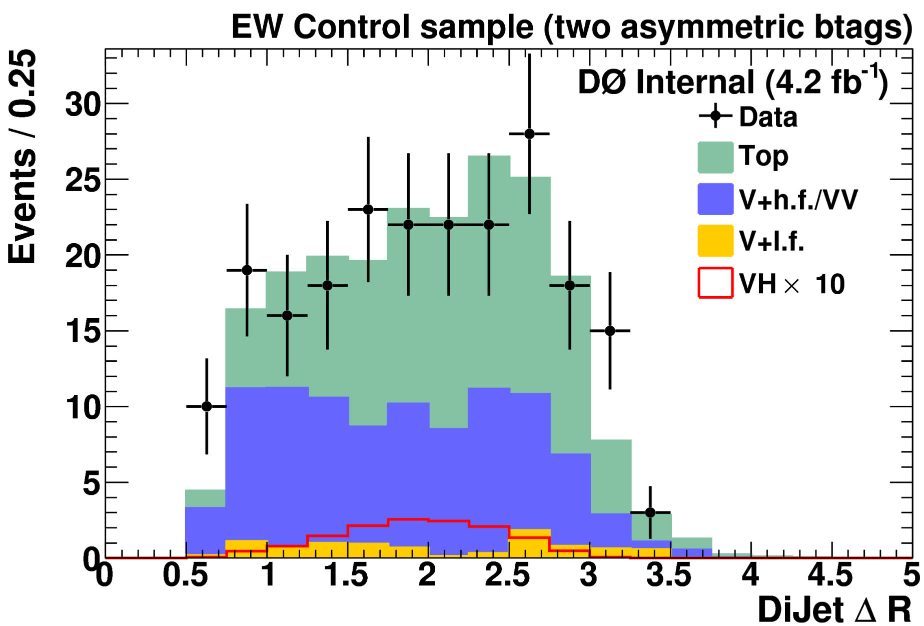


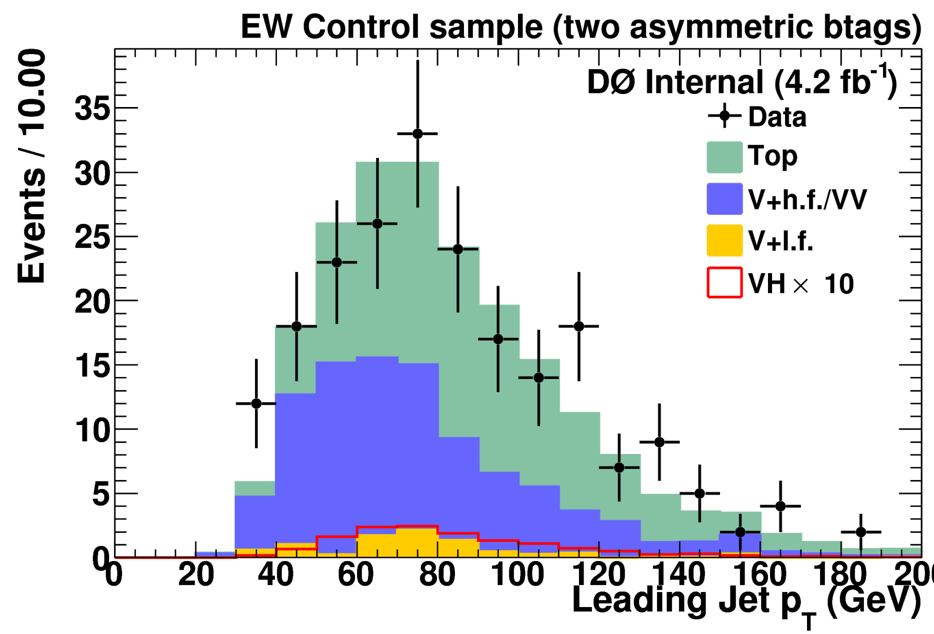


EW Sample with SFs

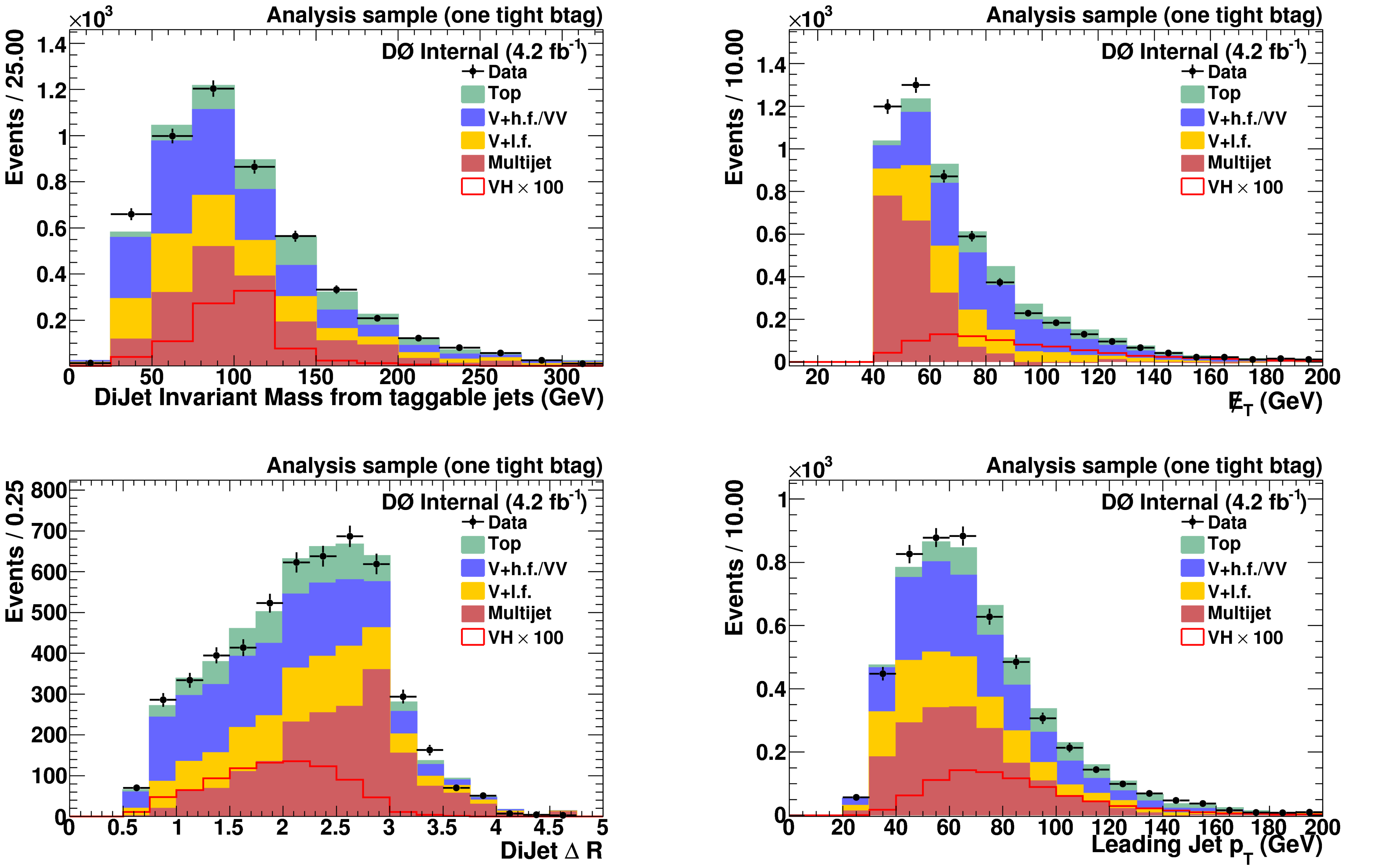




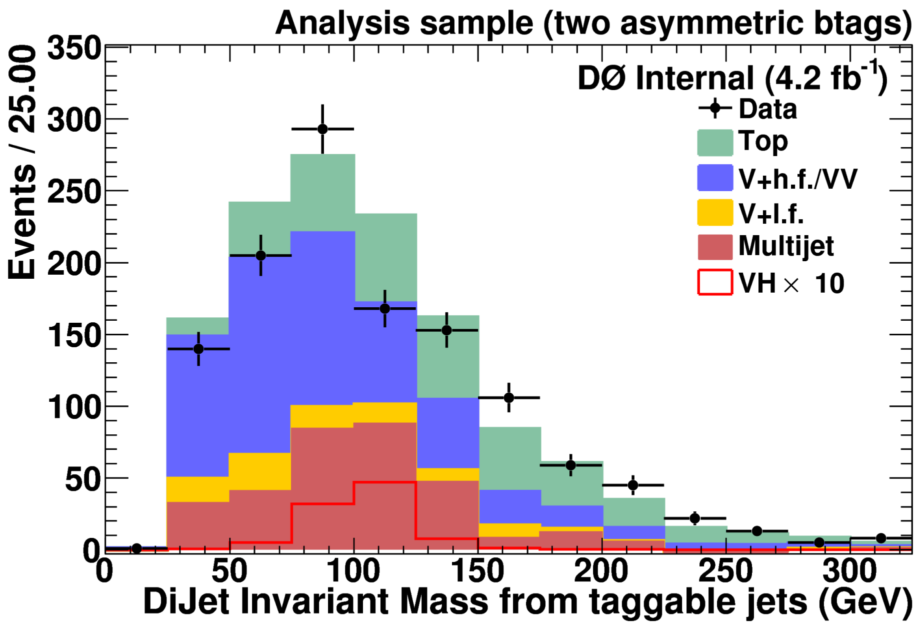


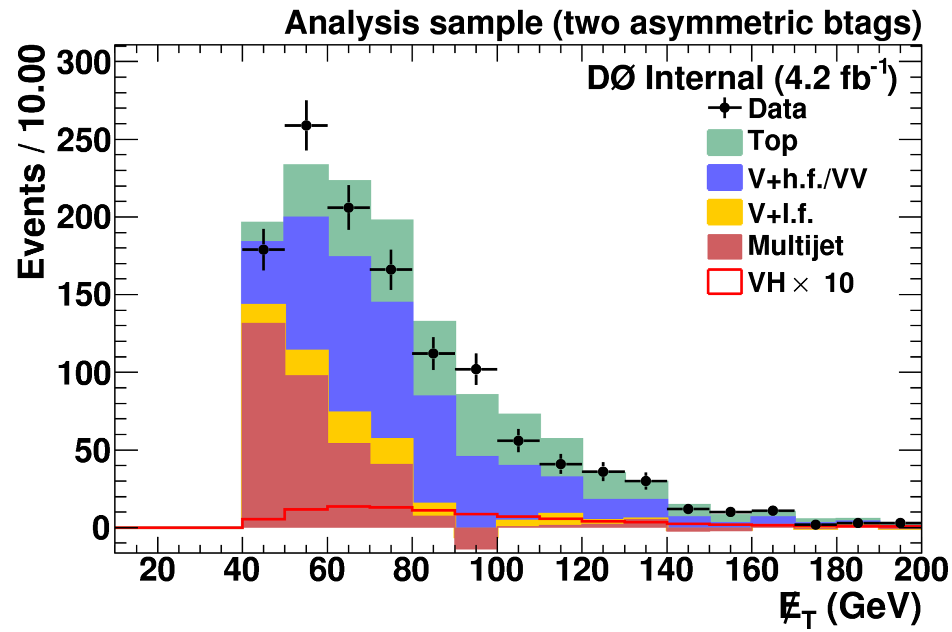


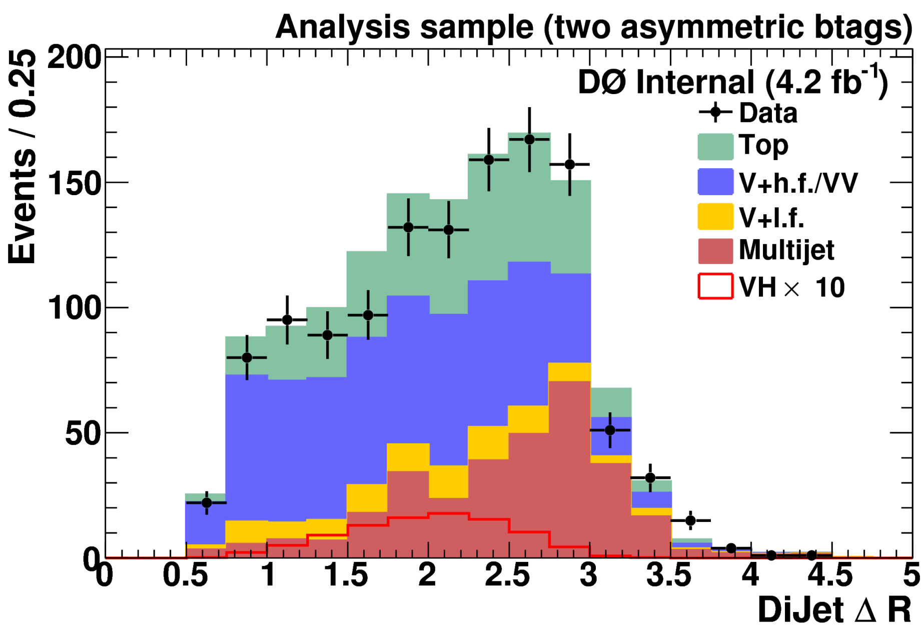
Analysis Sample before MJDT cut with SFs

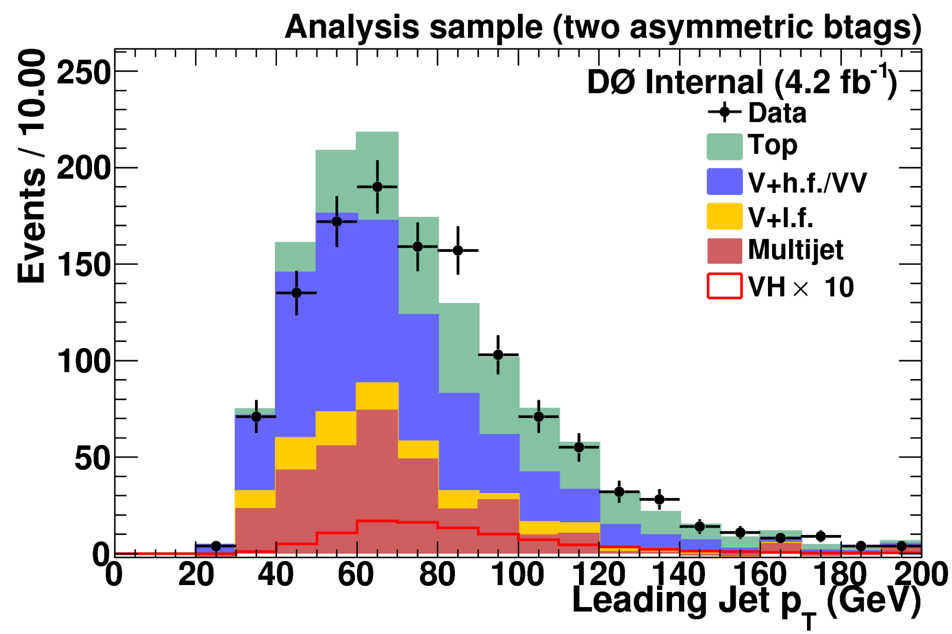


Analysis Sample before MJDT cut with SFs







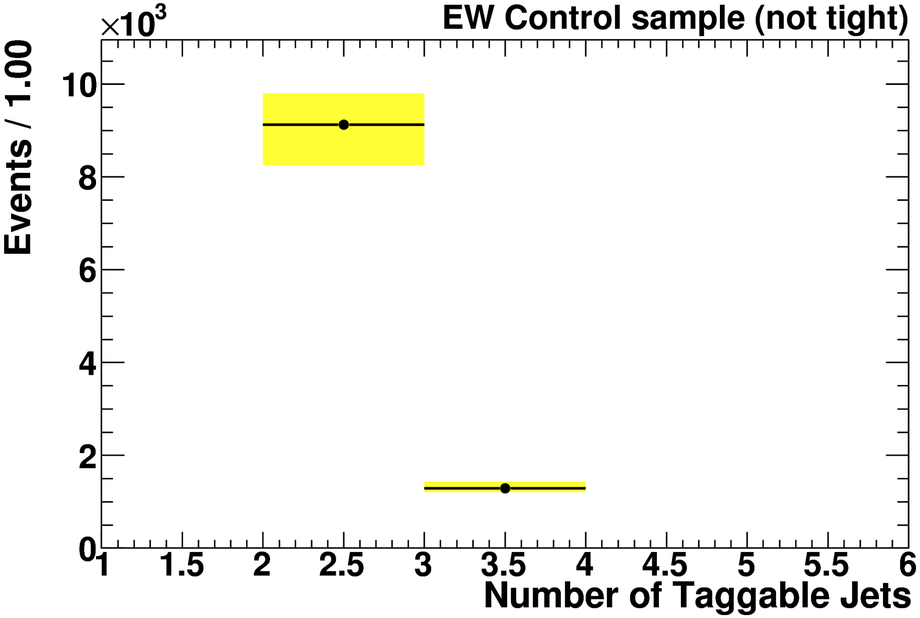


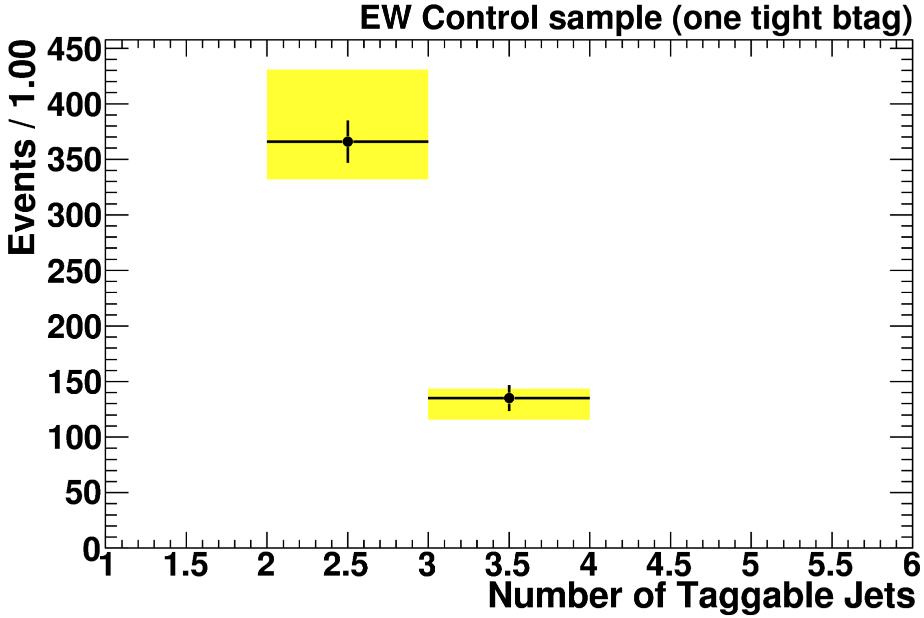
Normalization Systematics

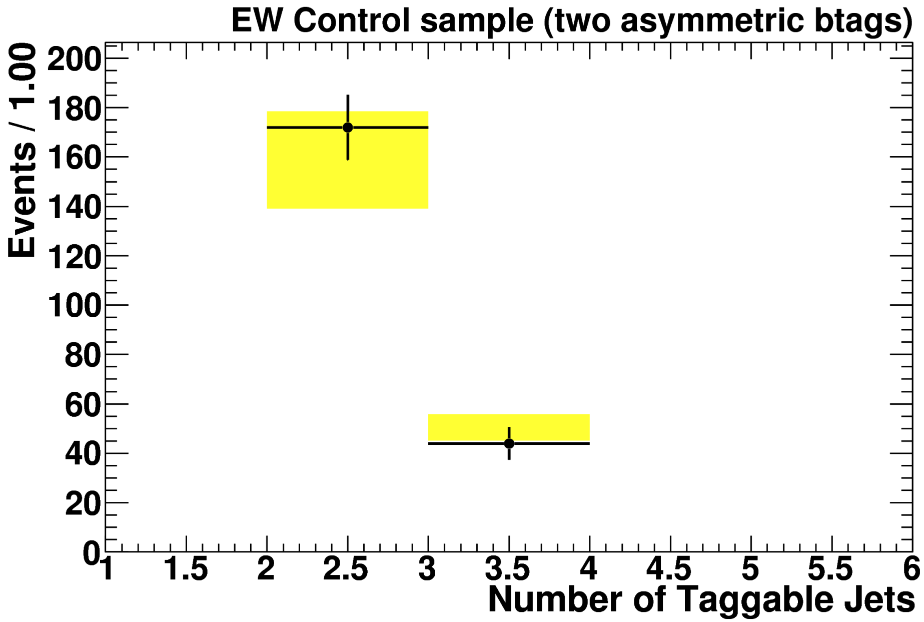
From this slide on, we show data vs. background with the following normalization uncertainties:

- 6% Luminosity
- 6% V+jets
- 20% V+heavy flavor
- 10% Top

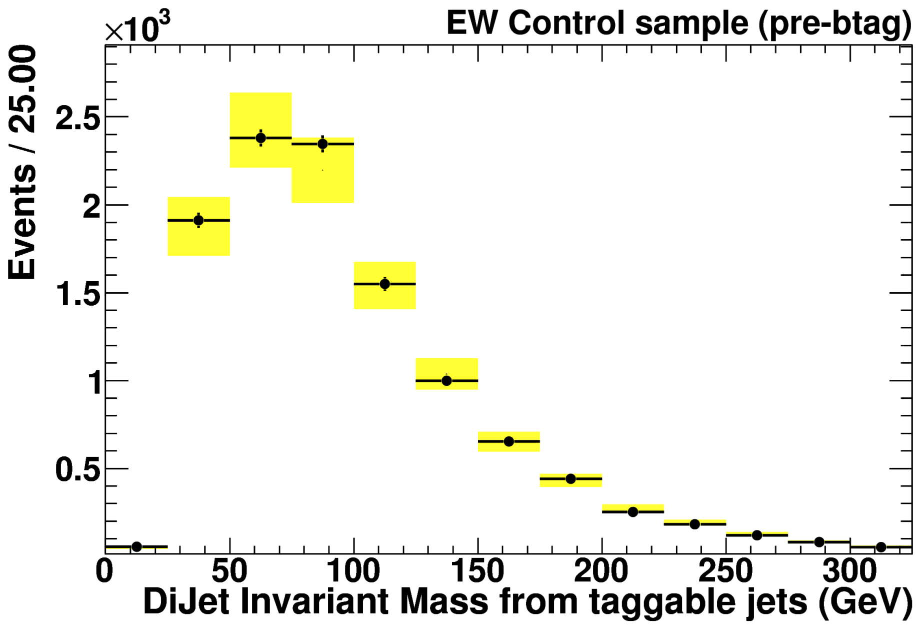
EW Sample (ALL SF = 1)

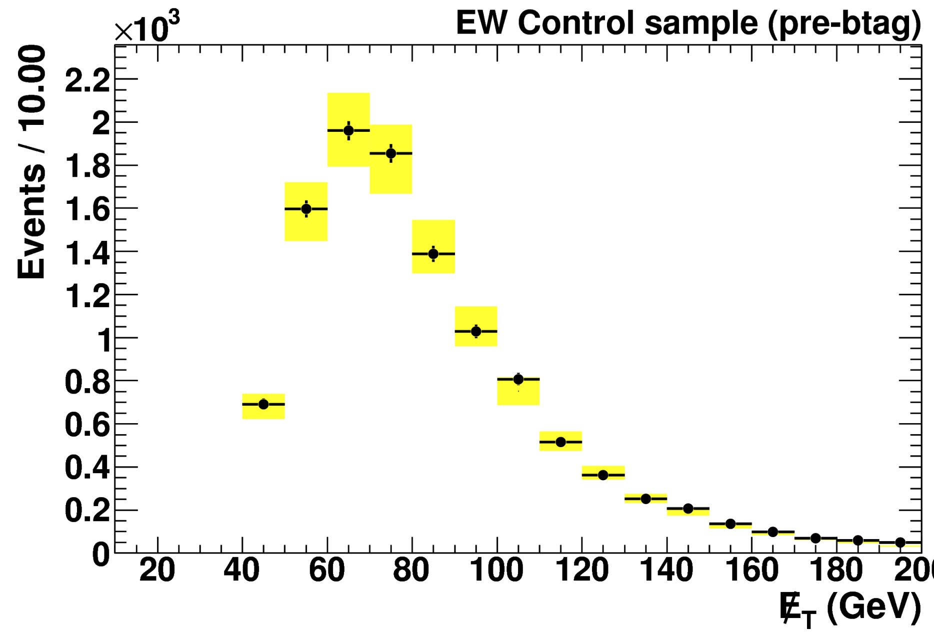


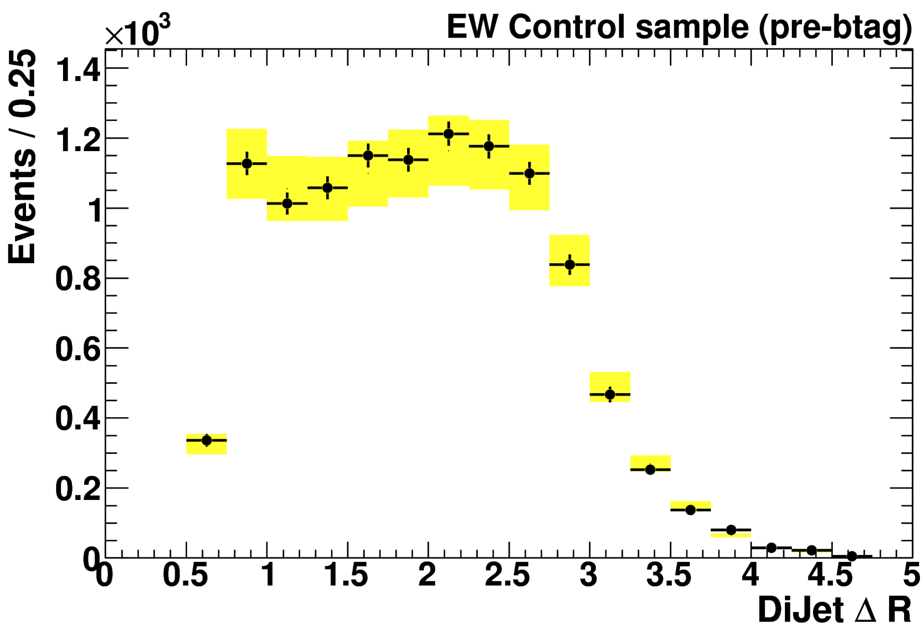


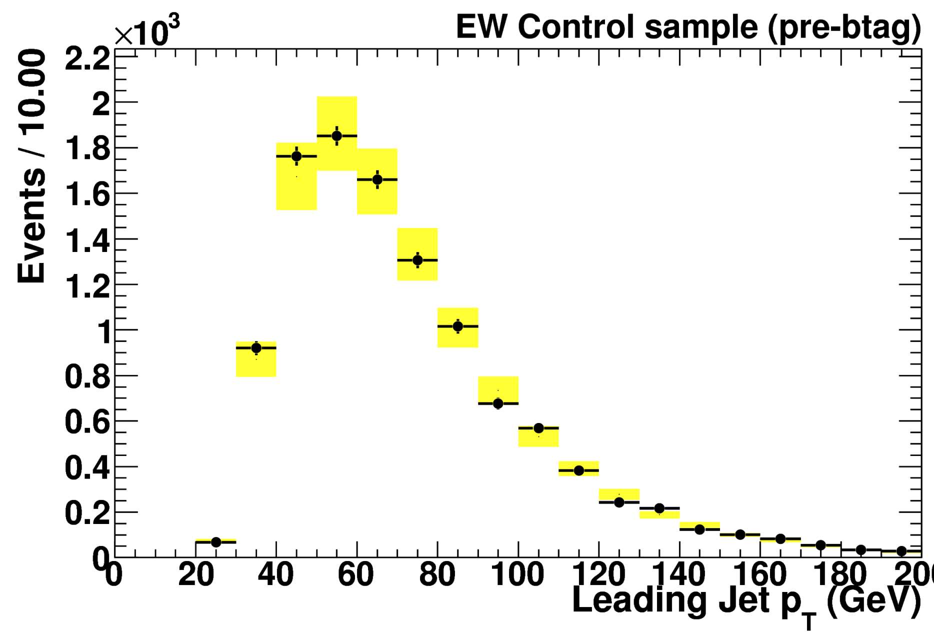


EW Sample (ALL SF=1)

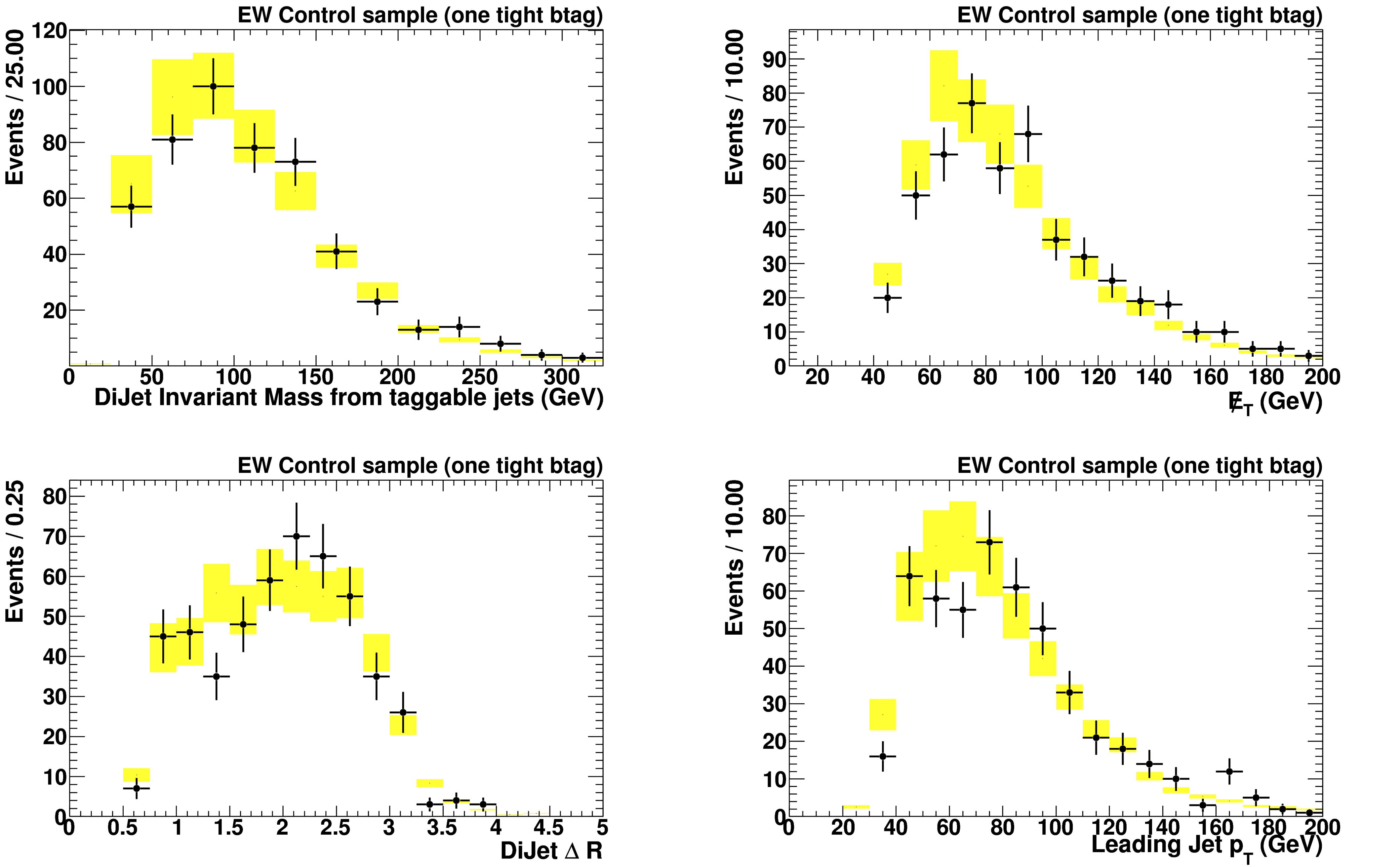








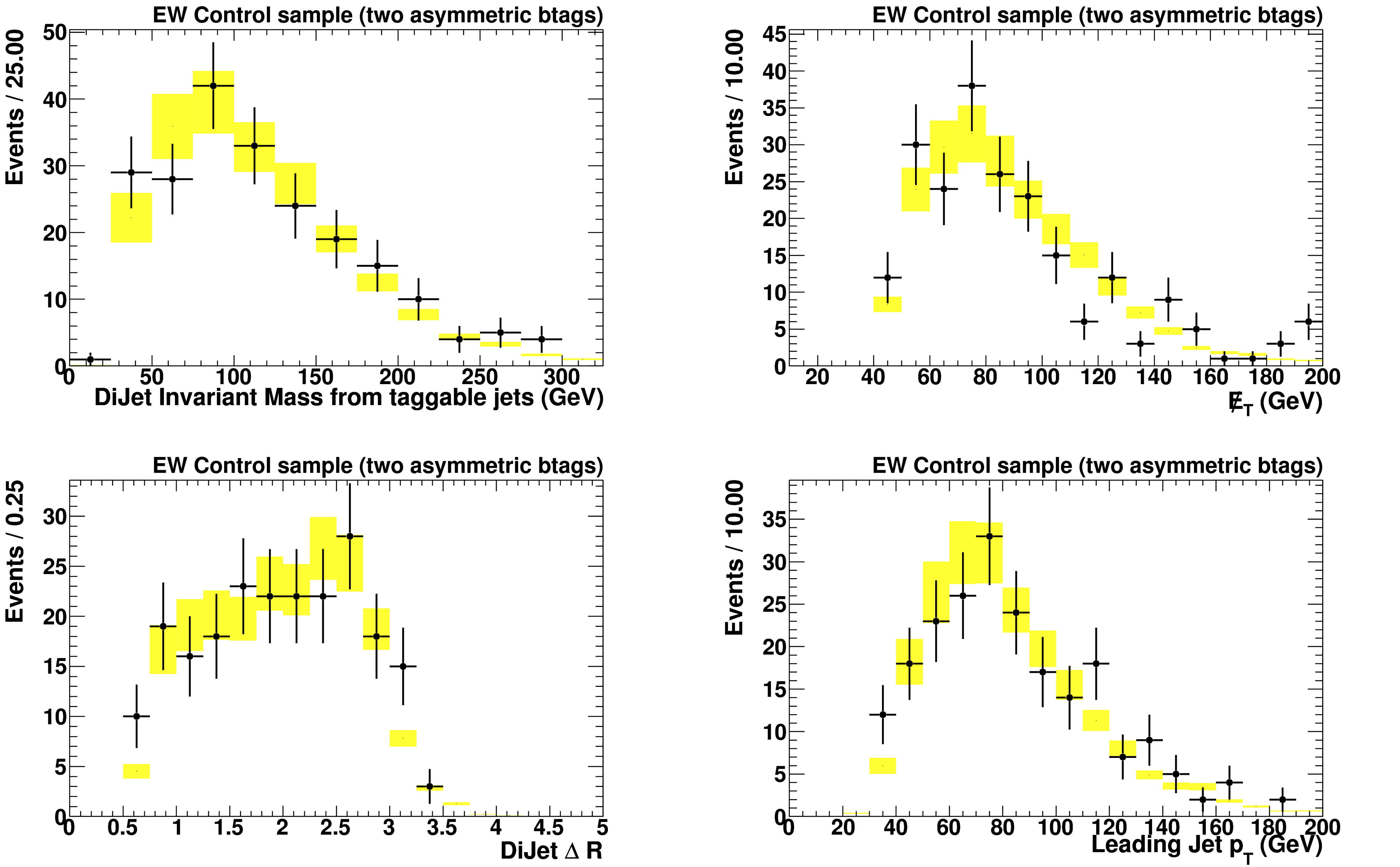
EW Sample (ALL SF=1)



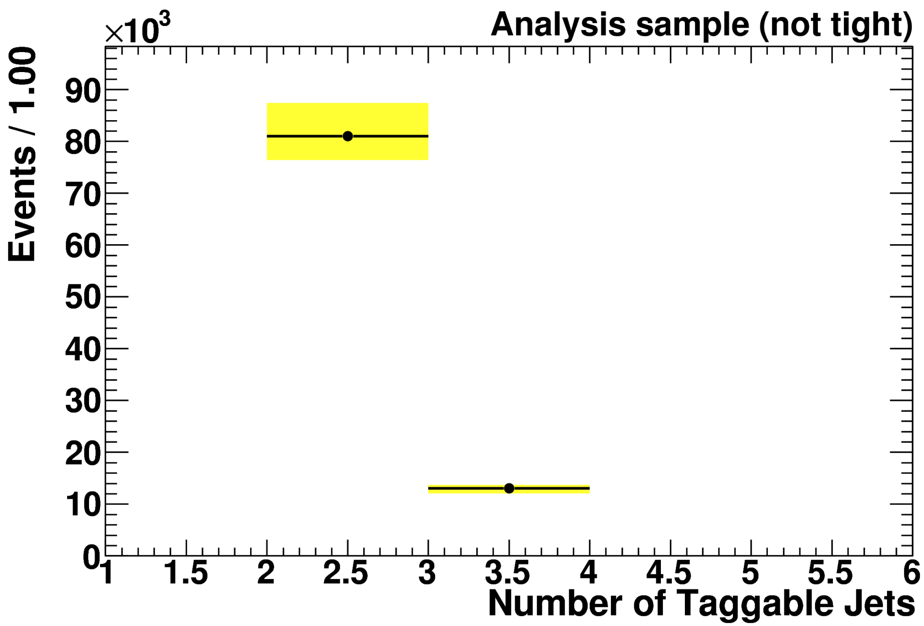
D0 France 2010-May-03

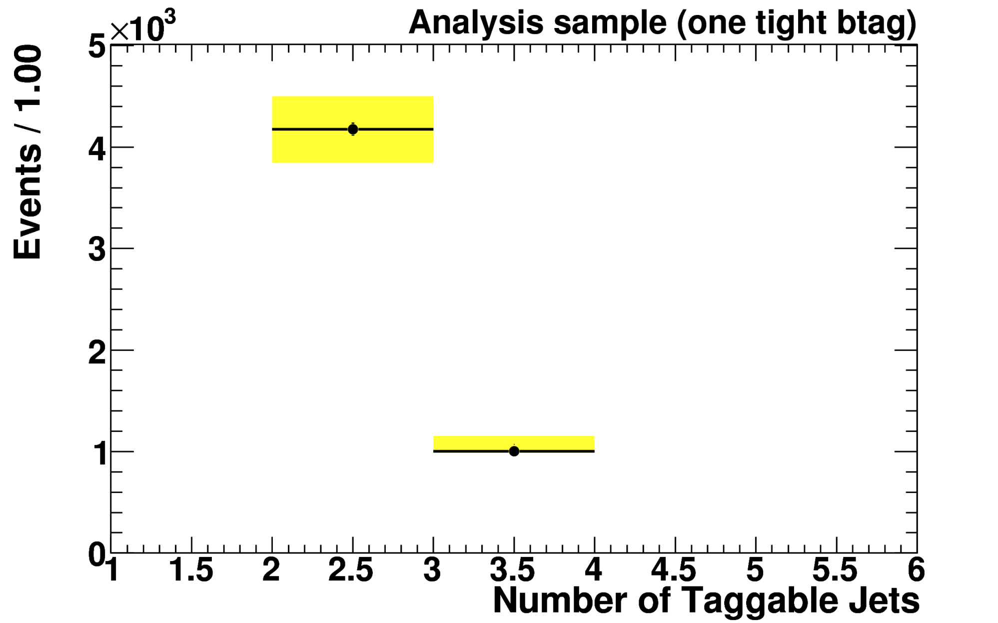
M. Rangel

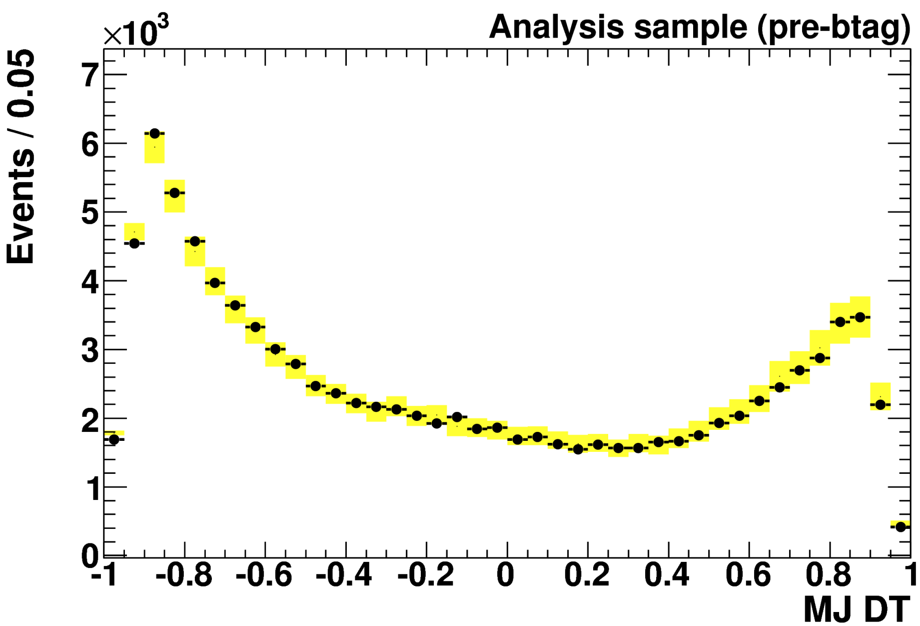
EW Sample (ALL SF=1)

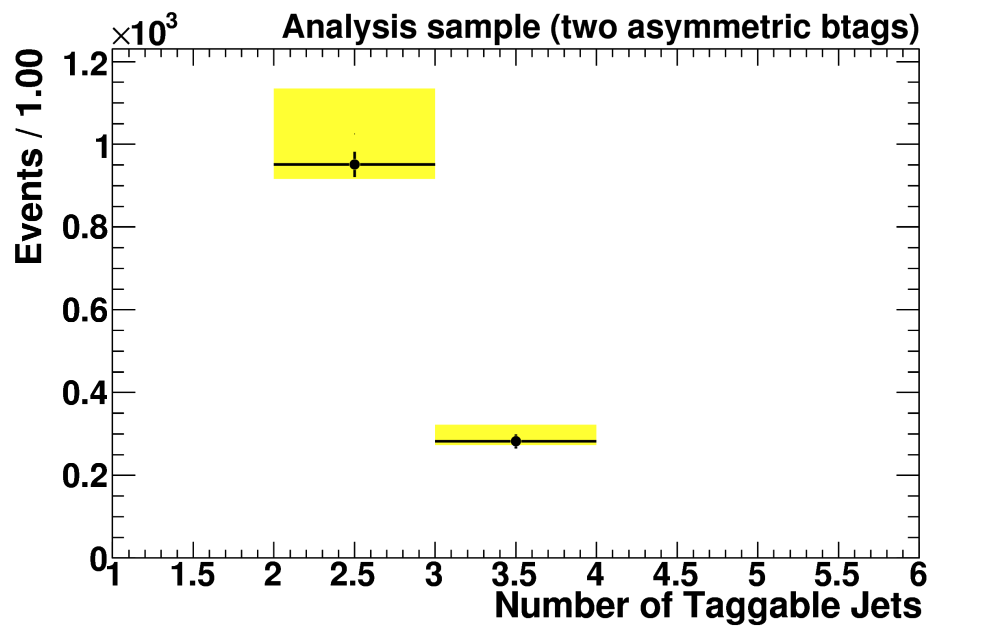


Analysis Sample before MJDT cut (ALL SF = 1)

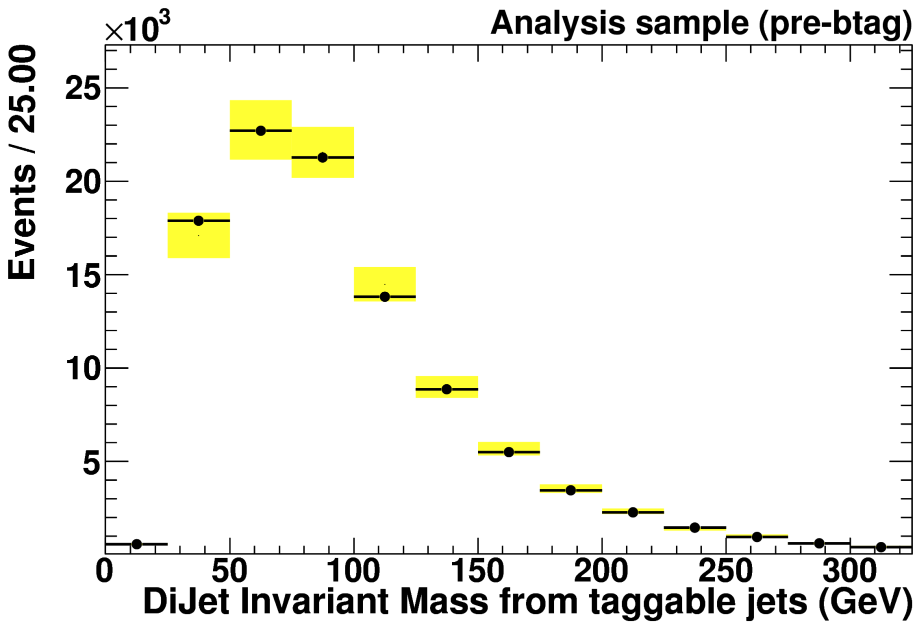


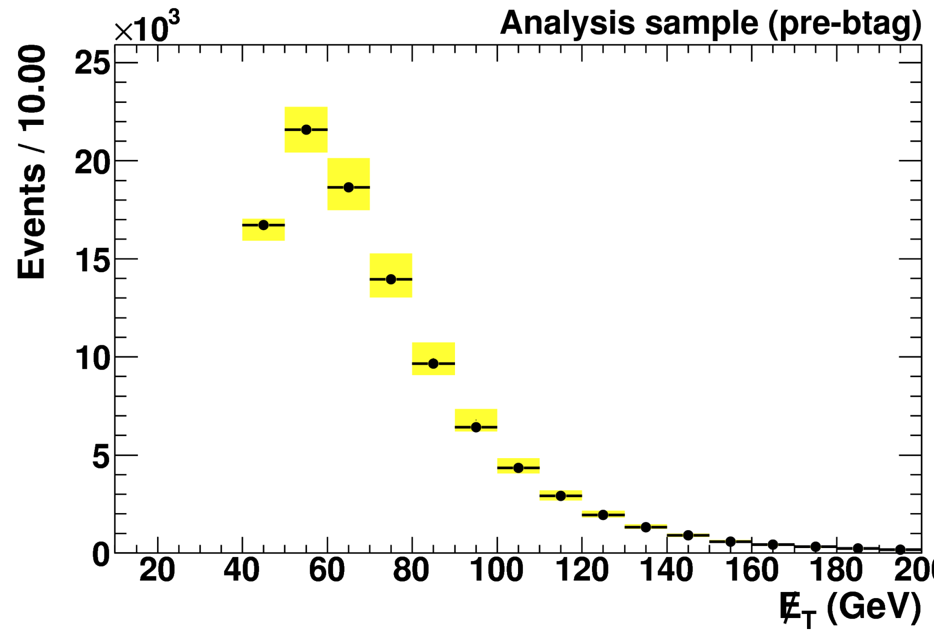


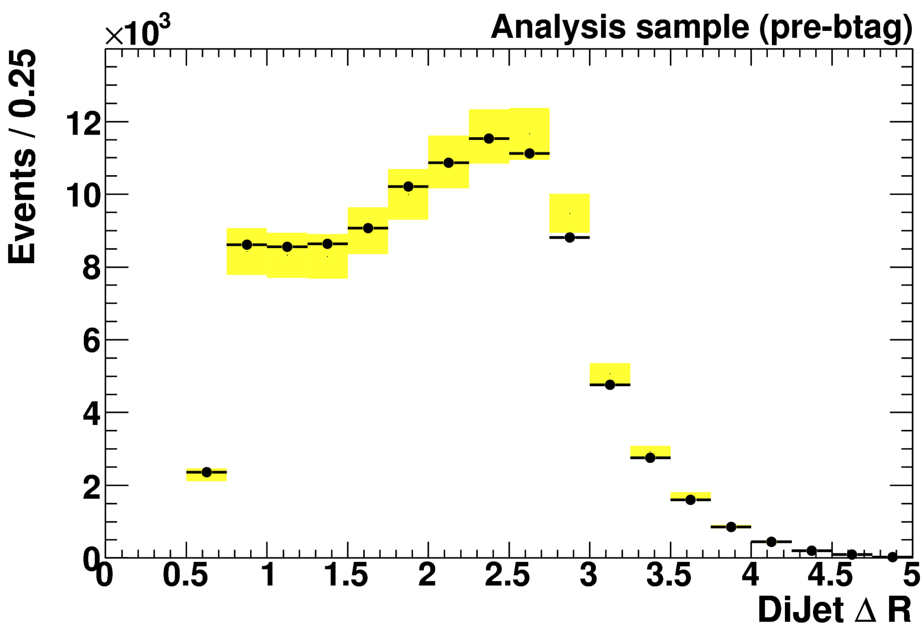


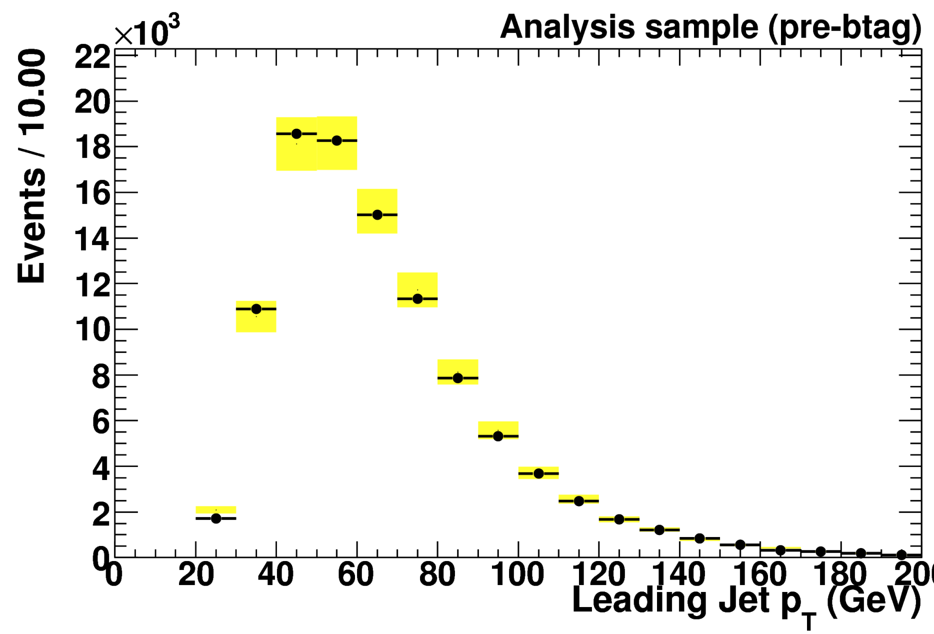


Analysis Sample before MJDT cut (ALL SF = 1)

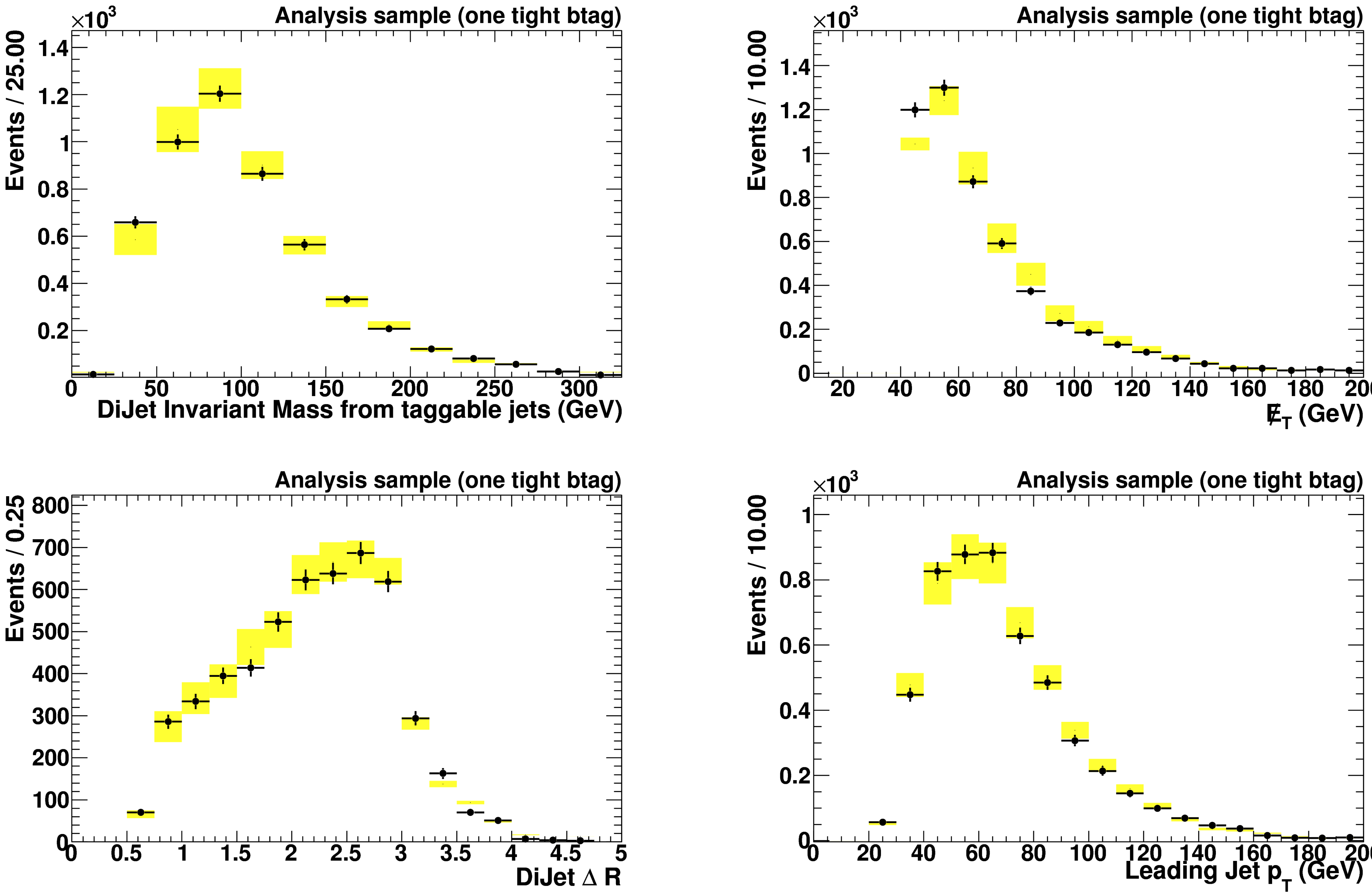




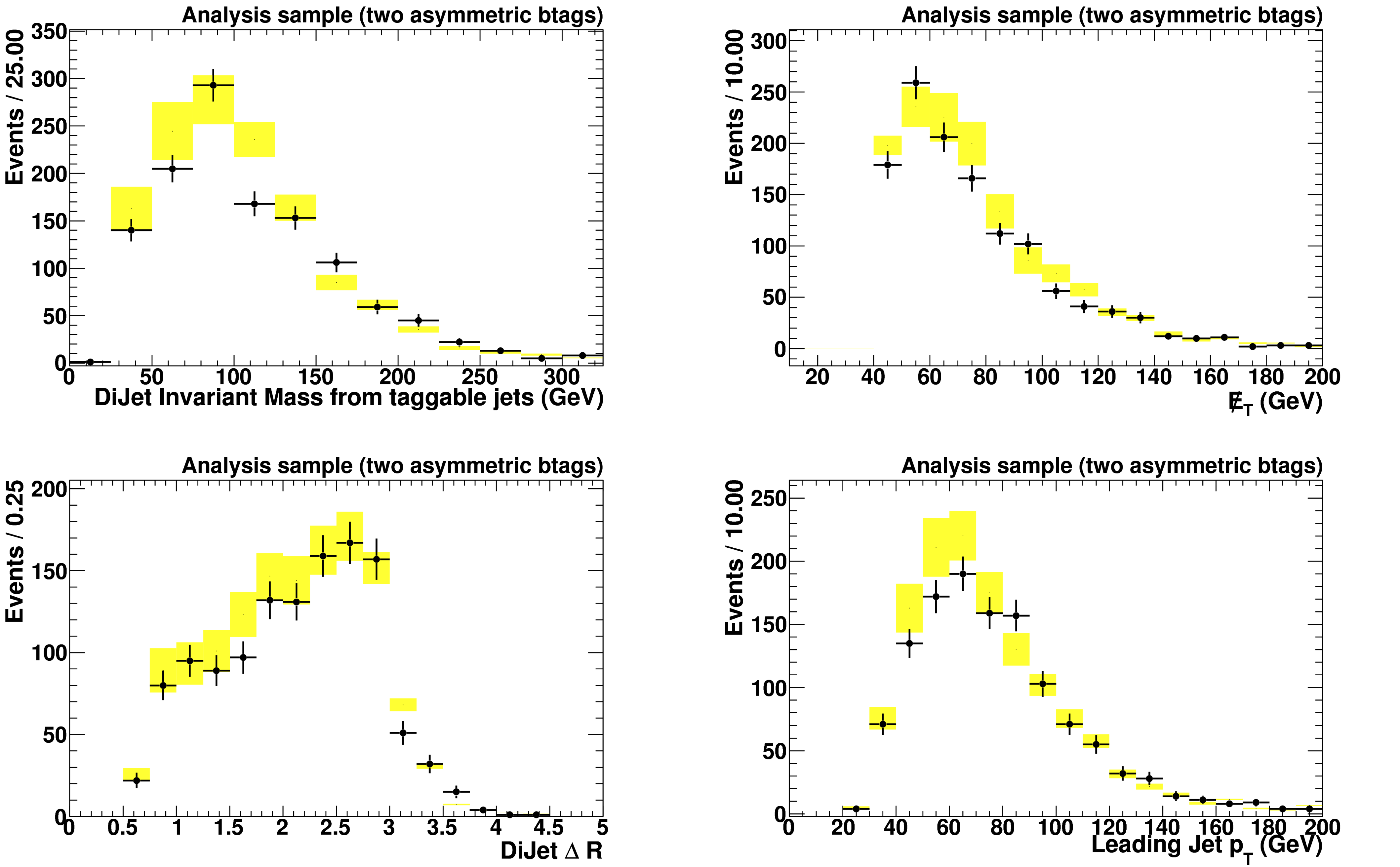




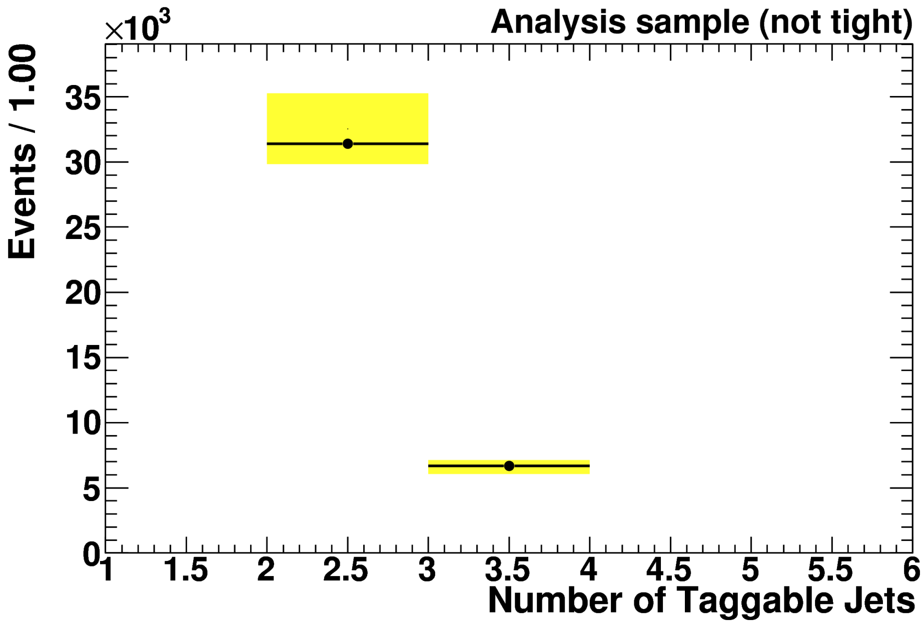
Analysis Sample before MJDT cut (ALL SF = 1)

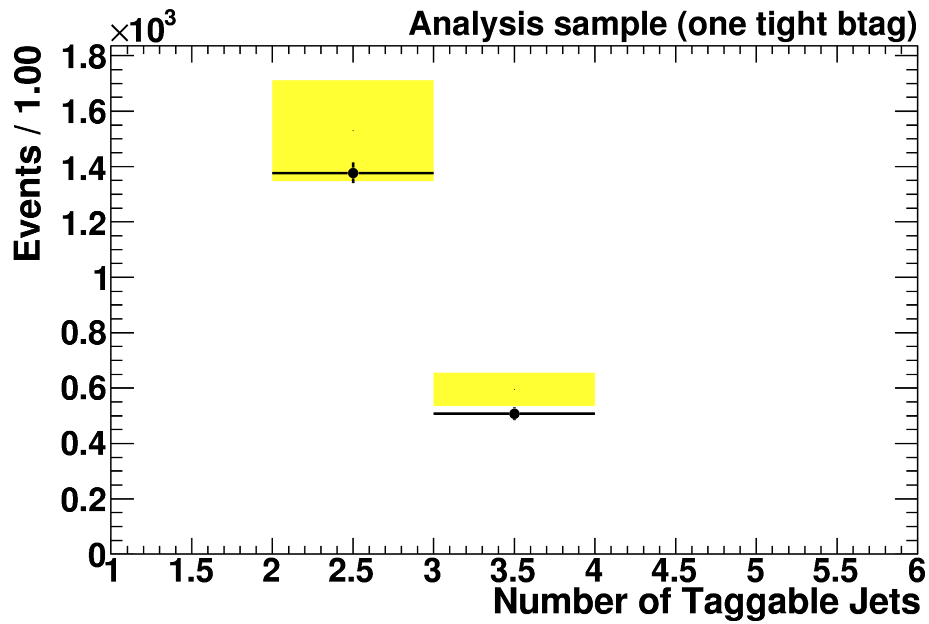


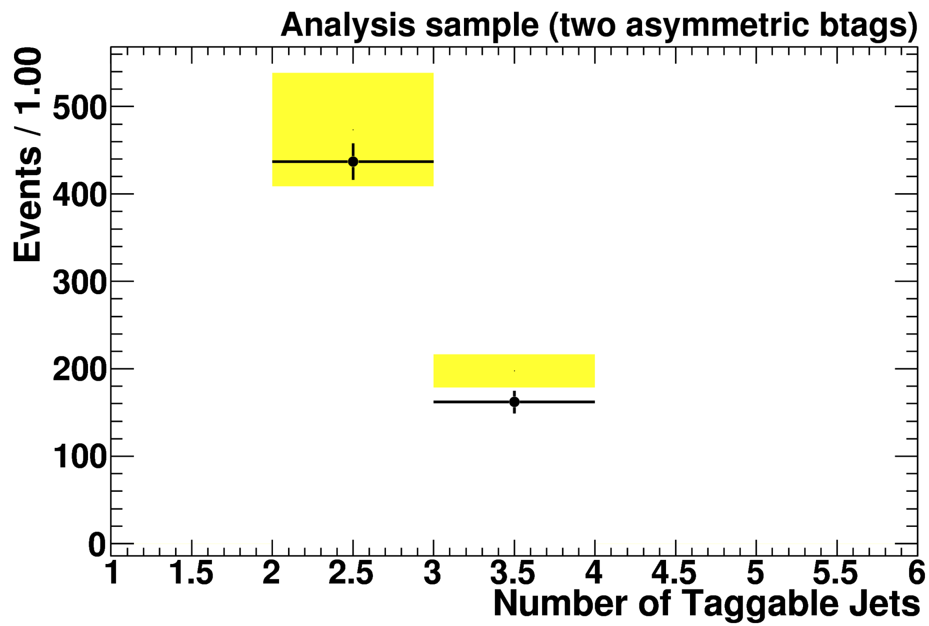
Analysis Sample before MJDT cut (ALL SF = 1)



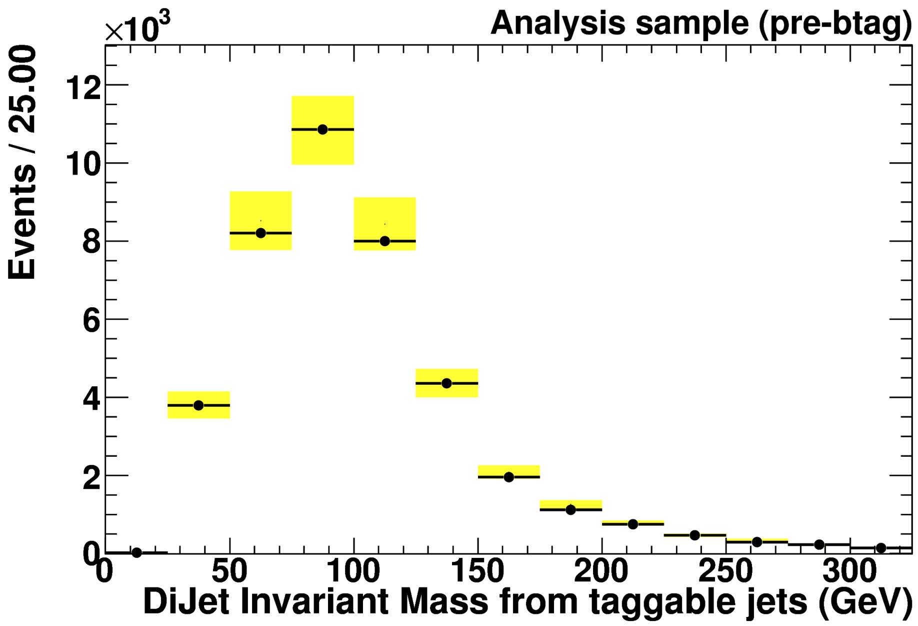
Analysis Sample after MJDT cut (ALL SF = 1)

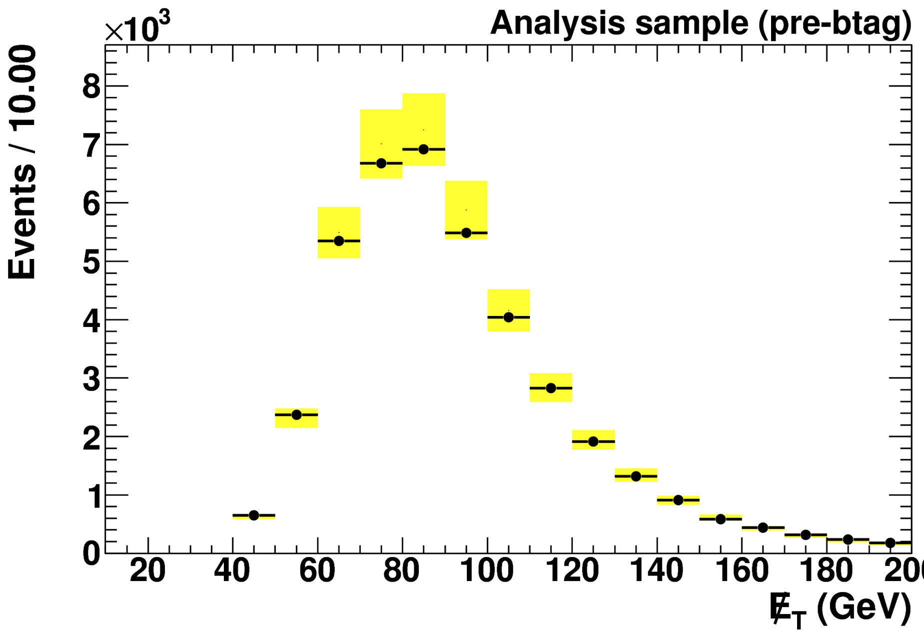


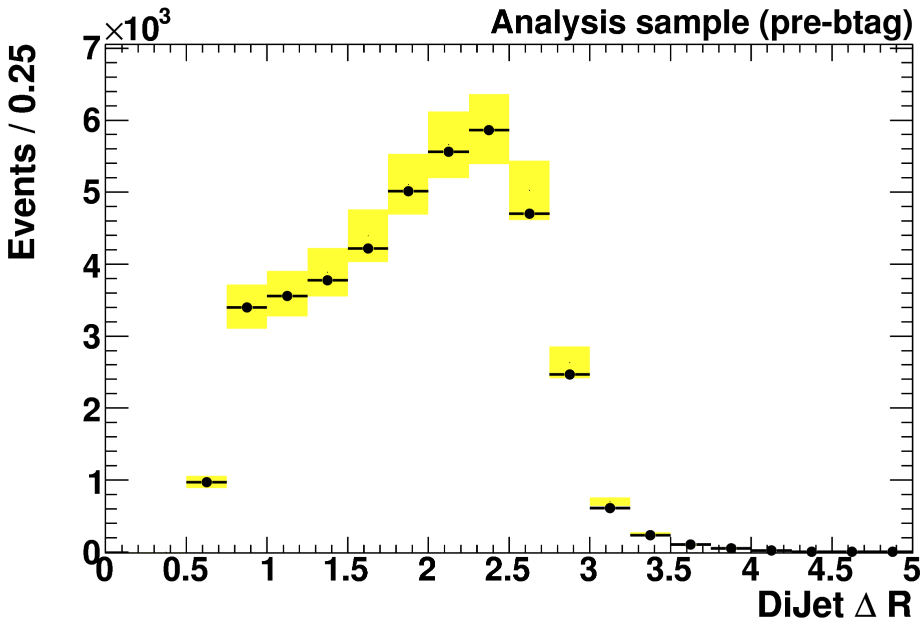


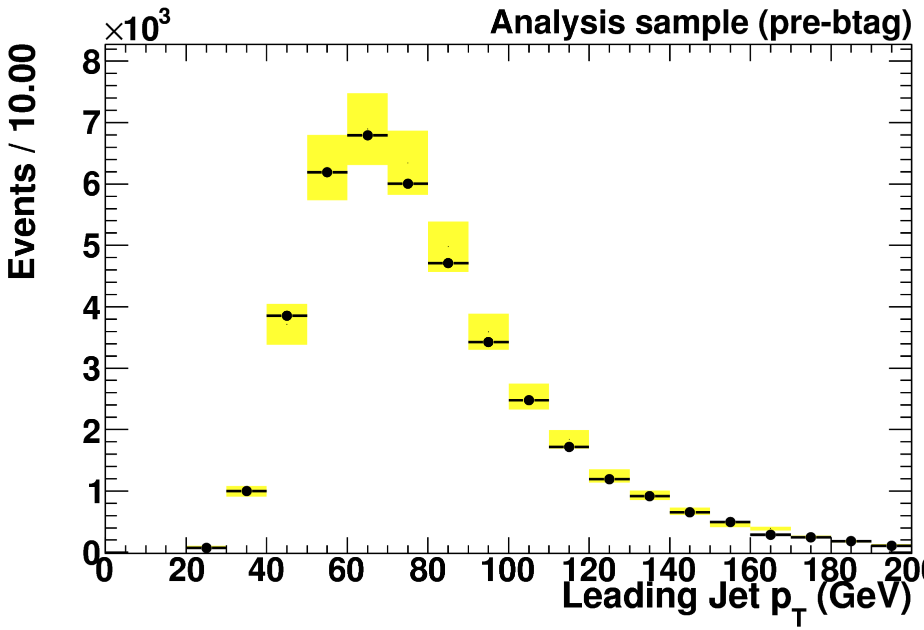


Analysis Sample after MJDT cut (ALL SF = 1)

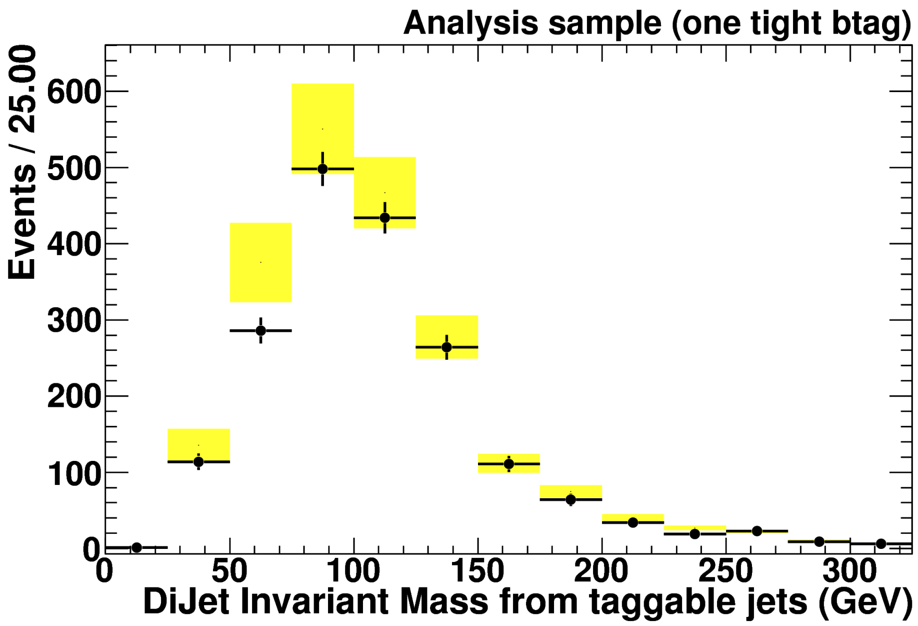


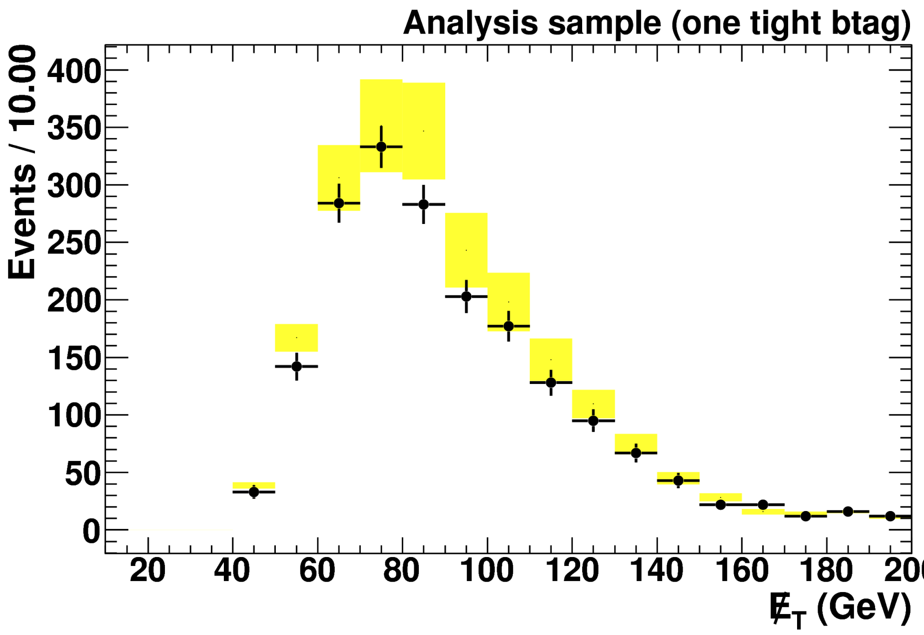


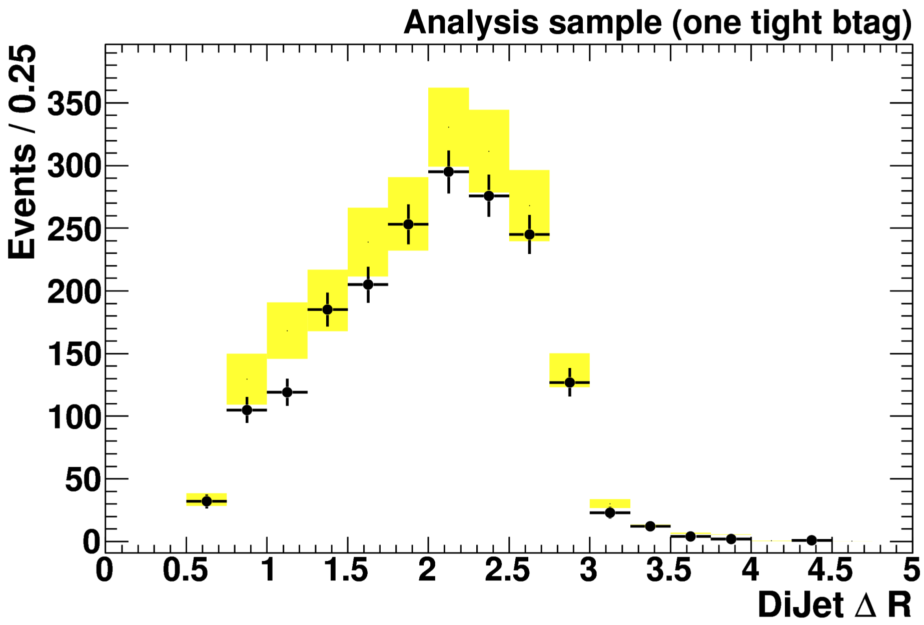


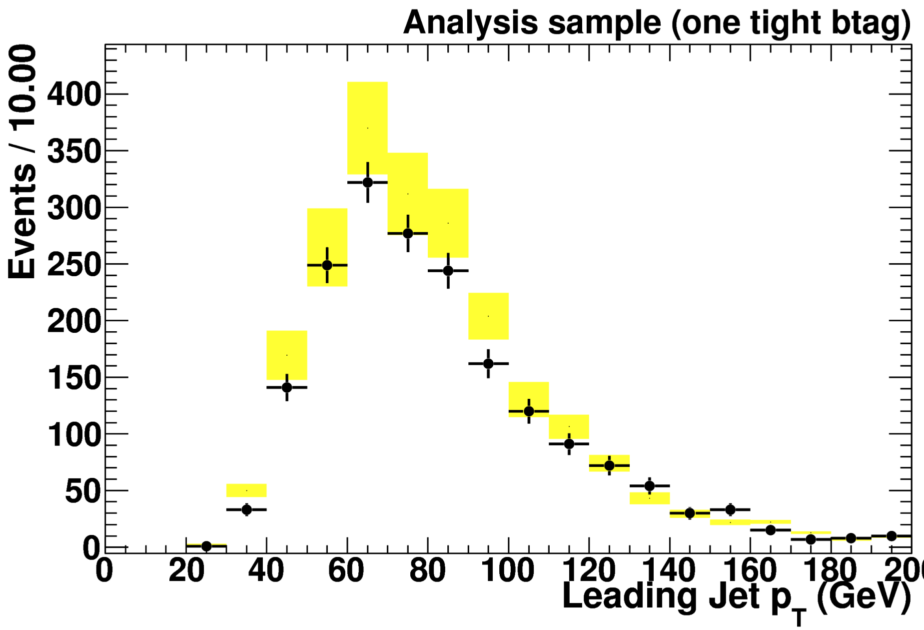


Analysis Sample after MJDT cut (ALL SF = 1)









Analysis Sample after MJDT cut (ALL SF = 1)