Resultat preliminaire du WH→lνbb a ASPEN

Gregorio Bernardi, Jonathan Brown, Duncan Brown, Yuji Enari, Ken Herner, Nils Huske,

Sebastien Greder, Jeremie Lellouch, Michiel Sanders,

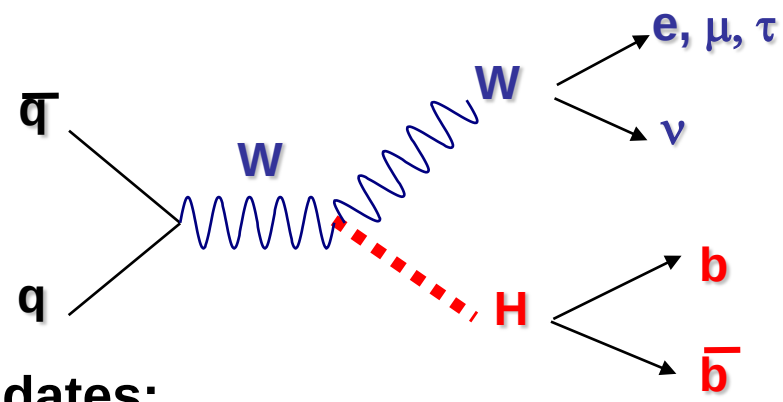
Isabelle Ripp-Baudot, Jianming Qian, Chun Xu

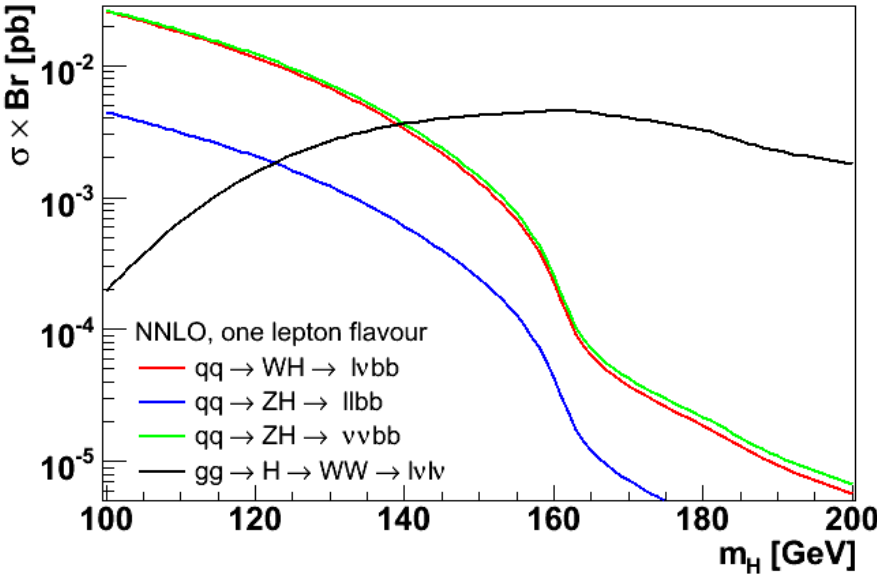
April 1st 2008

Reunion DØ France

WH analysis

Search for WH associated production
One of main channel at low mass





Updates:

- RunIIa + RunIIb 2.7 fb^{-1}
- Approved as preliminary result
- This will be good reference version

Feature of this analysis

Well established selection criteria
- good description at pretag level
- bID: direct tagging.
  - taggability is updated with vertex confirmation.

Multivariate technique: NN with ME discriminant
Almost all systematic error are re-evaluated.

Data and MC samples for p20

Data

Not use due to re-caffed data set

- Electron: non-recaffed sample (1609 pb^{-1})
  - CAF_EMinclusive_PASS4_p21.08.00 **+ p17 (1036 pb^{-1})**
  - ~~CAF_EMinclusive_PASS4_p21.10.00_p20.12.01~~
  - ~~CAF_EMinclusive_PASS4_p21.10.00_p20.12.02~~ **= 2.64 fb^{-1}**
- Muon: (1731 pb^{-1})
  - CAF_MUinclusive_PASS4-p21.10.00_p21.12.00 **+ p17 (1046 pb^{-1})**
  - CAF_MUinclusive_PASS4-p21.10.00_p21.12.01
  - CAF_MUinclusive_PASS4-p21.10.00_p21.12.02_summer2008 **= 2.78 fb^{-1}**

(3% from Cal noise effect need to be scaled)

MC: non-recaffed sample. (CAF-MCv2)

- W / Z + jet
  - Alpgen+Pythia, HF skimmed
  - W+ lp, Wcc+lp, Wbb+lp
  - p20.08.xx, p20.09.xx
- Top
  - Alpgen tt bar + lp
  - p20.08.xx p20.09.xx
- Single top
  - p20.09.xx.
- Di boson
  - Pythia
  - WW, WZ, ZZ
    p20.08.xx, p20.09.xx
- Higgs
  - Pythia WH, ZH
    p20.08.xx, p20.09.xx

Selection Criteria on p20

Event topology:
Lepton + ≥ 2 jets + MET
p20 Selection is same as p17 analysis

Lepton: pt > 15 GeV
mu: $|\eta|<2$
e: $|\eta|<1.1$, $1.5<|\eta|<2.5$

Jet:
Jet1 pt > 25 GeV
Jet2 pt > 20 GeV
(Jet3 pt > 20 GeV)
} $|\eta| < 2.5$

Met:
MET > 20 GeV
(25 GeV for EC electron)

2nd lepton veto: no 2nd loose lepton
HT cut: HT>60
Triangle cut: WtrMass>40-0.5MET

MUID: topP14Tight
MediumNseg3, medium track,
Iso: $\Delta R>0.5$, etSHalo<0.08,
etSTrHalo<0.06

EMID: top_tight
Track pT>8 GeV,
EMfrac > 0.9, HMx7<50
Iso<0.15, Lhood>0.85

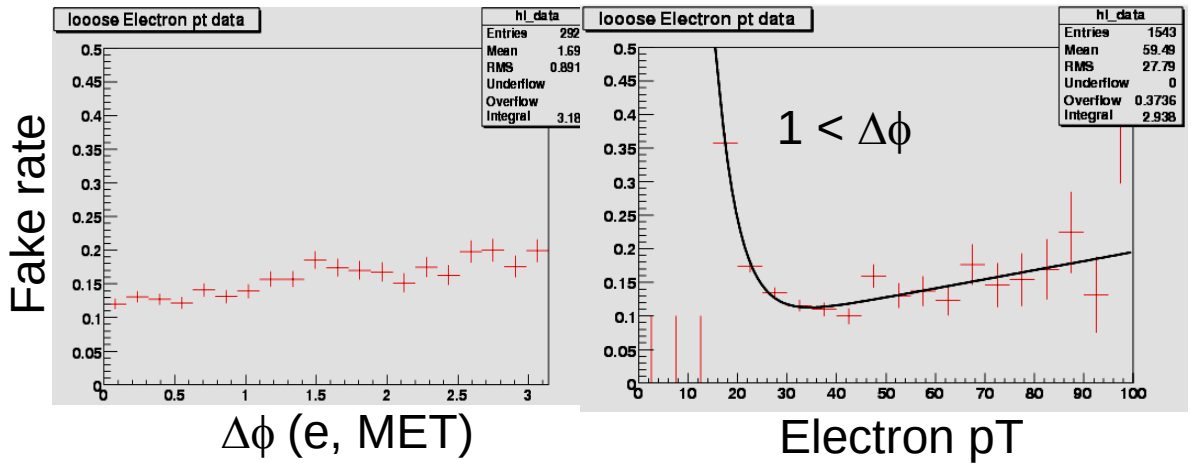
Jet: JCCB
JES, full JSSR is applied.

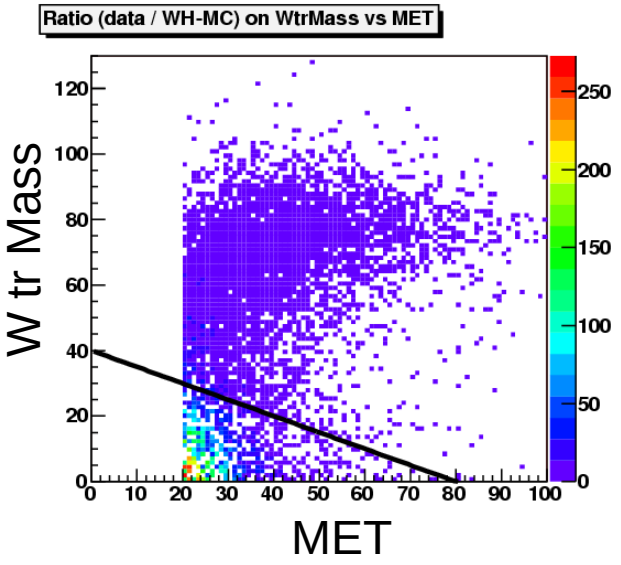
Trigger requirement:
mu: all trigger
e : single EM

QCD estimation

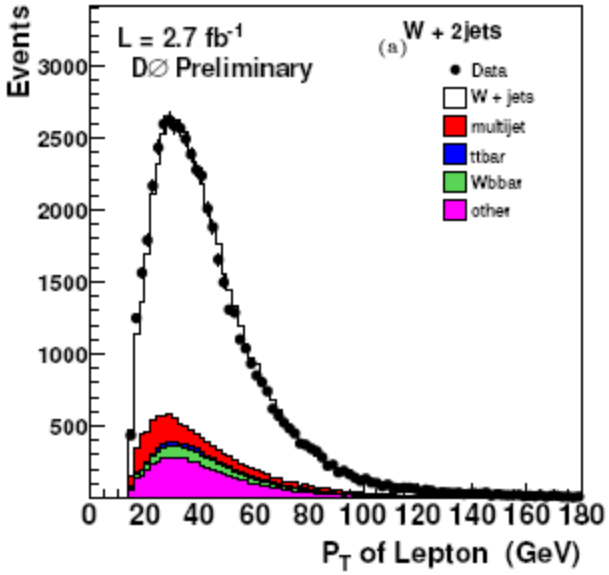
- Estimate from DATA using Matrix Method
  - Determine # of QCD and shape by bin by bin Matrix Method.
- Fake rate estimation
  - Apply exact same cut including trigger requirement except MET.
  - Apply correction to deal with real electron in MET<10 GeV using alpgen MC (normalized by luminosity).
  - Dependence:

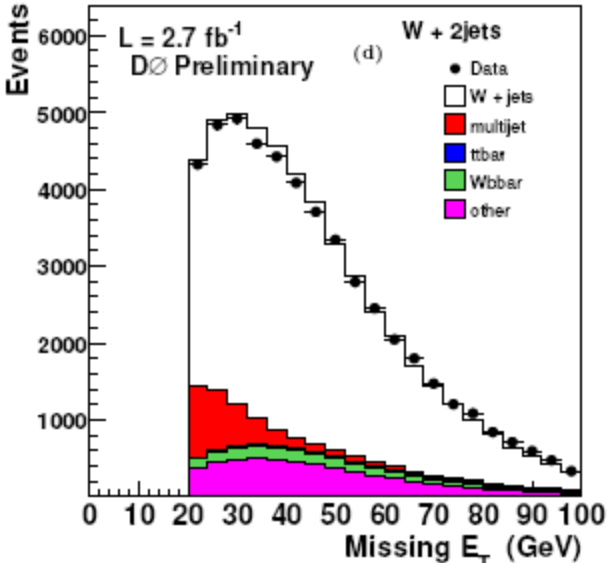
    Pt, $\Delta\phi$(lepton,MET), Trigger list, Jet multiplicity, CC, EC.
- Need to apply tighter triangle cut

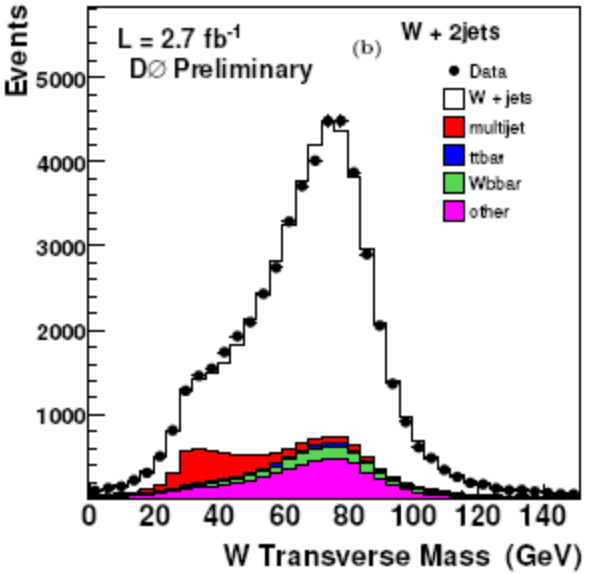


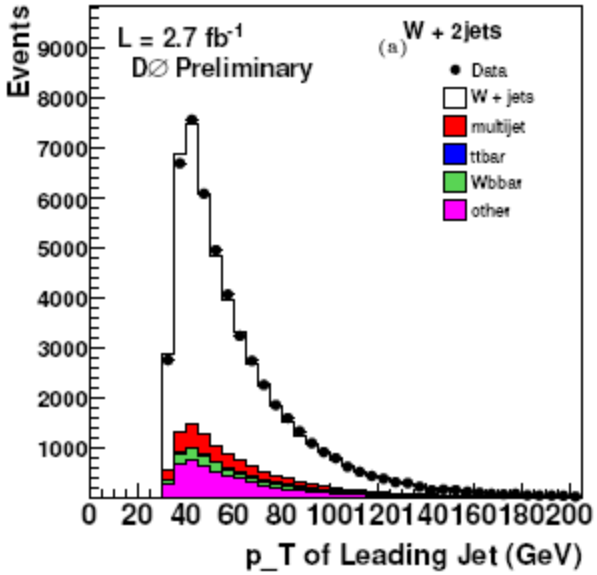


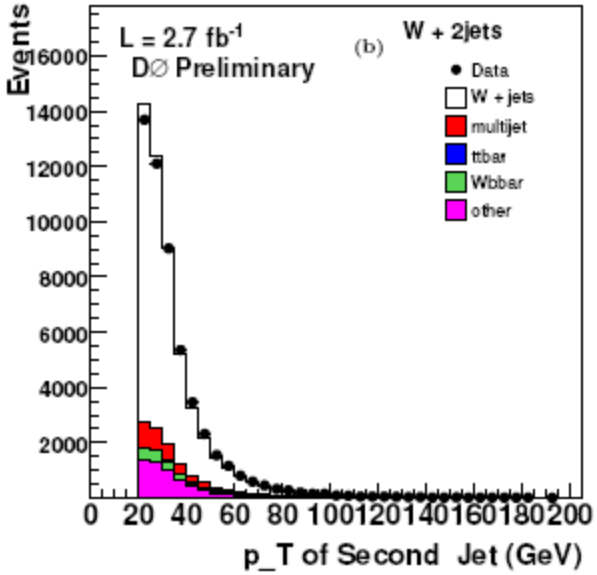
Combined plot p17+p20, mu + electron











MC Correction
- **Trigger**
- **LeptonID**
- **Jet ID**
- **Luminosity**
- **Z Vertex**
- **Jets for ALPGEN (η_{j1},η_{j2},$\Delta\eta_{jj}$, $\Delta\phi_{jj}$)**

Taggability with vertex confirmation

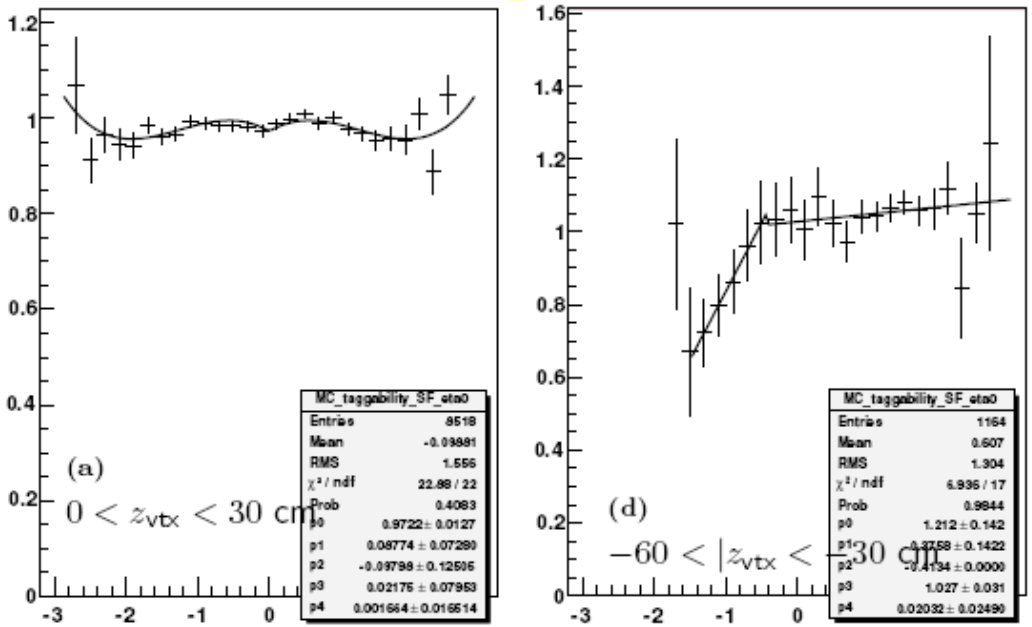
- Updated taggability requirement, and propaged to evaluation on Scale factors.

$$Taggability = \frac{Taggable \oplus Good}{Good}$$

$$Taggability = \frac{Taggable \oplus Good \oplus Ntrack}{Good \oplus Ntrack}$$

Vertex confirmation

Taggability Scale factor for MC



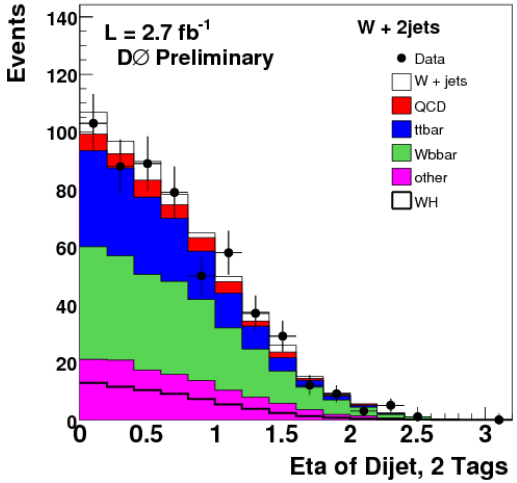
Recent study Murilo:

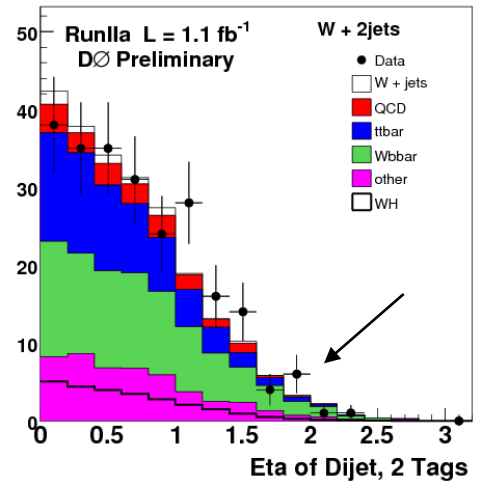
Scale factor for vertex conf. is not 1.

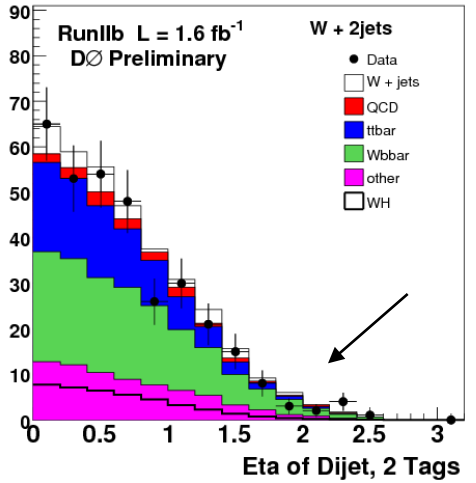
At high eta region, we may need ~ 30% correction.

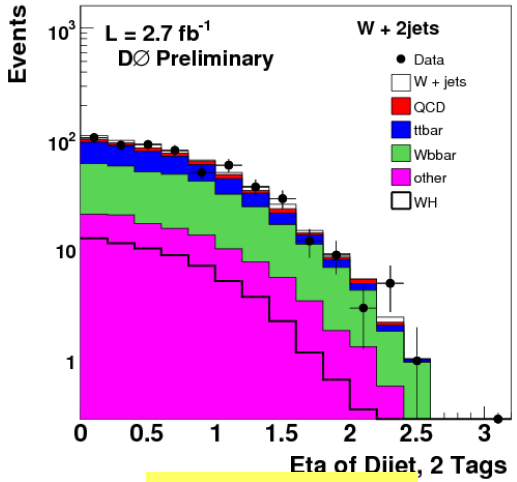
→ Is this true for bjets?

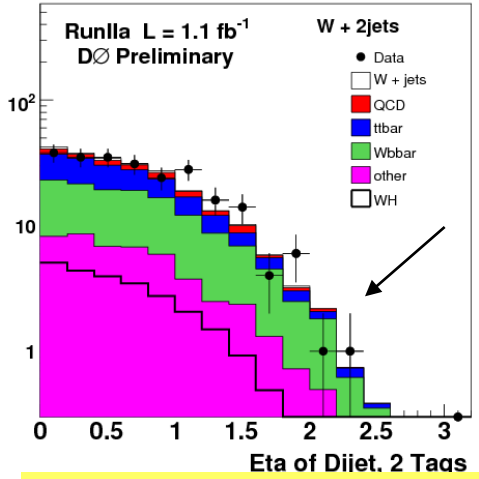
Agreement at high eta in double tag sample

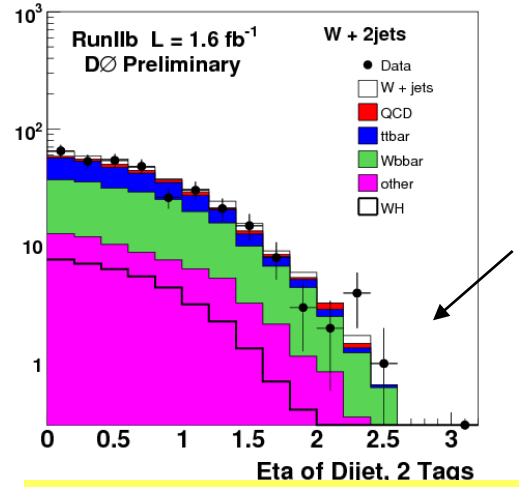












p17 + p20

p17: no Vertex conf.

p20: with Vertex conf.

Good agreement at high η:

Inefficiency due to vertex confimation on b-jet seems much smaller.

B tagging

- Using NN tagger.

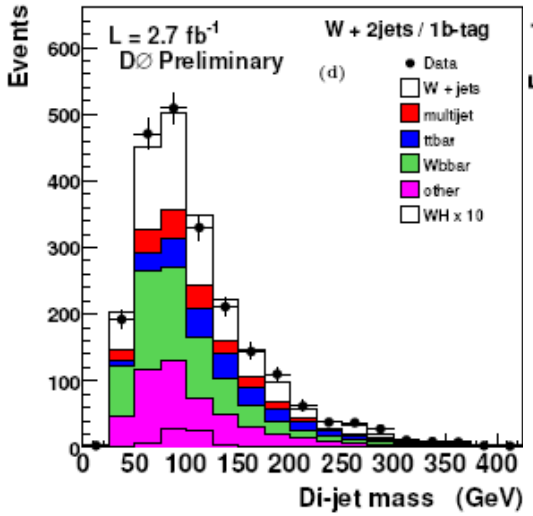
  Treatment for MC event

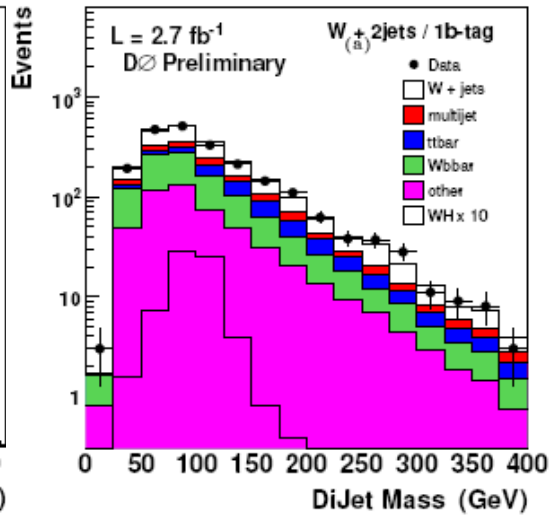
  – Use direct tagging approach
    - Better description of detector part
    - Various correlation can be taken account.
      ex. Jets at low ΔR region, etc..
  – Scale factor is obtained from TRF data / TRF MC.

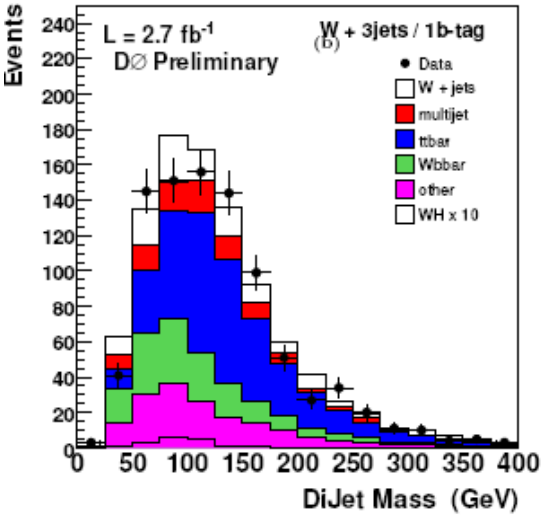
- Operating point
  – Tight for single tag (mistag: ~ 0.5%)
    OldLoose for double tag (mistag: ~1.5%)

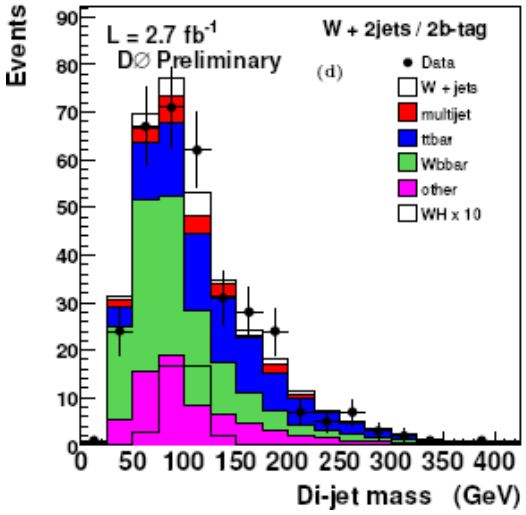
    → make orthogonal sample by direct tagging.
    Possible migrations due to data/MC scale factor is also taken account.

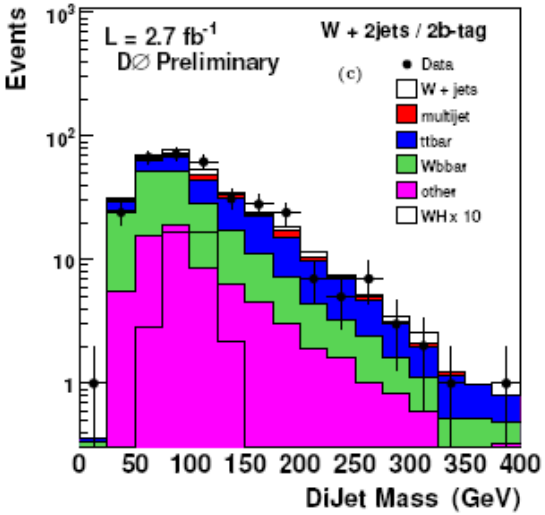
Tagged dijet mass, 2jet, 3jet

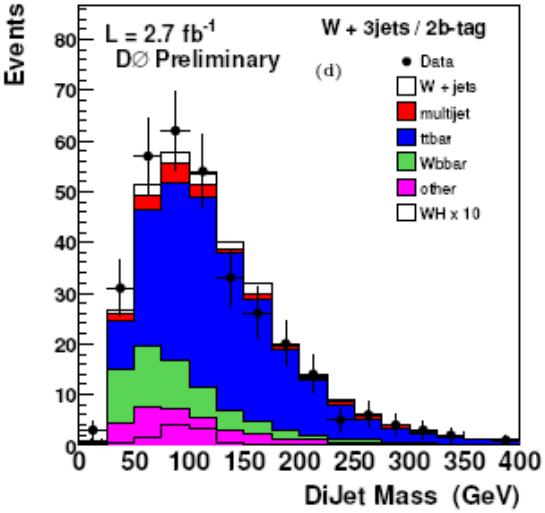












Matrix Element approach

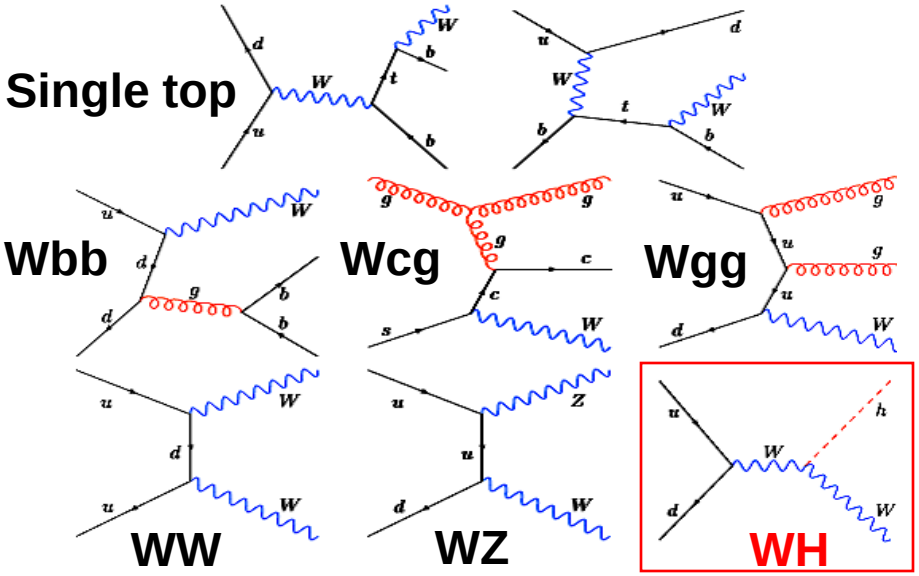
- Calculate the differential cross section of each physics process with 4 momentum for all object.

Detector response (Transfer function)

$$d\sigma(\vec{x}) = \sum_{i,j} \int d\vec{y} \left[\underbrace{f_i(q_1, Q^2) dq_1 \times f_j(q_2, Q^2) dq_2}_{p.d.f} \times \underbrace{\frac{\partial \sigma_{hs,ij}(\vec{y})}{\partial \vec{y}}} \times \overbrace{W(\vec{x}, \vec{y})} \times \underbrace{\Theta_{parton}(\vec{y})} \right]$$

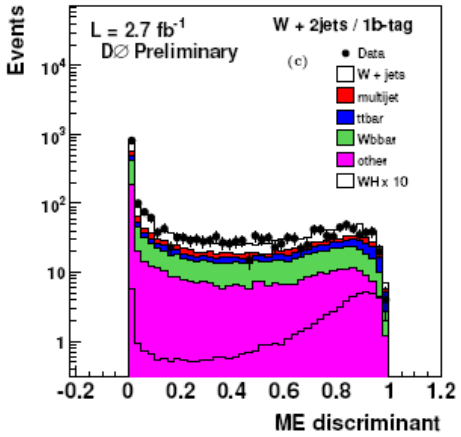
Parton level cut

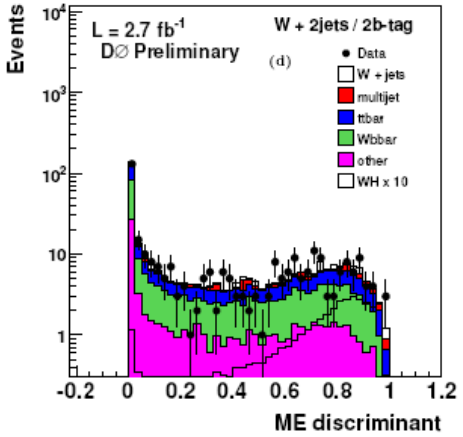
Matrix Element for hard scatter collision



Discriminant is defined as

$$D_{WH}(\vec{x}) = \frac{P_{WH}(\vec{x})}{P_{WH}(\vec{x}) + P_B(\vec{x})}$$

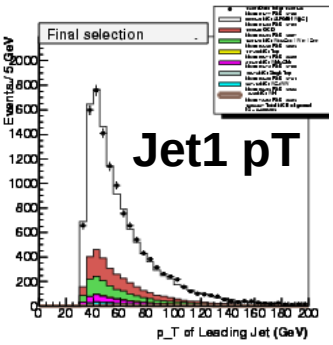


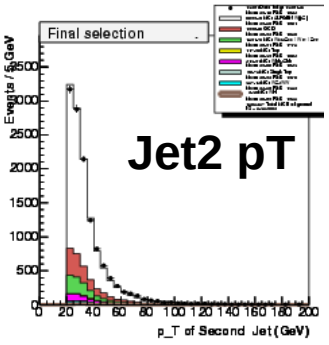


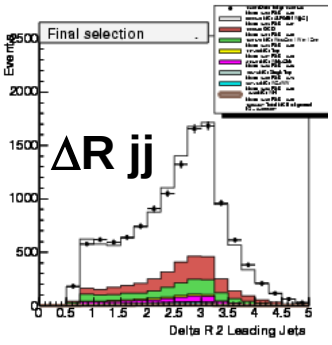
Neural Net

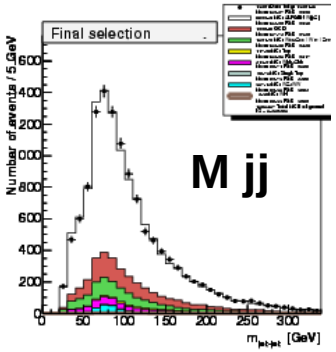
- NNet: 1 hidden layer, same number of node with input variable.
- Training (number of epoch with error is minimal.)
  - sample: BG: Wbb, Signal: WH(W→eν or μν)
  - Train NN each mass point, p17 and p20 separately.
  - Train NN separately for single tag event and double tag event
  - Train NN separately for muon and electron
- Input variables

→ 8 x mass point

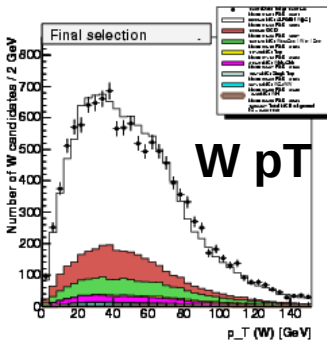


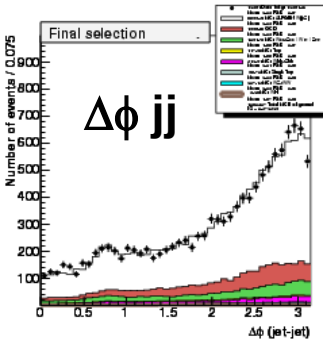


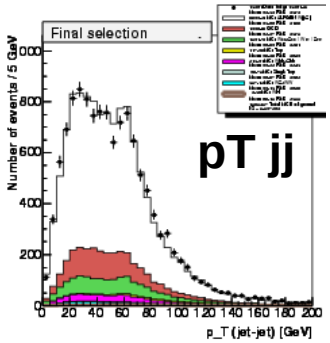


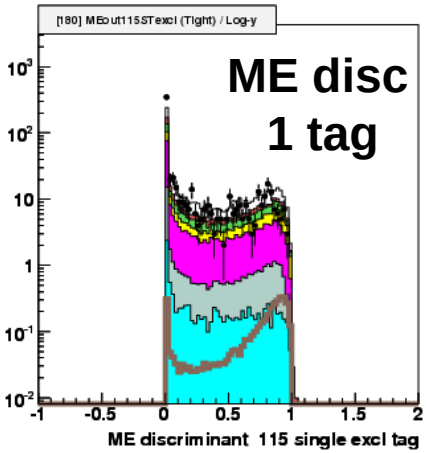


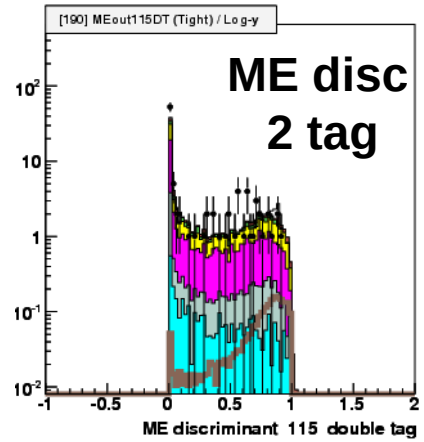
+



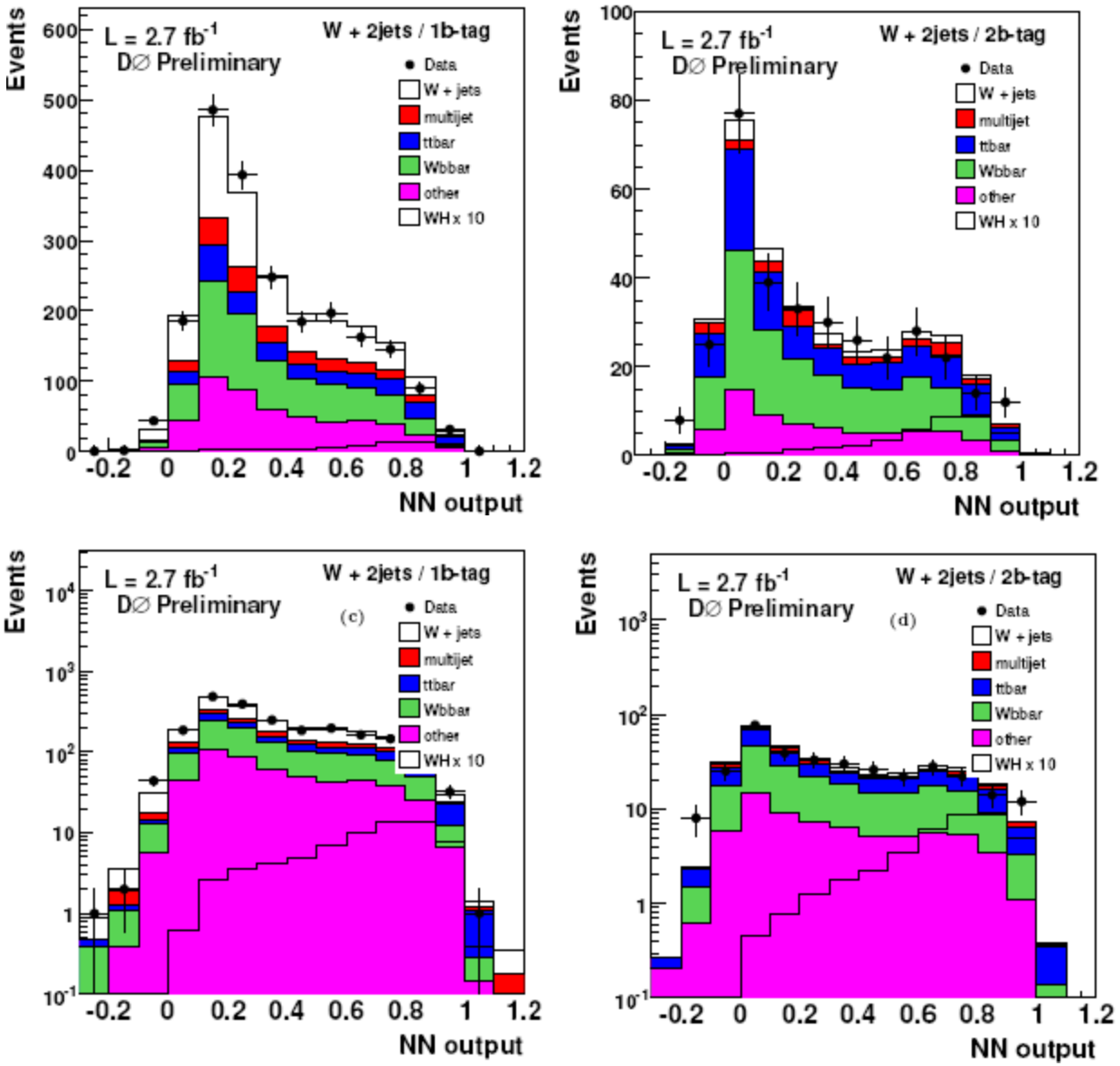








NN output


L = 2.7 fb-1
DØ Preliminary
W + 2jets / 1b-tag
Data
W + jets
multijet
ttbar
Wbbar
other
WH x 10
Events
NN output
L = 2.7 fb-1
DØ Preliminary
W + 2jets / 2b-tag
Data
W + jets
multijet
ttbar
Wbbar
other
WH x 10
Events
NN output
L = 2.7 fb-1
DØ Preliminary
(c)
W + 2jets / 1b-tag
Data
W + jets
multijet
ttbar
Wbbar
other
WH x 10
Events
NN output
L = 2.7 fb-1
DØ Preliminary
(d)
W + 2jets / 2b-tag
Data
W + jets
multijet
ttbar
Wbbar
other
WH x 10
Events
NN output

Systematic uncertainties

- Luminosity : 6.1%
- Trigger: 3~4% for electron, 5% for muon

- Electron ID: 4.2% (ID : 4%, Smearing: 2%)
- Muon ID: 4.1% (ID: 3.0%, trk match: 2%, Smearing: 2%)
- Jet: 3.9% (ID 3%, JES : 2~5%), 2~6% JSSR, Multiplicity: 4.2%
- QCD: 26% from matrix method (including stat. error in tagged samples).

- B-tagging:
  b-jet: 3 %, c-jet: 6% per jet including taggability.
  double tag : 25%, single: 15% for light jet.

- Cross section: ttbar 10%, diboson:6%, WH: 6%
- Alpgen K-factor: Determined from data
  - Error : Wbb 20%   - Wjj: ~5% (from QCD error)

- Checked following dependences
  - Lepton ID, trigger , Jes, JSSR, Vertex reweighting, Luminosity reweighting, ALPGEN reweighting, b-ID, Wpt shape
- Shape uncertainties from reweighting function.
  - Off/on reweighting function on dijet mass distribution.

Inputs to collie

- Combine RunIIa, RunIIb, e, mu, 2jet(NNoutput), 3jet (dijet mass), 1-tag, 2tag. → Total 16 inputs.

	electron 2jet-1b IIa IIb	electron 2jet-2b IIa IIb	electron 3jet-1b IIa IIb	electron 3jet-2b IIa IIb	muon 2jet-1b IIa IIb	muon 2jet-2b IIa IIb	muon 3jet-1b IIa IIb	muon 3jet-2b IIa IIb
DZero_Lumi	x x	x x	x x	x x	x x	x x	x x	x x
Lumi	x x	x x	x x	x x	x x	x x	x x	x x
JESID	x x	x x	x x	x x	x x	x x	x x	x x
bTag_Tagga_HF	x x	x x	x x	x x	x x	x x	x x	x x
bTag_Tagga_LF_WH	x x	x x	x x	x x	x x	x x	x x	x x
EMID	x x	x x	x x	x x				
Bkgd_Xsec_EW	x x	x x	x x	x x	x x	x x	x x	x x
Jet_NN_MMLM	x x	x x			x x	x x		
Bkgd_Xsec_NN_HF_WH	x	x			x	x		
Bkgd_WbbRW_NN_shape	x x	x x			x x	x x		
Bkgd_QCDev	x x	x x	x x	x x				
Bkgd_WjjRW_NN_shape	x x	x x			x x	x x		
Bkgd_Xsec_Top	x x	x x	x x	x x	x x	x x	x x	x x
Bkgd_Xsec_singletop	x x	x x	x x	x x	x x	x x	x x	x x
Jet_DJ_MMLM			x x	x x			x x	x x
Bkgd_Xsec_DJ_HF_WH			x	x			x	x
Bkgd_WbbRW_DJ_shape			x x	x x			x x	x x
Bkgd_WjjRW_DJ_shape			x x	x x			x x	x x
Bkgd_Xsec_NN_HF_WH_IIb	x	x			x	x		
Bkgd_Xsec_DJ_HF_WH_IIb			x	x			x	x
MUTrigger_WH					x x	x x	x x	x x
MUID					x x	x x	x x	x x
Bkgd_QCDmv					x x	x x	x x	x x
WHNLO	x x	x x	x x	x x	x x	x x	x x	x x

Results with CLFIT2 (with full systematic)

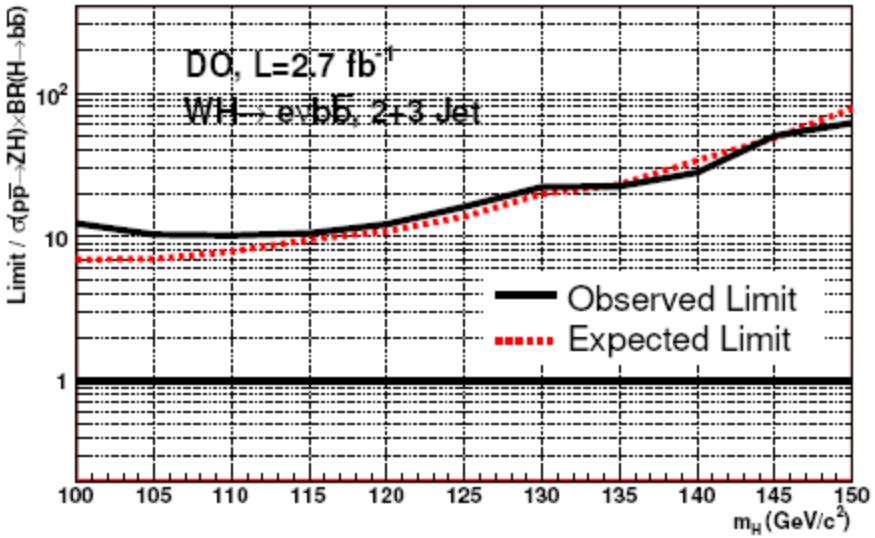
Limit ratio to SM expectation at M_H=115 GeV: Expected (Observed)

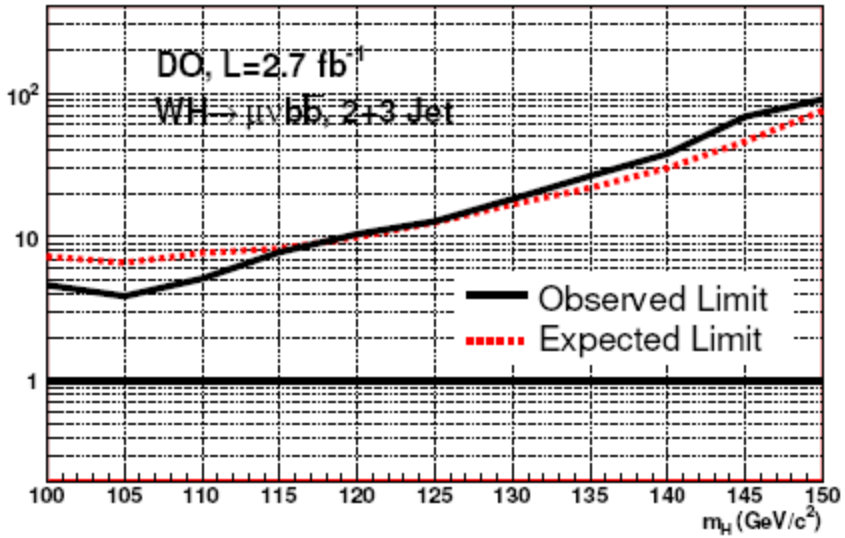
Analysis	elec. cut-b.	elec. NN	elec. ME	elec. NN+ME	muon cut-b.	muon NN	muon ME	muon NN+ME
2007/2j/IIa	20.3 (31.8)	–	–	–	21.7 (20.8)	–	-	
IIa / 2 jet	18.0 (23.8)	15.6 (22.2)	15.9 (16.3)	14.6 (15.0)	17.7 (11.5)	15.3 (10.1)	14.0 (12.3)	14.4 (12.3)
IIa / 2+3	17.0 (25.8)	15.1 (23.1)	15.6 (17.1)	14.4 (15.6)	17.0 (11.5)	14.9 (10.3)	13.7 (12.9)	14.0 (12.5)
IIb / 2 jet	15.9 (18.0)	12.8 (14.3)	13.0 (9.8)	12.6 (11.8)	13.7 (17.1)	12.0 (11.7)	10.4 (13.5)	10.2 (12.1)
IIb / 2+3	15.3 (18.9)	12.8 (15.2)	13.0 (10.7)	12.4 (12.6)	12.9 (13.8)	11.6 (9.9)	10.6 (11.6)	9.9 (10.5)
IIa+b / 2 jet	11.9 (15.7)	10.2 (13.9)	10.4 (8.9)	10.0 (10.1)	10.7 (9.6)	9.8 (7.2)	8.6 (9.1)	8.7 (8.9)
IIa+b / 2+3	11.2 (17.2)	9.8 (14.5)	9.8 (9.6)	9.6 (10.5)	10.3 (8.3)	9.3 (6.5)	8.2 (8.2)	8.3 (7.8)

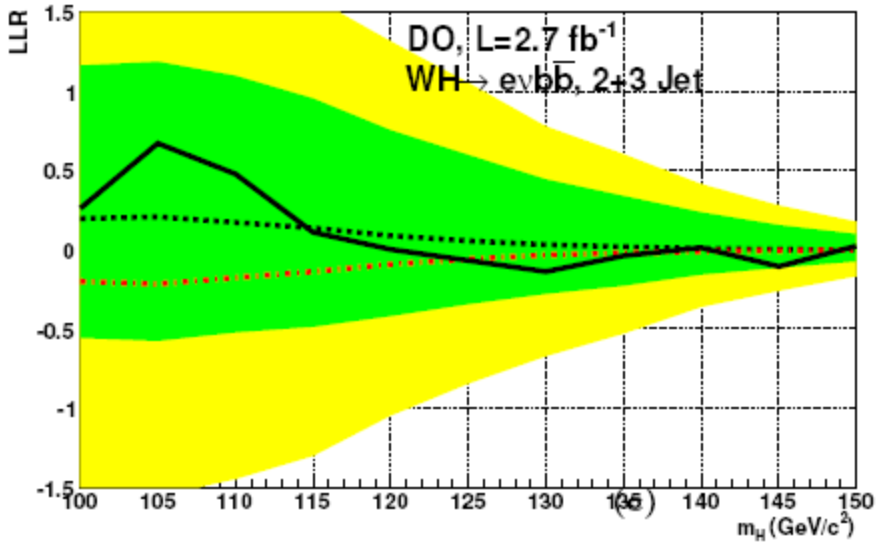
Final result with 2.7 fb-1 :

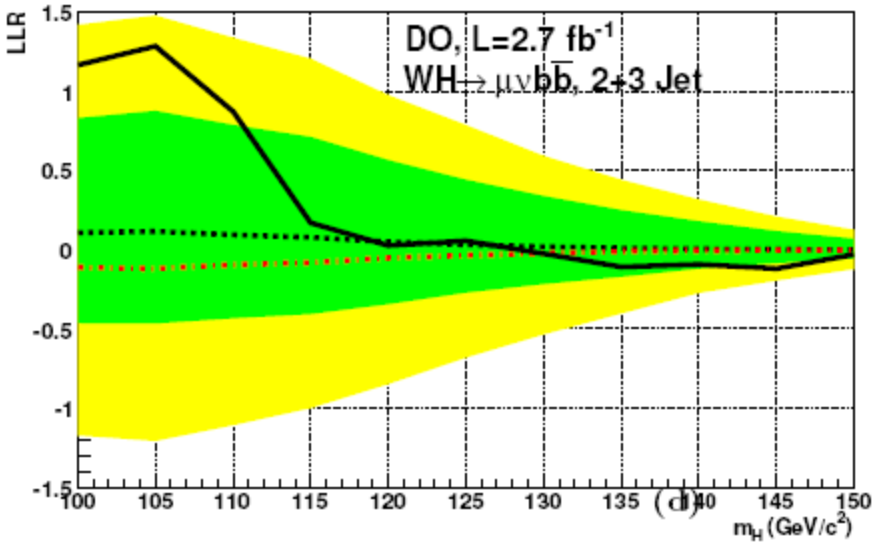
Analysis	$e+\mu$, cut-b.	$e+\mu$, NN	$e+\mu$, ME	$e+\mu$, NN+ME
3-2007/2 jet /1.0 fb^{-1}	14.2 (18.8)	– (–)	– (–)	– (–)
IIa / 2 jet	12.4 (10.7)	10.9 (10.8)	10.6 (10.1)	10.5 (10.1)
IIa / 2+3	11.8 (12.2)	10.5 (11.7)	10.1 (10.9)	10.0 (10.6)
IIb / 2 jet	10.3 (13.0)	8.9 (9.0)	8.3 (7.8)	8.1 (8.7)
IIb / 2+3	9.7 (11.3)	8.7 (8.2)	8.1 (7.3)	7.8 (7.6)
IIa+b / 2 jet	7.9 (8.8)	7.2 (7.3)	6.6 (6.3)	6.6 (7.4)
IIa+b / 2+3	7.5 (8.5)	6.9 (7.1)	6.4 (6.1)	6.4 (6.7)

Result for electron and muon

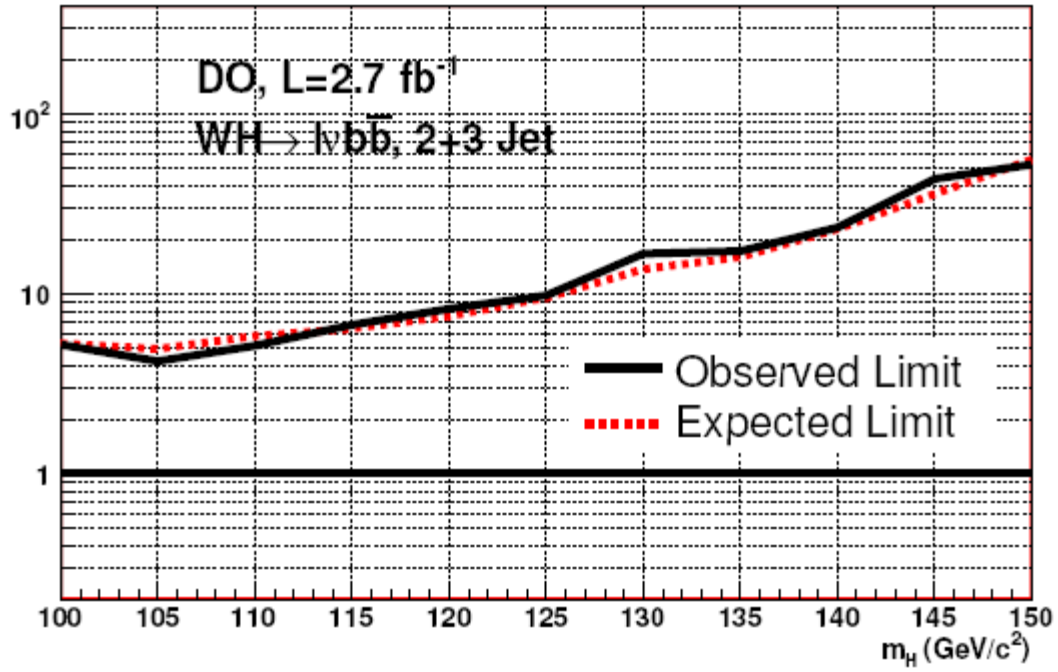


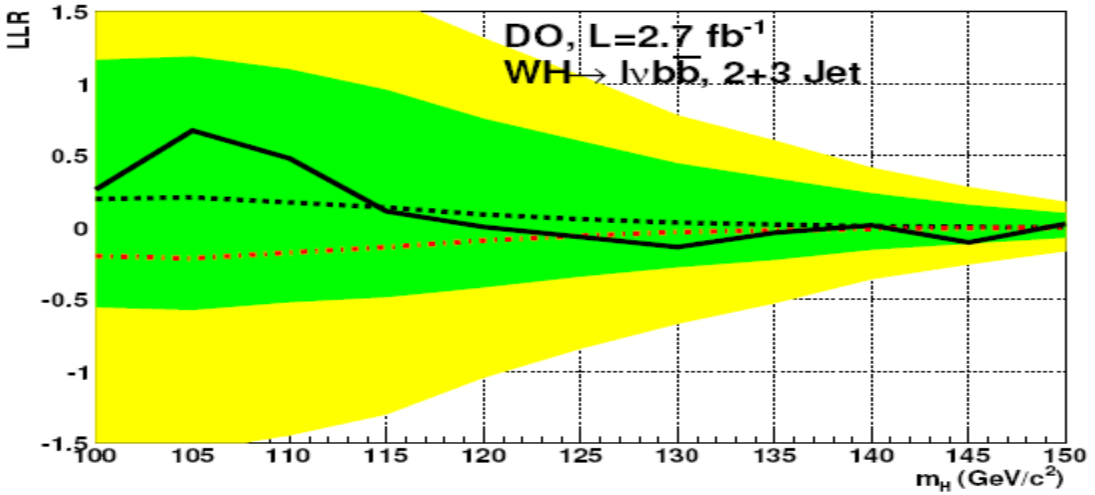






Combined result





Final result with NN+ME:

Expected: 6.4

Observed: 6.7

@ M_H = 115 GeV

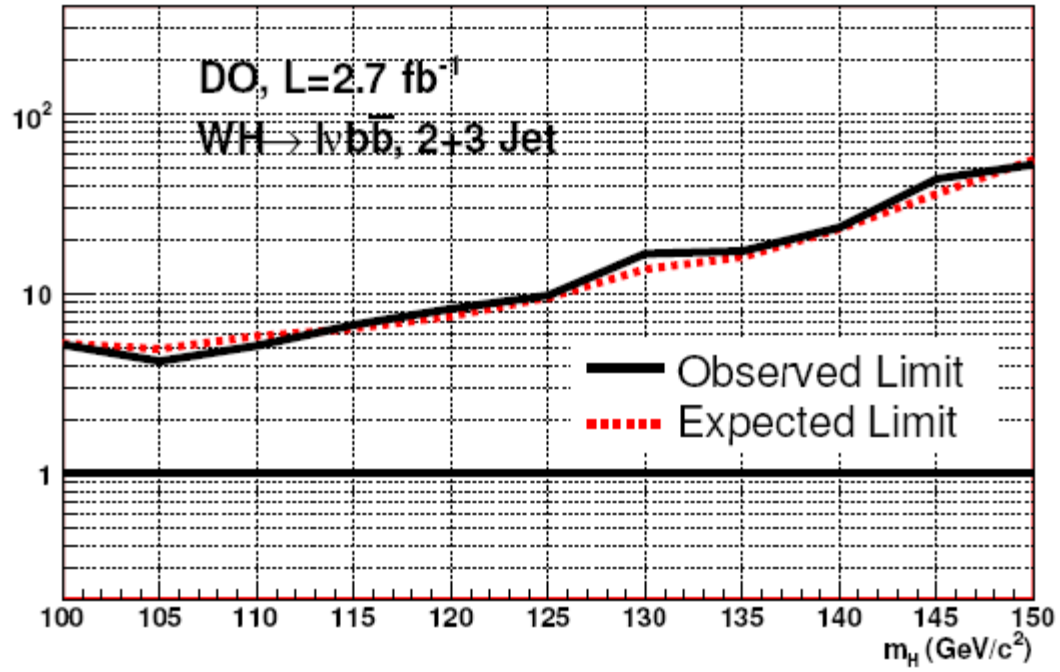
CDF result with 2.7 fb^{-1}

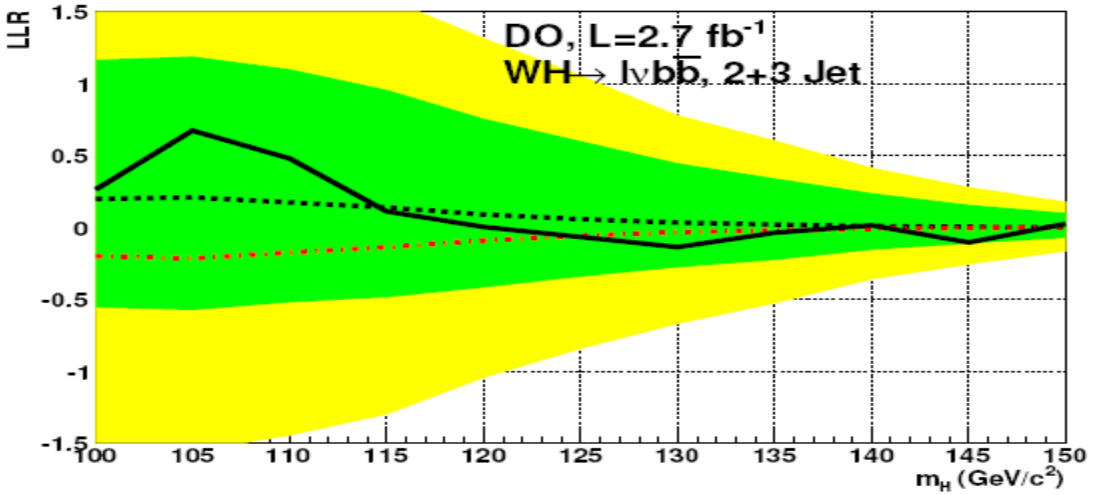
	exp	Obs
BDT+ME	5.2	6.2
NN	5.8	5.2

Combined with NEAT :
exp: 4.8 obs: 5.6

No excess or deficit is observed.

Combined result





Final result with NN+ME:
with Baysian method:
Expected: 5.8
Observed: 5.7
@ M_H = 115 GeV

CDF result with 2.7 fb^{-1}

	exp	Obs
BDT+ME	5.2	6.2
NN	5.8	5.2

Combined with NEAT :
exp: 4.8 obs: 5.6

No excess or deficit is observed.

Conclusion

- This analysis has now been approved by EB, and released at ASPEN, also used in Tevatron combination.

- On going:
  - Update with full data set.
  - Re-optimizing
    - event selection.
    - B-tagging part.
  - Improve on Multivariate technique.

Back up

Latest factor on this analysis

- K' x S (factor of top of ALPGEN cross section)

	channel	2 jet	3 jet
p17	Electron (CC)	1.65 ± 0.01	1.62 ± 0.03
	Muon	1.63 ± 0.01	1.80 ± 0.03
p20	Electron (CC)	1.85 ± 0.01	2.05 ± 0.03
	Muon	1.52 ± 0.01	1.73 ± 0.03

→ Assign 10% systematic error on this.

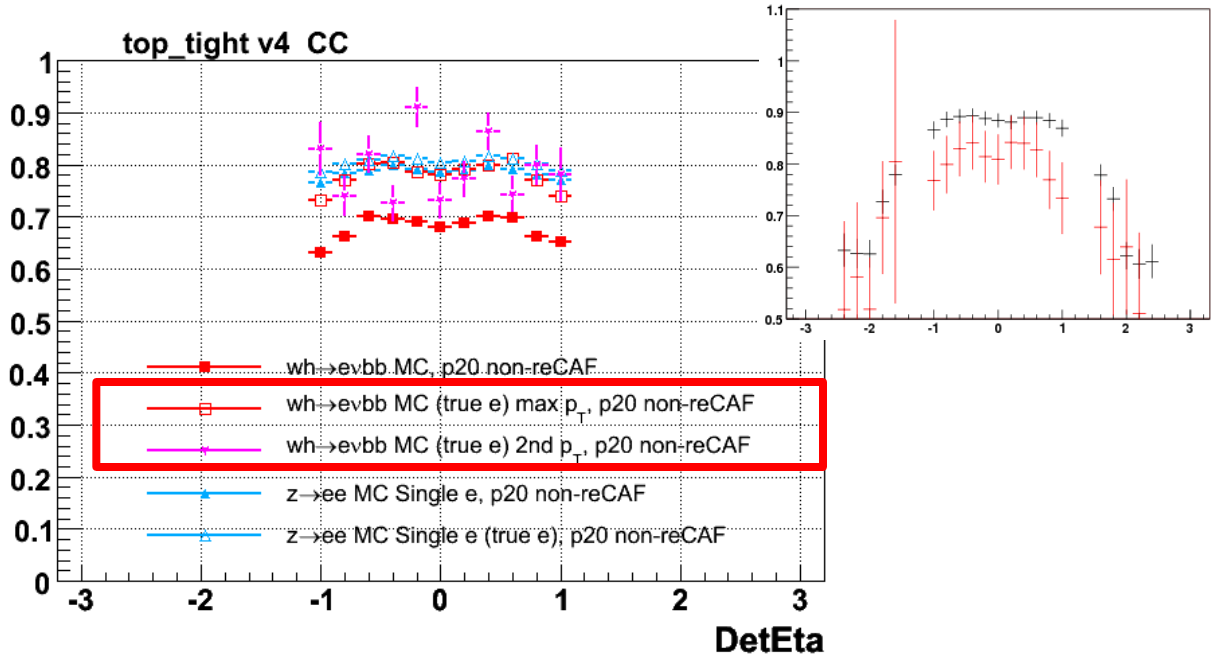
- HF factor (Additional factor on HF MCs.)

data	btag OP	VeryTight	Tight	oldLoose
p17	Electron	0.99 ± 0.02	0.99 ± 0.02	0.97 ± 0.03
	Muon	1.02 ± 0.02	1.02 ± 0.02	1.01 ± 0.03
p20	Electron	1.29 ± 0.02	1.36 ± 0.02	1.34 ± 0.03
	Muon	1.45 ± 0.02	1.52 ± 0.02	1.38 ± 0.03

→ Assign 20% systematic error on this.

About EMID on non-recaffed data set

- At last meeting, I reported EMID efficiency in SPC file may not correct. It's confirmed by EMID group.



Red point: WH
Black: in SPC file

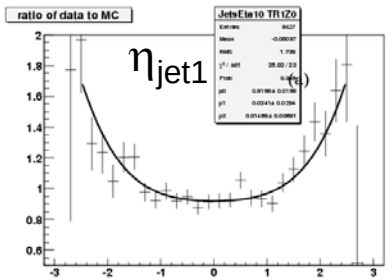
- But, it turned out its is due to definition of denominator is not same. So that data efficiency in SPC file is also low.
- Efficiency ratio (correction factor to MC) is correct.

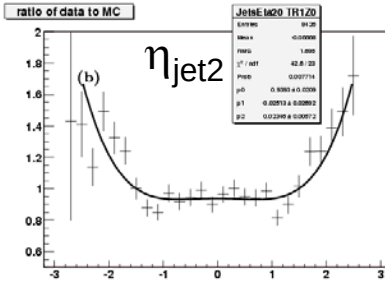
  → We boost 15% due to this observation, it turned out this correction is not valid, so we removed this 15% correction.

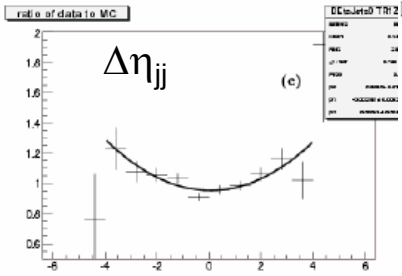
ALPGEN mis-modeling correction

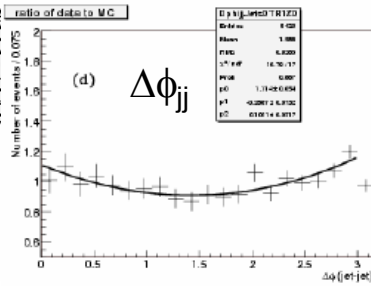
ALPGEN is not perfect.
Need to correct mis-modeling.
Functions is obtained from data.
Func = (Data-QCD-TOP,etc) / (W/Z+jet MC).

Function from p17 is used on p20

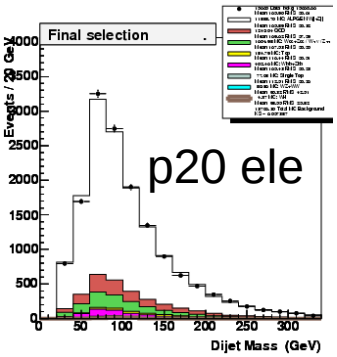


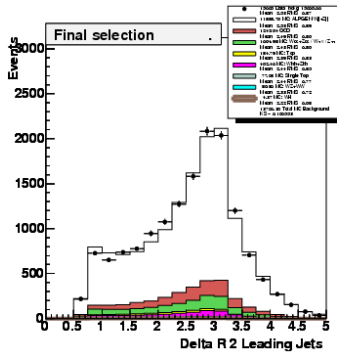


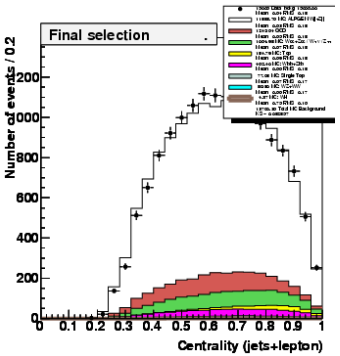


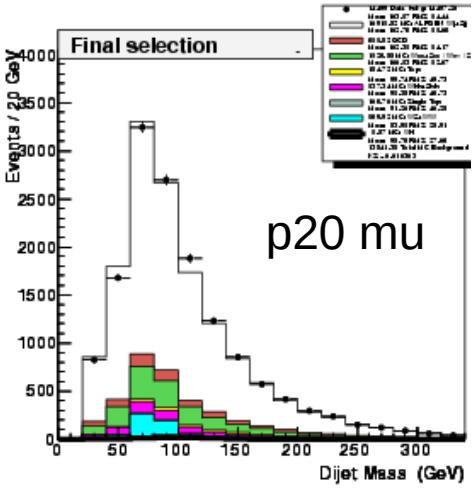


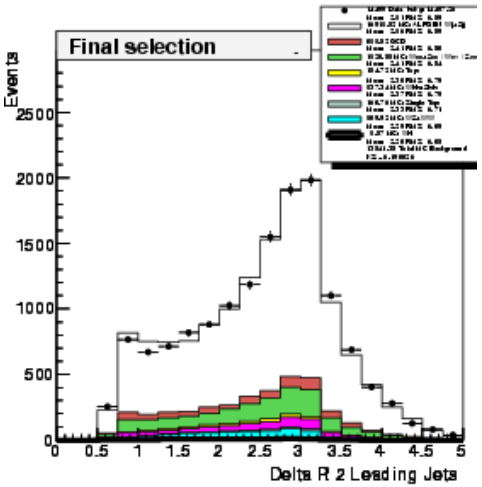
After reweigting

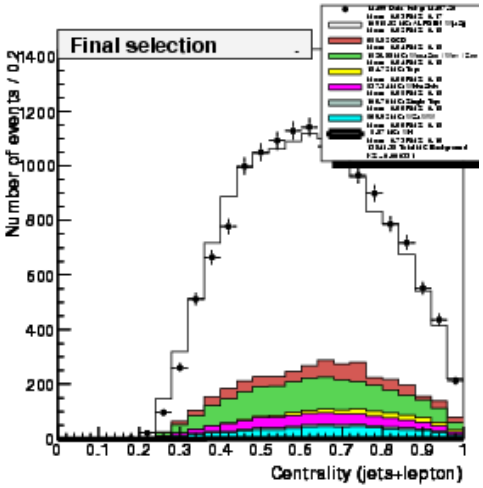








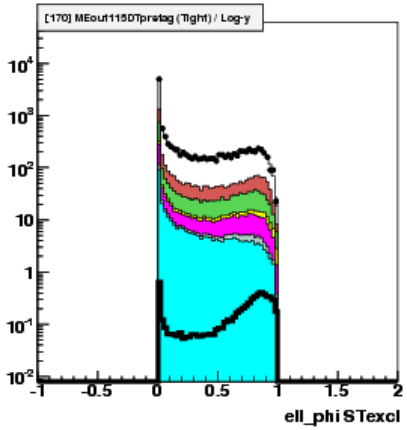




Procedure to get ME result

- Main issue is ME needs integrate all events.
  - Integration time: 1~2 min / event.
  - In total, 2 ~ 4 weeks to integrate all RunIIb sample.

  → We started integration as soon as all smearing was fixed.
  Before fixing correction factor or b-tagging.



Integration:
ASCI file → ASCI file

Updated event weight

CAF job:
CAF tree → topo tree → asci file

ME discriminant:
ASCI file → topo tree
CAF tree

ME discriminant, updated event weight

NNet analysis

In each path, need to match event by event. Smearing factor needs to be all same.