Higgs boson search in the ZH→ννbb channel and development of a soft muon b-tagger algorithm

A. Duperrin, **D. Jamin**
(CPPM Marseille)

05/03/2010
DØ France

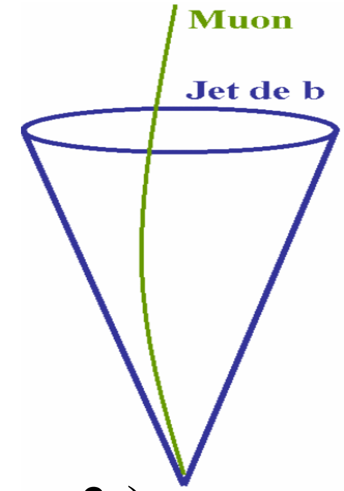
INTRODUCTION

My PhD (defense end of September) thesis contains two parts :

- update and improve the soft muon b-tagging (**SLTNN**) for Run IIb dataset

- implementation and results with SLTNN in the **$ZH \rightarrow \nu\nu bb$** channel.

SLT-NN



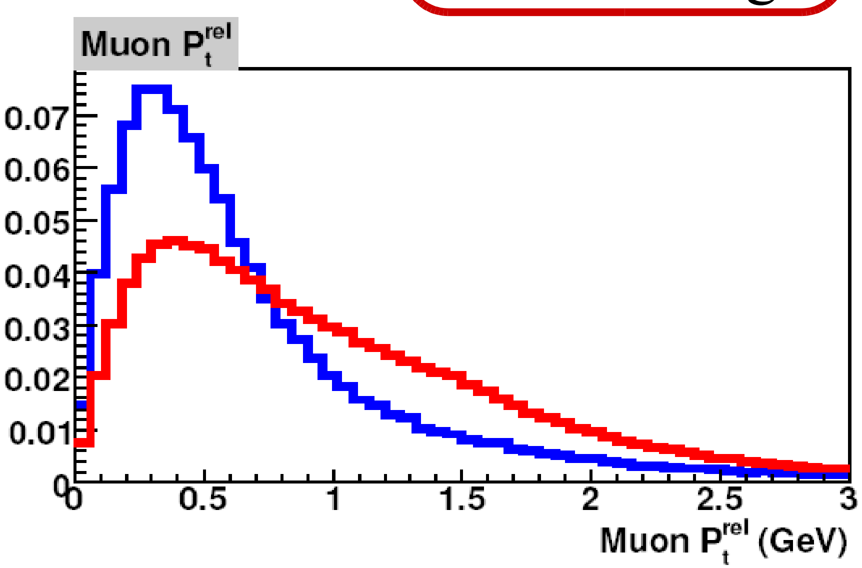
- "<u>Muon Tagging" (SLT) :</u> muon pT > 4 GeV
  muon | η | < 2
  ΔR (muon , Jet) < 0.5

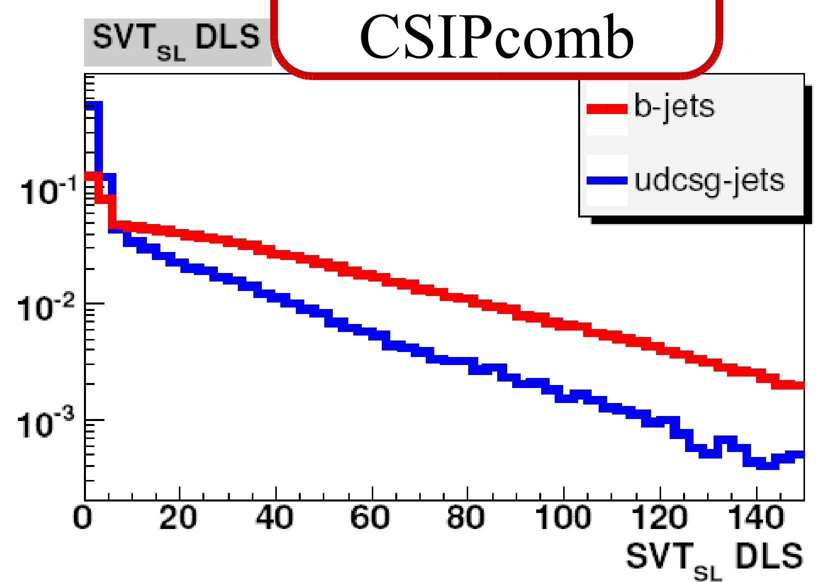
3 types of muon quality are studied in parallel :

Loose, Medium, Medium3 (Medium with nseg = 3)

- <u>Neural Network :</u> **5 Muon variables and 6 Secondary Vertex variables**

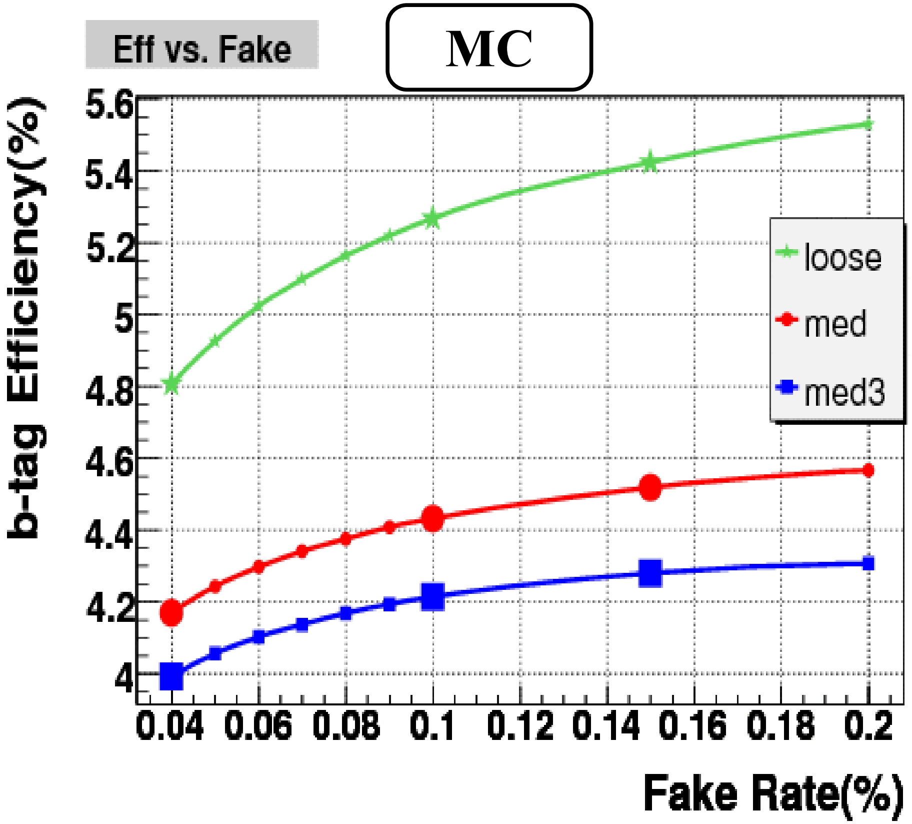
Muon variables	Secondary Vertex variables
muon pT	SV DLS
muon pTrel	SV Mass
muon ΔR	SV χ^2/dof
muon χ^2/dof	SV Ntracks
muon IP sig	SV Nvtx
	CSIPcomb



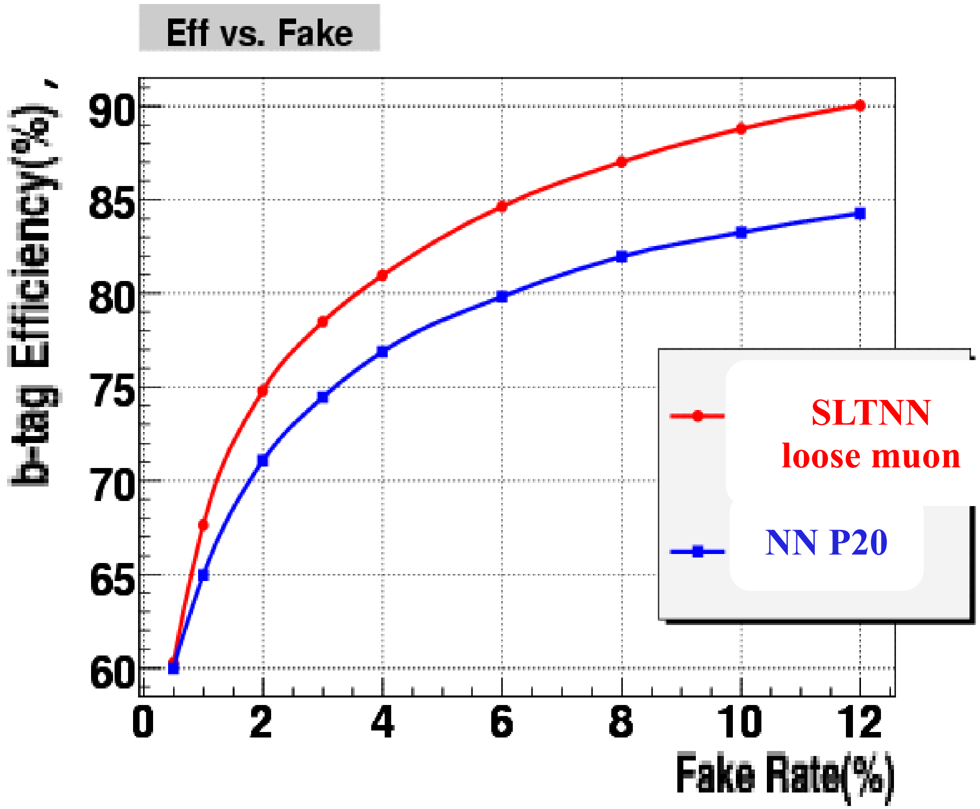


SLT-NN results

For b jets



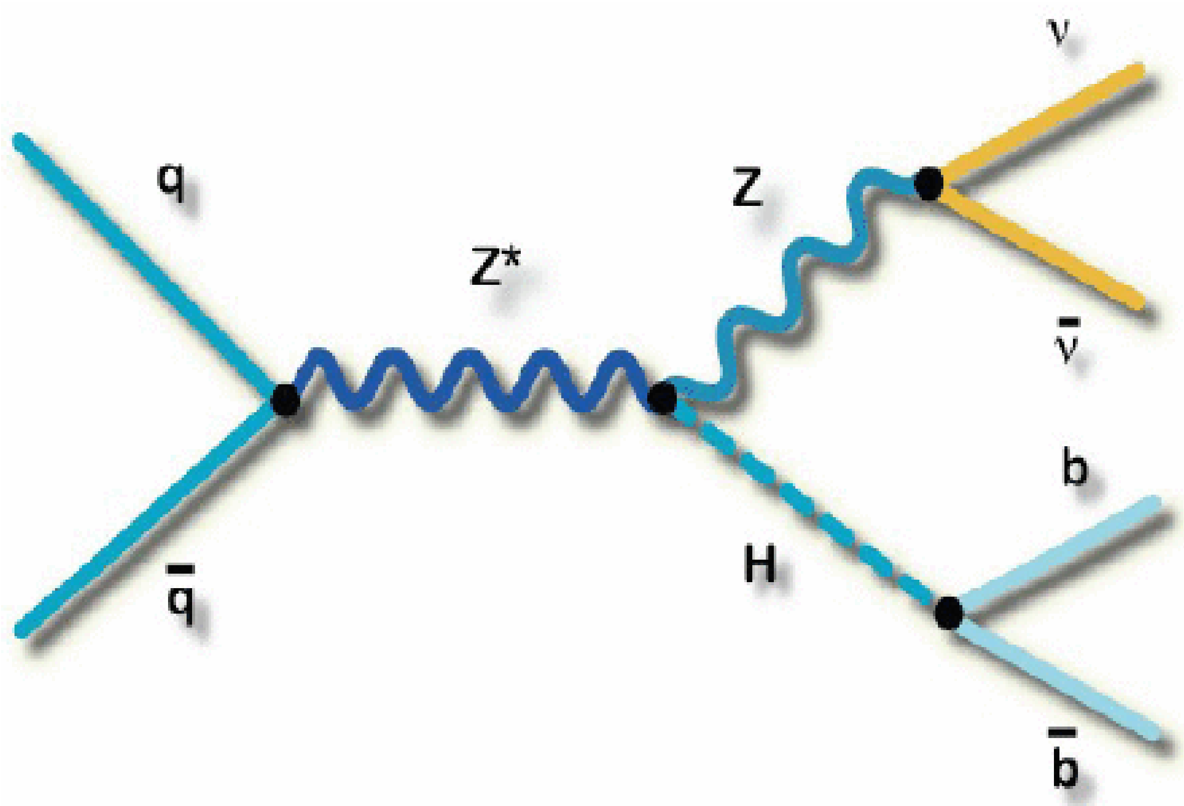
For b jets associated to a muon



SLTNN is part of the official b-id tools release:

- **Provide SF and TRF** (as well as direct b-tagging)
- **Systematics are computed**

Analysis Method



• see Murilo's talk later today about the details on Physics background and analysis procedures

• in this talk, we concentrate on the SLTNN results in the ZH→ννbb analysis

NN vs. SLTNN performances comparison

Dataset used for this study is 3 fb-1 of run IIb data

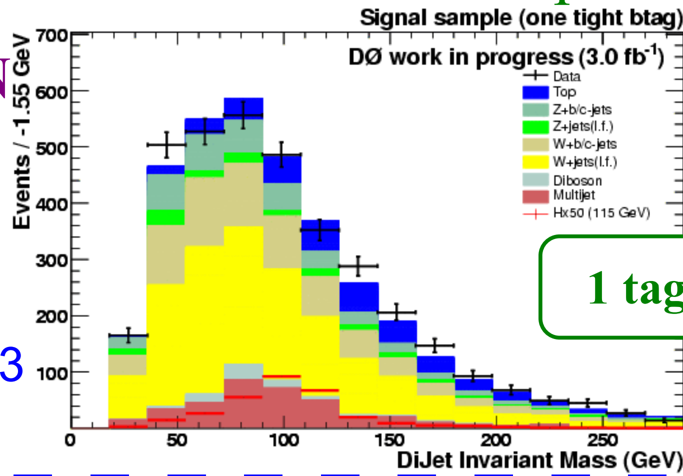
(it's a feasability study, the conclusion would remain the same with more data)

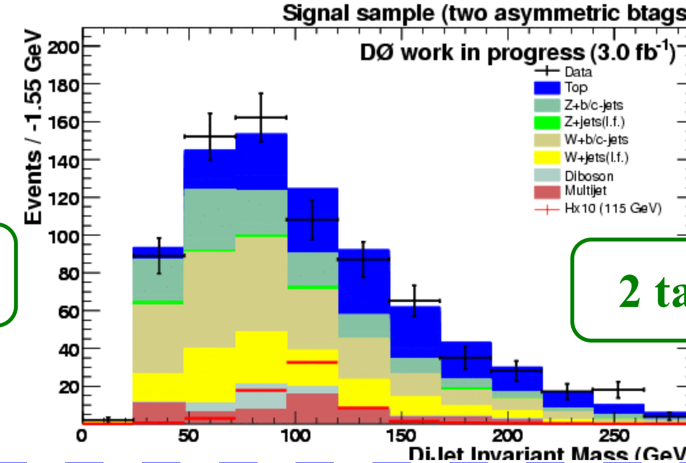
NN vs. SLTNN results

This is the bb invariant mass plot after all selection (including b-tagging)

Standard NN (all jets)

N data = 3588
N bkg = 3527
N sig = 5.9
S/sqrt(B) = 0.113
(0.103)

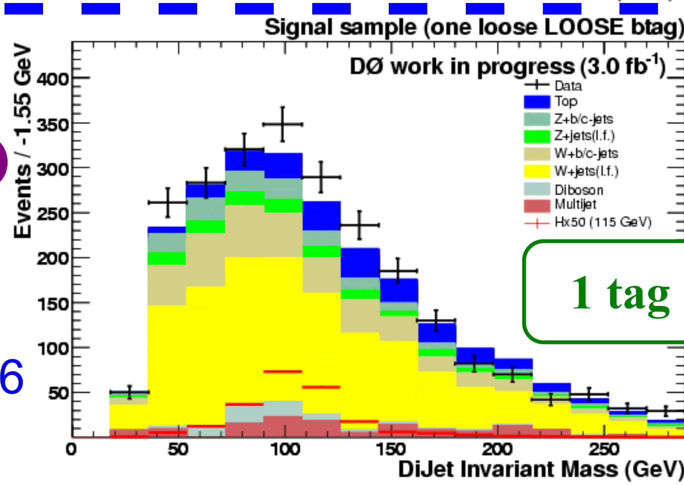


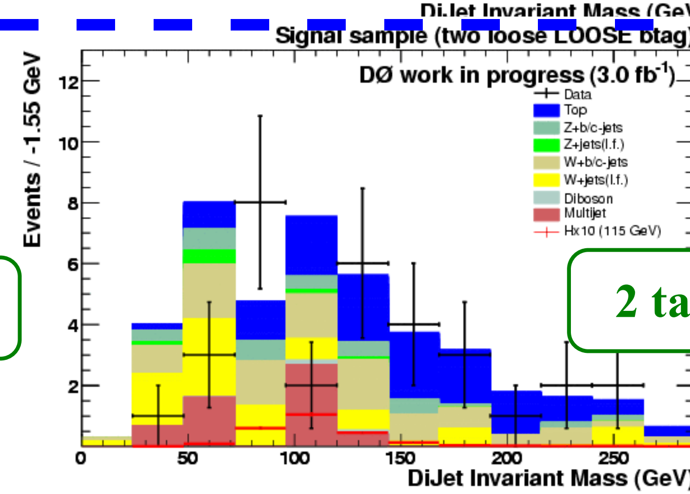


N data = 783
N bkg = 795
N sig = 6.3
S/sqrt(B) = 0.290
(0.231)

SLTNN (muon in jets)

N data = 2466
N bkg = 2380
N sig = 4.3
S/sqrt(B) = 0.106
(0.091)





N data = 32
N bkg = 44
N sig = 0.2
S/sqrt(B) = 0.048
(0.037)

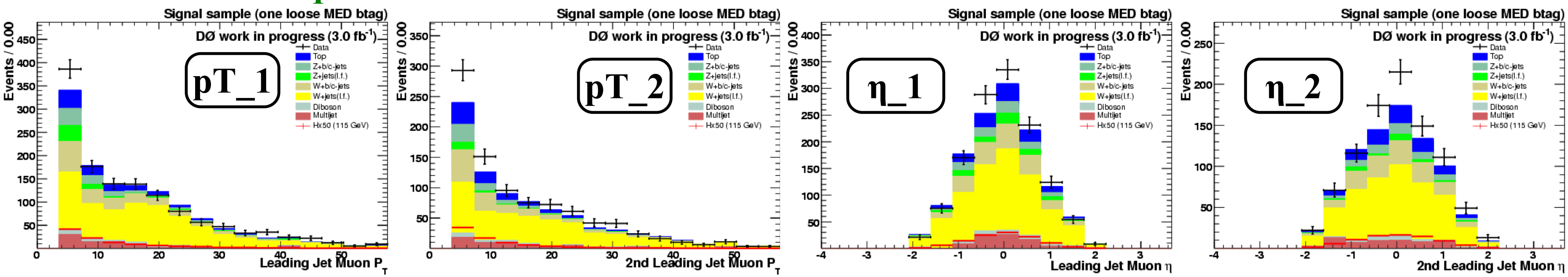
Conclusion:

- there is a slight disagreement for 1 tag SLTNN
- as expected, on the all jets samples, the NN sensitivity is better than SLTNN (the semileptonic branching fraction is 40% only to start with)

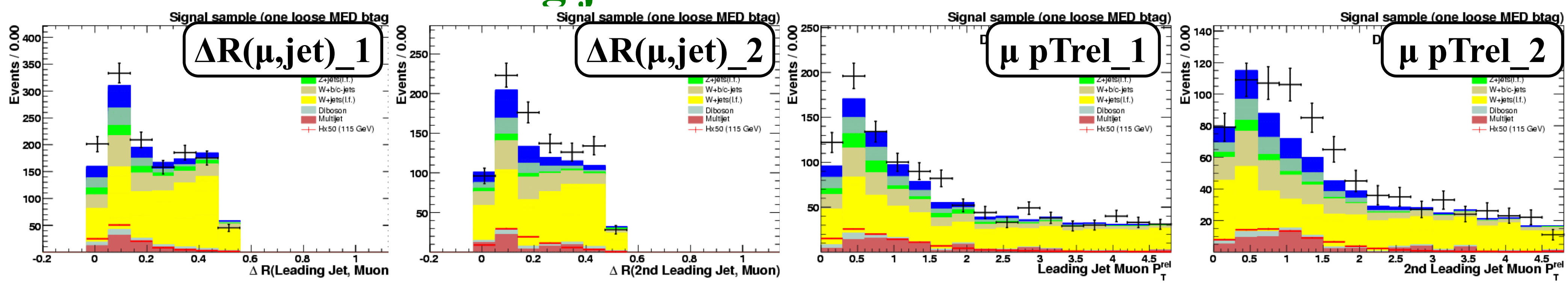
→ Let's now restrict to muon in jet sample only (leaving the non-muonic jets to NN)

Intermede: data/MC closer look

- for low P_T values and $\eta \sim 0$

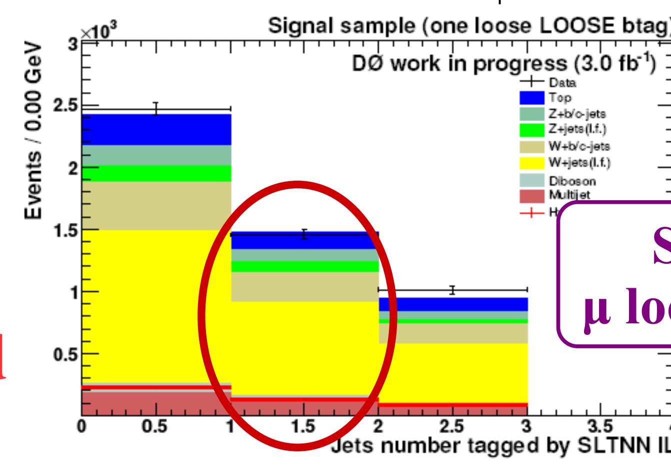


- for Next-To-Leading jets



Conclusion:

- most of the data/MC problem from the second leading jet
- just for illustration, using events which have the leading jet tagged (i.e. loose 40% of the stat)

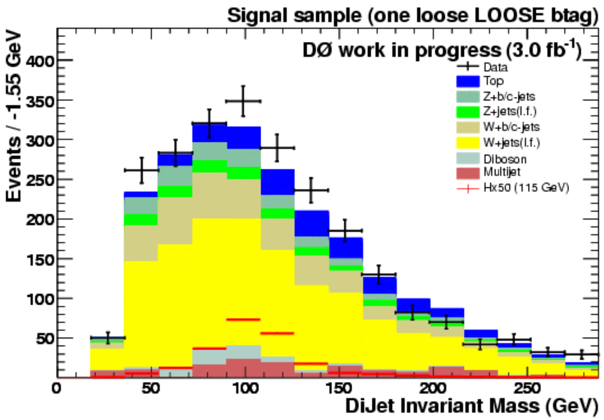


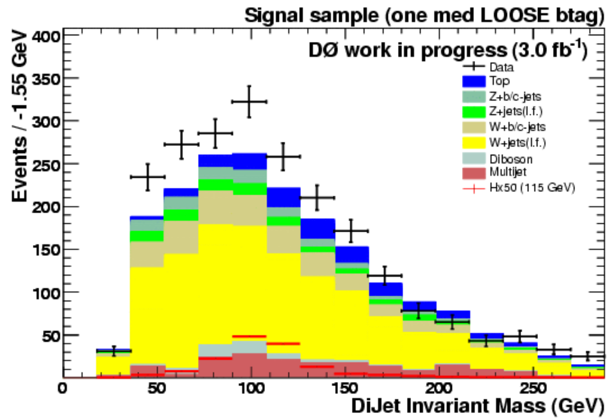
SLTNN 1 tag
μ loose OP LOOSE

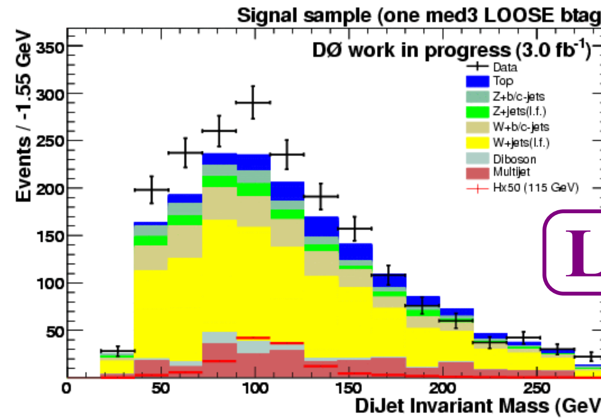
1 SLTNN tag muonic jet (either jet 1 or jet 2)

loose | med | med3

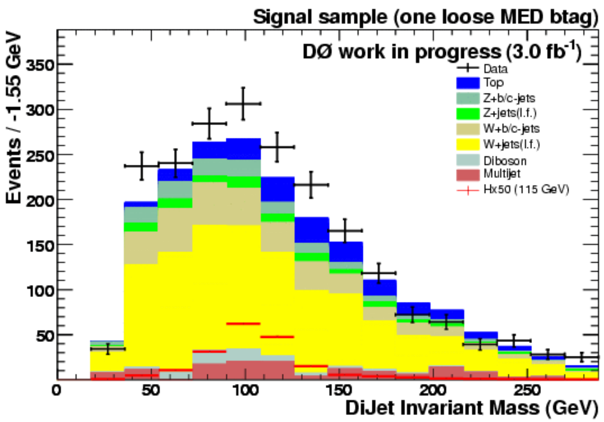
μ / OP

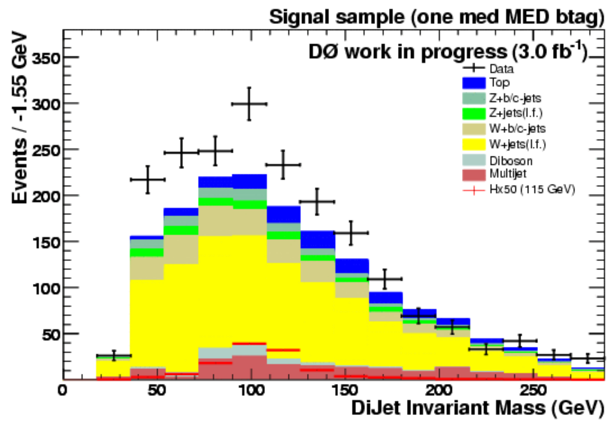


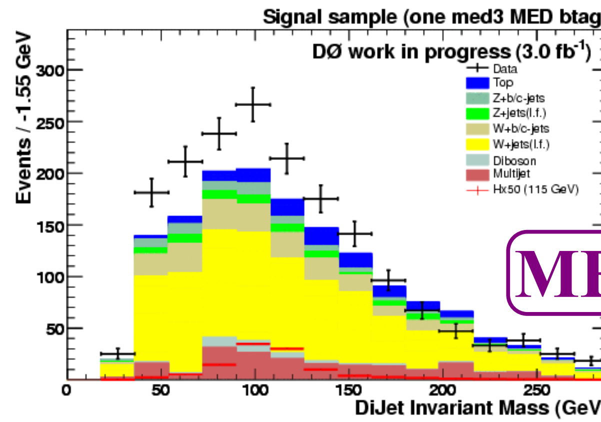




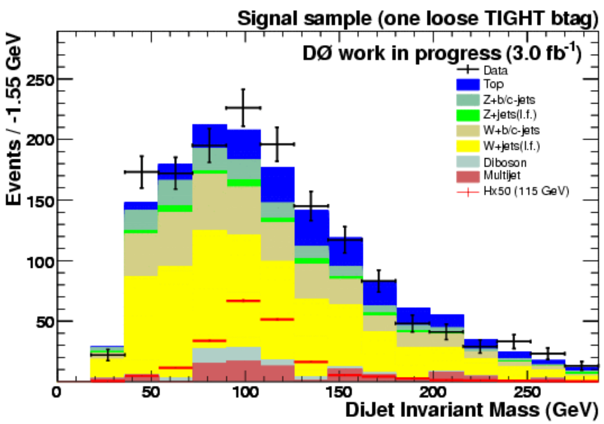
LOOSE

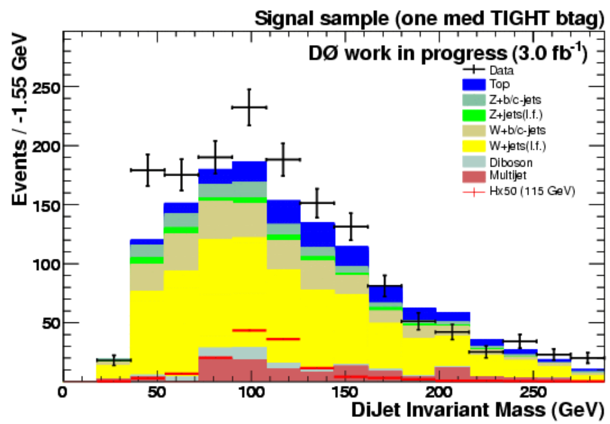


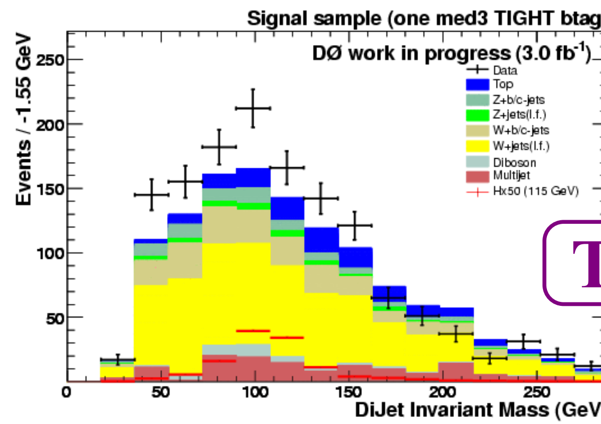




MEDIUM





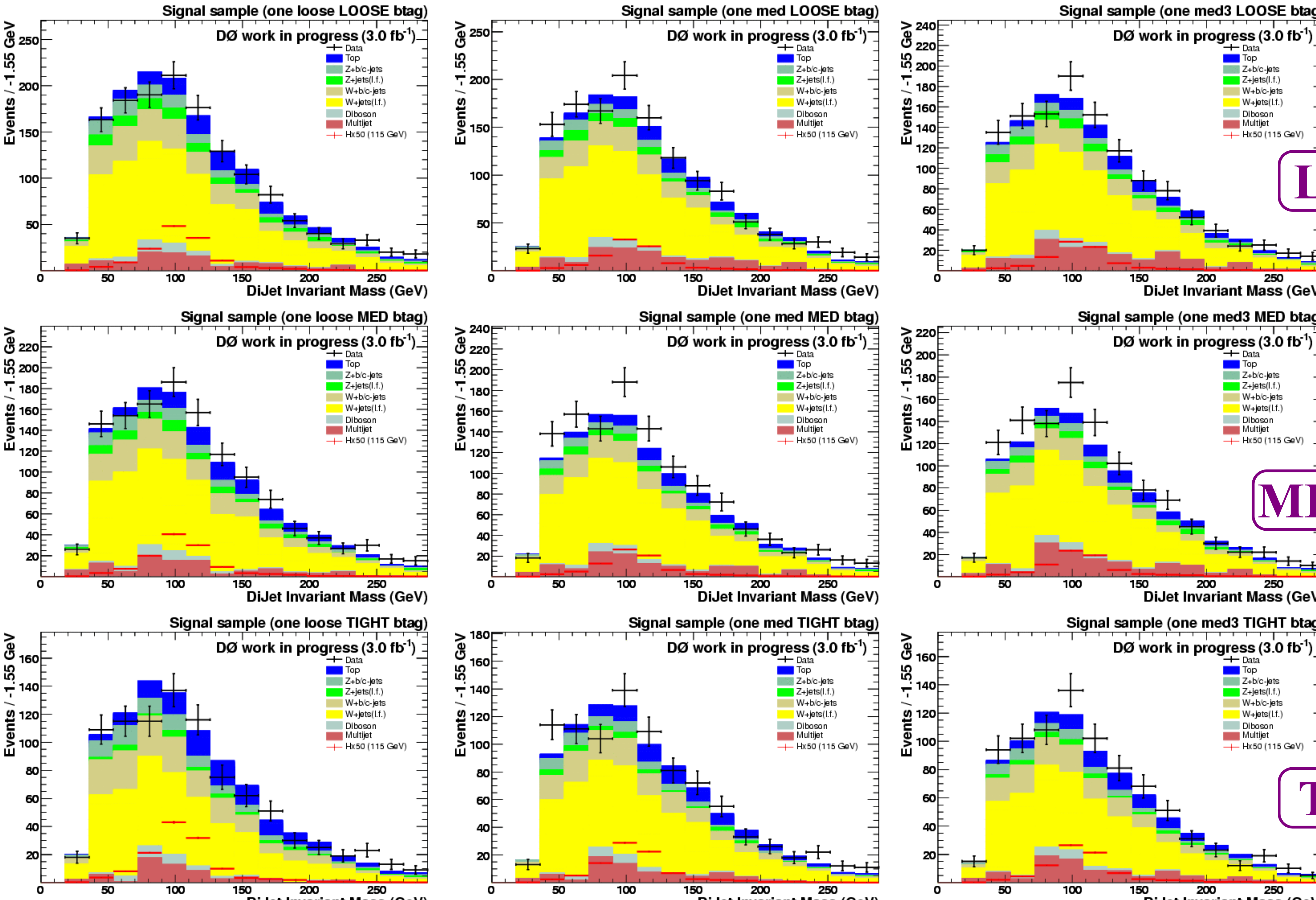


TIGHT

1 SLTNN tag muonic jet using jet 1

μ / OP

loose | med | med3



LOOSE

MEDIUM

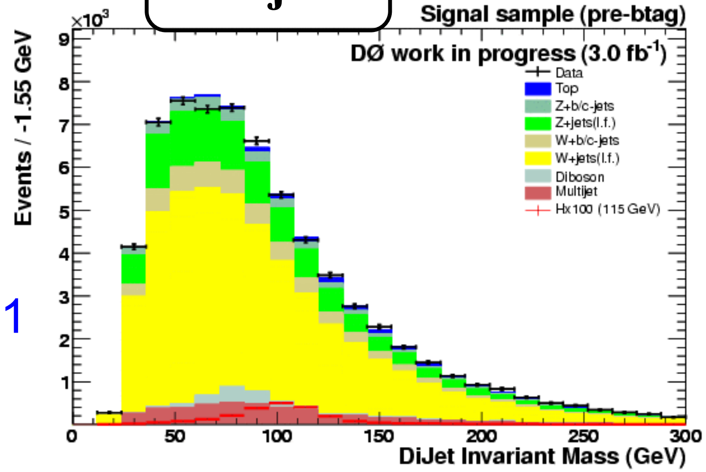
TIGHT

NN / SLTNN comparison strategy

All jets

2 jets events

Pre-bTag
N data = 68k
N bkg = 68k
N sig = 21.2
S/sqrt(B) = 0.101
(0.084)

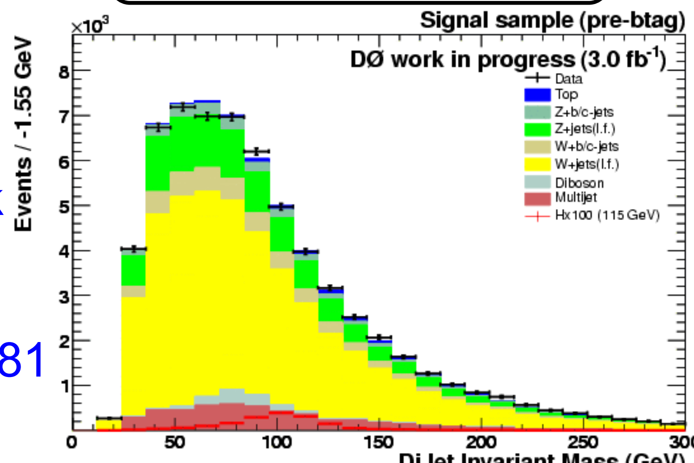


In this page, NN and SLTNN have not yet been used

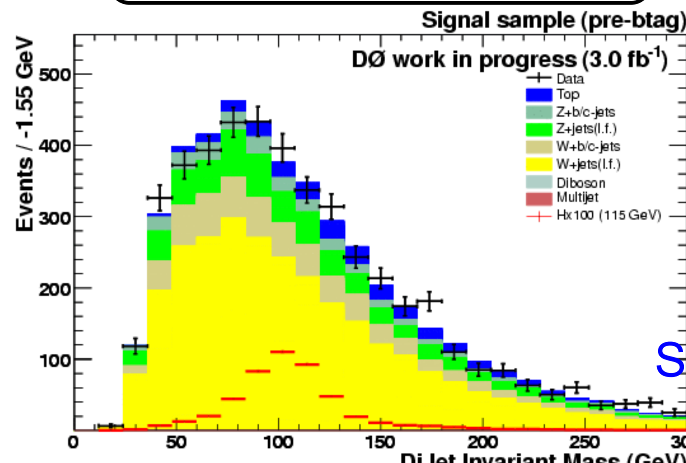
here the events have been sorted according Delta(mu,jet)<0.5 in two categories

NON-muonic
(ΔR (muon , Jet) > 0.5)

Pre-bTag
N data = 64k
N bkg = 64k
N sig = 16.6
S/sqrt(B) = 0.081
(0.068)



muonic
(at least 1 jet has a muon)



Pre-bTag
N data = 4626
N bkg = 4626
N sig = 4.7
S/sqrt(B) = 0.084
(0.071)

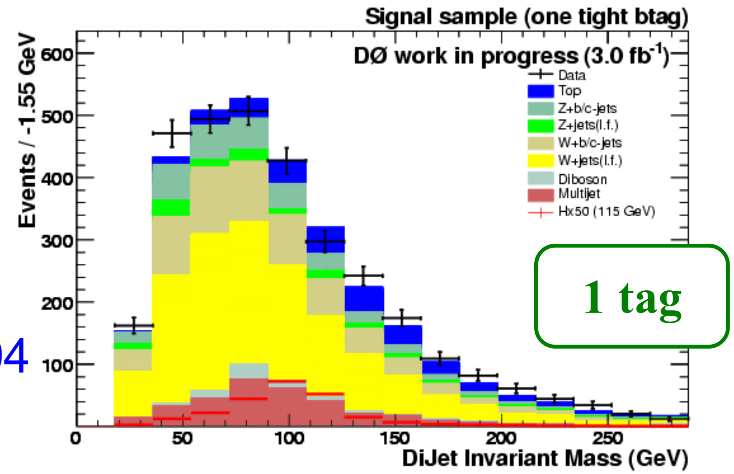
this sample will be for the NN analysis

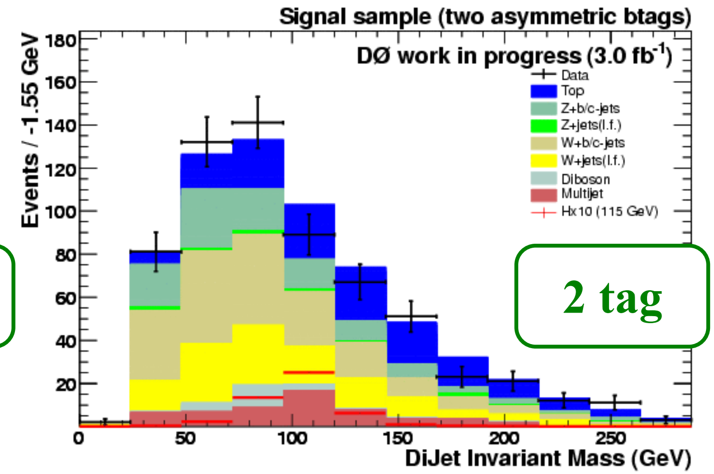
let's compare NN and SLTNN for this "muonic sample"

NN results

NON-muonic sample

N data = 3186
N bkg = 3129
N sig = 4.7
S/sqrt(B) = 0.094
(0.086)

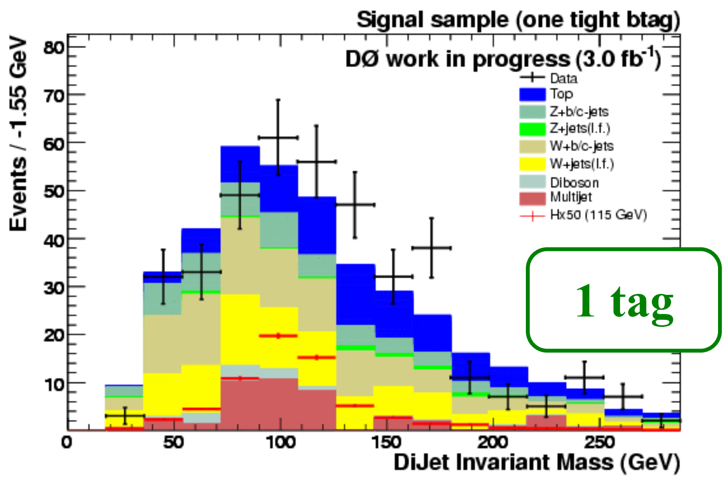


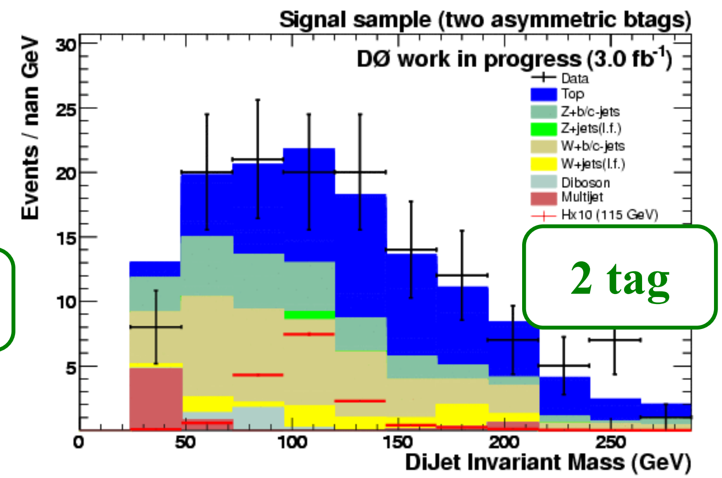


N data = 646
N bkg = 658
N sig = 4.8
sqrt(B) = 0.242
(0.193)

muonic sample

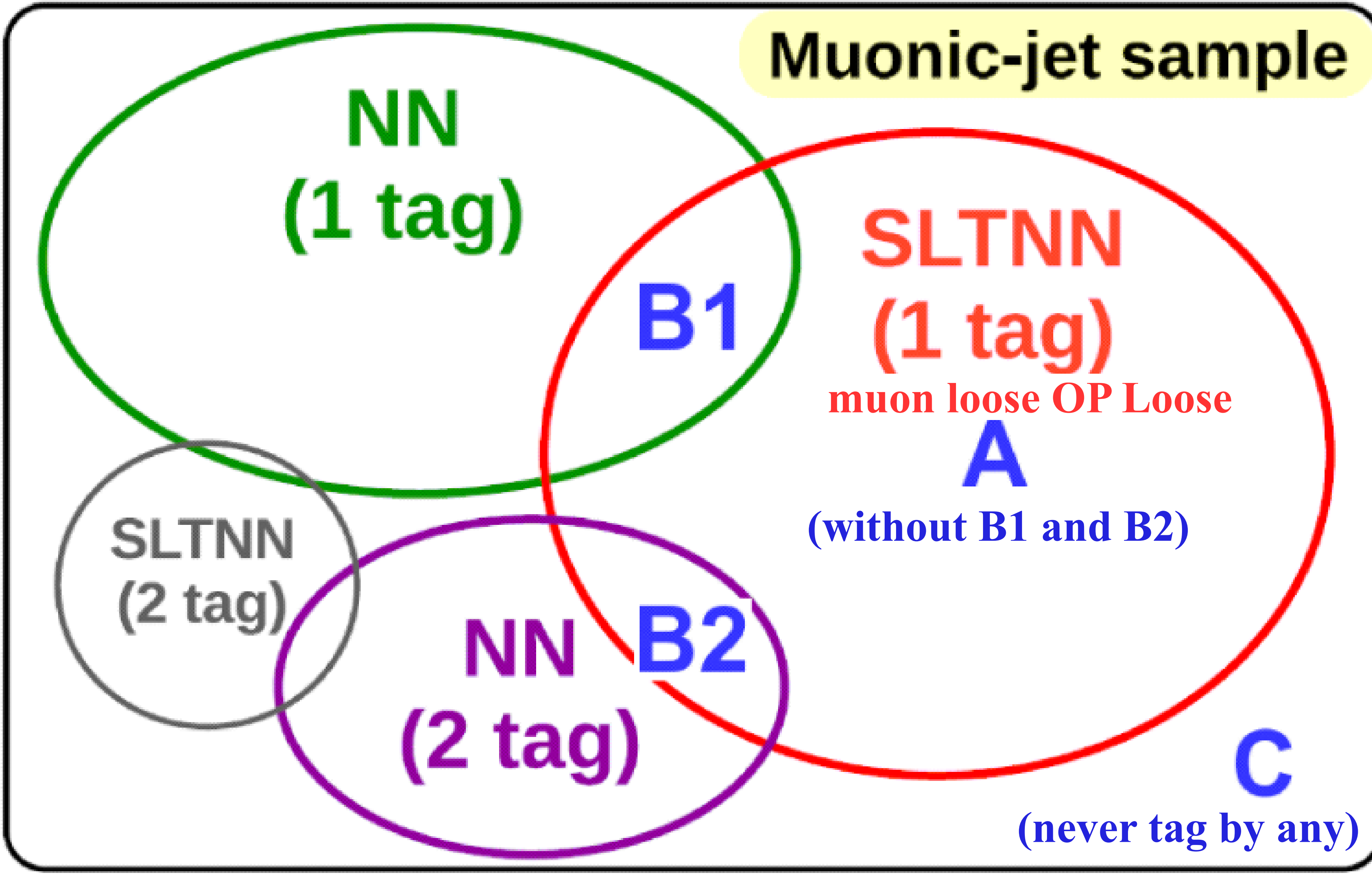
N data = 407
N bkg = 401
N sig = 1.3
S/sqrt(B) = 0.07
(0.065)



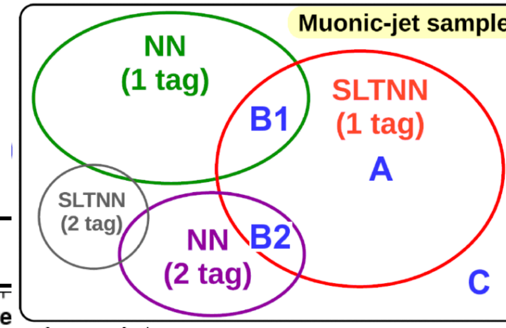


N data = 138
N bkg = 139
N sig = 1.5
/sqrt(B) = 0.167
(0.132)

→ good agreement from NN on the muonic and non-muonic samples

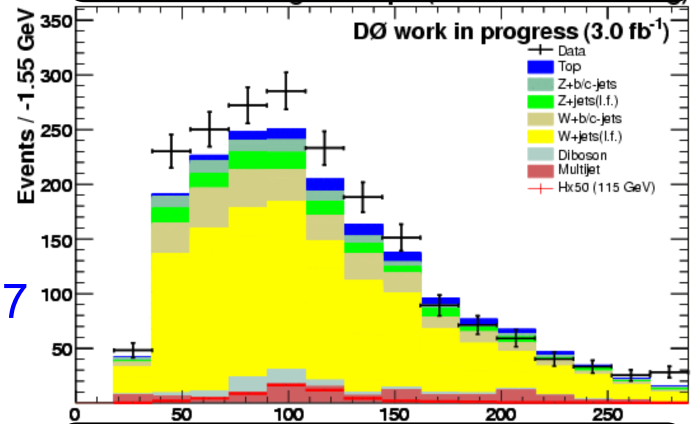

Muonic-jet sample
NN
(1 tag)
SLTNN
(1 tag)
muon loose OP Loose
B1
A
(without B1 and B2)
SLTNN
(2 tag)
NN
(2 tag)
B2
C
(never tag by any)

NN / SLTNN in the muonic sample (1)



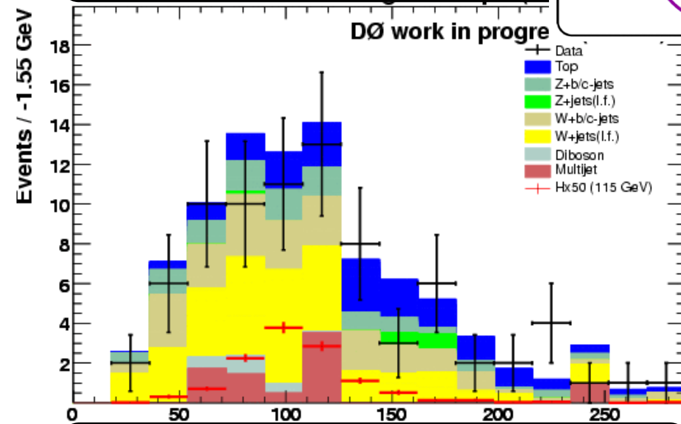
Case A

N data = 2050
N bkg = 1877
N sig = 1.0
S/sqrt(B) = 0.027
(0.025)



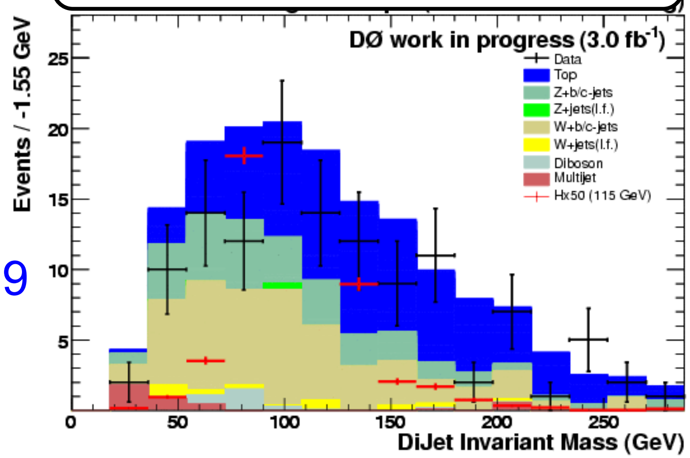
Case B1

N data = 81
N bkg = 91
N sig = 0.2
S/sqrt(B) = 0.029
(0.025)



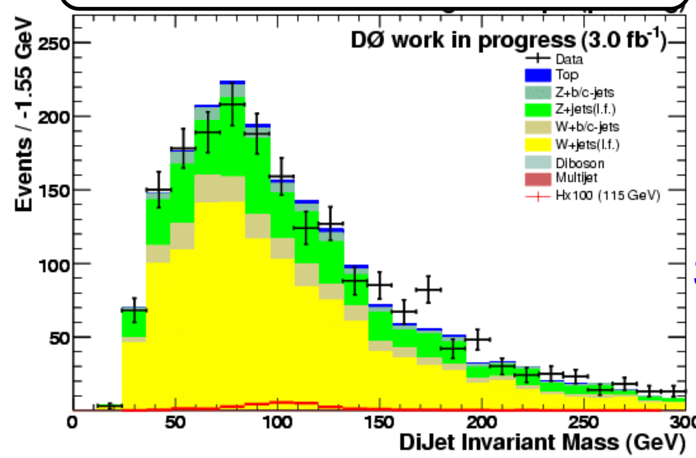
Case B2

N data = 124
N bkg = 166
N sig = 2.0
S/sqrt(B) = 0.209
(0.164)



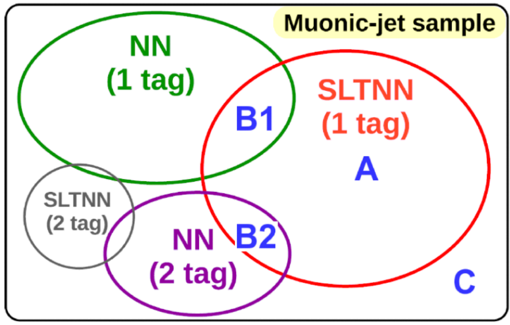
Case C

N data = 2014
N bkg = 2014
N sig = 0.3
S/sqrt(B) = 0.007
(0.006)



- Case A (1 tag only by SLTNN) → 1 signal event recovered
  → overwhelming background!
  → Data/MC seen before
- B1 and B2 are covered by either SLTNN or NN → nothing to gain there
  → most efficient case b.t.w

NN / SLTNN in the muonic sample (2)



All SLTNN OP case A

Sample	SLTNN 1 tagged events only								
	lL	lM	lT	mL	mM	mT	m3L	m3M	m3T
Top	82.2	67.7	61.6	55.0	46.4	50.1	48.5	41.4	46.0
Di-boson	57.3	45.4	32.6	35.1	32.2	29.4	29.0	27.2	26.2
Zjj	117.5	85.5	34.7	46.3	46.2	36.5	37.0	37.7	32.7
Zbb/cc	83.8	63.4	49.1	43.6	36.7	37.3	36.2	30.8	34.4
Wjj	1169.0	1000.4	662.2	795.9	778.4	689.1	673.6	674.0	590.3
Wbb/cc	246.7	199.0	161.7	157.2	139.7	139.1	137.6	124.5	125.7
Phy. Bkgd	1756.5	1461.4	1002.0	1133.1	1079.6	981.5	961.8	935.6	855.3
Instr. Bkgd	120.5	112.4	66.7	108.7	111.2	88.8	119.9	118.9	112.5
Tot. Bkgd	1877.0	1573.8	1068.7	1241.8	1190.8	1070.3	1081.7	1054.5	967.8
DATA	2050	1776	1175	1439	1420	1238	1238	1228	1106
signal (115 GeV)	1.0	0.9	0.8	0.7	0.6	0.6	0.6	0.5	0.6
$S/\sqrt{B}$	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02

- best DATA/MC agreement for μ loose OP LOOSE and TIGHT
  → ~1.0 signal events gain (in this cases)
- low sensitivity

CLFast Limit results

Sample	Limit	Higgs boson mass (GeV)										
		100	105	110	115	120	125	130	135	140	145	150
Std analysis	Expected	3.97	4.31	4.84	5.56	6.52	7.97	10.11	13.30	18.36	27.14	42.57
	Observed	4.26	4.18	4.35	4.72	5.47	6.70	8.98	12.67	18.78	30.29	51.17
Std analysis + case A (lL)	Expected	3.94	4.37	4.82	5.40	6.57	7.94	10.06	13.35	18.41	26.91	42.84
	Observed	4.77	4.77	4.86	5.36	6.16	7.54	10.30	14.34	21.28	34.02	56.77
Std analysis + case A (lT)	Expected	3.94	4.28	4.82	5.52	6.52	7.88	10.04	13.27	18.21	27.14	42.32
	Observed	4.64	4.64	4.78	5.22	6.09	7.42	9.92	13.93	20.42	33.08	55.09

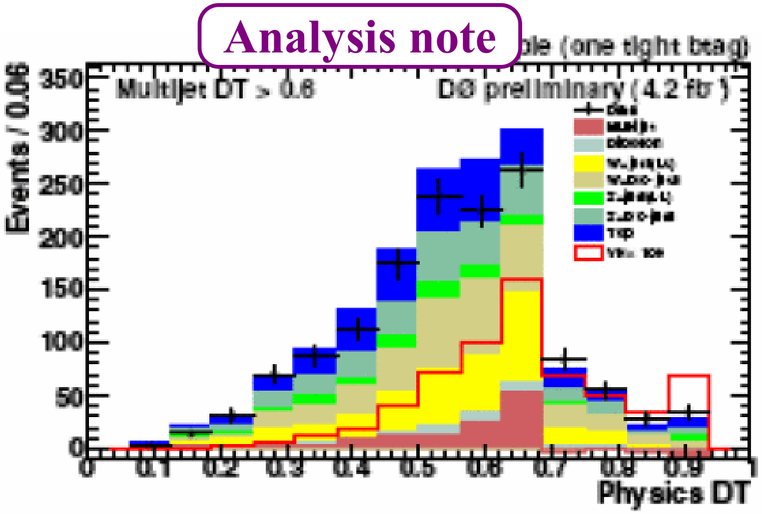
Sad conclusion: ~1% gain in the expected limit when adding the case A (SLTNN)

We also tried two other ideas :

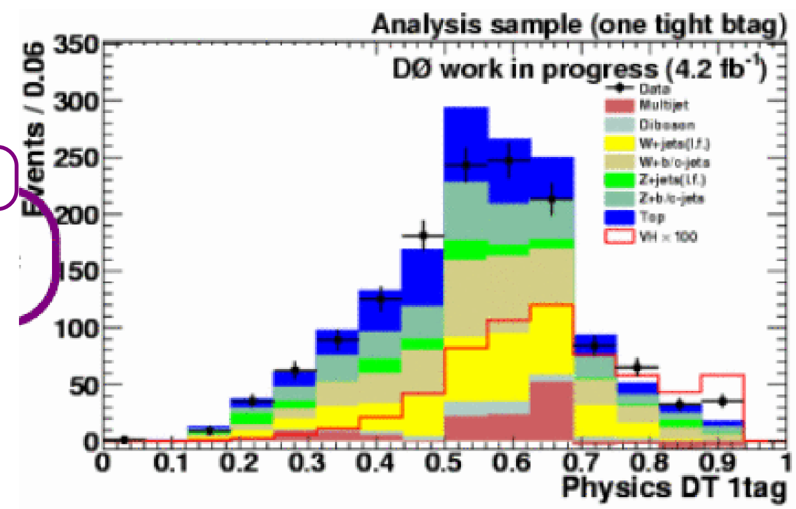
- **add the SLTNN info in the selection decision tree**
- **use muon+jets triggers (orthogonal sample to jet+met)**

Adding SLTNN in the DT selection tree

- Add new information in Physics DT : number of jets tagged by SLTNN
- Physics DT ouput (Multi-jet DT cut 0.6):



Add new var



→ **CLFast expected limit (1 tag) : 9.0** (9.6 initially)

Conclusion:

- **6% improvement in sensitivity**
- **Problem: NN/SLTNN correlation hard to evaluate**

muon+jets triggers

- The ZH→ννbb analysis **uses jet+met triggers**

- Let's add muon+jets triggers (and veto jet+met events)
- This orthogonal sample would have to worry about the correlation between NN and SLTNN (we use only NN there..)

<u>• Problem:</u>
single muon and muon+jets triggers efficiencies are derived for **muon pT > 15 GeV and isolated muon**

→ cannot be used

→ project: redirived the muon+jets efficiencies for non isolated muon (too late for me)

CONCLUSION

- SLTNN gives a small improvement (limit)

- Using the SLTNN information in the selection decision tree gain 6% there are correlation beetween NN and SLTNN which would be difficult to correctly evaluate

- adding the mu+jets triggers is under studies: need to redirive the trigger efficiencies for non isolated muon

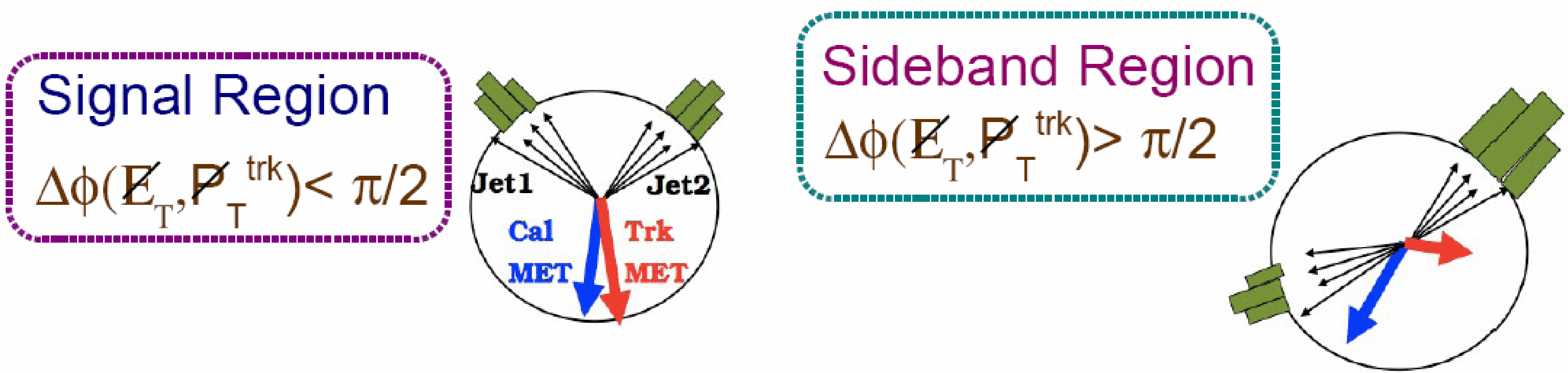
- I am now writing my PhD thesis. The tentative date for a defense : September 30, 2010

BACK-UP

Multijet Modeling

In signal like events, MET from calorimeter and tracks is aligned.
It is not the case for multijet events (jet energies mis-measured)

→ modeling multijet events in **DATA sideband region**
(region where MET from calorimeter and tracks is not aligned)

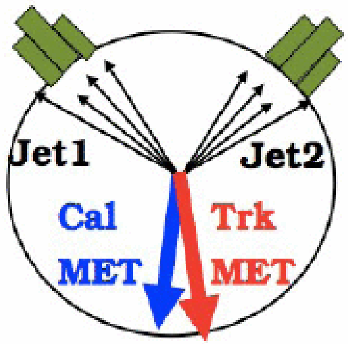


- Physics MC sideband contributions are substracted from DATA sideband

- We define a SF for remaining multijet bkgd to match DATA in signal region

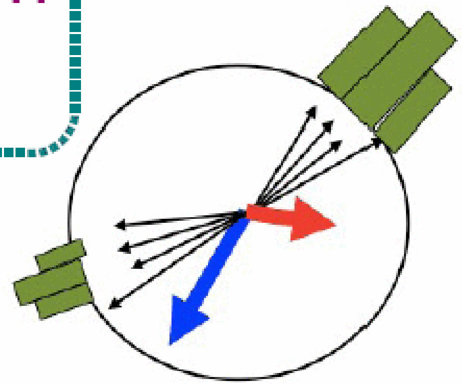
Signal Region

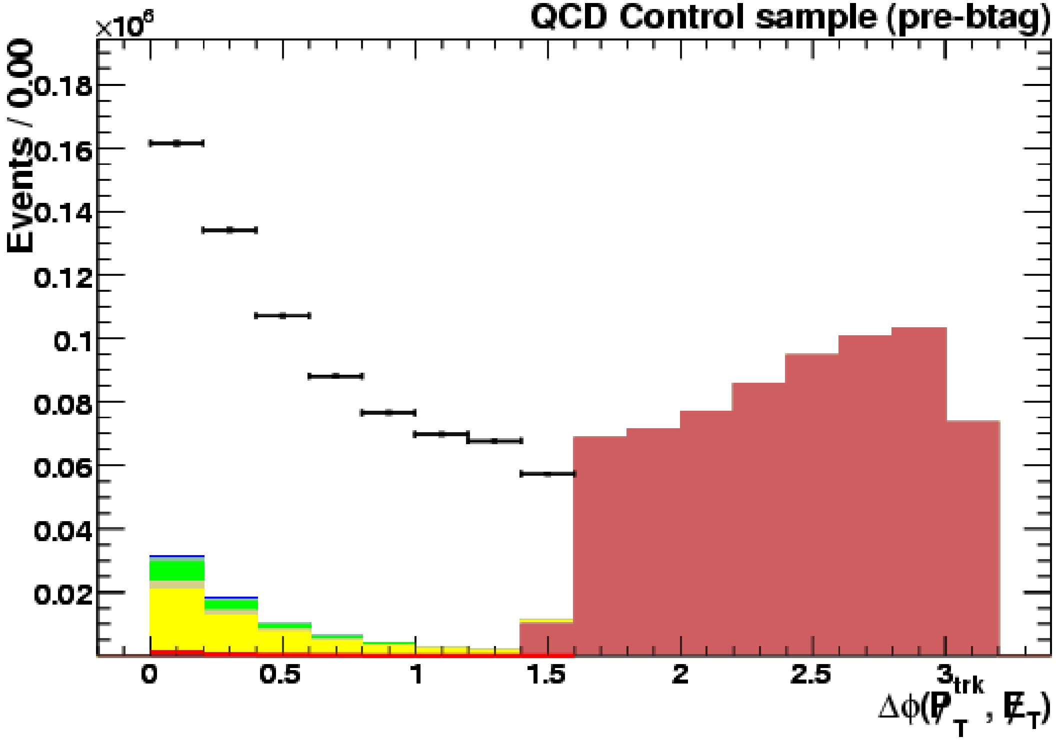
$\Delta\phi(\not{E}_T, \not{P}_T^{trk}) < \pi/2$

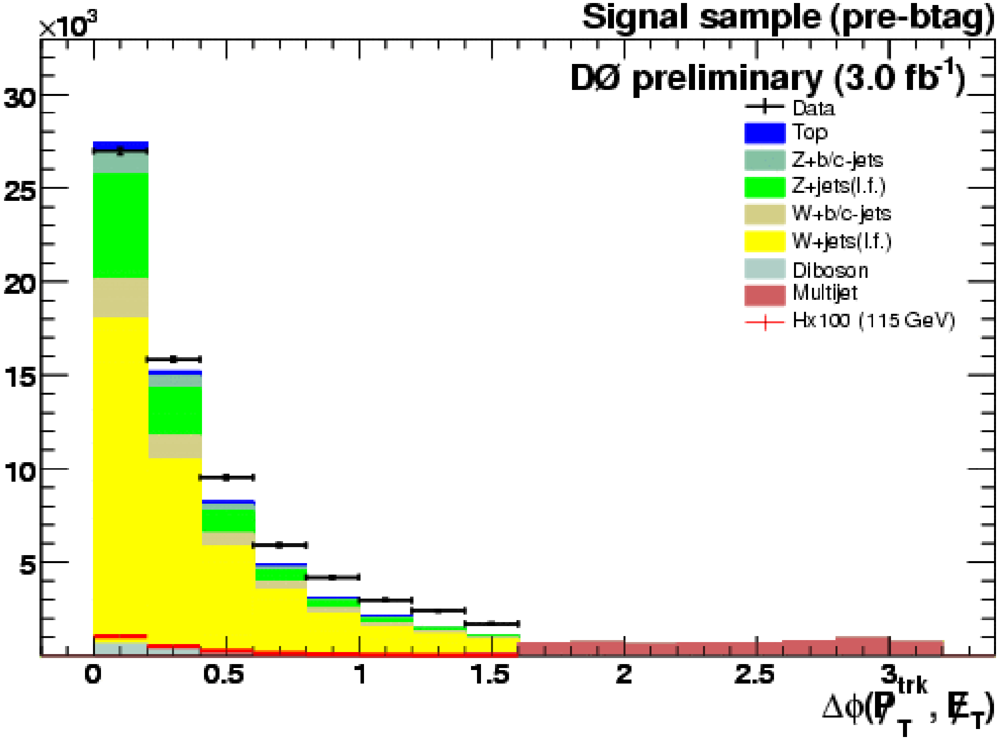


Sideband Region

$\Delta\phi(\not{E}_T, \not{P}_T^{trk}) > \pi/2$







- analysis : so far 3 fb-1 (moriond 2009 dataset)

All plots :
http://marwww.in2p3.fr/~jamin/Higgs/moriond09_version_mucorr/hznunubb_09032009/

SF's :
(in addition to K-factors)

0.97 # EW control sample preBtag
→ put to 1.0 in the following

SF	Sample	Btag	Group
1.30 #	QCD enriched	preBtag	
1.41 #		oneTightBtag	NN
1.44 #		twoAsymmBtag	NN
1.33 #		onelooseLOOSEBtag	SLTNN
1.44 #		twolooseLOOSEBtag	SLTNN
1.13 #	analysis	preBtag	
1.22 #		oneTightBtag	NN
1.24 #		twoAsymmBtag	NN
1.15 #		onelooseLOOSEBtag	SLTNN
1.24 #		twolooseLOOSEBtag	SLTNN

implement muon ID correction for mu-tagged jets in the analysis

Muon correction :

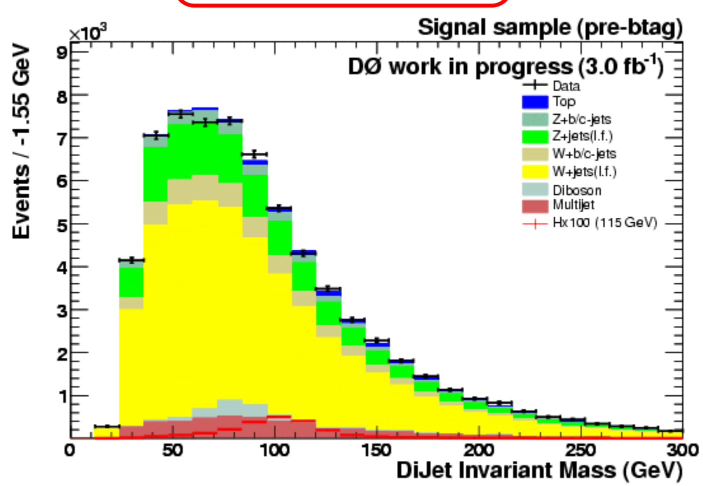
from Phi and dEta of muon
z and dEtaCFT of track
compute muon ID SF thanks to muid_eff package -> gives DATA and MC efficiencies

- 2 jets with muon inside : SF[jetA]*SF[jetB]
- 1 jet with muon inside : SF[jet with muon]
- 2 jets with muon inside but only one muon reconstructed :

½ { (SF[A]*(1-effDATA[B])/(1-effMC[B]) + SF[B]*(1-effDATA[A])/(1-effMC[A]) }

All jets

Pre-bTag
N data = 68k
N bkg = 68k
N sig = 21.2
S/sqrt(B) = 0.101
(0.084)



• Add new information in Physics DT : number of jets tagged by SLTNN