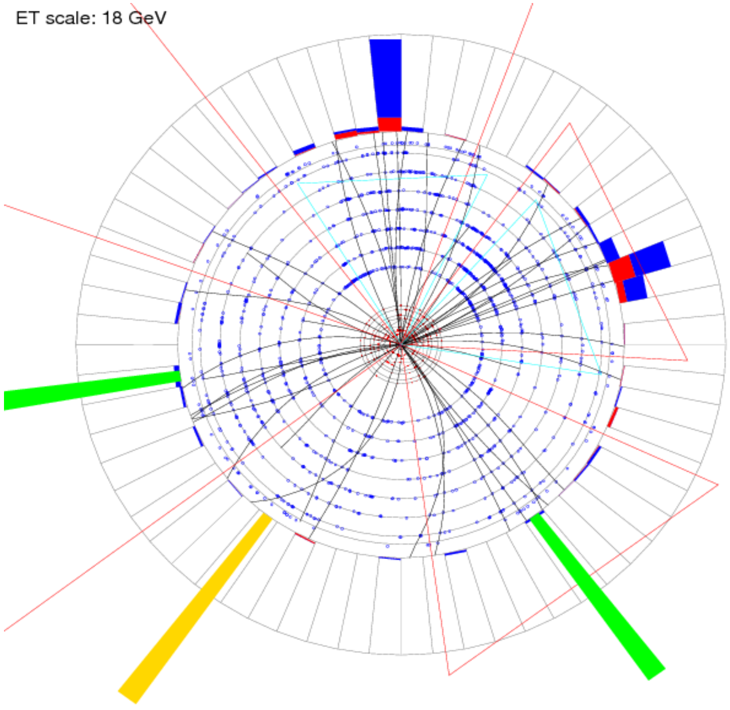


p17 ZH -> mumu bb w/ Neural Net Event Selection

Andy Haas
Columbia University

Higgs Workshop
April 12, 2007

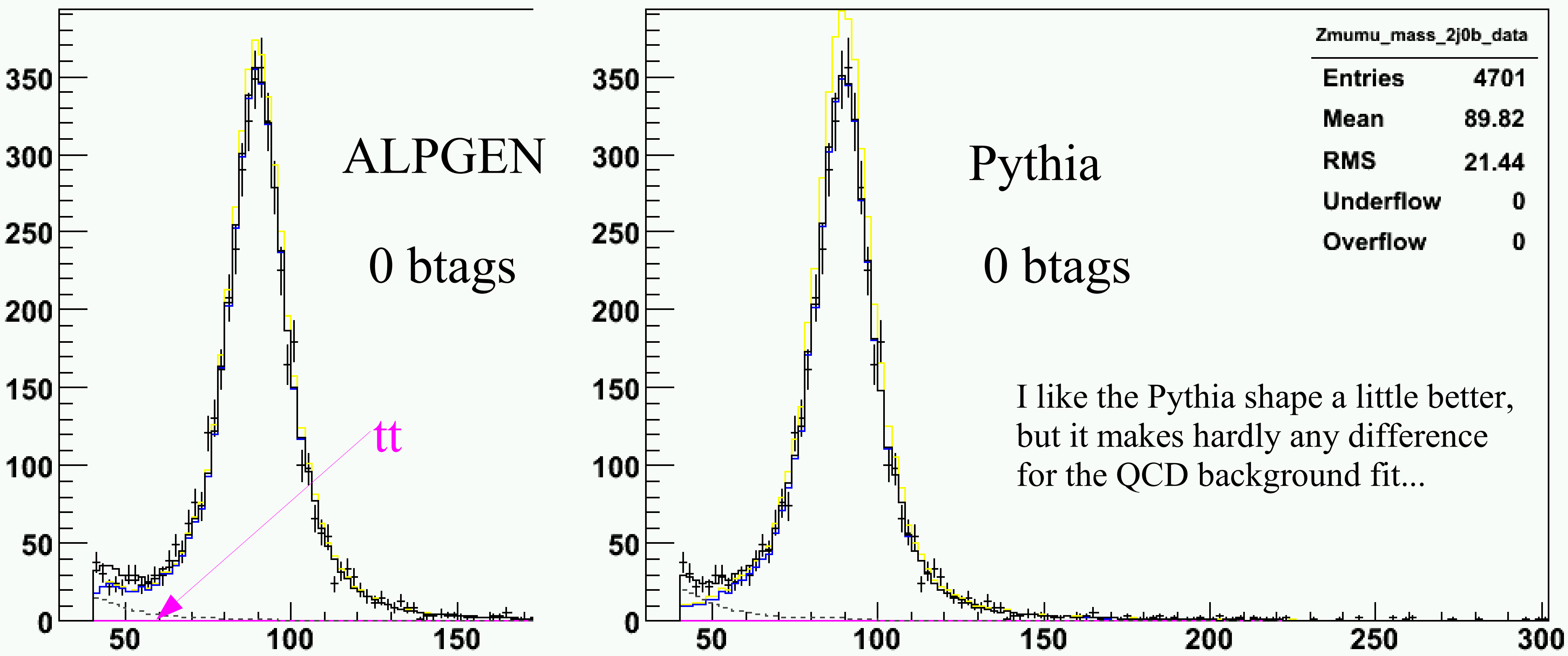
Analysis Updates

- Now using all available MC samples (mostly backgrounds)
- Keeping track of all these numbers and typing them into a table has been lots of fun :)
- Use ALPGEN as main MC source for Z+jets, but have Pythia as a cross-check

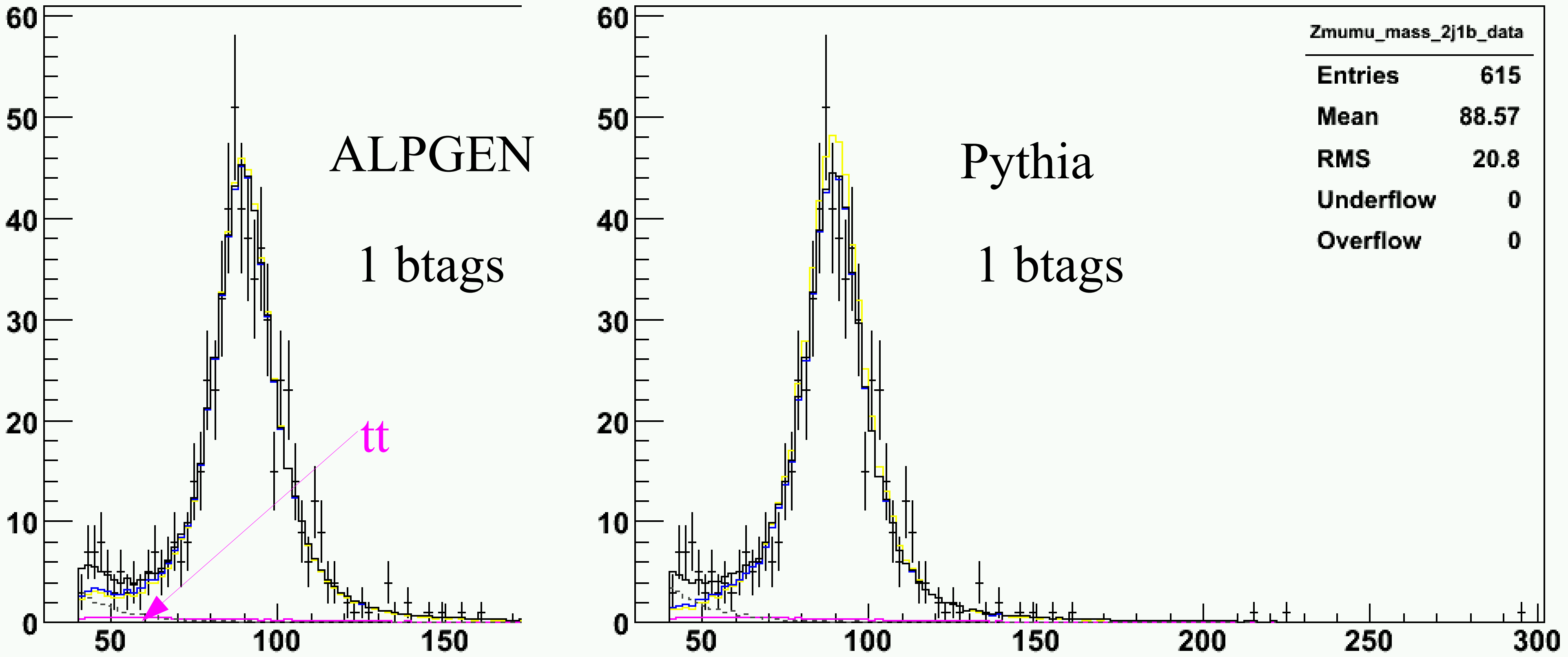
Sample	Cross section x BR	SAM req-id's	N. Events
ZH($\to\mu^+\mu^- b\bar{b}$), m_H=105 GeV	4.0 fb	28575	50000
ZH($\to\mu^+\mu^- b\bar{b}$), m_H=115 GeV	2.8 fb	28576	50000
ZH($\to\mu^+\mu^- b\bar{b}$), m_H=125 GeV	1.8 fb	28577	50000
ZH($\to\mu^+\mu^- b\bar{b}$), m_H=135 GeV	1.1 fb	28578	52250
ZH($\to\mu^+\mu^- b\bar{b}$), m_H=145 GeV	0.50 fb	32426	49750
ZH($\to\mu^+\mu^- b\bar{b}$), m_H=155 GeV	0.20 fb	32427	50000
Z($\to\mu^+\mu^-$)+0lp excl. (mZ=15-60 GeV)	335 * 1.3 fb	32900, 32901, 32902	650250
Z($\to\mu^+\mu^-$)+1lp excl. (mZ=15-60 GeV)	38.5 * 1.3 fb	32903, 32904, 32905	512500
Z($\to\mu^+\mu^-$)+2lp excl. (mZ=15-60 GeV)	10.1 * 1.3 fb	32906, 32907	205000
Z($\to\mu^+\mu^-$)+3lp incl. (mZ=15-60 GeV)	4.2/1.5 * 1.3 fb	32908, 32909	100500
Z($\to\mu^+\mu^-$)+0lp excl. (mZ=60-130 GeV)	139 * 1.3 fb	28762, 28884, 38410, 38409, 38408, 38364, 37409, 37408, 37407, 37004, 37003, 29279, 29278, 29273, 29272	1469750
Z($\to\mu^+\mu^-$)+1lp excl. (mZ=60-130 GeV)	41.6 * 1.3 fb	28763, 28885, 38362, 37416, 37005	874116
Z($\to\mu^+\mu^-$)+2lp excl. (mZ=60-130 GeV)	10.3 * 1.3 fb	28764, 38368, 37006	413500
Z($\to\mu^+\mu^-$)+3lp incl. (mZ=60-130 GeV)	5.3/1.5 * 1.3 fb	28767, 38372, 37007	317000
Z($\to\mu^+\mu^-$)+0lp excl. (mZ=130-250 GeV)	0.90 * 1.3 fb	32621	108000
Z($\to\mu^+\mu^-$)+1lp excl. (mZ=130-250 GeV)	0.36 * 1.3 fb	32623	100250
Z($\to\mu^+\mu^-$)+2lp excl. (mZ=130-250 GeV)	0.097 * 1.3 fb	32624	100750
Z($\to\mu^+\mu^-$)+3lp incl. (mZ=130-250 GeV)	0.053/1.5 * 1.3 fb	32648	99750
Z($\to\mu^+\mu^-$)+0lp excl. (mZ>=250 GeV)	0.072 * 1.3 fb	33738	106250
Z($\to\mu^+\mu^-$)+1lp excl. (mZ>=250 GeV)	0.036 * 1.3 fb	33739	105750
Z($\to\mu^+\mu^-$)+2lp excl. (mZ>=250 GeV)	0.011 * 1.3 fb	33740	102250
Z($\to\mu^+\mu^-$)+3lp incl. (mZ>=250 GeV)	0.0066/1.5 * 1.3 fb	33741	105750
Z($\to\mu^+\mu^-$)+2b+0lp excl.	0.97 * 1.5 fb	32356, 32806, 32807	342000
Z($\to\mu^+\mu^-$)+2b+1lp excl.	0.36 * 1.5 fb	32357	52750
Z($\to\mu^+\mu^-$)+2b+2lp incl.	0.21/1.5 * 1.5 fb	32257	25000
$t\bar{t}$ inclusive (M_{top}=170-175 GeV)	7.0 fb	35833, 35834, 35835, 35437, 35438, 35439, 34873, 34874, 33810, 33811	1615030
WZ inclusive	3.6 fb	30488, 30489, 33685, 33684, 42212, 38491	724250
ZZ inclusive	1.423 fb	30486, 30487, 33687, 33686, 42213, 38492	711000
Pythia Z($\to\mu^+\mu^-$) (mZ=15-60 GeV)	479139.25 fb	40687 40686 40685 40684 40683, 40682 40681 40680 40679 40678, 36595 36594 36593 36592 36591	3107250
Pythia Z($\to\mu^+\mu^-$) (mZ=60-130 GeV)	255203.33 fb	38912 38911 38910 38909 38908, 38907 38906 38905 38904 38903, 38902 38901 38900 38899 38898	2821282
Pythia Z($\to\mu^+\mu^-$) (mZ=130-250 GeV)	1875.78 fb	41252 41251	413521
Pythia Z($\to\mu^+\mu^-$) (mZ=250-500 GeV)	476. fb	41256	102750
Pythia Z($\to\mu^+\mu^-$) (mZ>=500 GeV)	24.7 fb	41259	51500

New QCD Fit

Use QCD shape from non-isolated data
Smear muons properly in MC to describe Z shape better
-> Add "ApplyMuonSmear" processor
Consider tt contribution while fitting
Apply re-weighting for ALPGEN bugs in 0,1 lp samples for mZ distributions



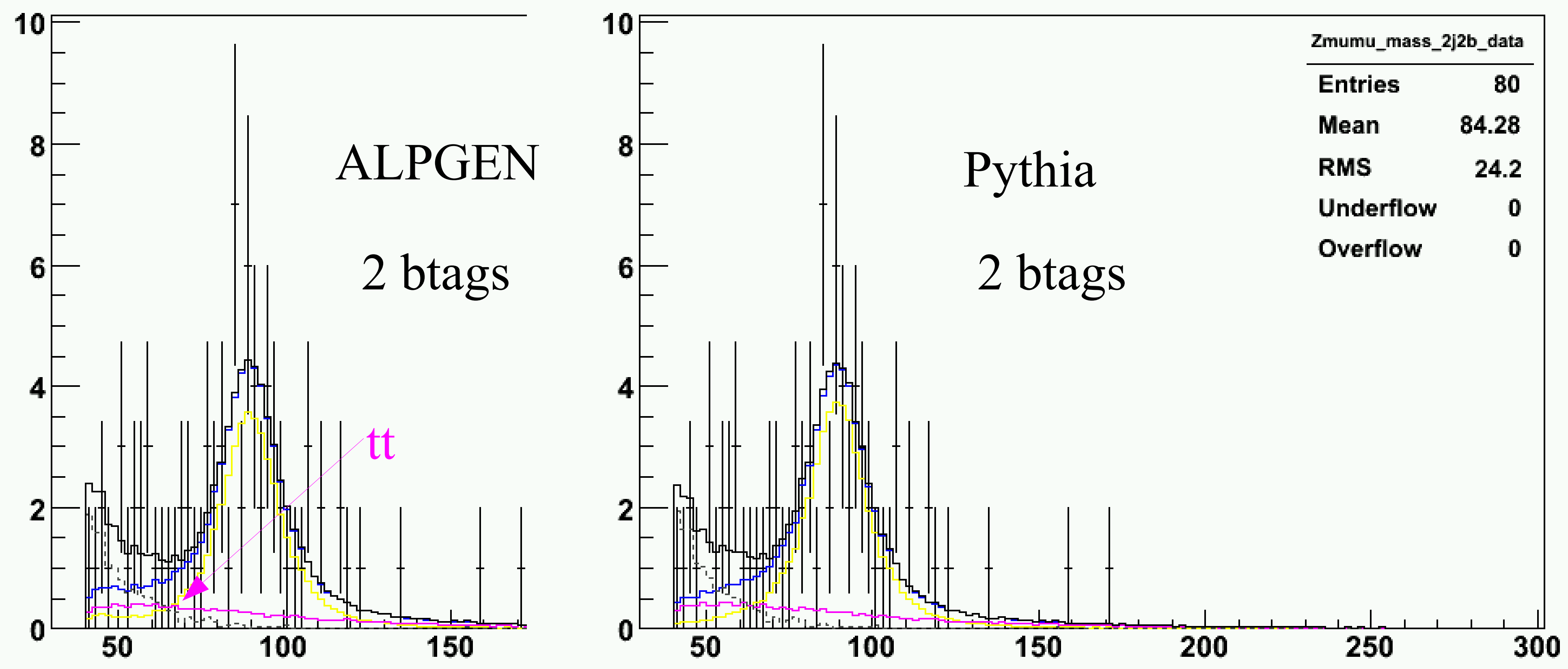
New QCD Fit



ALPGEN
1 btags
tt
Pythia
1 btags
Zmumu_mass_2j1b_data
Entries 615
Mean 88.57
RMS 20.8
Underflow 0
Overflow 0

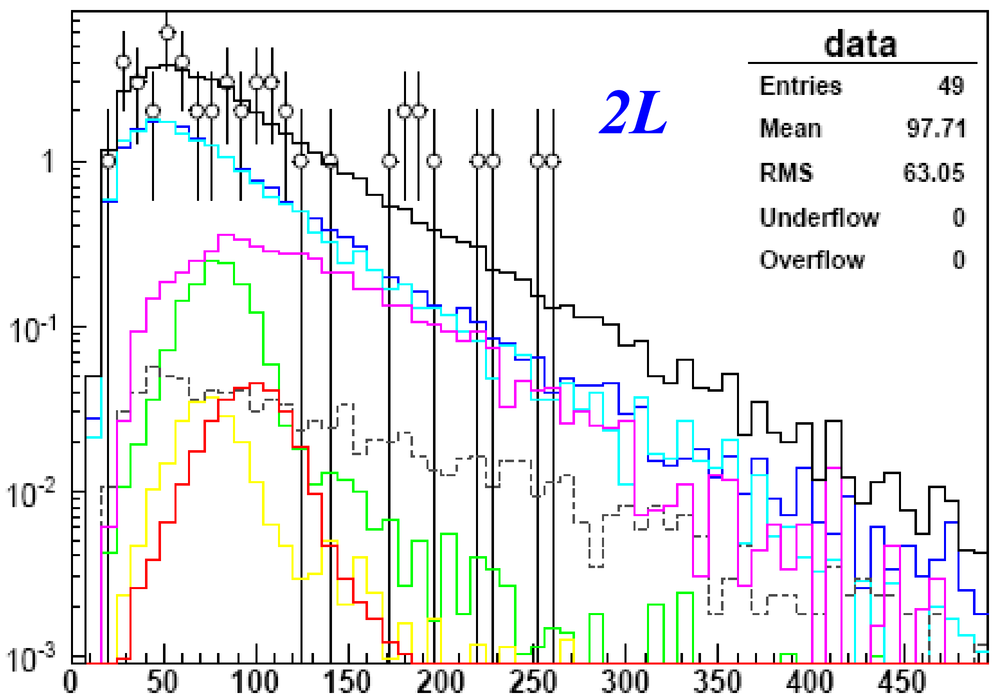
New QCD Fit

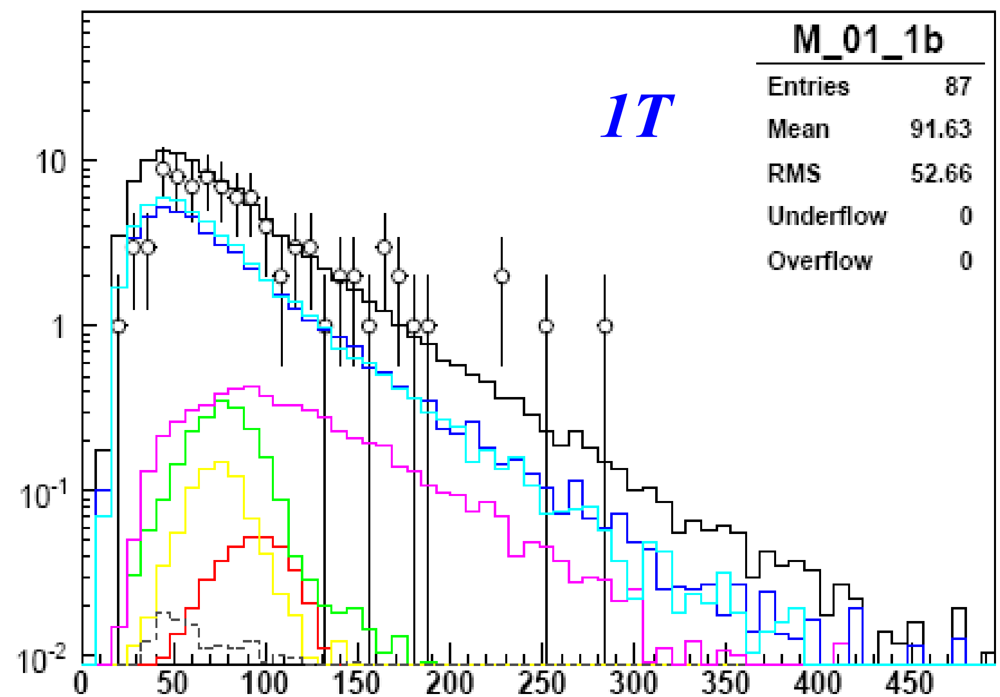
Result is that QCD remains an almost negligible background inside Z mass window (70-110 GeV)



Mjj / 2L and 1T

- Separate data into two orthogonal samples:
  - “2Loose”: >=2 L4 b-tagged jets
  - “1Tight”: >=1 VERYTIGHT b-tagged jets && !>=2 L4 b-tagged jets
- Using TRFs for b-tagging in MC, so it was some exercise in probability theory to calculate the probability for an event to have 1 VT tag given that it does not have >=2 L4 tags





Mjj Limits

m_H (GeV)	Data	QCD	Total Bgnd.	Sig.	Eff(%).	Exp.(pb) (Exp./SM)	Obs.(pb)
ZZ	47	0.10	53.22	1.76	0.11	21.885 (15.38)	18.420
105	50	0.11	57.86	0.47	10.66	7.215 (60.71)	5.988
115	47	0.11	55.33	0.35	11.24	6.693 (80.46)	5.360
125	44	0.11	52.44	0.23	11.55	6.214 (116.21)	4.925
135	44	0.12	52.19	0.15	12.02	5.997 (183.50)	4.744
145	46	0.12	47.14	0.06	11.56	5.759 (387.70)	5.581
155	41	0.12	42.16	0.03	11.71	5.215 (877.69)	5.034
ZZ	22	0.32	20.72	1.28	0.08	14.795 (10.40)	16.711
105	21	0.32	20.06	0.37	8.63	4.406 (37.08)	4.689
115	20	0.34	20.28	0.28	9.50	3.975 (47.78)	3.975
125	16	0.34	19.86	0.19	10.01	3.589 (67.11)	2.955
135	17	0.36	18.70	0.12	10.53	3.313 (101.38)	3.099
145	16	0.35	17.21	0.06	10.39	3.296 (221.87)	3.072
155	15	0.36	16.43	0.02	10.75	3.051 (513.43)	2.837

1T

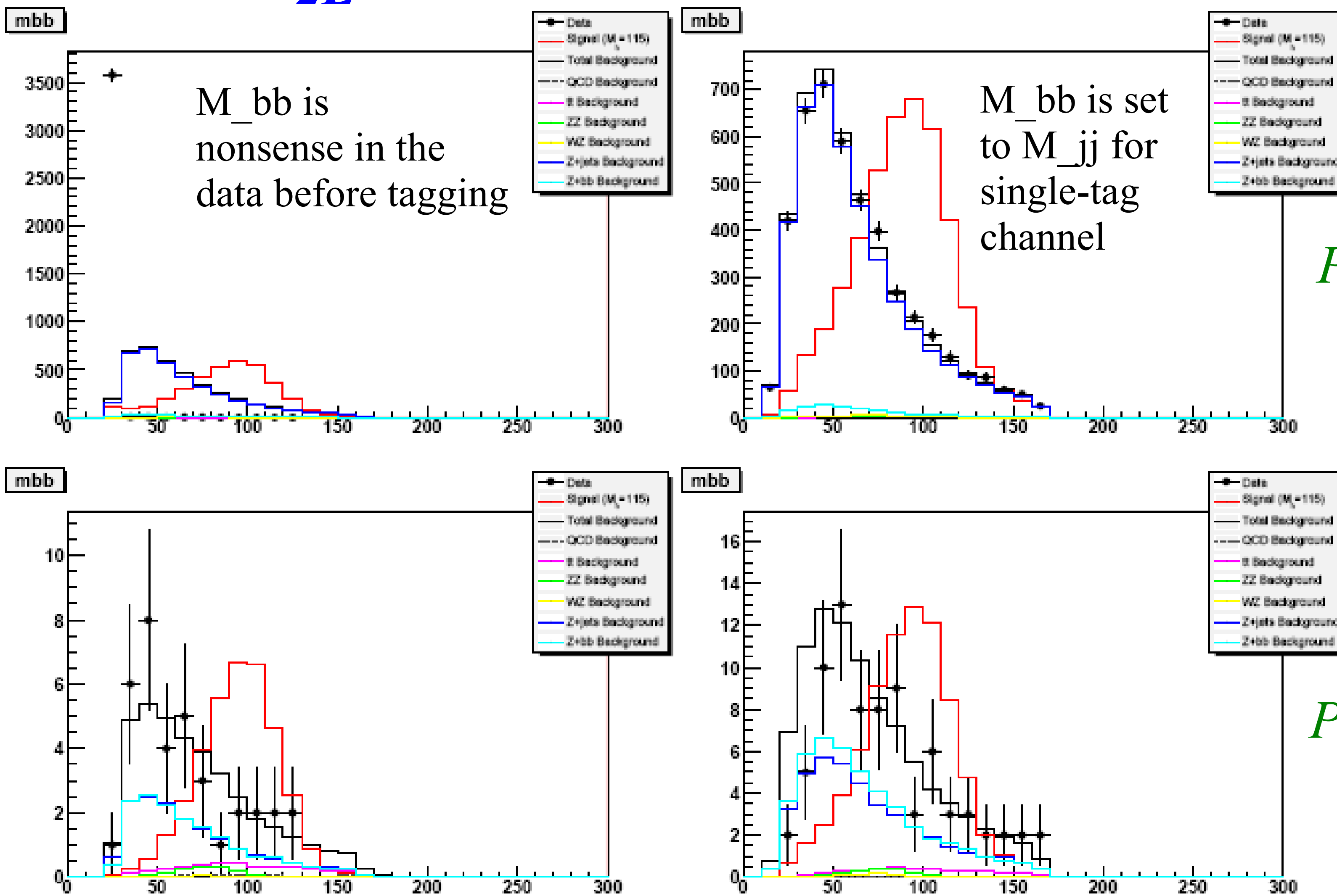
2L

Train Neural Networks

1. M_{bb}: The invariant mass of a randomly chosen b-tagged jet pair. For the 1T channel, this variable was set to the invariant mass of the b-tagged jet and the highest p_T non-b-tagged jet. Note that the 4-vector invariant mass calculation was used throughout this analysis, which takes into account the mass of each individual jet.
2. $p_T(1)$: The p_T of the leading p_T jet.
3. $p_T(2)$: The p_T of the second-leading p_T jet.
4. ZΔR: ΔR between the two muons in the Z candidate.
5. $|\Delta\eta|$: The absolute value of the difference in η between the two highest p_T jets.
6. $|\Delta\phi|$: The absolute value of the difference in ϕ between the two highest p_T jets.
7. $\Delta R(Z, jet1)$: The ΔR between the Z candidate and the highest p_T jet.
8. $|\eta_Z|$: The $|\eta|$ of the Z candidate.
9. $\not{E}_T$: The missing E_T of the event (useful against $t\bar{t}$).
10. SE_T: The scalar E_T of the event.

NN Inputs

2L **1T**

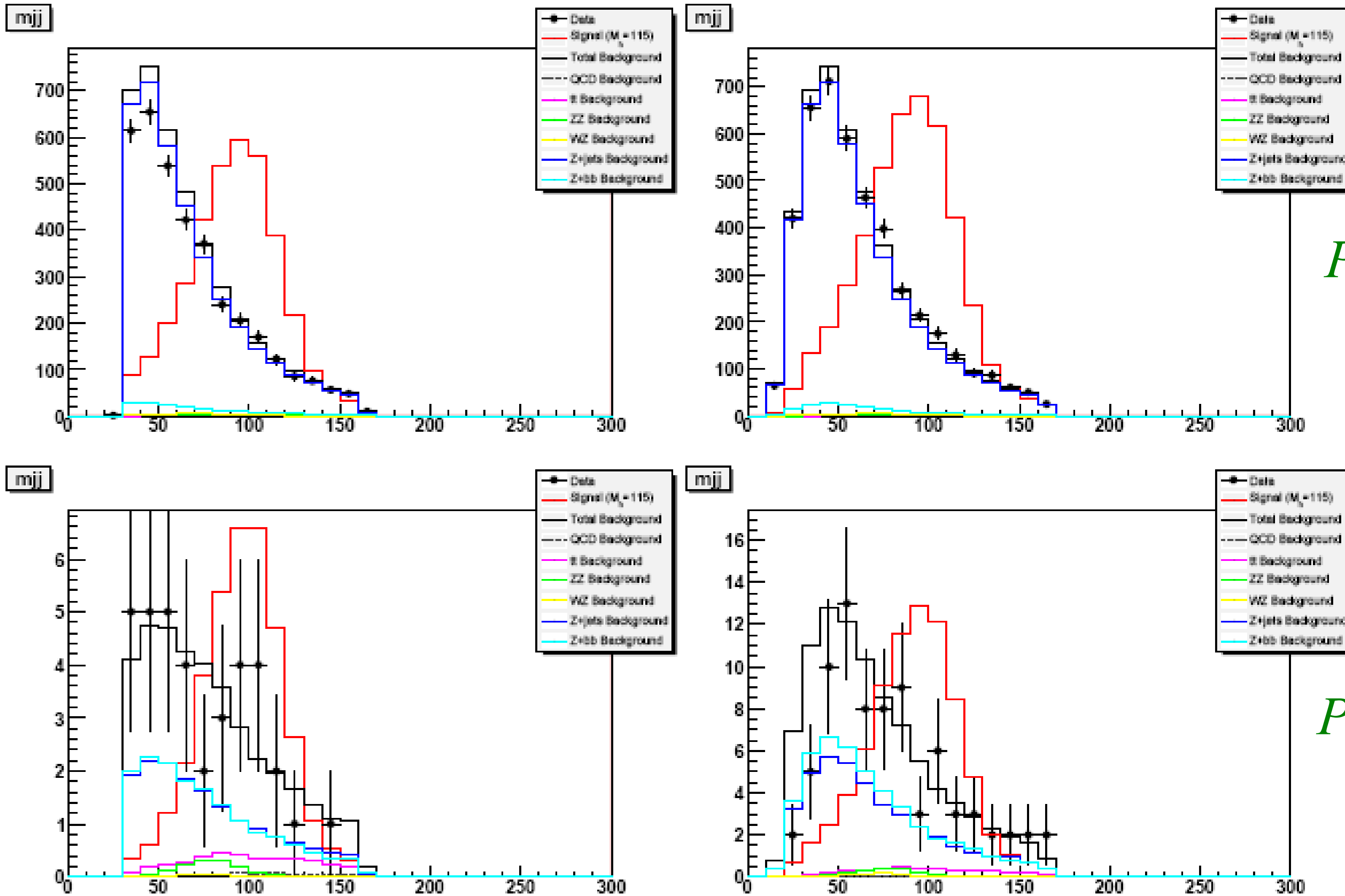


Pre-tag

Post-tag

NN Inputs

2L *1T*



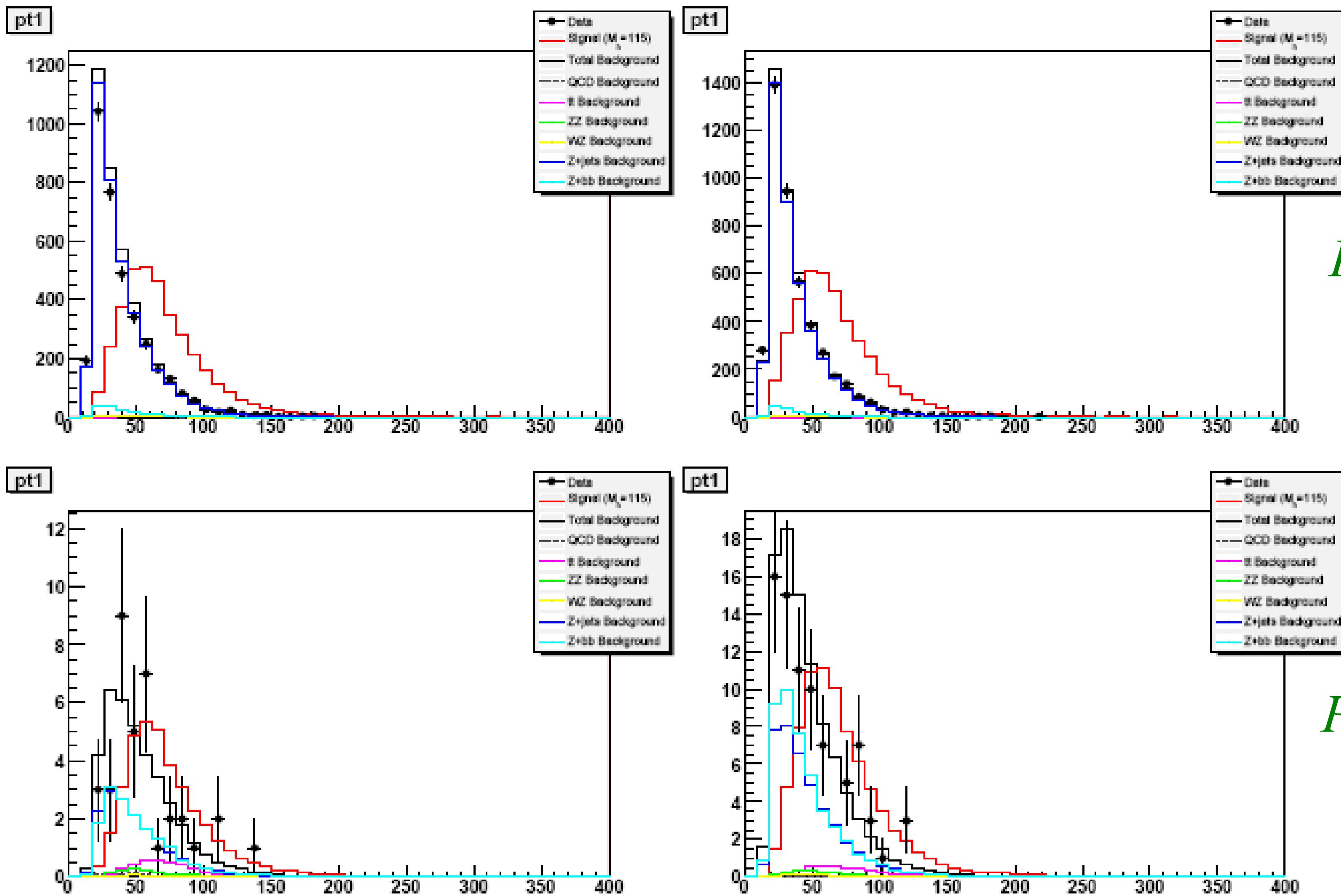
Pre-tag

Post-tag

M_jj is *not* used in the (2L) NNs – M_bb performs significantly better!

NN Inputs

2L *1T*

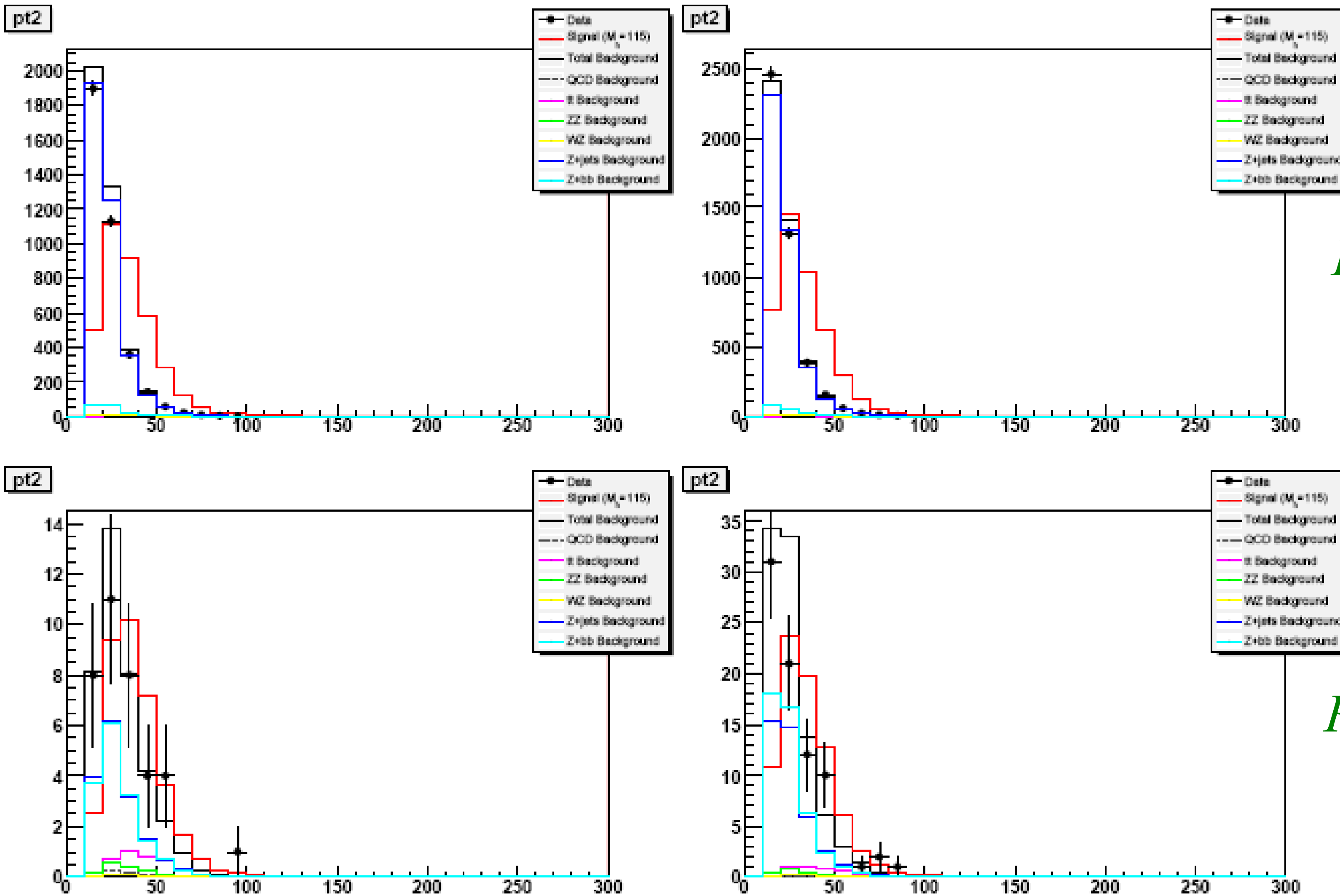


Pre-tag

Post-tag

NN Inputs

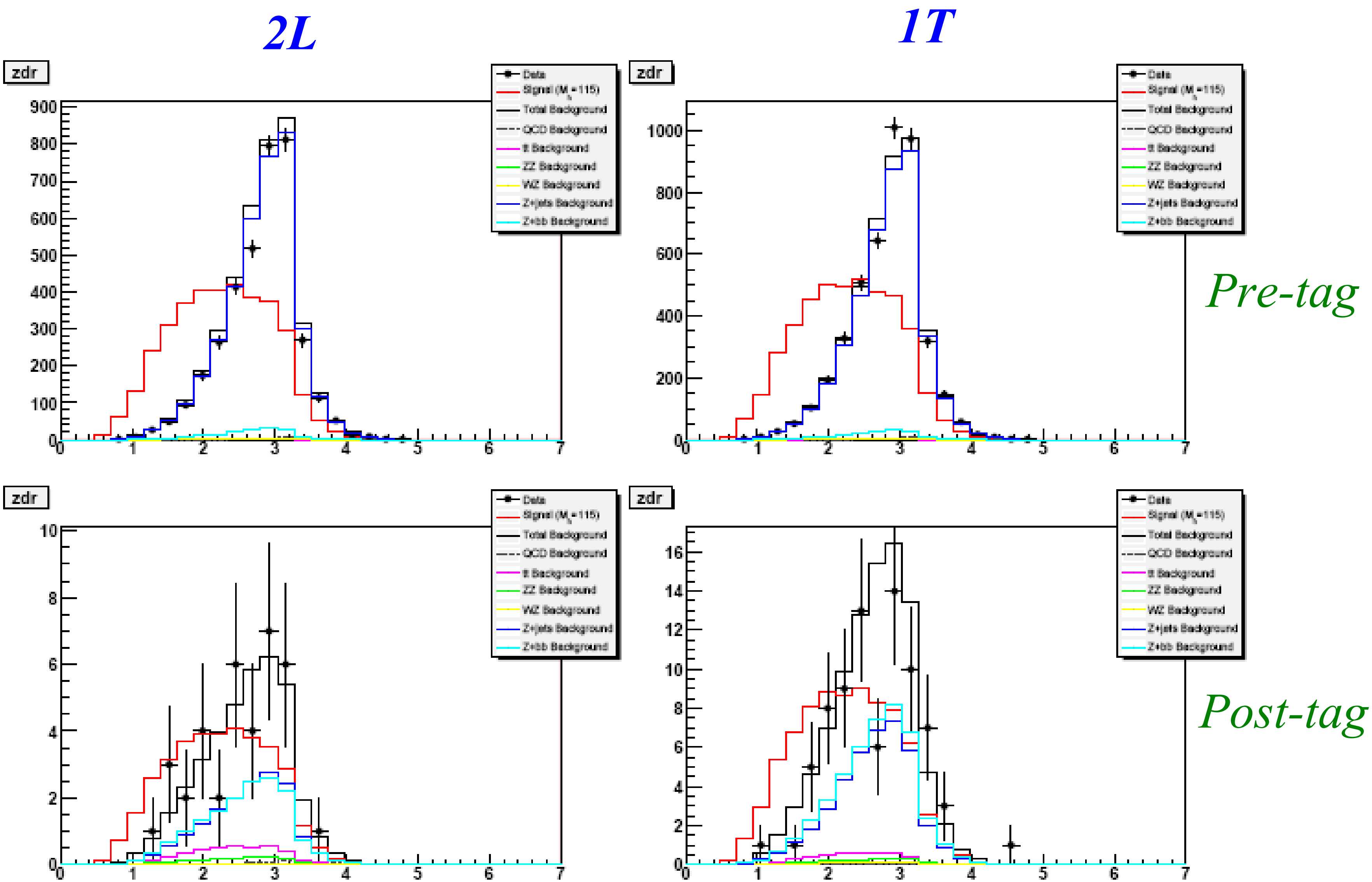
2L **1T**



Pre-tag

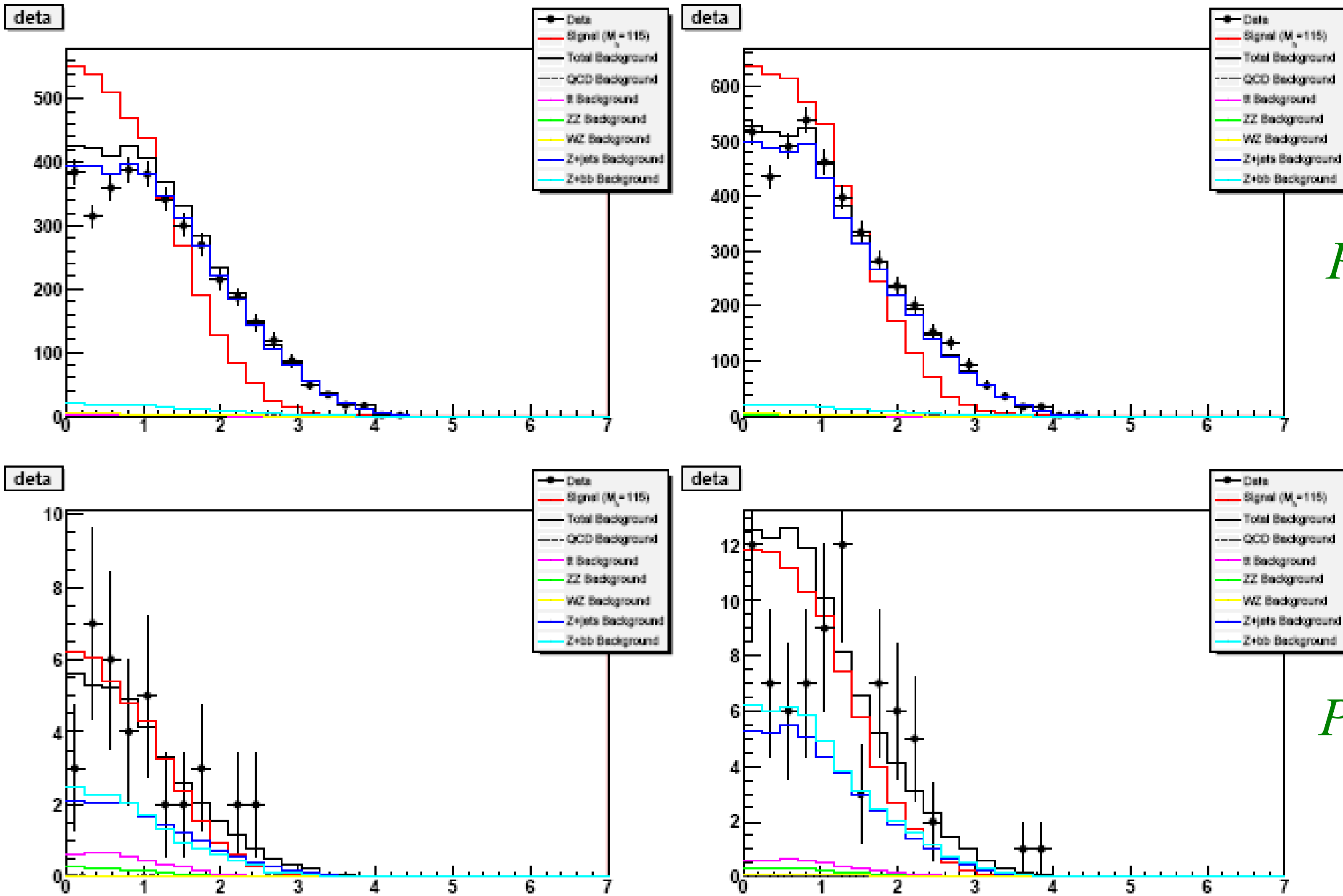
Post-tag

NN Inputs


2L
1T
Pre-tag
Post-tag
zdr
Data
Signal (M_h=115)
Total Background
QCD Background
tt Background
ZZ Background
WZ Background
Z+jets Background
Z+bb Background

NN Inputs

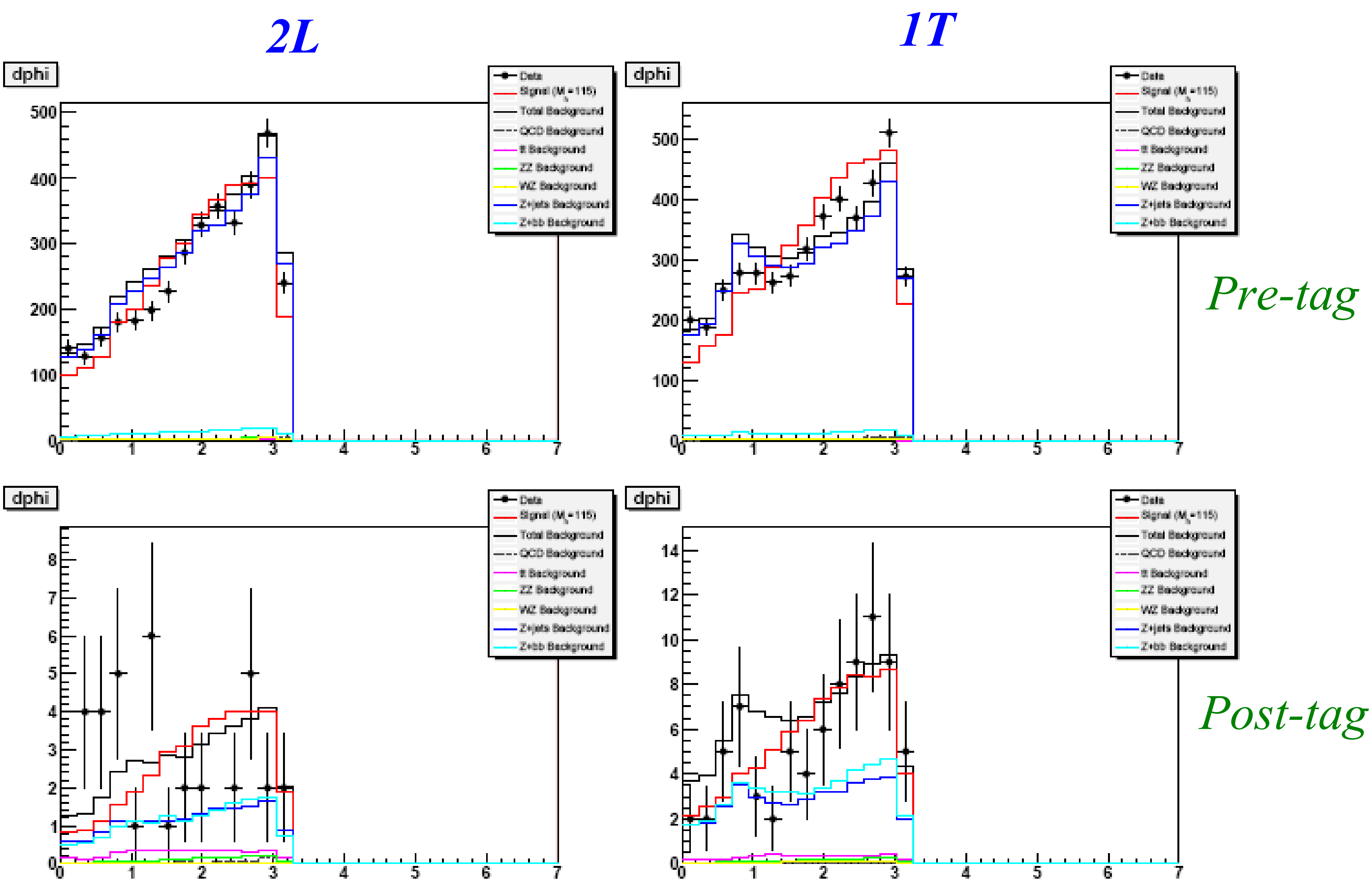
2L *1T*



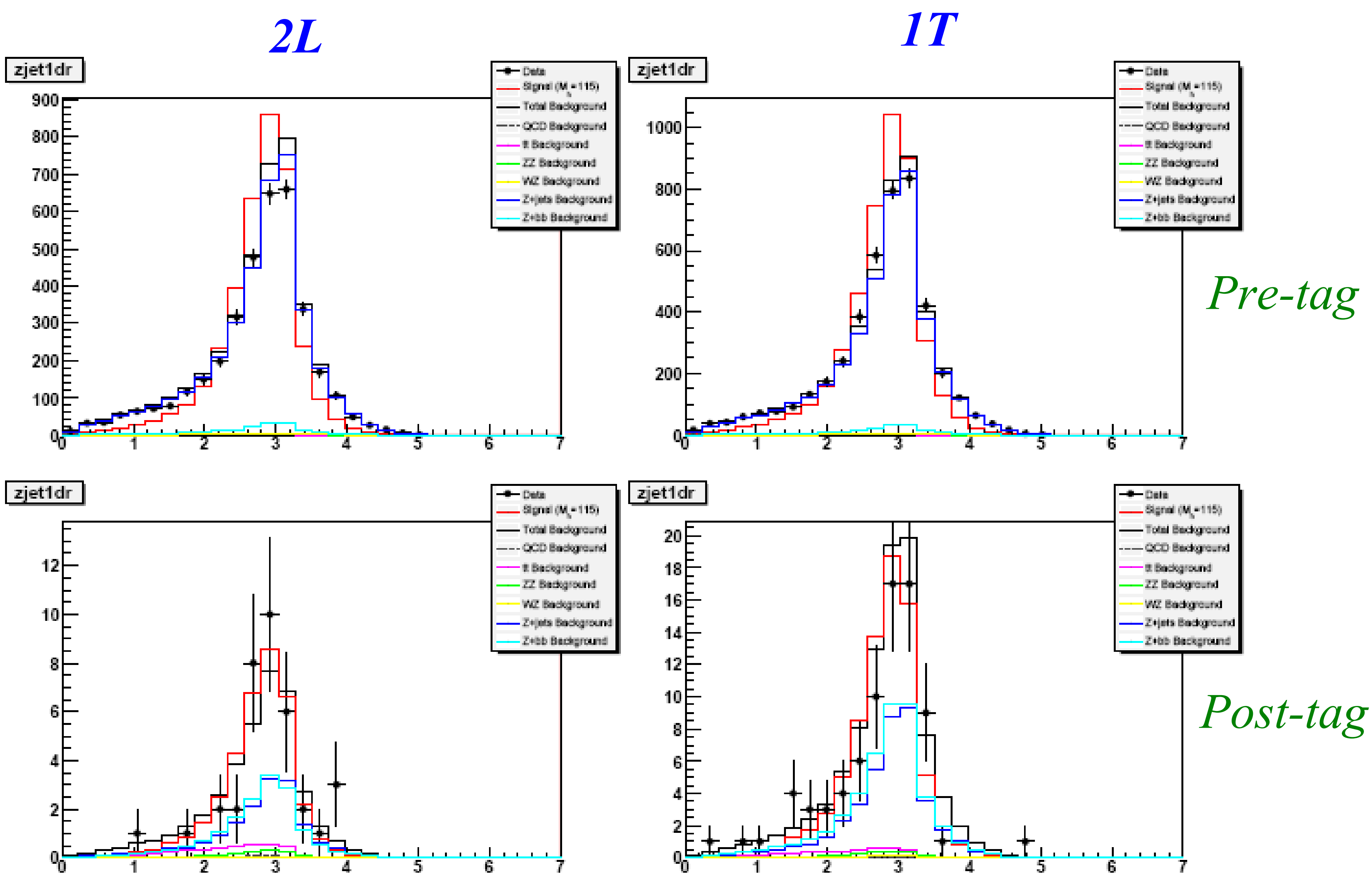
Pre-tag

Post-tag

NN Inputs

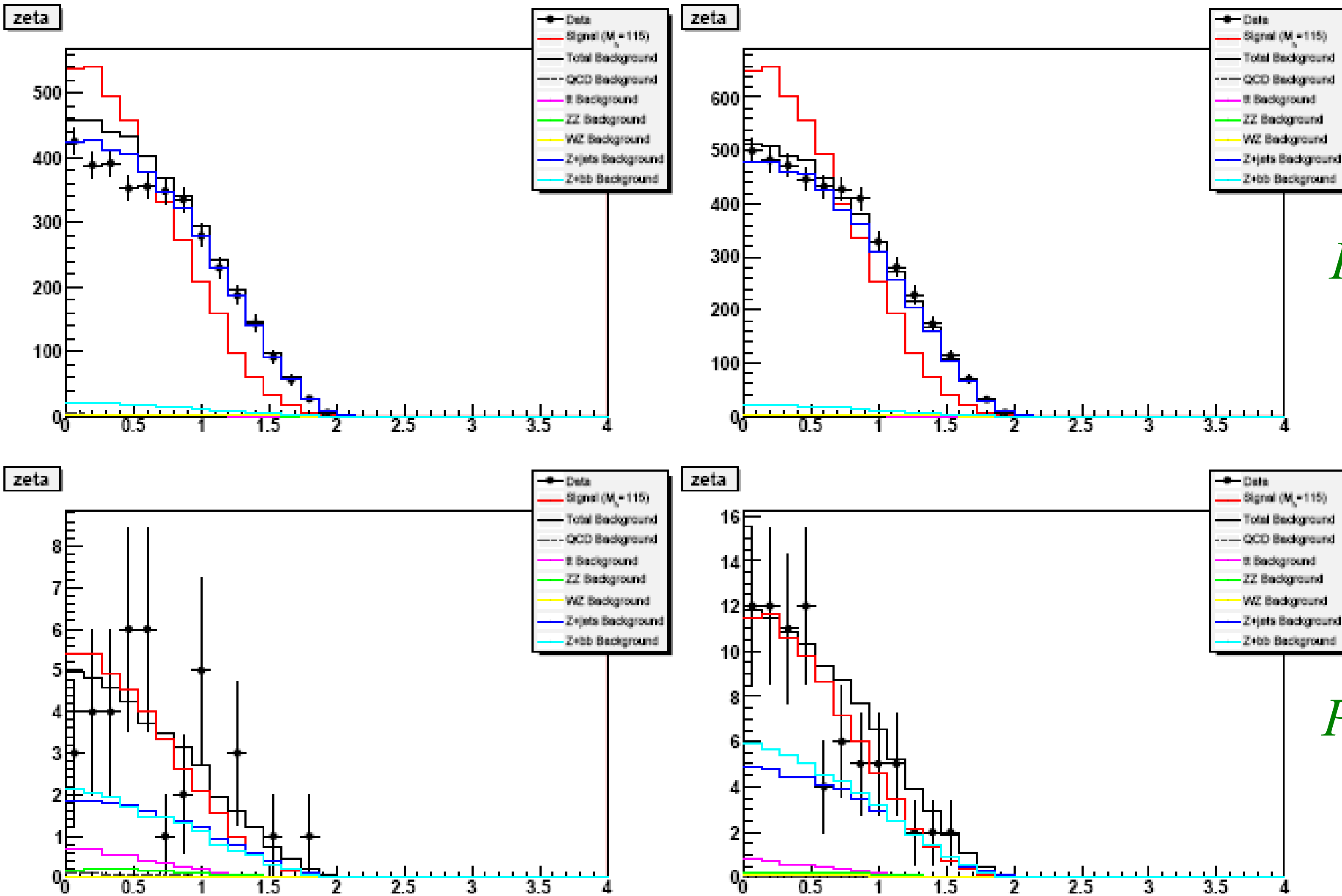

2L
1T
Pre-tag
Post-tag
dphi
Data
Total Background
QCD Background
tt Background
ZZ Background
WZ Background
Z+jets Background
Z+bb Background

NN Inputs


2L
1T
Pre-tag
Post-tag
zjet1dr
Data
Total Background
QCD Background
ZZ Background
WZ Background
Z+jets Background
Z+bb Background

NN Inputs

2L *1T*

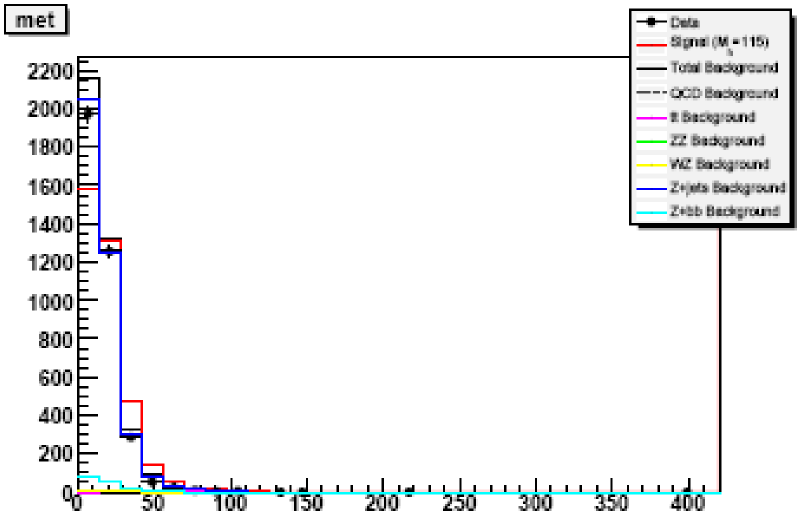


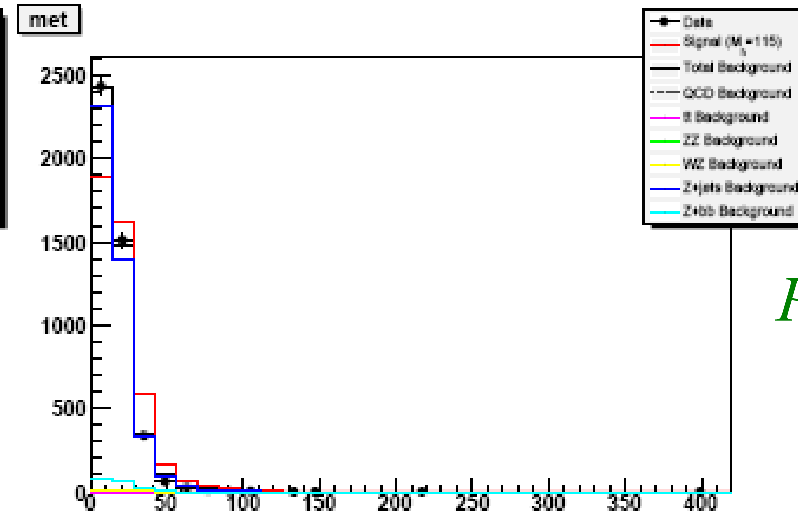
Pre-tag

Post-tag

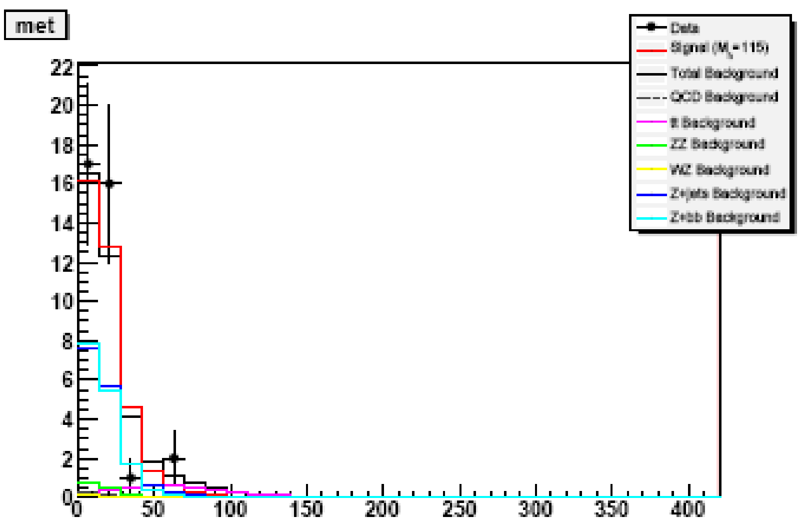
NN Inputs

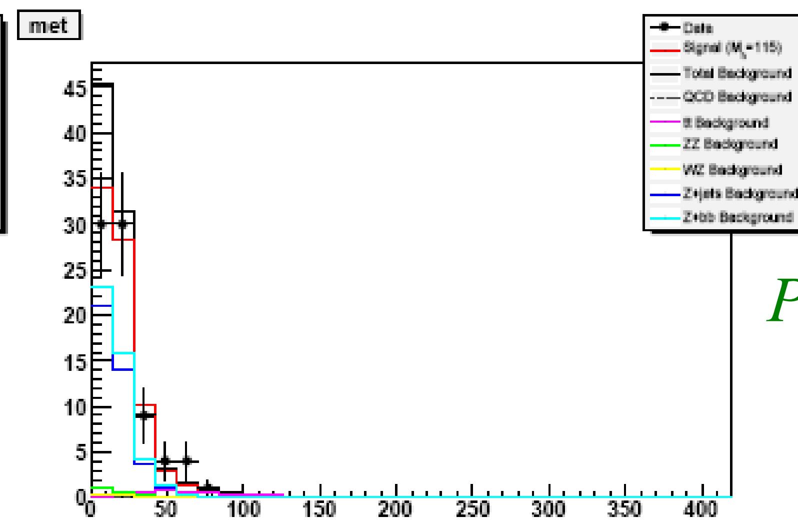
2L ***1T***





Pre-tag





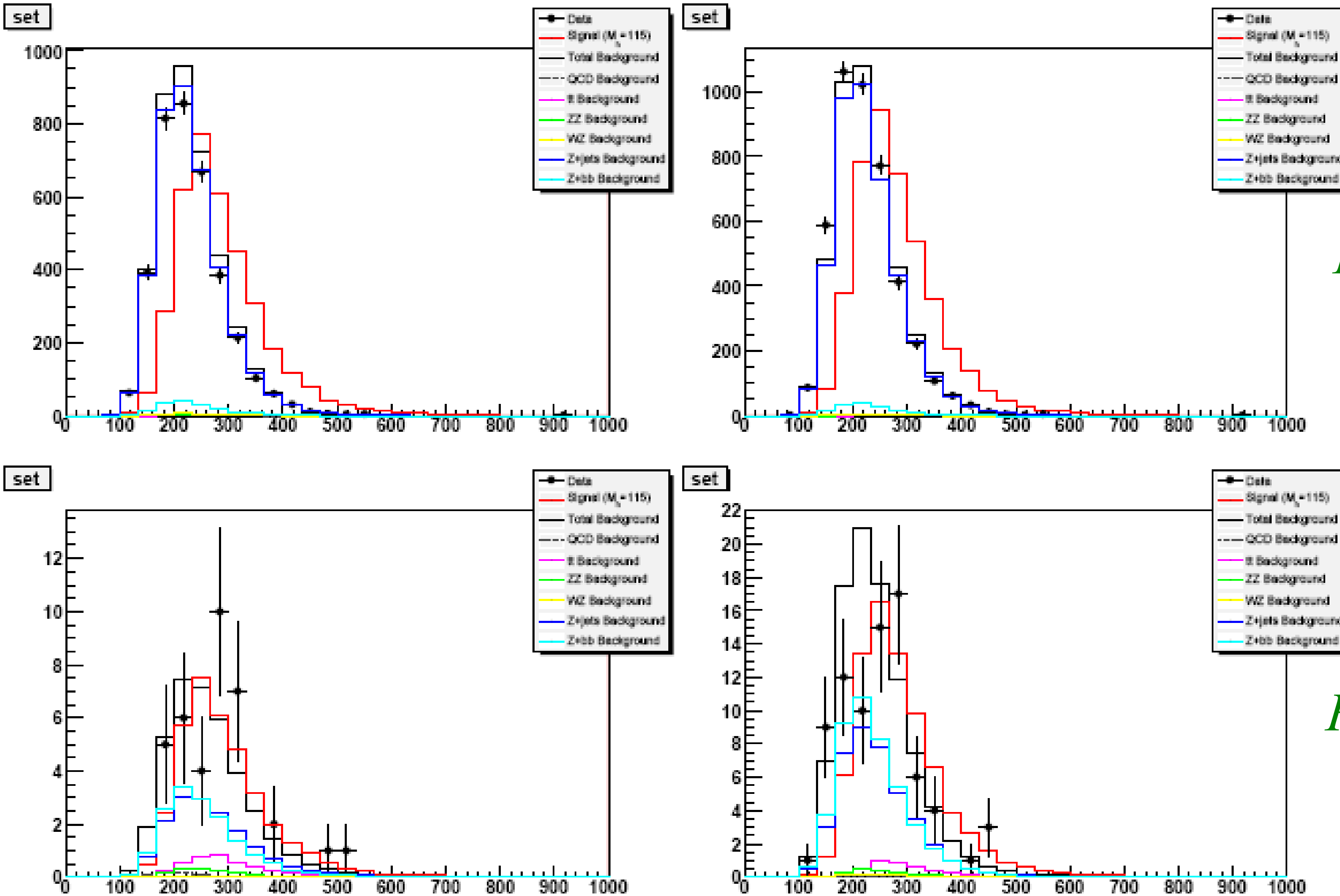
Post-tag

NN Inputs

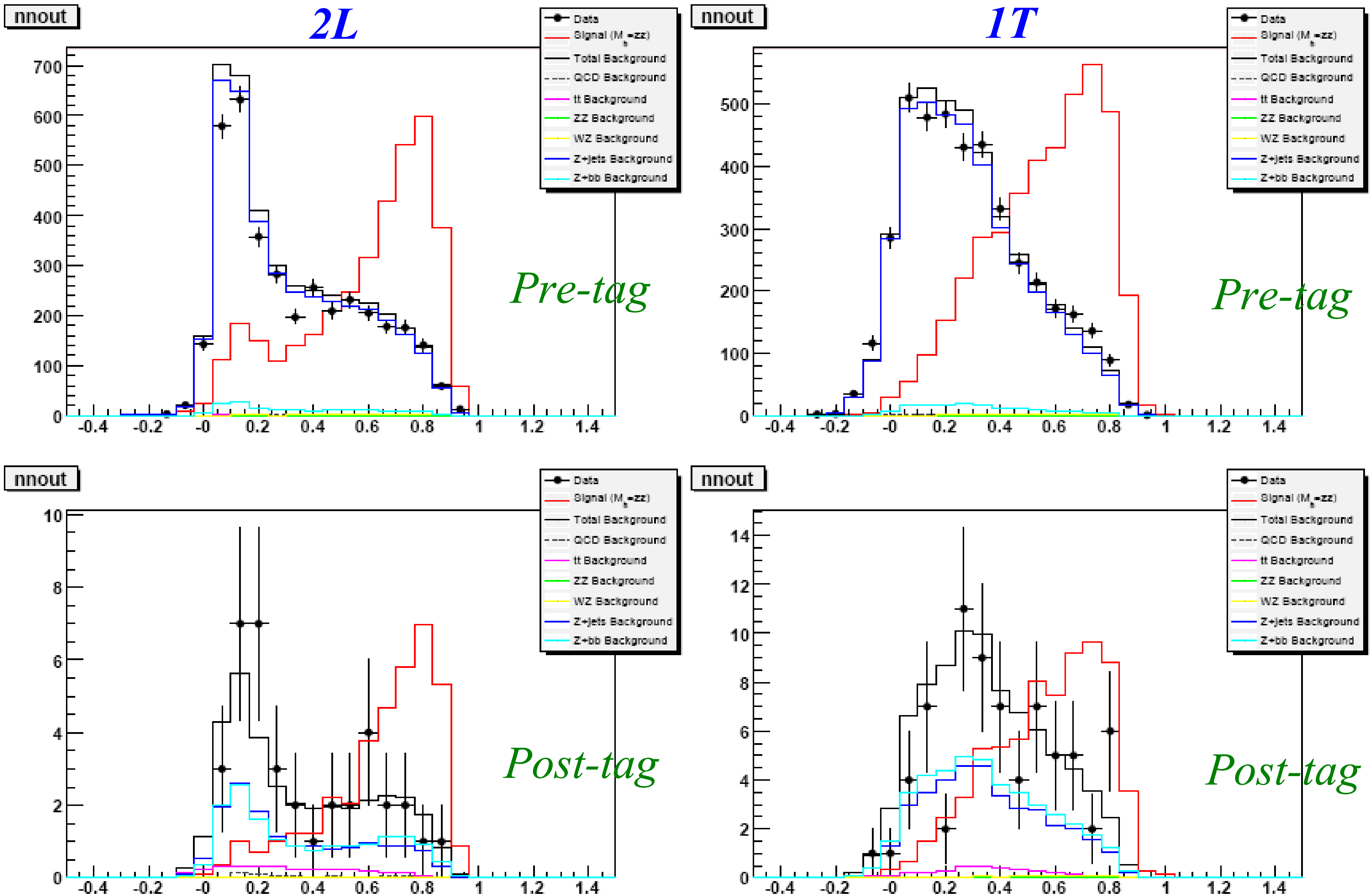
2L *1T*

Pre-tag

Post-tag

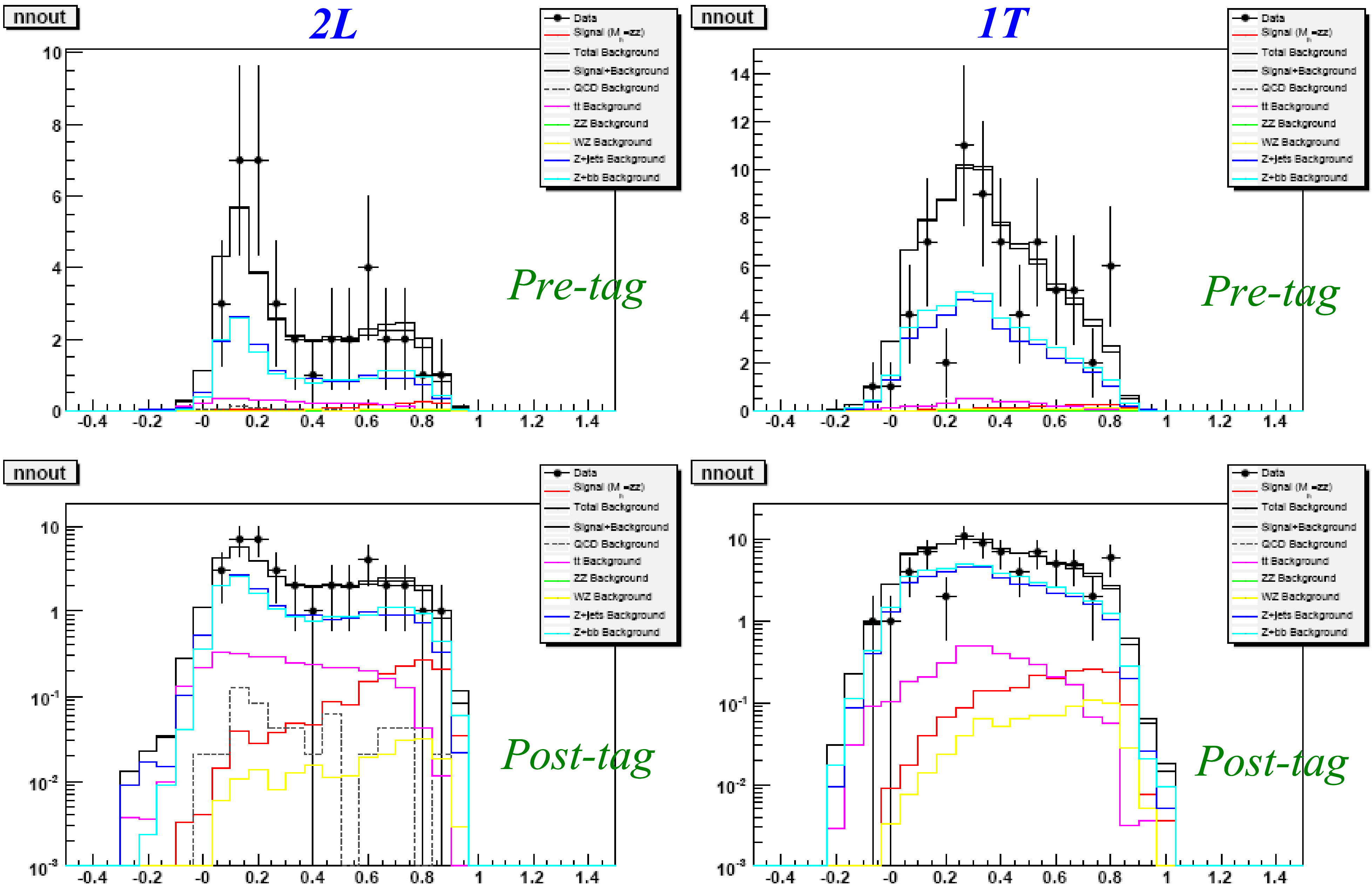


NN Output: ZZ

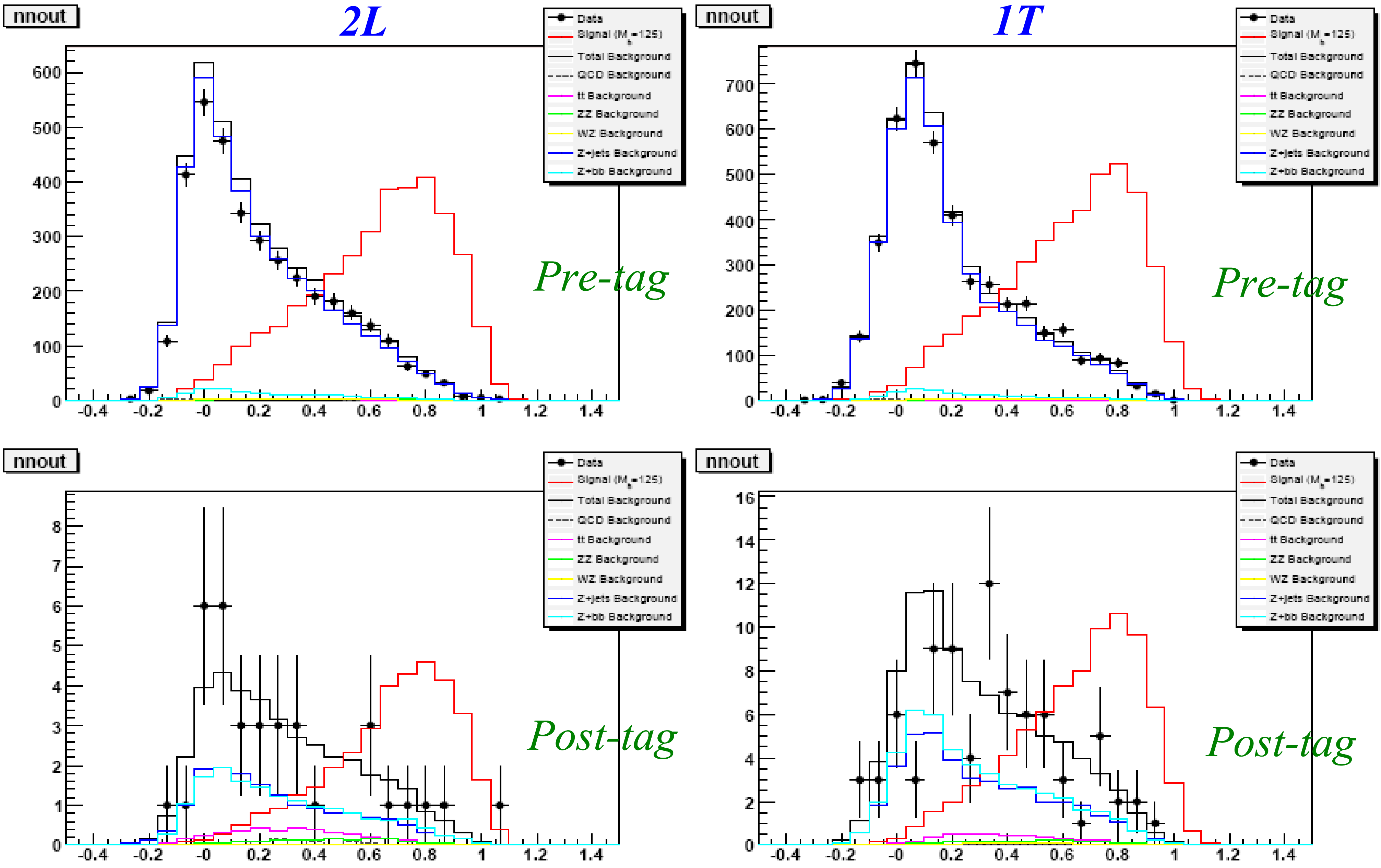


nnout
2L
Data
Signal (M_h=zz)
Total Background
QCD Background
tt Background
ZZ Background
WZ Background
Z+jets Background
Z+bb Background
Pre-tag
nnout
1T
Pre-tag
nnout
Post-tag
nnout
Post-tag

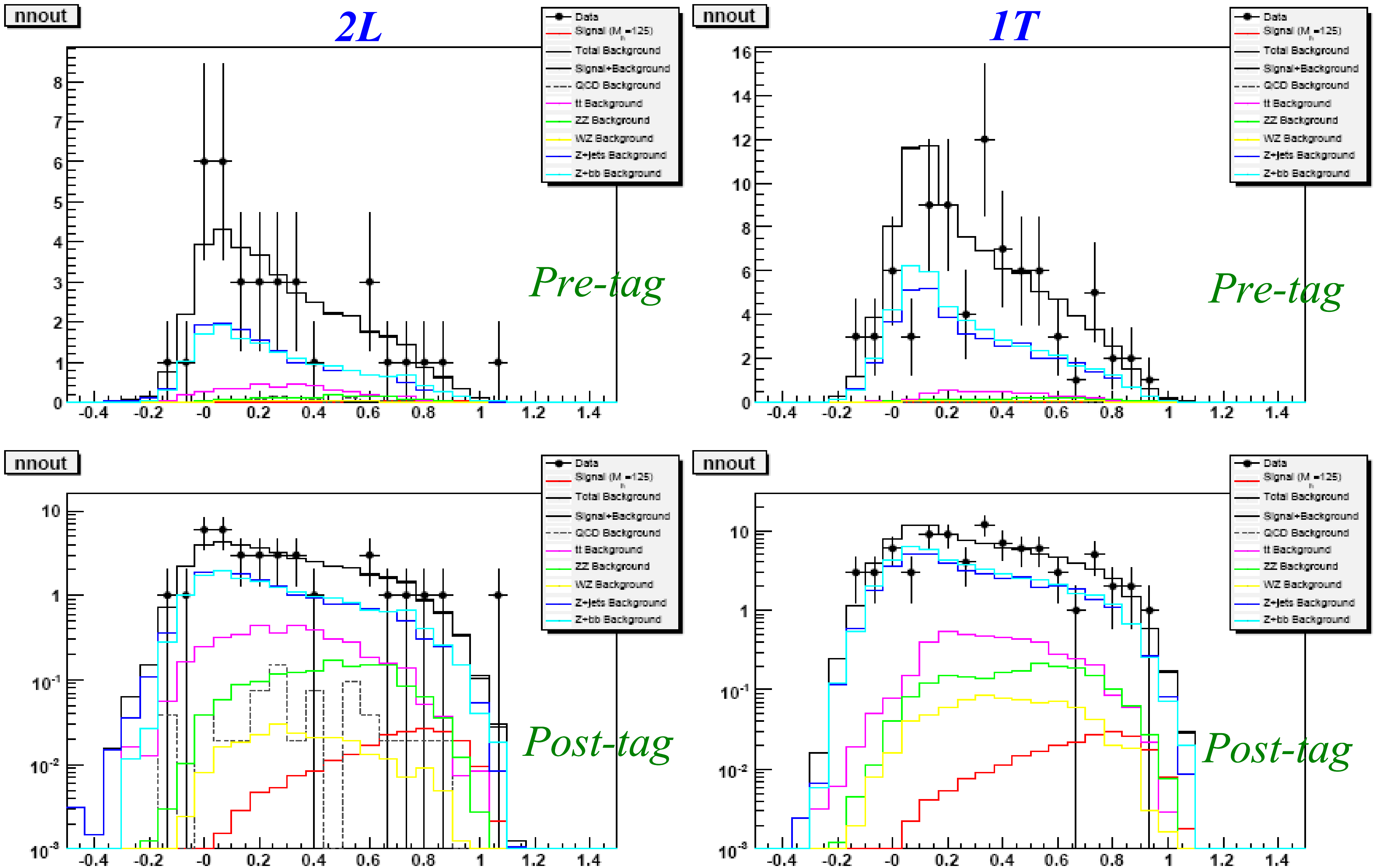
NN Output: ZZ


nnout
2L
1T
Pre-tag
Post-tag
Data
Signal (M_h=zz)
Total Background
Signal+Background
QCD Background
tt Background
ZZ Background
WZ Background
Z+jets Background
Z+bb Background

NN Output: mH=125



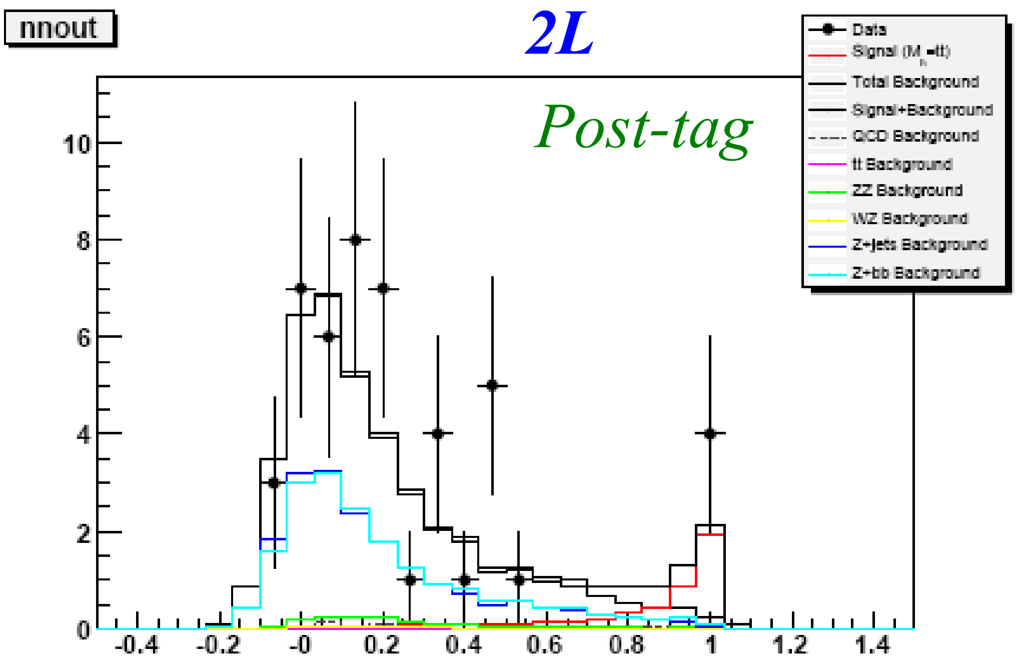
NN Output: mH=125

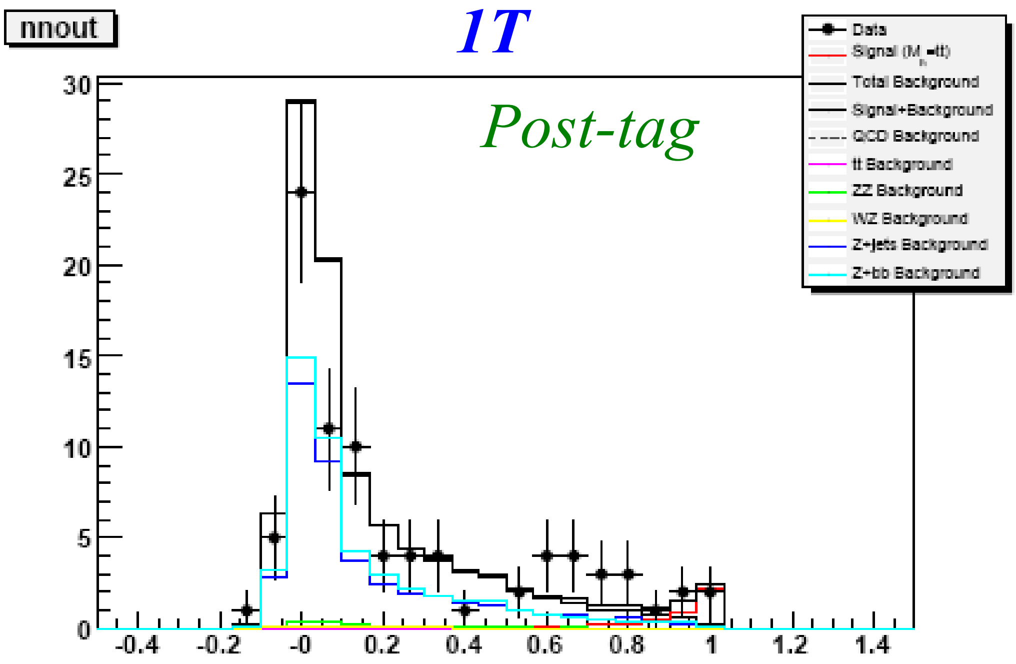


NN Output: tt Cross-check

Train NN against tt as the signal.

Some evidence for tt, consistent with expectations (I input xs=7.0pb).

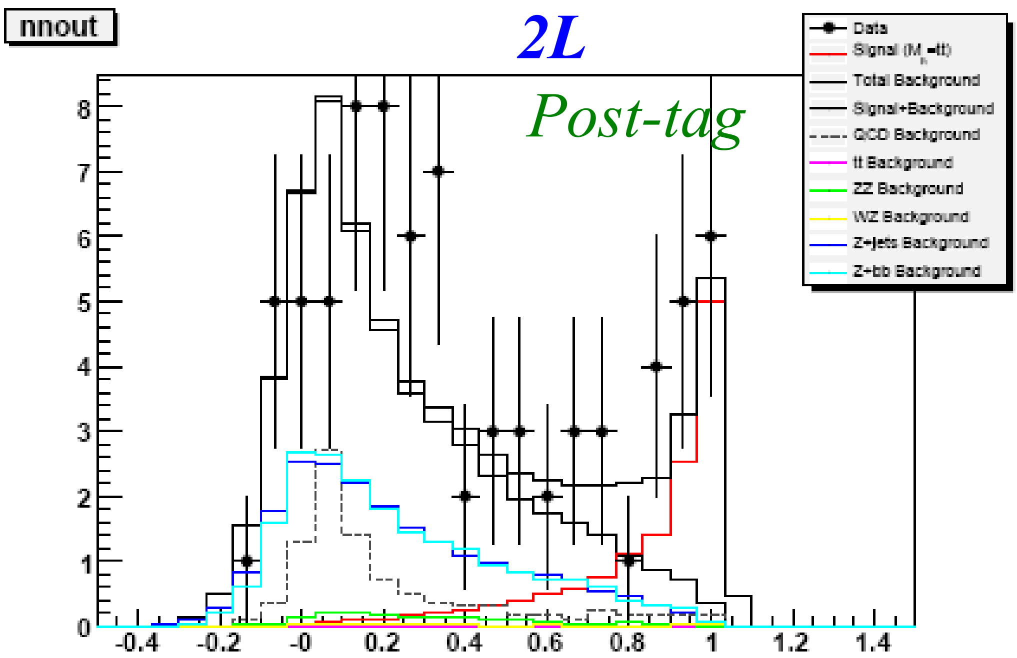


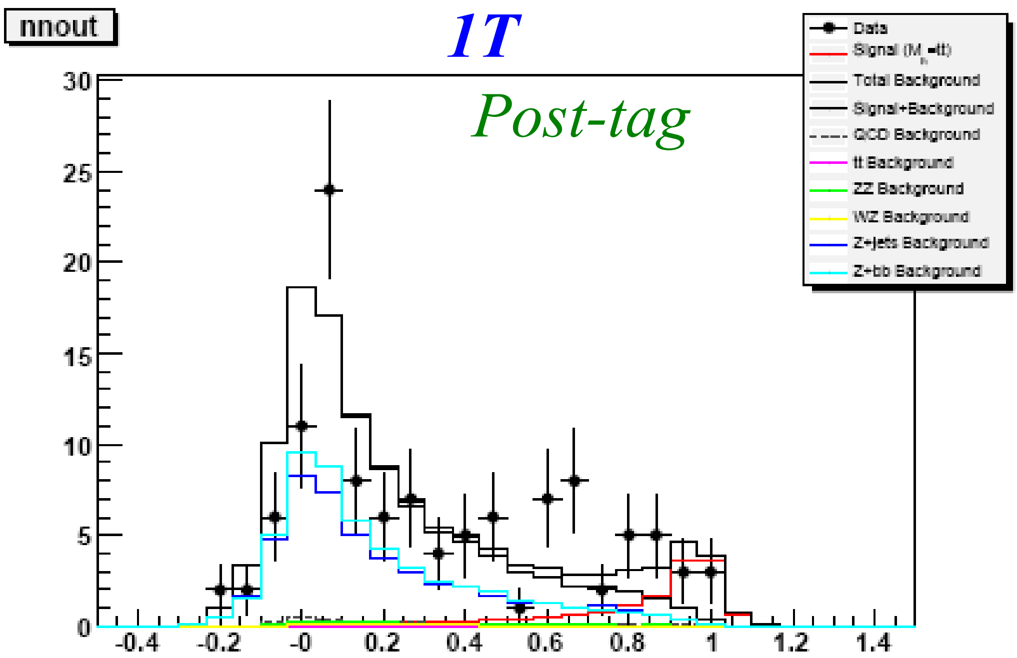


tt Cross-check w/ Wider Acceptance

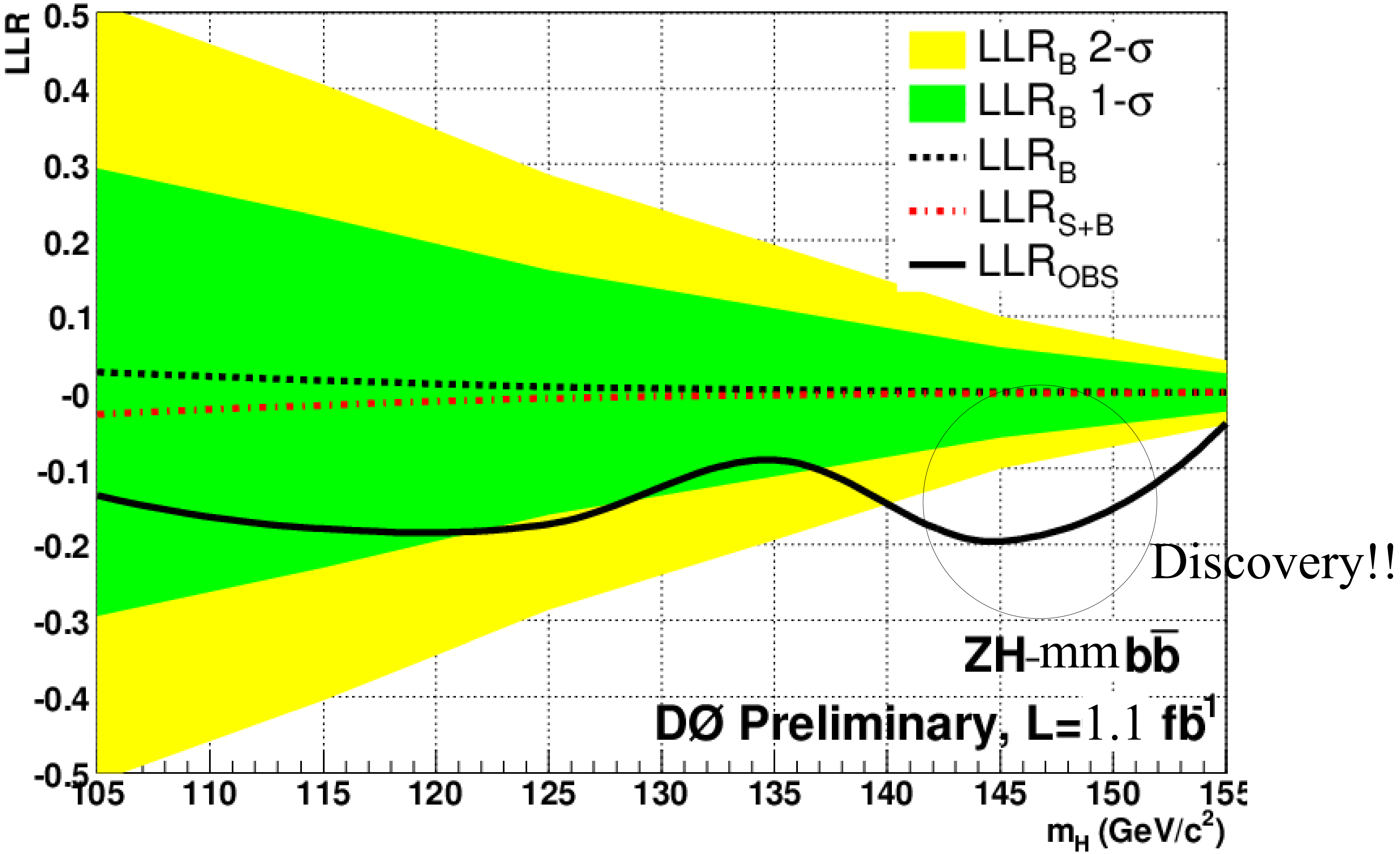
Widen Z mass window, 70-110 -> 40-500, accept more tt signal.

Good evidence for tt, consistent with expectations (I input xs=7.0pb).



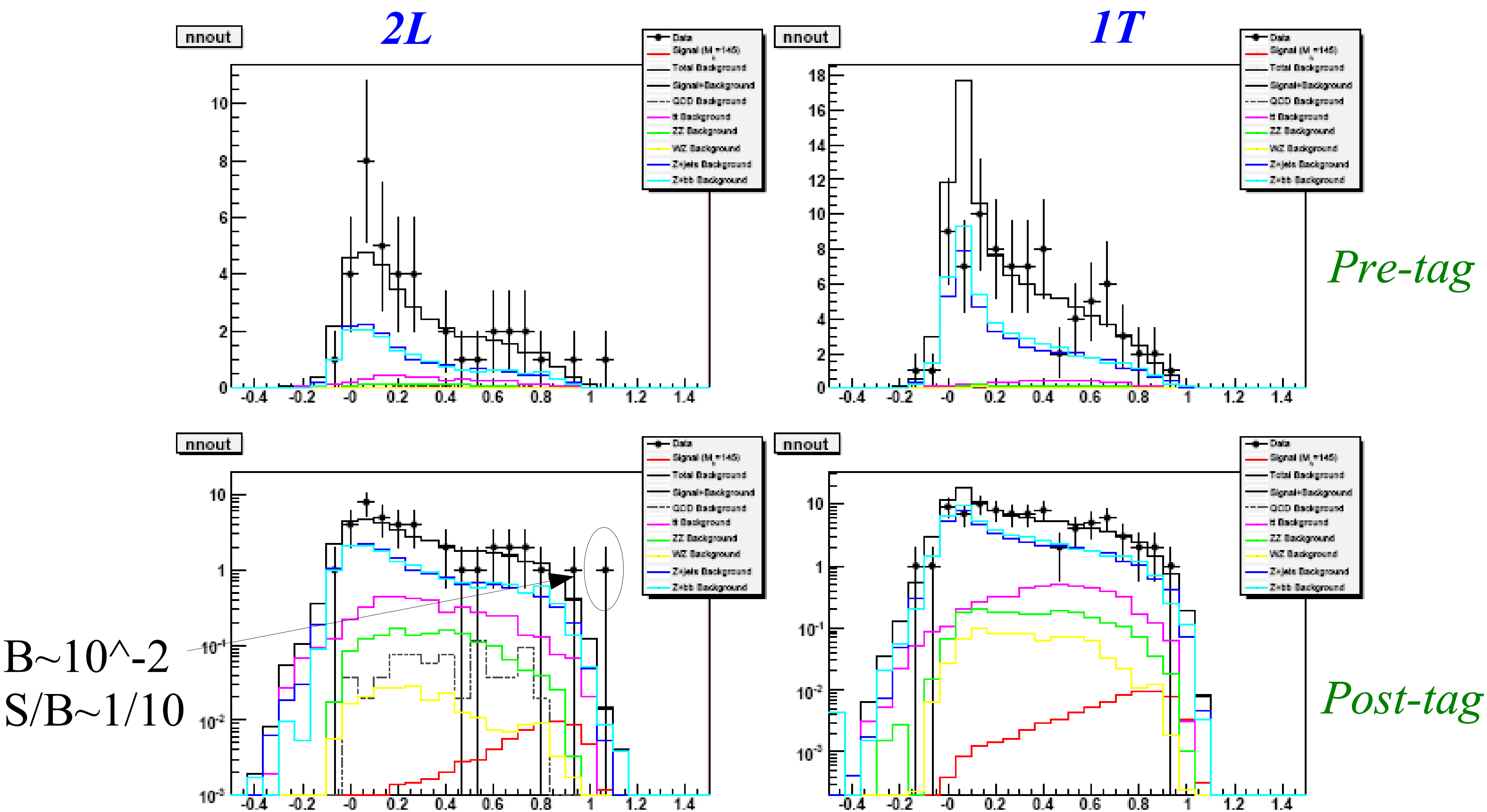


Wade's LLR's

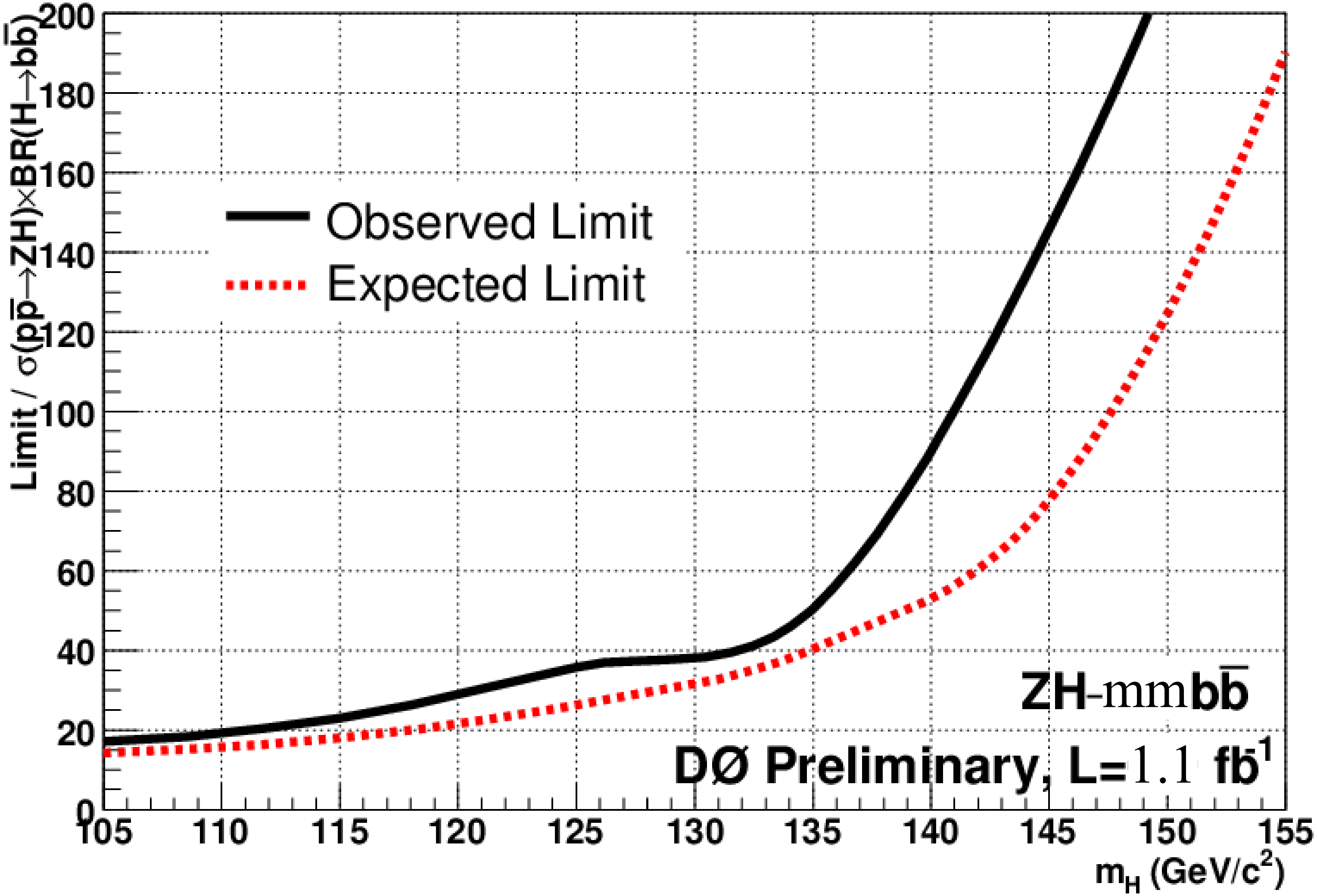


LLR
0.5
0.4
0.3
0.2
0.1
-0
-0.1
-0.2
-0.3
-0.4
-0.5
105
110
115
120
125
130
135
140
145
150
155
m_H (GeV/c^2)
LLR_B 2-σ
LLR_B 1-σ
LLR_B
LLR_{S+B}
LLR_{OBS}
Discovery!!!
ZH-mm $b\bar{b}$
DØ Preliminary, L=1.1 fb^{-1}

NN Output: mH=145



Limit / SM from Wade



Conclusions

- Results using combination of full 1.1/fb, single/double b-tags, and NN are equivalent to ~4x the luminosity of last summer's conference result

- Need more work on systematics, now that a NN is involved
  - Use agreement of un-tagged data/MC NN outputs in signal region?
- Have updated the analysis note
  - About 90 pages long (mostly plots)! :)

Analysis: Limit/SM (mh=115)

Nominal Mjj (ICHEP'06): 42

Updated P17 Mjj: 36

p17 NN 2Tag: 27

p17 NN 1Tag: 35

p17 NN 1+2Tags: 20

p17 NN 1+2Tags:
e+mu combined: 14 (estimated)

CDF (1/fb – Moriond'07)
e+mu combined, MET filter: 16

We'll be ahead of CDF in ZH->llbb!