Measurement of WW+WZ production in lνjj Final States

Gregorio Bernardi
(LPNHE-Paris, CRNS/IN2P3)

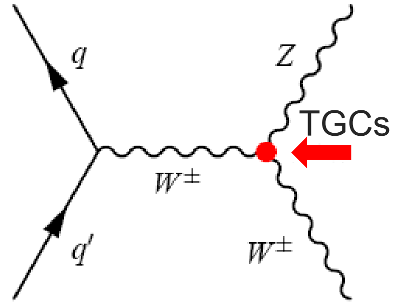
On behalf of the

DØ collaboration

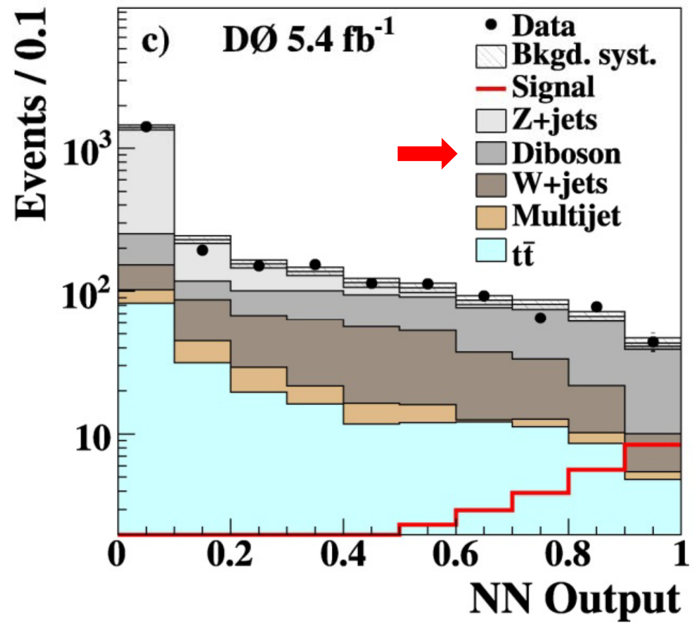
Thanks to all my Dzero colleagues, in particular to J. Sekaric

Motivation

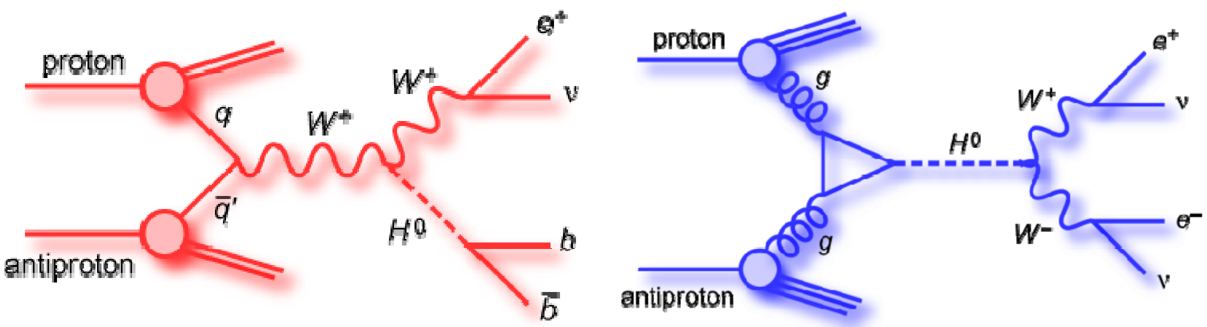
Probe of the EWK Symmetry Breaking mechanism
SM tests
Indirect searches for New Physics (NP)
(cross sections, kinematic distributions, gauge boson couplings)



Important background to Top, Higgs, SUSY
Good understanding is highly valuable



Proving ground for analysis techniques (MVA) and statistical treatment used in the Tevatron **Higgs searches**



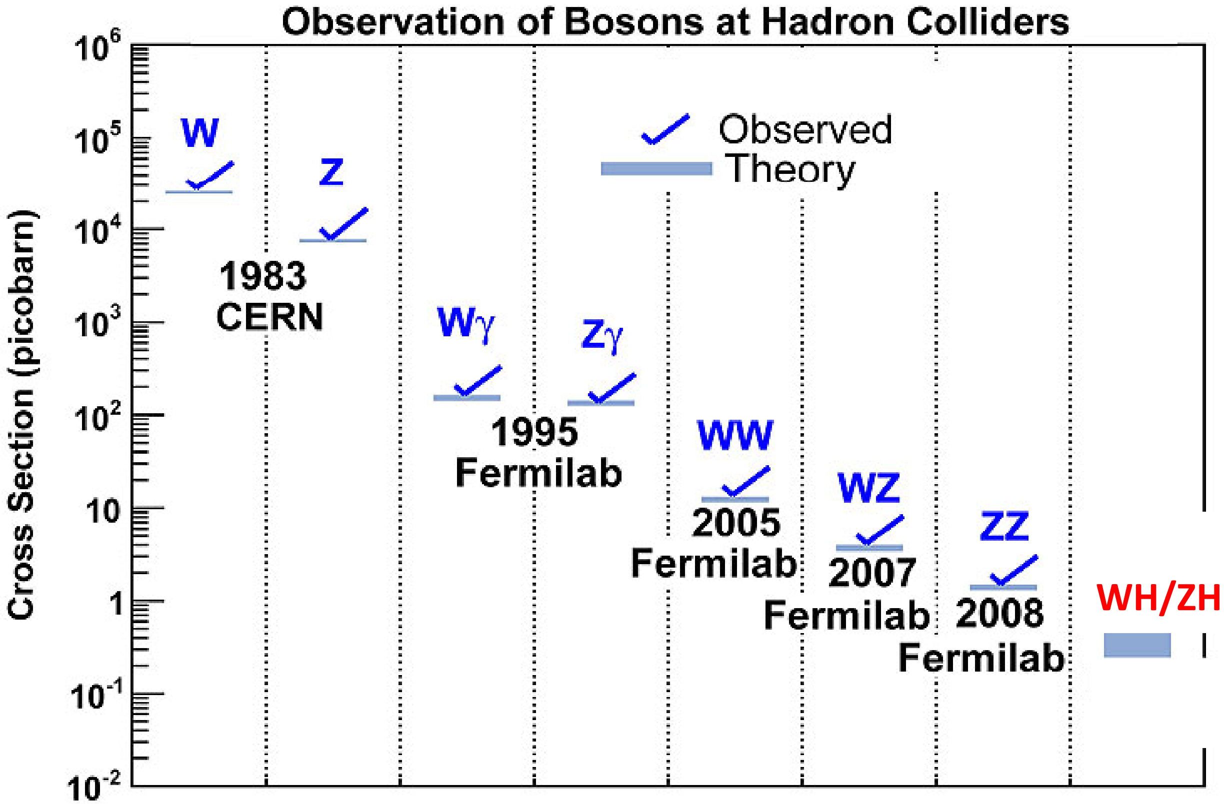
Same final states, similar challenges

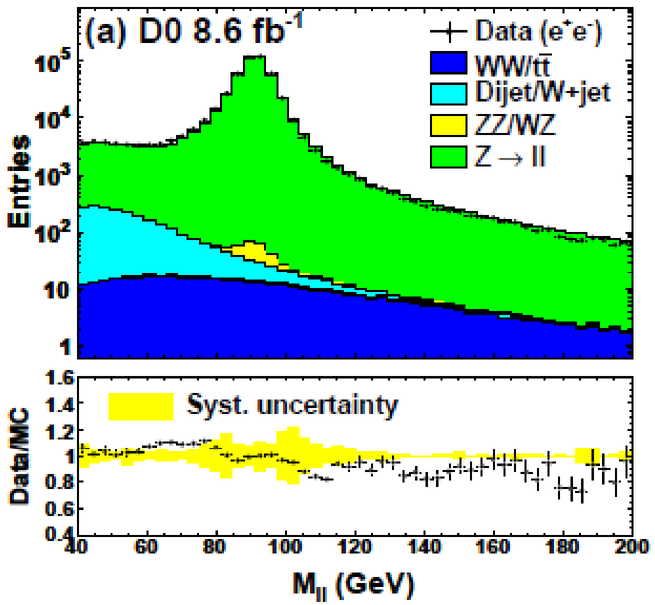
Dibosons

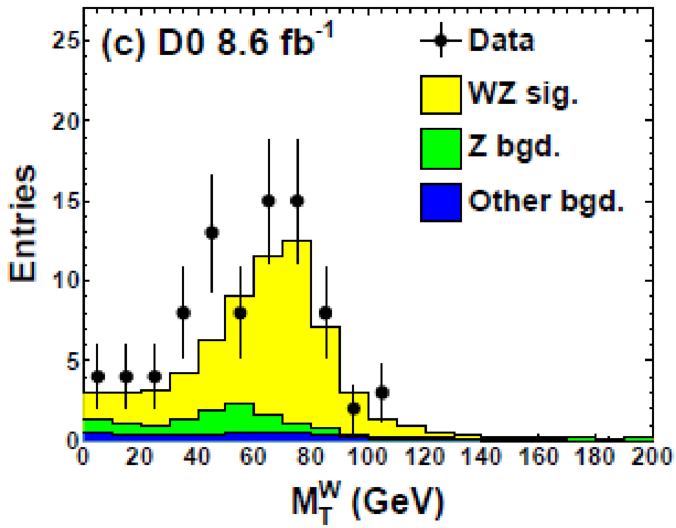
For SM tests, and H→WW
Fully leptonic final states

(l = e/μ)

- most analyzed
- small branching ratio
- clean signal (low backgrounds)
- observed at the Tevatron



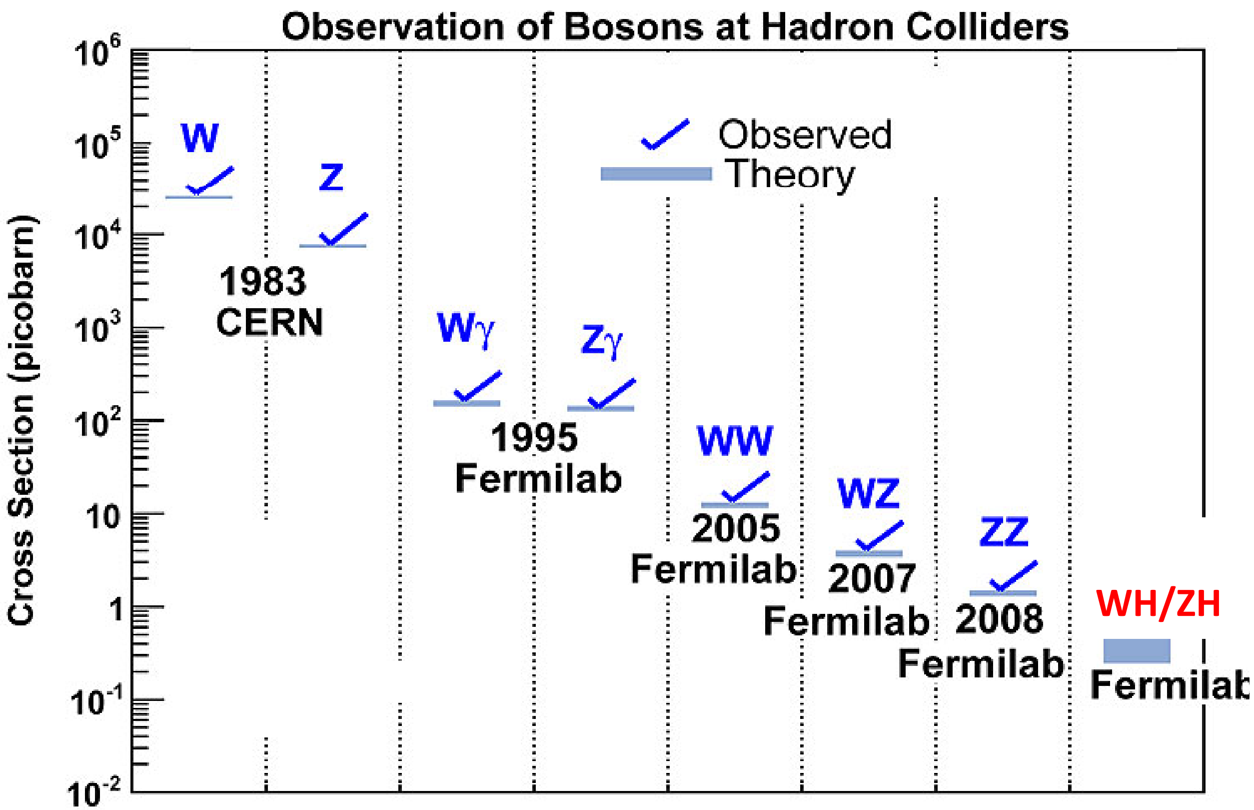


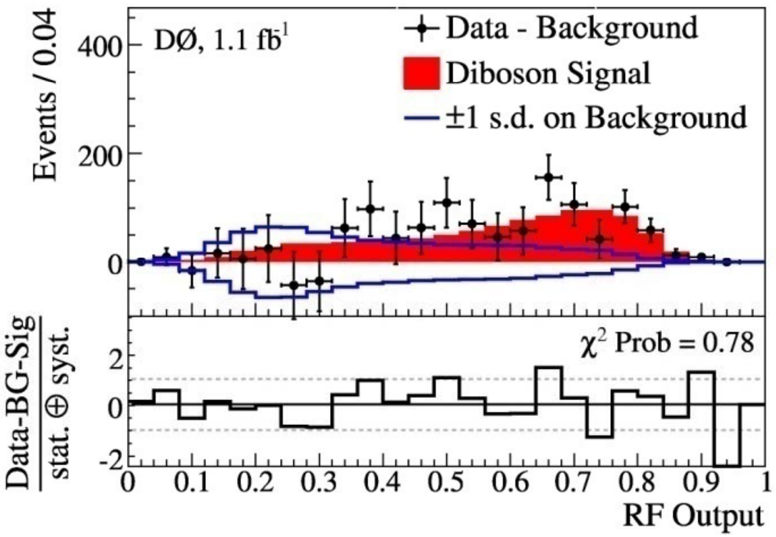


Dibosons

To understand VH Production:
lepton+jets final states

(l = e/μ) + jets
- 5-10 × leptonic BR
- large background from jets
- significant systematic effects
- dijet mass resolution:
  $|M_Z - M_W|$ **<** D0 detector resolution
- WW+WZ
  Evidence by DØ (4.4σ)   2009
  Observation by CDF in   2010





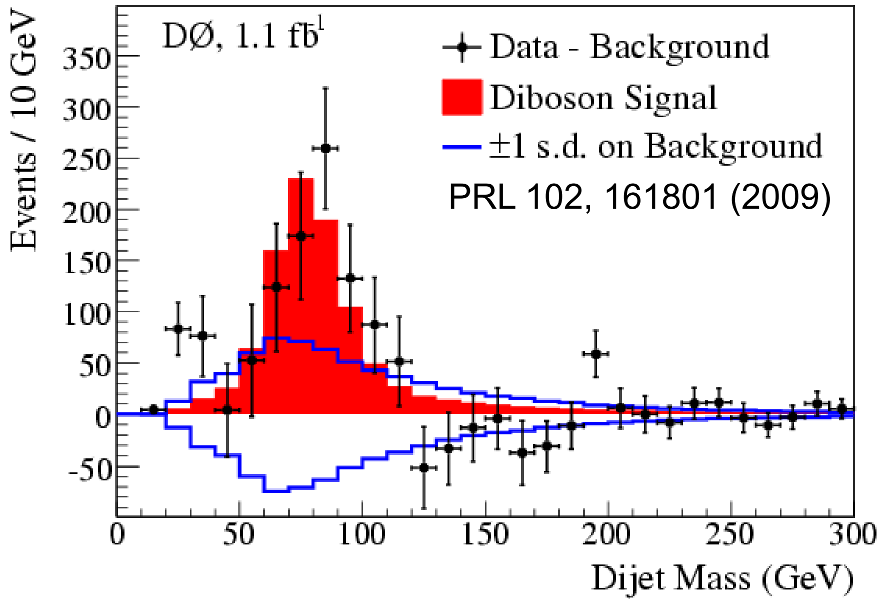
Cross Sections	WW	WW+WZ	WW part
	lνlν [pb]	lνjj [pb]	lνjj [pb]
CDF 4.3/fb	12.1 ± 1.8	18.1 ± 4.1	13.9 ± 3.2
DØ 1.1/fb	11.5 ± 2.2	20.2 ± 4.6	15.5 ± 3.5

SM@NLO σ(WW) = 11.7 ± 0.8 pb
SM@NLO σ(WZ) = 3.5 ± 0.3 pb

Benchmarks

- **WW+WZ→lνjj as a benchmark**

First evidence of the dijet mass resonance from WW+WZ



$$\sigma(WW+WZ) = 20.2 \pm 4.6 \text{ pb}$$

$SM@NLO: \sigma(WW+WZ) = 16.1 \pm 0.9$ pb

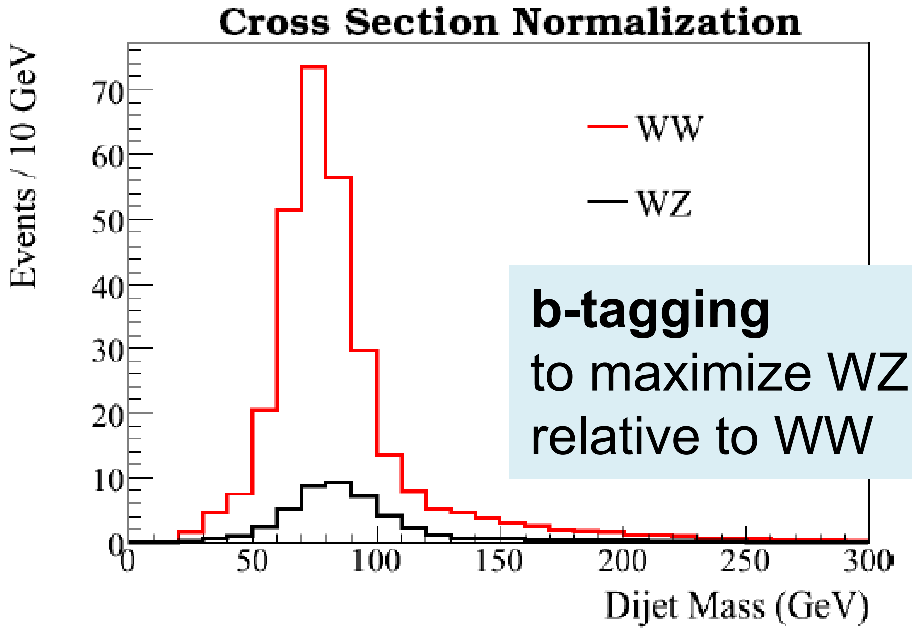
SM@NLO: (l = e+μ)

$\sigma_{WW} \times BR(W\to l\nu; W\to jj) \approx 3.5$ pb

$\sigma_{WZ} \times BR(W\to l\nu; Z\to jj) \approx 0.6$ pb

- **WZ→lνjj as a *benchmark***

For WH, ZH (WZ→lνbb)



b-tagging to maximize WZ relative to WW

Components of the WH/ZH searches: b-tagging, event selection, multivariate analysis, statistical treatment, systematic uncertainties, background modeling

$$\sigma_{WZ} \times BR(W\to l\nu; Z\to bb) \approx 4\sigma_{WH} \times BR(W\to l\nu; H@115\to bb)$$

Event Selection, Signal and Background

Data: **4.3 fb^{-1} of Run IIb Data**

Trigger Selection: electron(+jets) and inclusive trigger (muon)

Electron:

$p_T \geq 15$ GeV, $|\eta|_{DET} < 1.1$

Muon :

$p_T \geq 20$ GeV, $|\eta|_{DET} \leq 1.6$,

$\Delta R_{\mu\text{-jet}} > 0.5$

Global

MET ≥ 20 GeV, $M_T \geq 40 - 0.5 \cdot MET$

$M_T(\mu\nu) < 200$ GeV, $|PV_Z| < 60$ cm

Jets (2 or 3 jet bin):

Minimum 2 vertex confirmed jets

$p_T^{Jet1,Jet2} \geq 20$ GeV, $|\eta|_{DET} < 2.5$

- calibrated jets (data and MC)
- Resolution calibration in MC

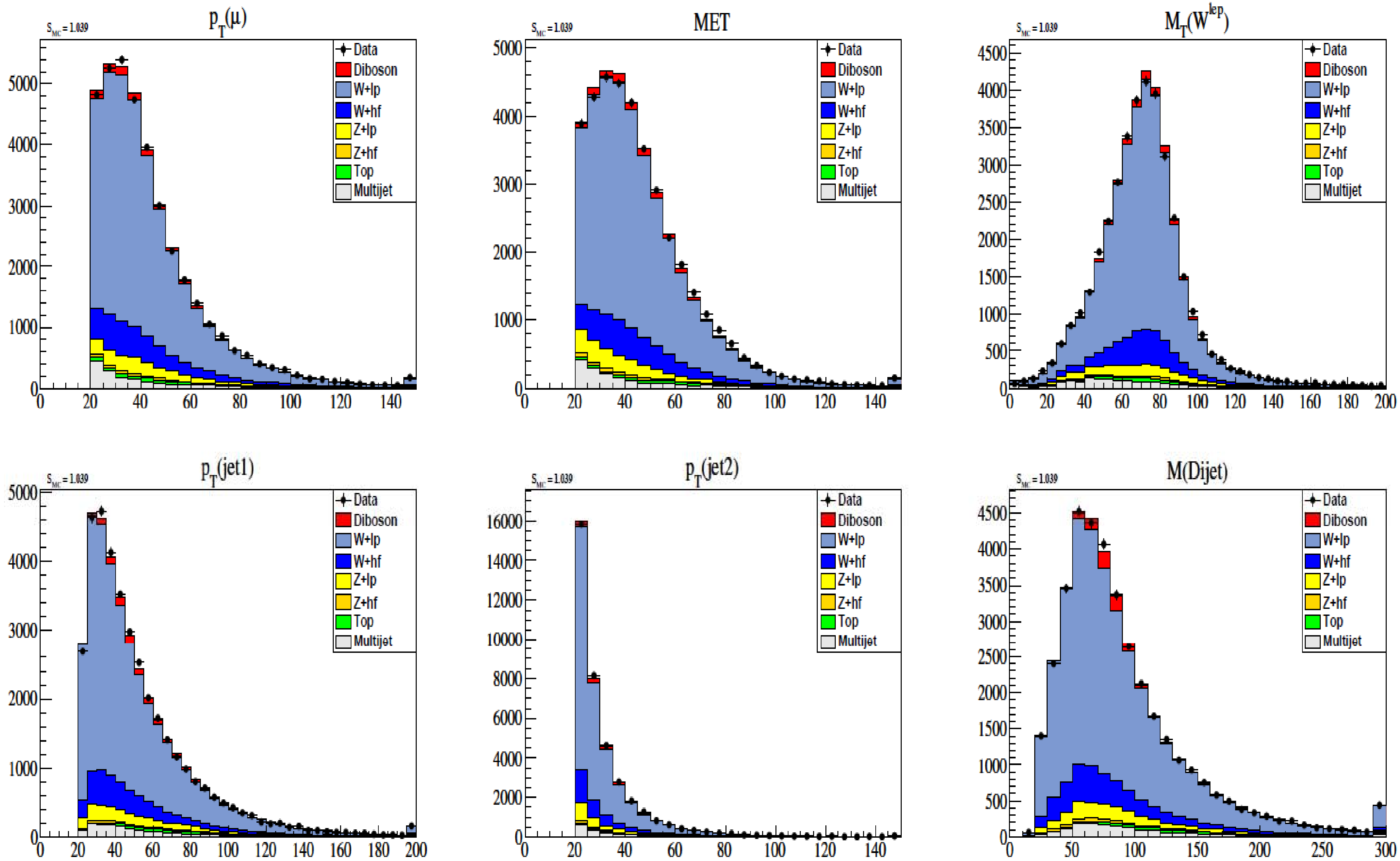
correction to jet/Z p_T imbalance and energy resolution due to the different quark/gluon sample composition

	Events
Data	100 k
W/Z+l.ight f.	70 k
W/Z +Heavy f.	10 k
Top	3 k
Multijet	13 k
Diboson	3.2 k

Event Source	Generator		σ(SM) / σ(WW)	σ(WW) = 11.7 pb
WW	Pythia		1.0	NLO
WZ	Pythia		0.3	NLO
ZZ	Pythia		0.1	NLO
W+light flavor jets	Alpgen	+Pythia	850	from FIT
W+heavy flavor jets	Alpgen	+Pythia	32	from FIT
Z+light flavor jets	Alpgen	+Pythia	32	NNLO
Z+heavy flavor jets	Alpgen	+Pythia	1.1	NNLO
Double-Top	Alpgen	+Pythia	0.6	NNLO
Single-Top	Comphep	+Pythia	0.2	NNLO

data/MC Comparison

Muon channel, 2 jet bin, normalized with a scale factor to match the data (MC×1.04)

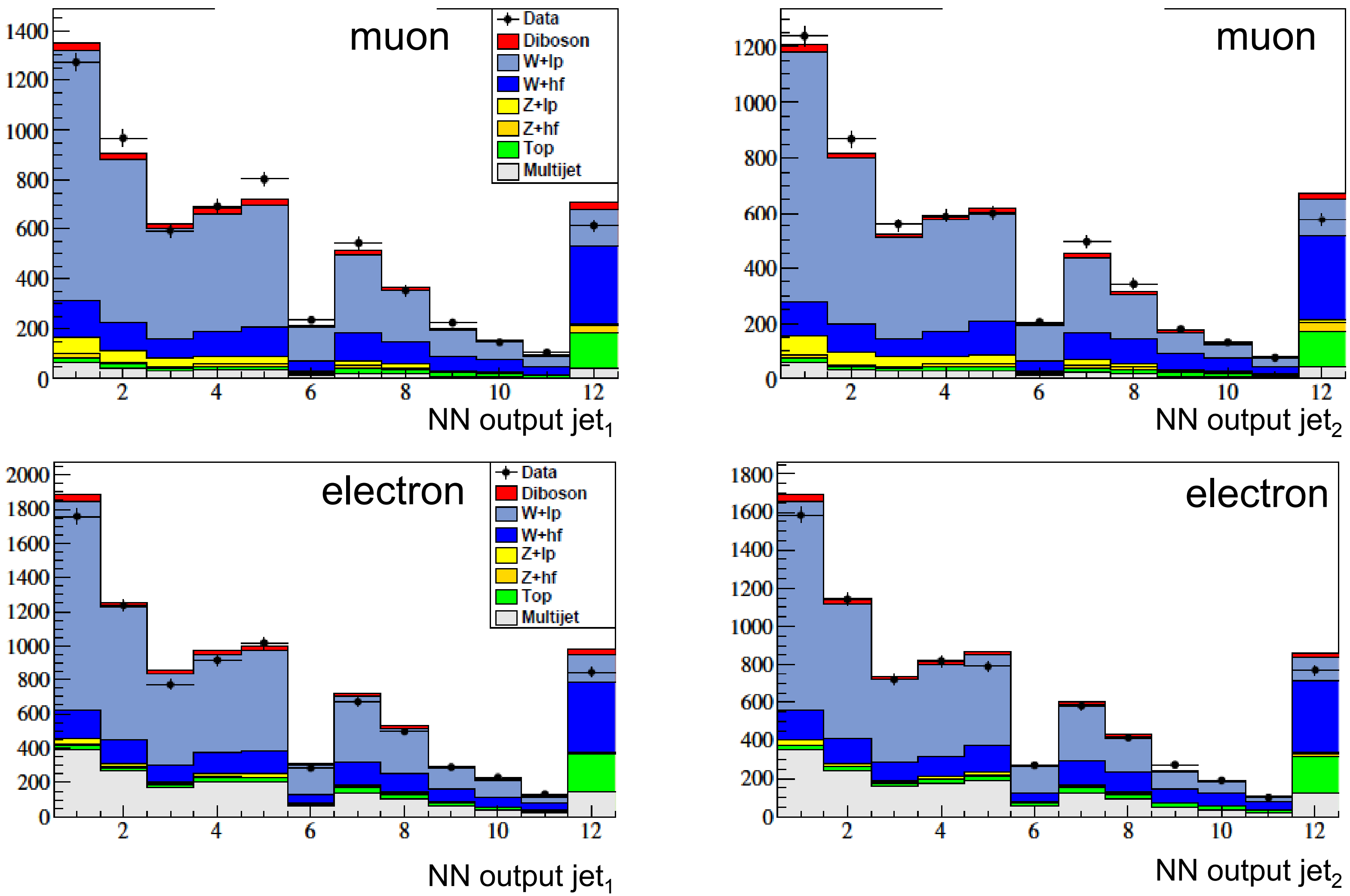


b-tagging

To separate bb-decays we use a Neural Network b-tagging algorithm

NN b-tag output: ≥ L6 for two leading jets (12 identification points)

< L6 or non-taggable: 0-tag bin



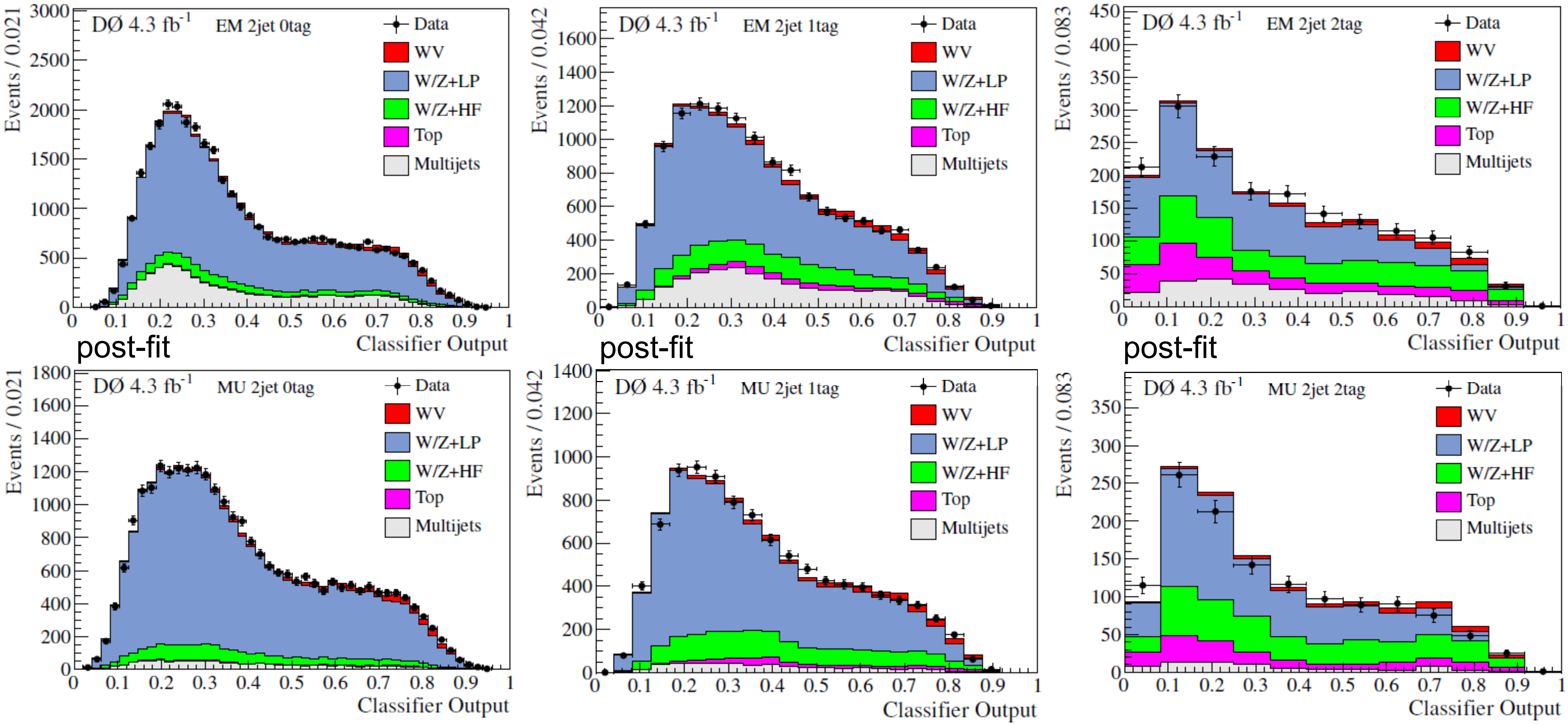
NN b-tag output as an input to a MVA (Random Forest)

Multivariate Analysis

Random Forest (RF)

Input: 15 well described variables to separate signal from background (including min. and max. of b-tag NN output, cf backup slides)

- **0-tag**: trained with all dibosons as signal
- **1-tag**: trained with WZ+ZZ as signal
- **2-tag**: trained with WZ+ZZ as signal



after a combined fit to data over all jet/tag sub-channels

Total Diboson cross section

Obtained from the fit to the **Dijet Mass Distributions** or to the Random Forest output, minimizing a χ^2 function with respect to variations in the systematic uncertainties

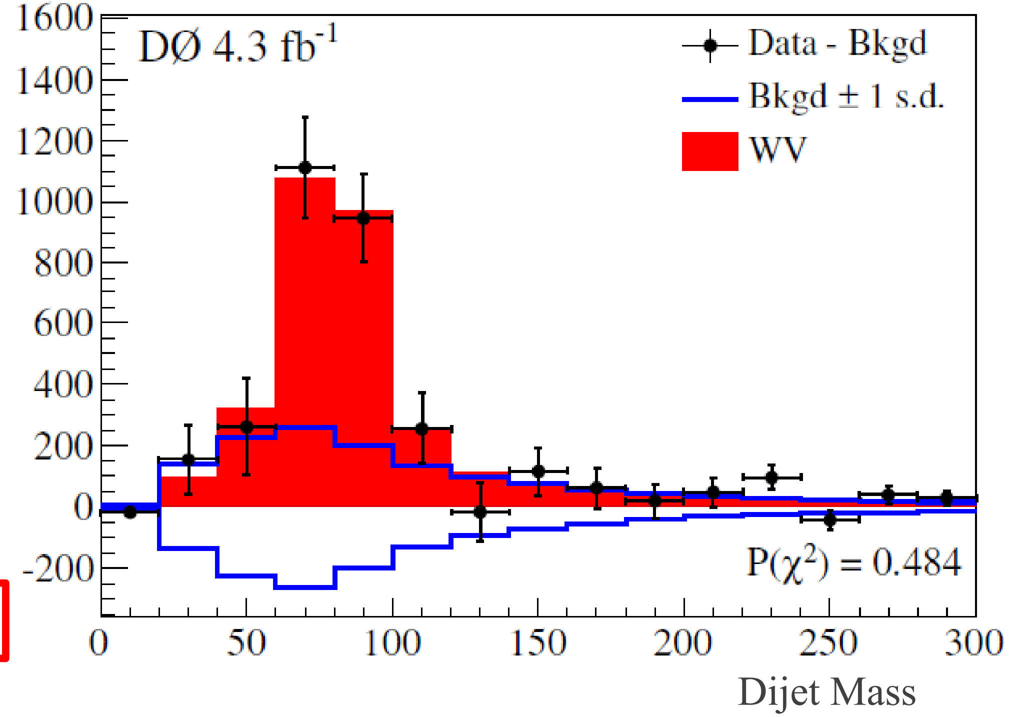
(simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3- jet bins)

- W+jets and diboson (WW+WZ) normalizations are free parameters

	Electron channel	Muon channel
Diboson signal	1725 ± 84	1465 ± 67
W/Z+light-flavor jets	37200 ± 1000	33520 ± 710
W/Z+heavy-flavor jets	5370 ± 610	4850 ± 490
$t\bar{t}$ + single top	1750 ± 130	1214 ± 86
Multijet	10600 ± 1000	1980 ± 380
Total predicted	56700 ± 720	43030 ± 620
Data	56698	43044

ZZ contribution to VV ≈ 1.5%

	Measured $\sigma(WV)$ [pb]
RF Output	$19.6 \pm 1.4\ (\text{stat})\ ^{+2.9}_{-2.7}\ (\text{syst})$
Dijet Mass	$18.3 \pm 1.5\ (\text{stat})\ ^{+3.5}_{-3.3}\ (\text{syst})$



SM@NLO $\sigma(WW+WZ) = 16.1 \pm 0.9$ pb

Total Diboson cross section

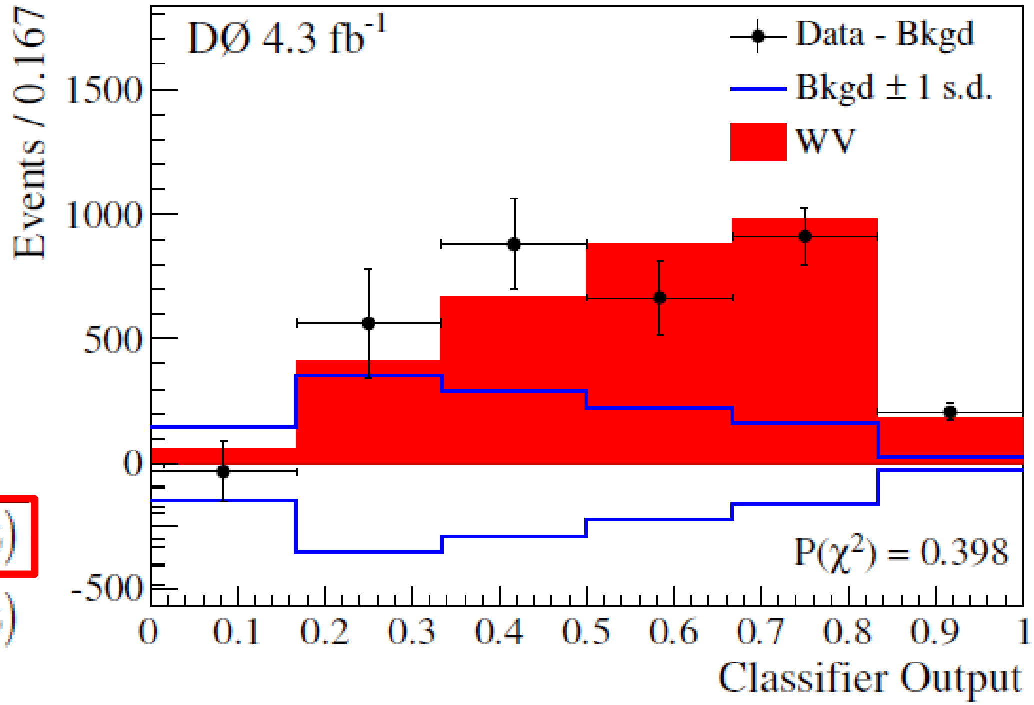
Obtained from the fit to the Dijet Mass Distributions or to the **Random Forest output**, minimizing a χ^2 function with respect to variations in the systematic uncertainties

(simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3- jet bins)

- W+jets and diboson (WW+WZ) normalizations are free parameters

	Electron channel	Muon channel
Diboson signal	1725 ± 84	1465 ± 67
W/Z+light-flavor jets	37200 ± 1000	33520 ± 710
W/Z+heavy-flavor jets	5370 ± 610	4850 ± 490
$t\bar{t}$ + single top	1750 ± 130	1214 ± 86
Multijet	10600 ± 1000	1980 ± 380
Total predicted	56700 ± 720	43030 ± 620
Data	56698	43044

	Measured $\sigma(WV)$ [pb]
RF Output	19.6 ± 1.4 (stat) $^{+2.9}_{-2.7}$ (syst)
Dijet Mass	18.3 ± 1.5 (stat) $^{+3.5}_{-3.3}$ (syst)



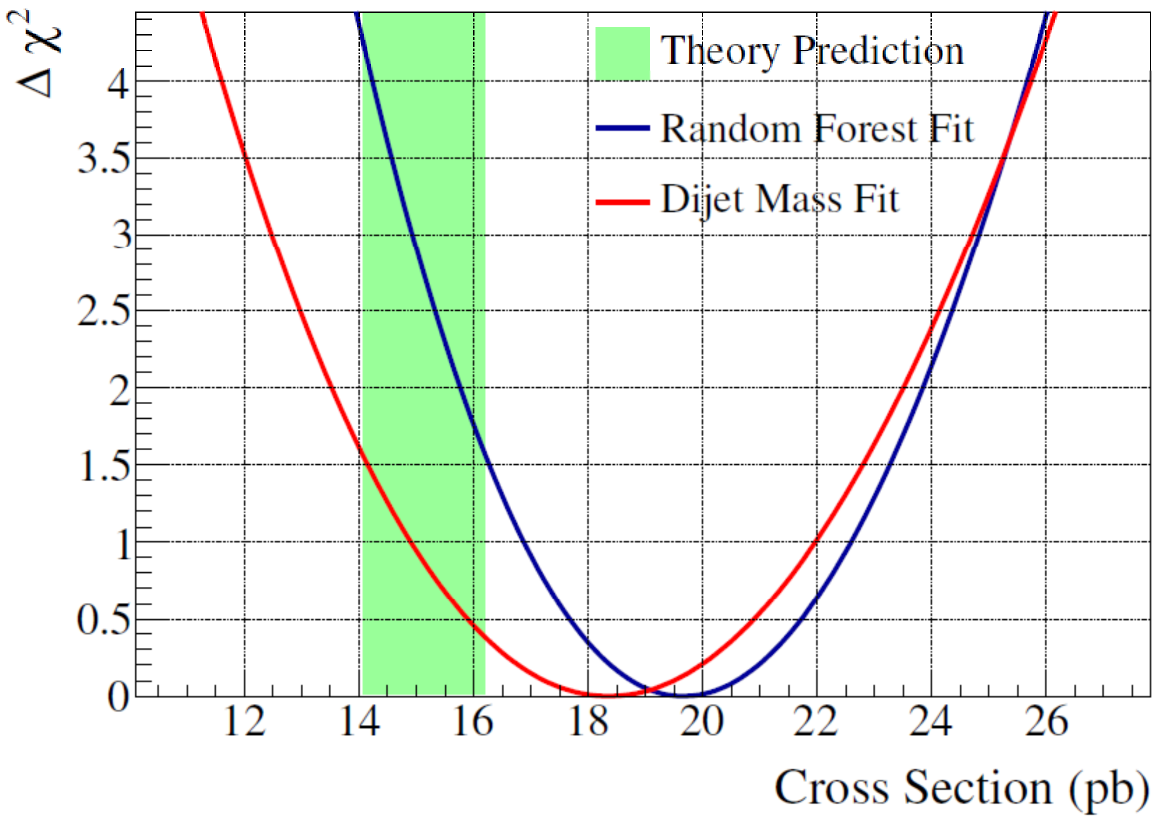
SM@NLO: $\sigma(WW+WZ) = 16.1 \pm 0.9$ pb

Significance

$$\sigma_{RF}(WW+WZ) = 19.6\,{}_{+3.2\,-3.0}\ \text{pb} \Rightarrow \text{significance}: 7.9\sigma\ (\text{expected}: 5.9\sigma)$$

$$\sigma_{M_{JJ}}(WW+WZ) = 18.3\,{}_{+3.8\,-3.6}\ \text{pb} \Rightarrow \text{significance}: 5.6\sigma\ (\text{expected}: 4.6\sigma)$$

Significance gain from the Multi Variate Approach: 22%



Significance:

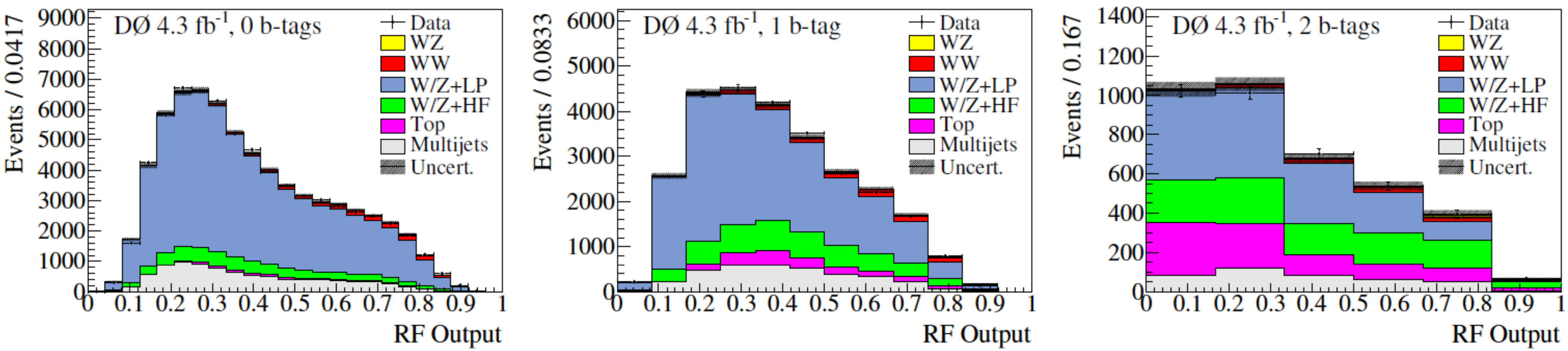
$$\Delta\chi^2 = \left| -2\ln\left(\frac{L(D;S+B,\theta_k)}{L(D;B,\theta_k)}\right)\right| = \left|\chi^2(D;S+B,\theta_k) - \chi^2(D;B,\theta_k)\right|$$

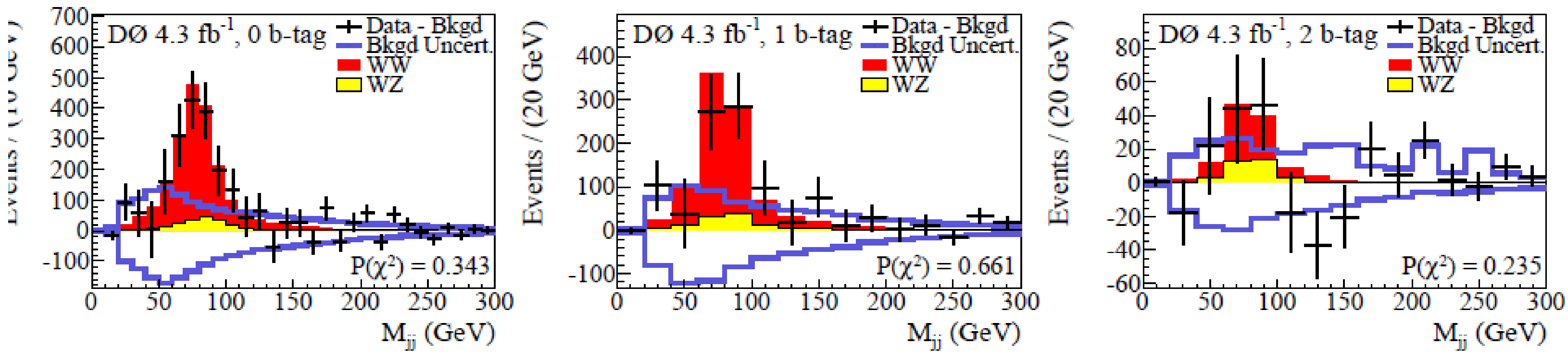
D: data or background-only

WW & WZ cross sections

From the fit to data of the Random Forest output (or the Dijet Mass Distributions), minimizing a χ^2 function with respect to variations in the systematic uncertainties

(simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3- jet bins)





WW & WZ cross sections

From the fit to data of the Random Forest output (or the Dijet Mass Distributions), minimizing a χ^2 function with respect to variations in the systematic uncertainties

(simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3- jet bins)

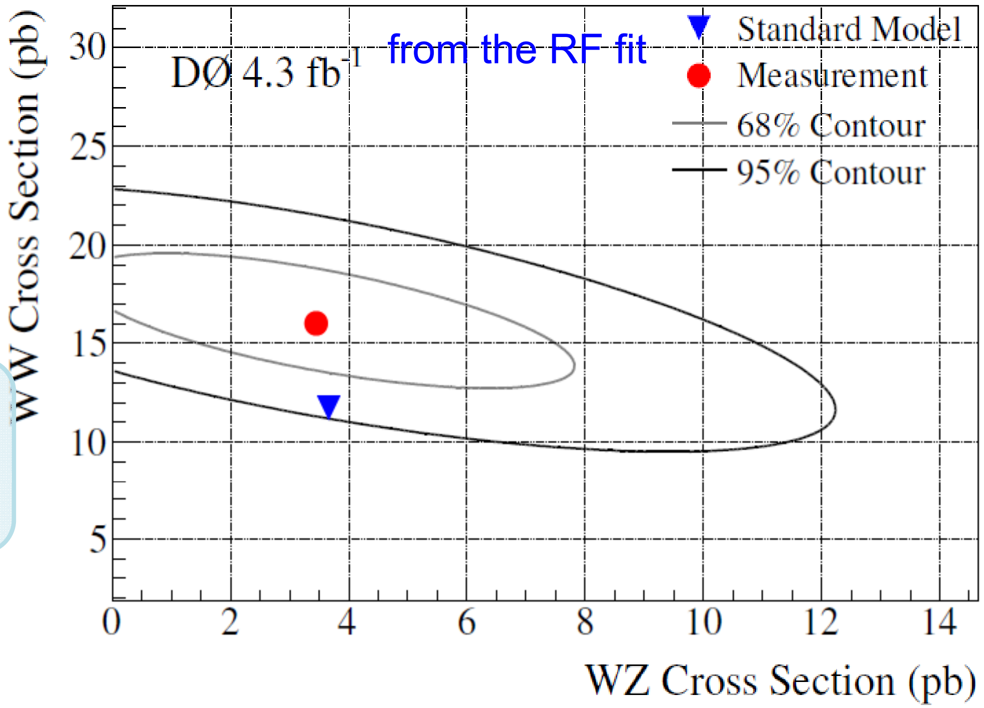
Taking W+jets, WW and WZ normalizations are free parameters:

$$\sigma_{RF}(WW) = 15.9^{+3.7}_{-3.2} \text{ pb}$$

$$\sigma_{RF}(WZ) = 3.3^{+4.1}_{-3.3} \text{ pb}$$

SM@NLO $\sigma(WW) = 11.7 \pm 0.8$ pb

SM@NLO $\sigma(WZ) = 3.5 \pm 0.3$ pb



Constraining WW :

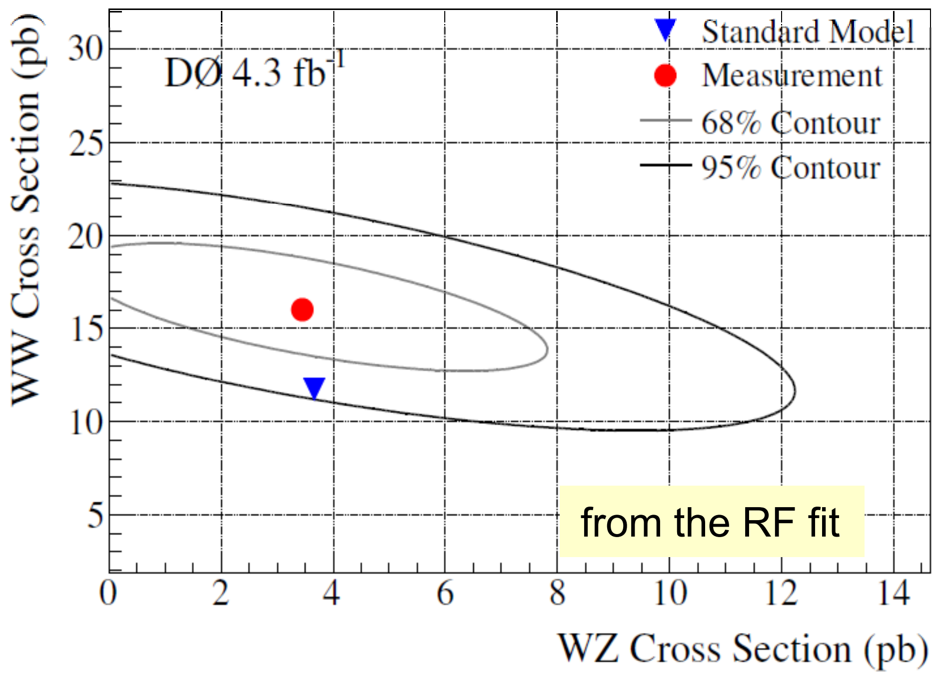
$$\sigma^{WW\ fixed}_{RF}(WZ) = 6.5 \pm 3.1 \text{ pb}$$

observed significance: 2.2σ (expected: 1.2σ)

Summary

- DØ analyzed lνjj final states (e,μ) using 4.3 fb^{-1} of RunIIb data
- Kinematic distributions in good agreement with the SM predictions
- Measured WW+WZ cross section in agreement with the SM, using similar techniques as in the Higgs WH analyis ➔yields 7.9σ observed significance
- Extraction of the WZ signal (2.2 σ observed)

Cross Sections	WW	WW+WZ	WW part
	lνlν [pb]	lνjj [pb]	lνjj [pb]
CDF 4.3/fb	12.1 ± 1.8	18.1 ± 4.1	
DØ 1.1/fb	11.5 ± 2.2	20.2 ± 4.6	
DØ 4.3/fb		**19.6** $^{+3.2}_{-3.0}$	**15.9** $^{+3.7}_{-3.2}$



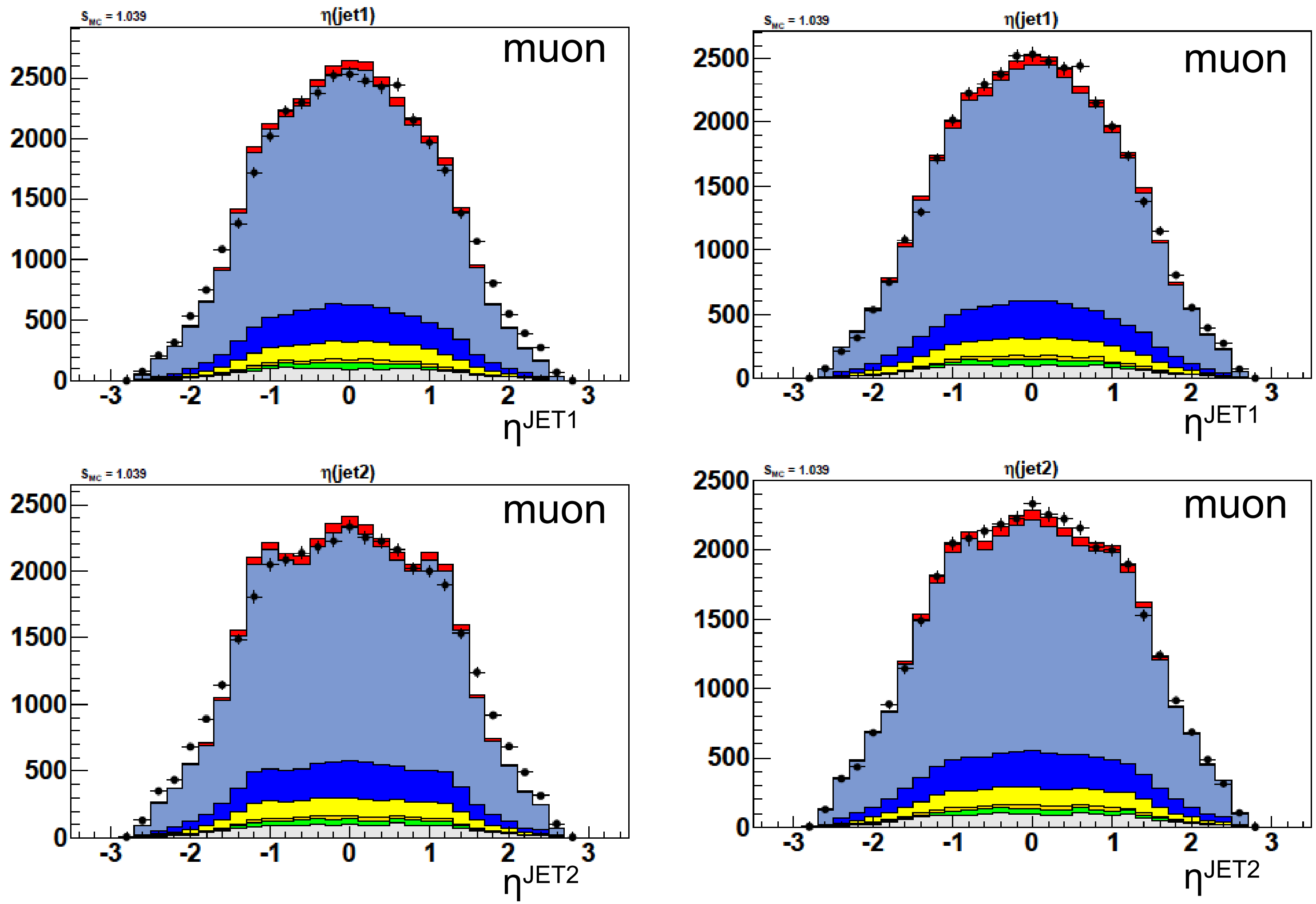
Backup Slides

Additional MC Corrections

ALPGEN jet η Reweighting

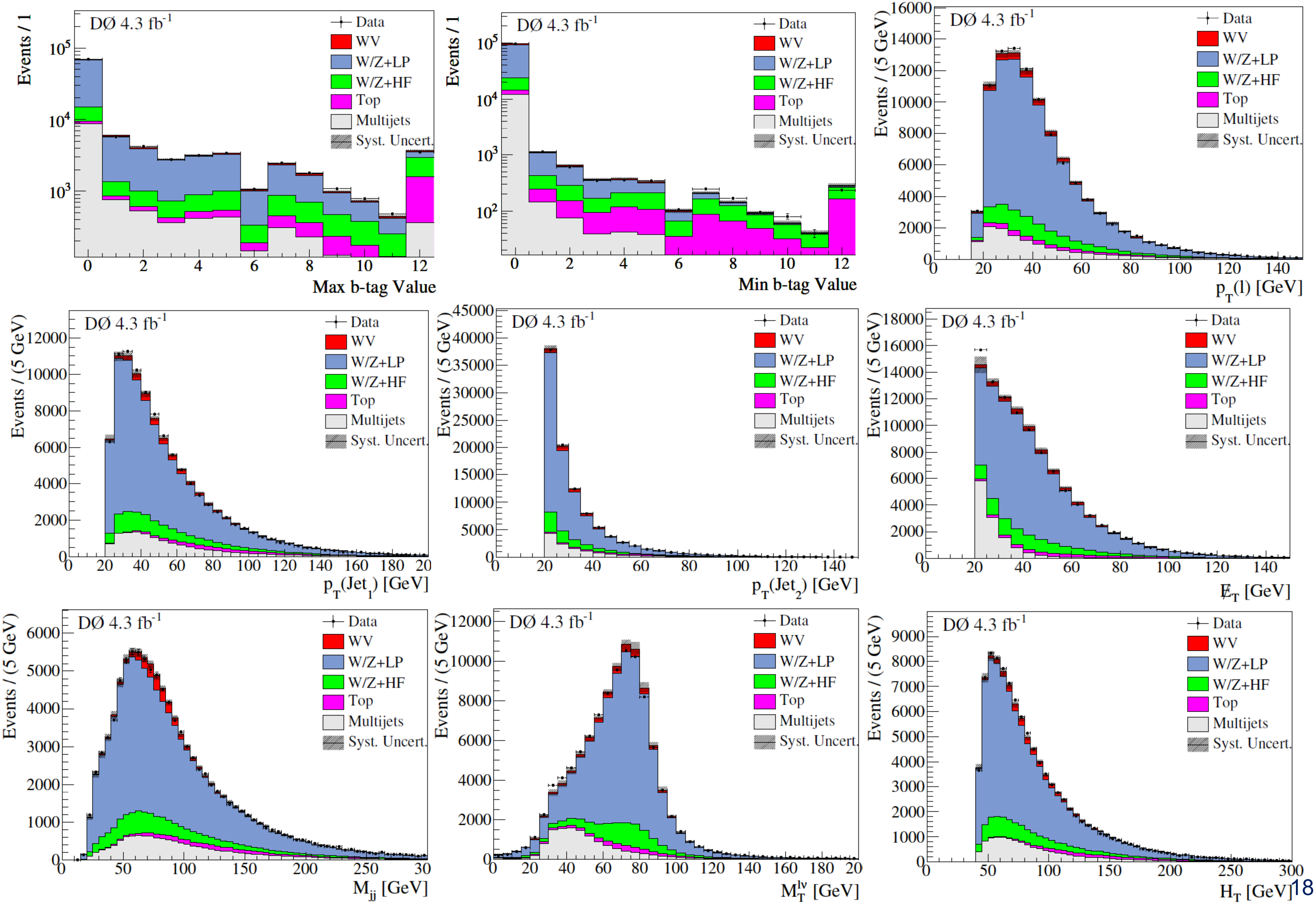
Correction to W/Z+jets MC samples [shape ~ (data-QCD-nonVjets)/Vjets]

Before correction *After correction*



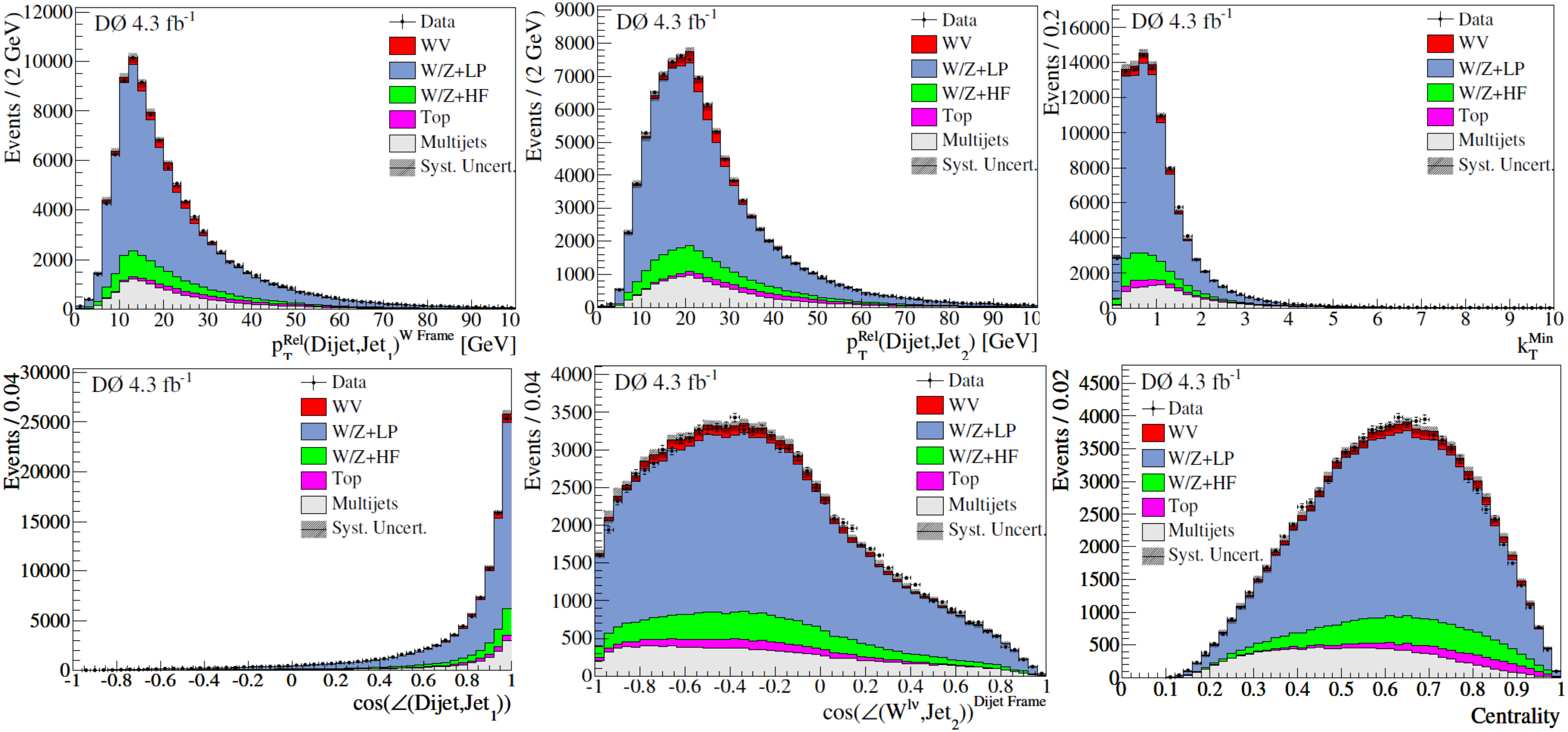
Backup: RF Input Variables (I)

post-fit



Backup: RF Input Variables (II)

post-fit



Modeling of SM processes

RunIIb1 MC

Event Source	Generator	σ(SM) / σ(WW)	σ(WW) = 11.7 pb
WW	Pythia	1.0	NLO
WZ	Pythia	0.3	NLO
ZZ	Pythia	0.1	NLO
W+light flavor jets	Alpgen +Pythia	850	from FIT
W+heavy flavor jets	Alpgen +Pythia	32	from FIT
Z+light flavor jets	Alpgen +Pythia	32	NNLO
Z+heavy flavor jets	Alpgen +Pythia	1.1	NNLO
Double-Top	Alpgen +Pythia	0.6	NNLO
Single-Top	Comphep +Pythia	0.2	NNLO

Standard MC Corrections

(to account for differences from data)

- Reconstruction and Identification efficiencies of leptons/jets
- Trigger selection
- Z boson p_T modeling (njet-dependent)
- W boson p_T modeling (inclusive)
- Luminosity reweighting, beam z-position reweighting

Multijet Background

(jet misidentified as a lepton)

- Estimated from (multijet enriched) data
  - Muon channel: Reverse muon isolation cuts
  - Electron channel: Loose electron quality criteria
- Corrected for contributions already accounted for by MC
- Normalization: template fit of $M_T(W\rightarrow l\nu)$

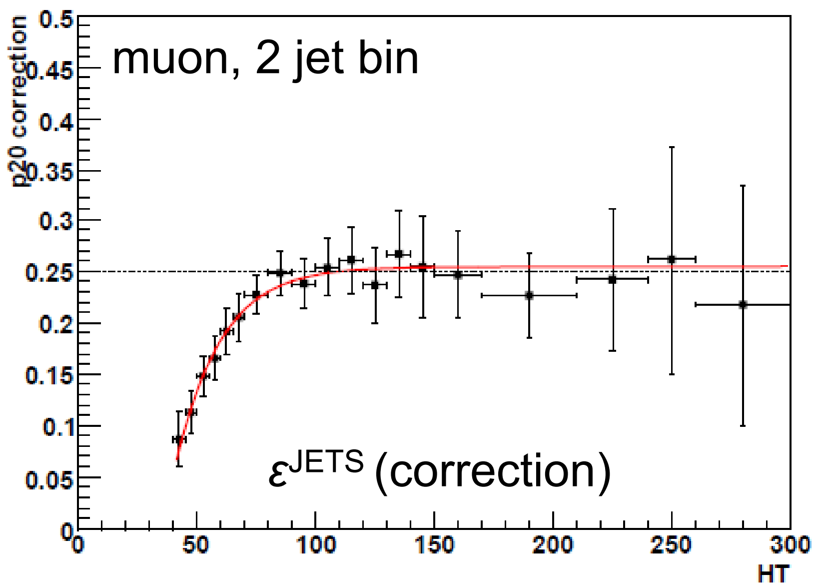
Additional MC Corrections

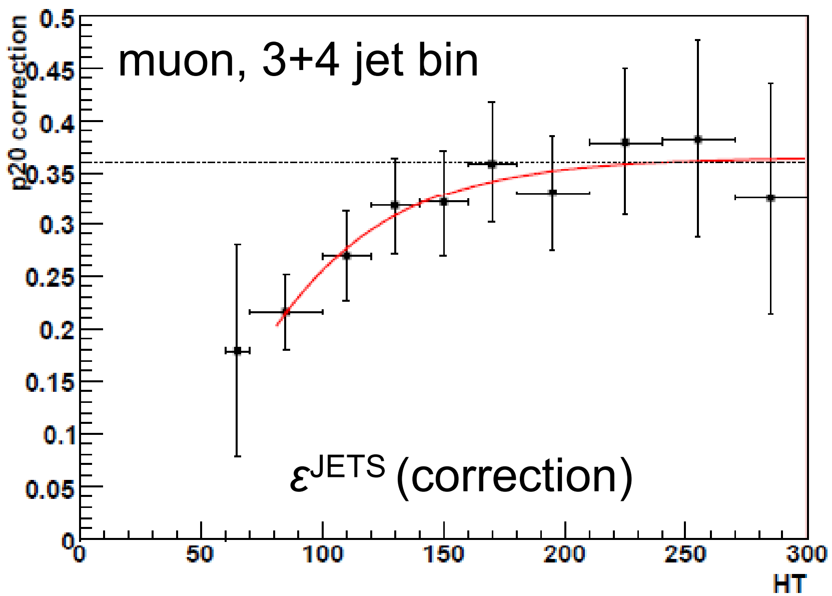
Inclusive Muon Trigger Modeling

- Inclusive trigger gives ~50% increase in statistics on Single Muon OR (additional data collected by jet triggers mainly)
- Modeling: bootstrap from SMOR parameterized in HT ($p_T^{j1}+p_T^{j2}$)
- Correction: data driven, as a function of HT

$$\varepsilon^{Inclusive} = \varepsilon^{SMOR} + \varepsilon^{JETS}\ (\text{correction})\ (\max \varepsilon^{JETS} < 1-\varepsilon^{SMOR})$$

$$\varepsilon^{JETS} = \left[\frac{N^{data}_{Inclusive} - QCD_{Inclusive} - N^{data}_{SMOR} - QCD_{SMOR}}{N^{MC}_{Inclusive}}\right]$$





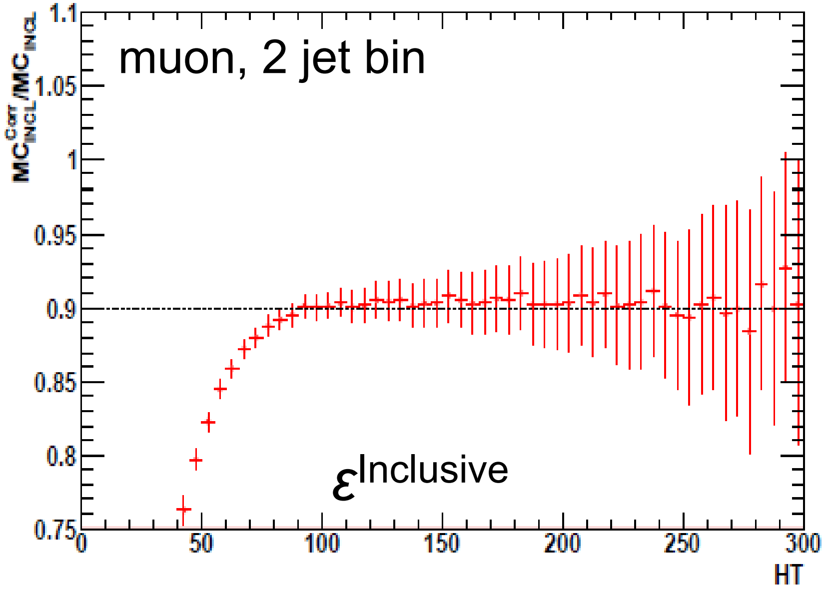
Additional MC Corrections

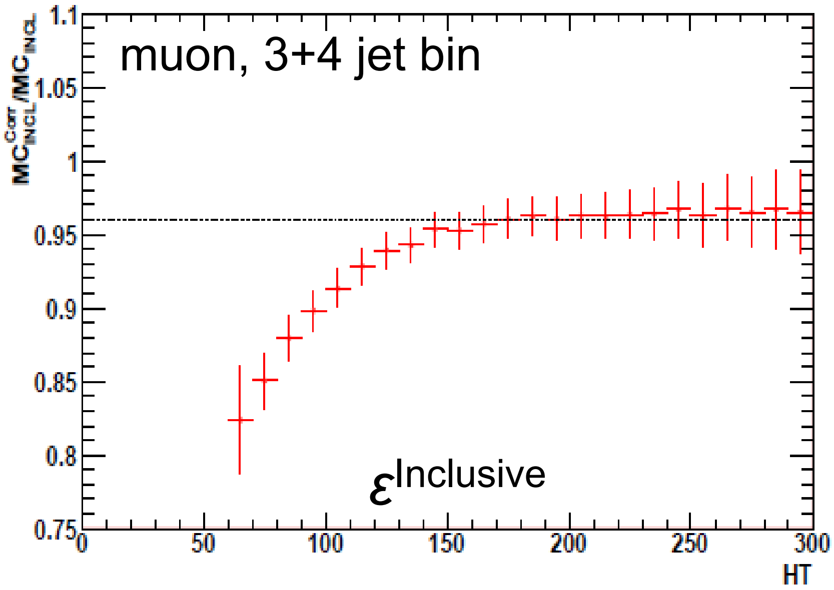
Inclusive Muon Trigger Modeling

- Inclusive trigger gives ~50% increase in statistics on Single Muon OR (additional data collected by jet triggers mainly)
- Modeling: bootstrap from SMOR parameterized in HT ($p_T^{j1}+p_T^{j2}$)
- Correction: data driven, as a function of HT

$$\varepsilon^{\text{Inclusive}} = \varepsilon^{\text{SMOR}} + \varepsilon^{\text{JETS}} \text{ (correction) } (\max \varepsilon^{\text{JETS}} < 1-\varepsilon^{\text{SMOR}})$$

$$\varepsilon^{\text{JETS}} = \left[\frac{N^{\text{data}}_{\text{Inclusive}} - QCD_{\text{Inclusive}} - N^{\text{data}}_{\text{SMOR}} - QCD_{\text{SMOR}}}{N^{\text{MC}}_{\text{Inclusive}}} \right]$$

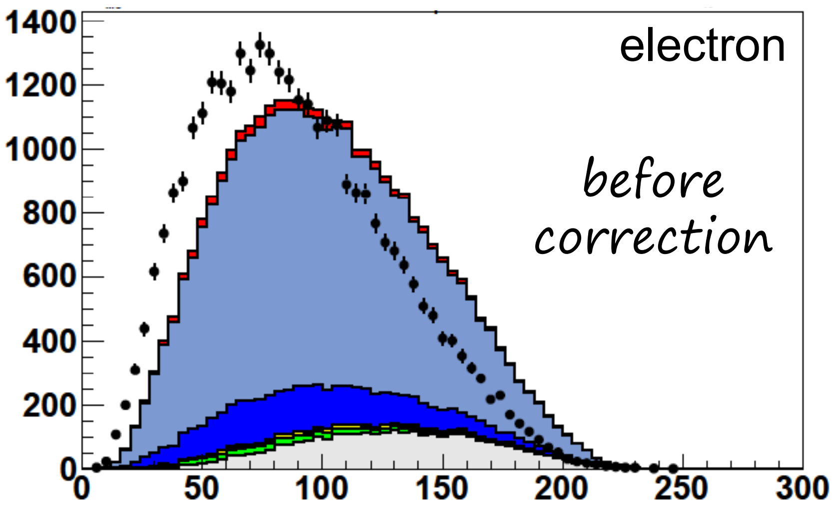


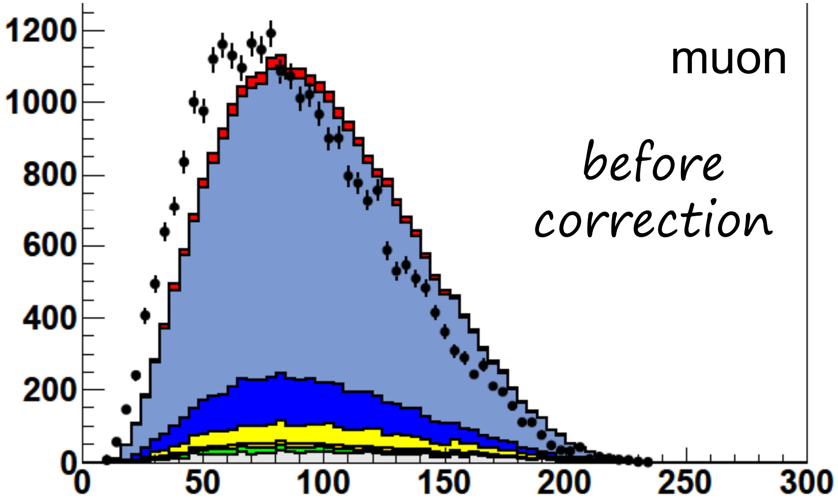


Additional MC Corrections (I)

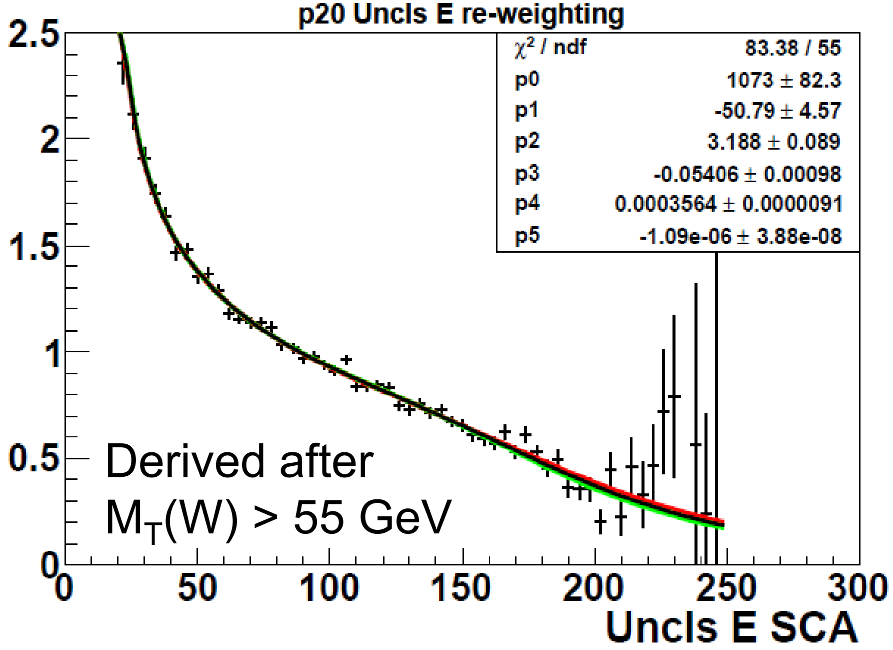
(1) Unclustered Energy (UE) Reweighting

- Sum of the energy deposits in EM, HAD, ICD not associated to a jet
- Overlaid MB, low p_T objects (jets, muons) not accounted for in missing E_T
- Correction to all MC samples [shape ~ (data-QCD)/MC]





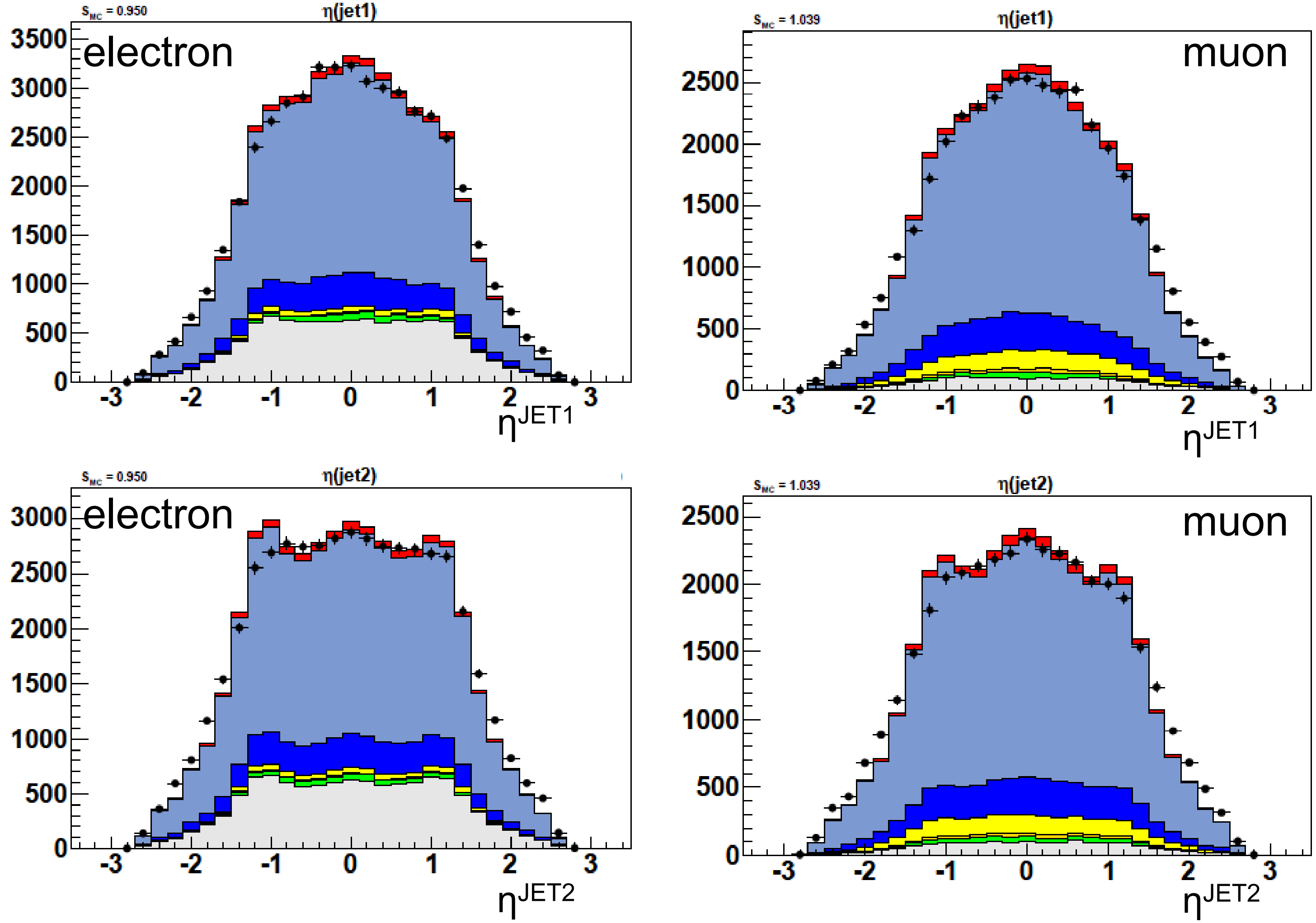
Derived for ±1σ
(fit systematics)



Additional MC Corrections (II)

(2) ALPGEN jet η Reweighting

Before correction

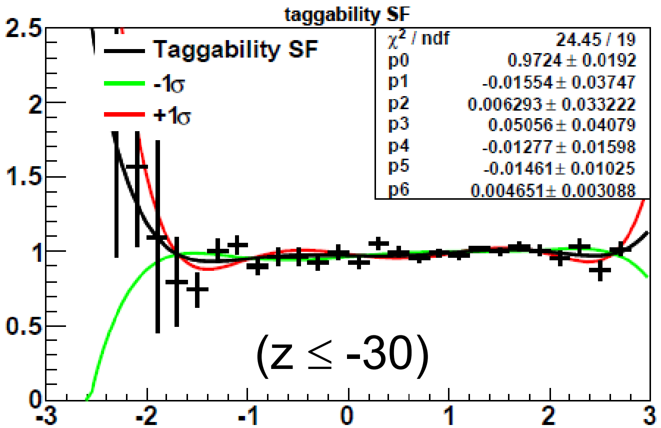


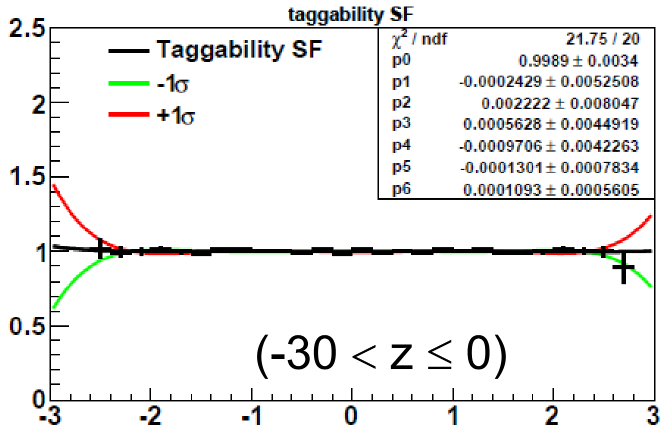
b-tagging

- To separate bb-decays we use NN b-tagging algorithm
- Taggable jets are good vertex confirmed jets ($\varepsilon_{data} < \varepsilon_{MC}$)

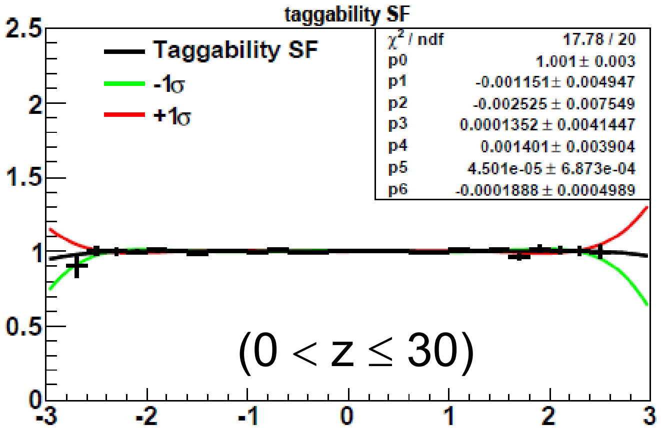
MC correction for each η_{jet} in different z_{PV} bins:
($z \le -30$), ($-30 < z \le 0$), ($0 < z \le 30$), ($30 < z$)

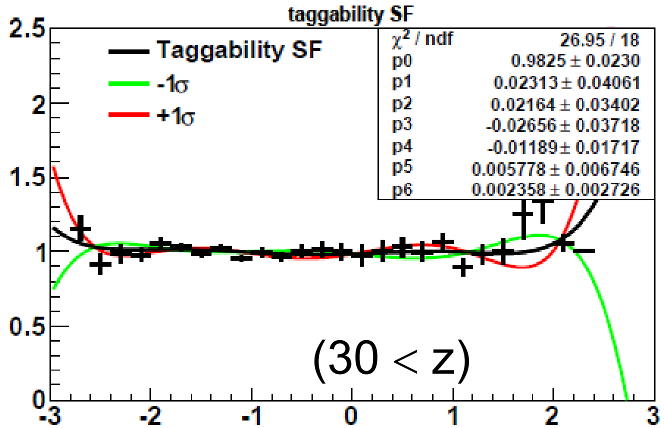
$$SF_{taggable}(z,\eta) = \frac{\varepsilon_{data}(z,\eta)}{\varepsilon_{MC}(z,\eta)}; \quad SF_{non\text{-}taggable}(z,\eta) = \frac{1-\varepsilon_{data}(z,\eta)}{1-\varepsilon_{MC}(z,\eta)}$$





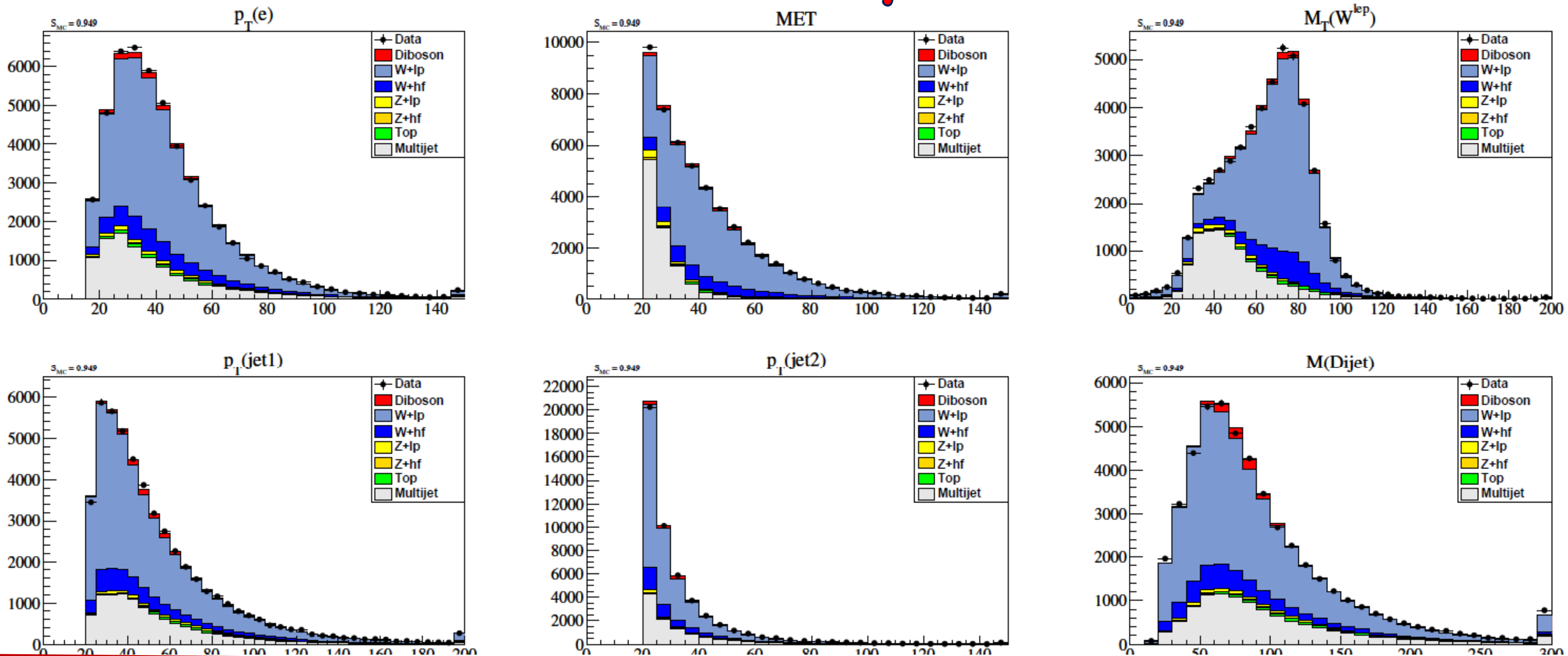
closure test



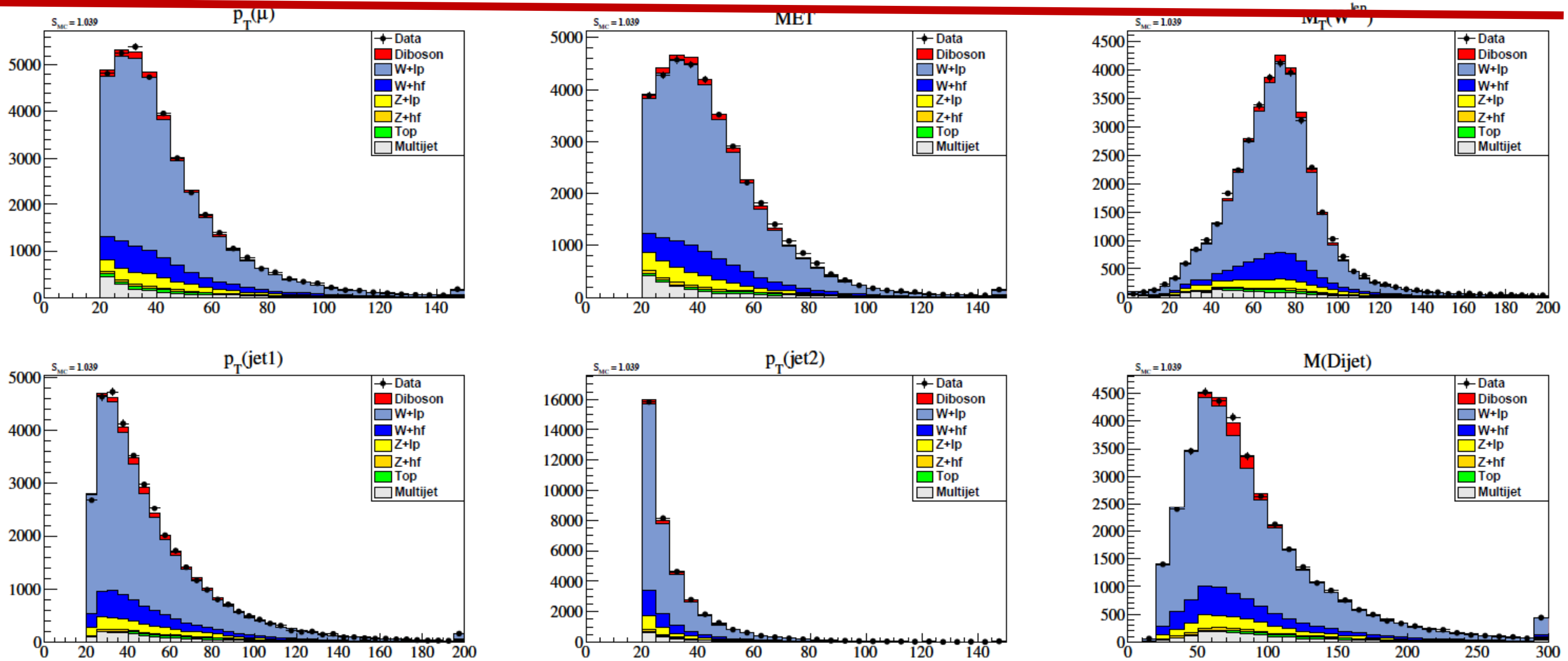


Pre-fit data/MC Comparison

Electron channel, 2 jet bin, normalized with a scale factor to match the data (MC×0.95)



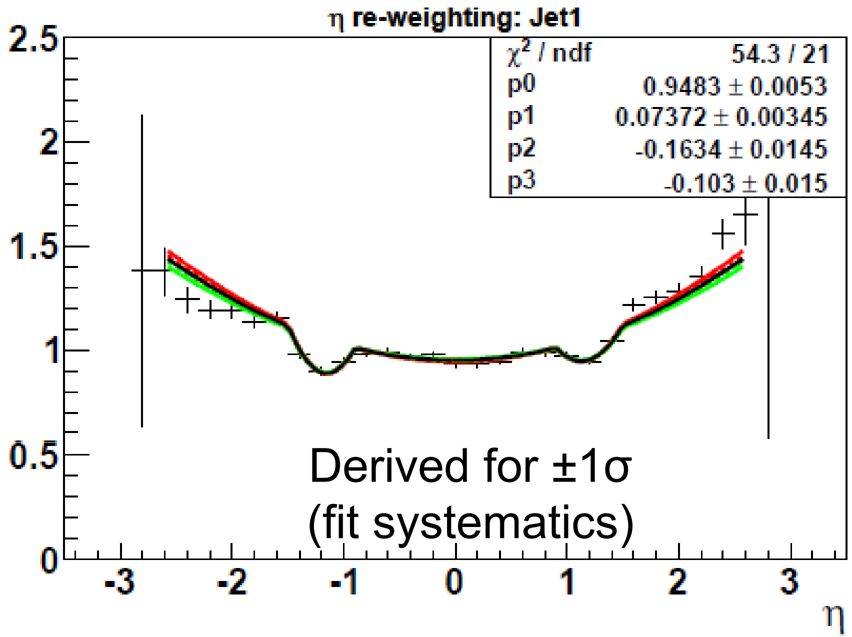
Muon channel, 2 jet bin, normalized with a scale factor to match the data (MC×1.04)

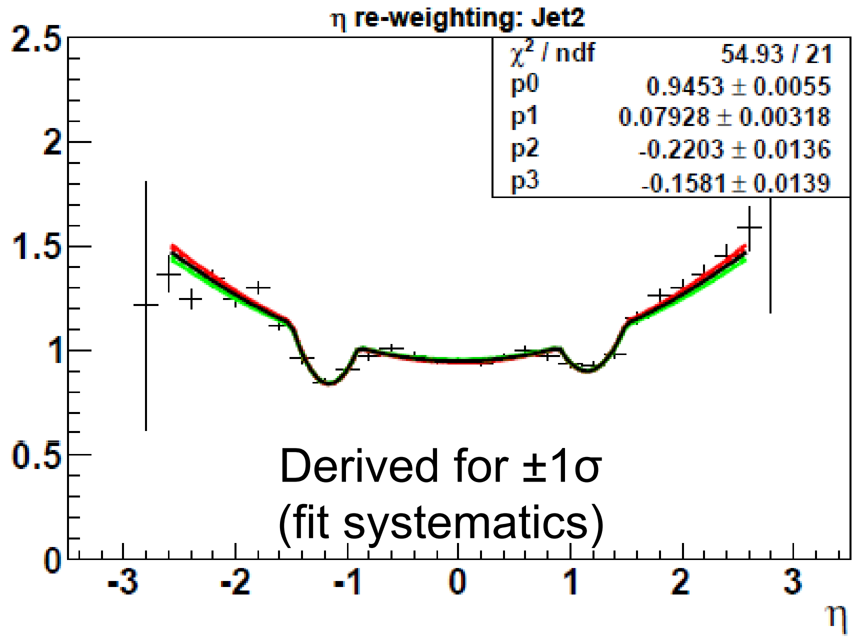


Additional MC Corrections (II)

(2) ALPGEN jet η Reweighting

- Correction to W/Z+jets MC samples [shape ~ (data-QCD-nonVjets)/Vjets]





Also derived for σ(VV)×0 and σ(VV)×2 (fit systematics)

Additional MC Corrections (cont'd)

(4) ALPGEN parton-jet p_T Matching Reweighting

- minimum p_T for jet clusters that are used for the MLM jet-parton matching procedure
- recommended: the generator level jet p_T cut + 20% (or 5 GeV if larger)

(5) Diboson Modeling

- LO-to-NLO correction at the generator level using MC@NLO+Herwig
- 2D reweighting (p_T VV - leading V) applied to WW and WZ events
- Systematic uncertainty is half the size of the correction

I. Each reweighting preserves the total normalization of the MC that is being reweighted

II. Two systematic uncertainties for each reweighting:
- From the fit uncertainty; vary the fit parameter that causes the most change in shape, using the covariance matrix to change the other parameters accordingly
- The uncertainty on the diboson cross section (0 and 2 × σ_{SM}), taken as correlated between reweightings