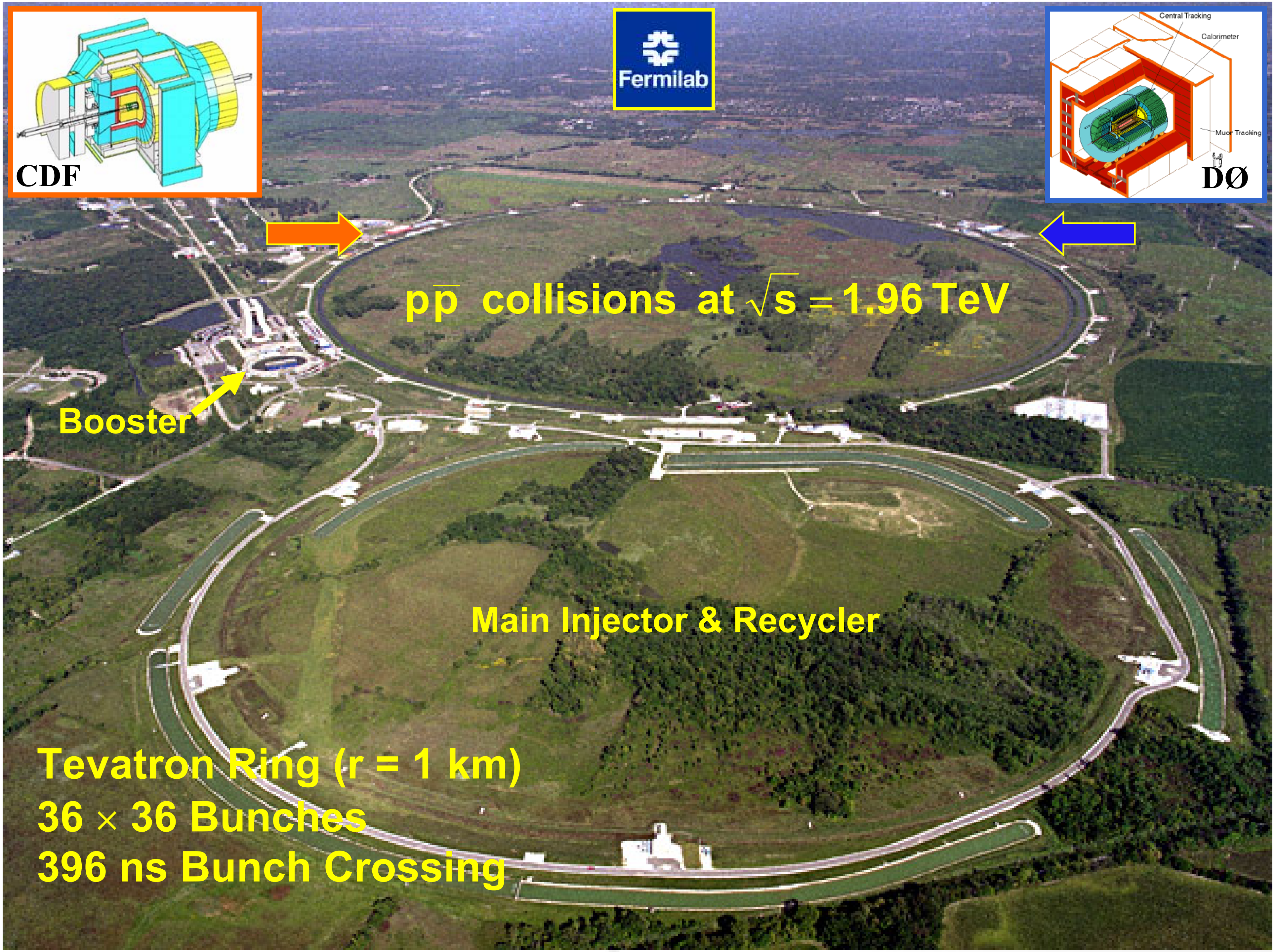
WW and WZ Production at the Tevatron

Jadranka Sekaric
(Florida State University, DØ)

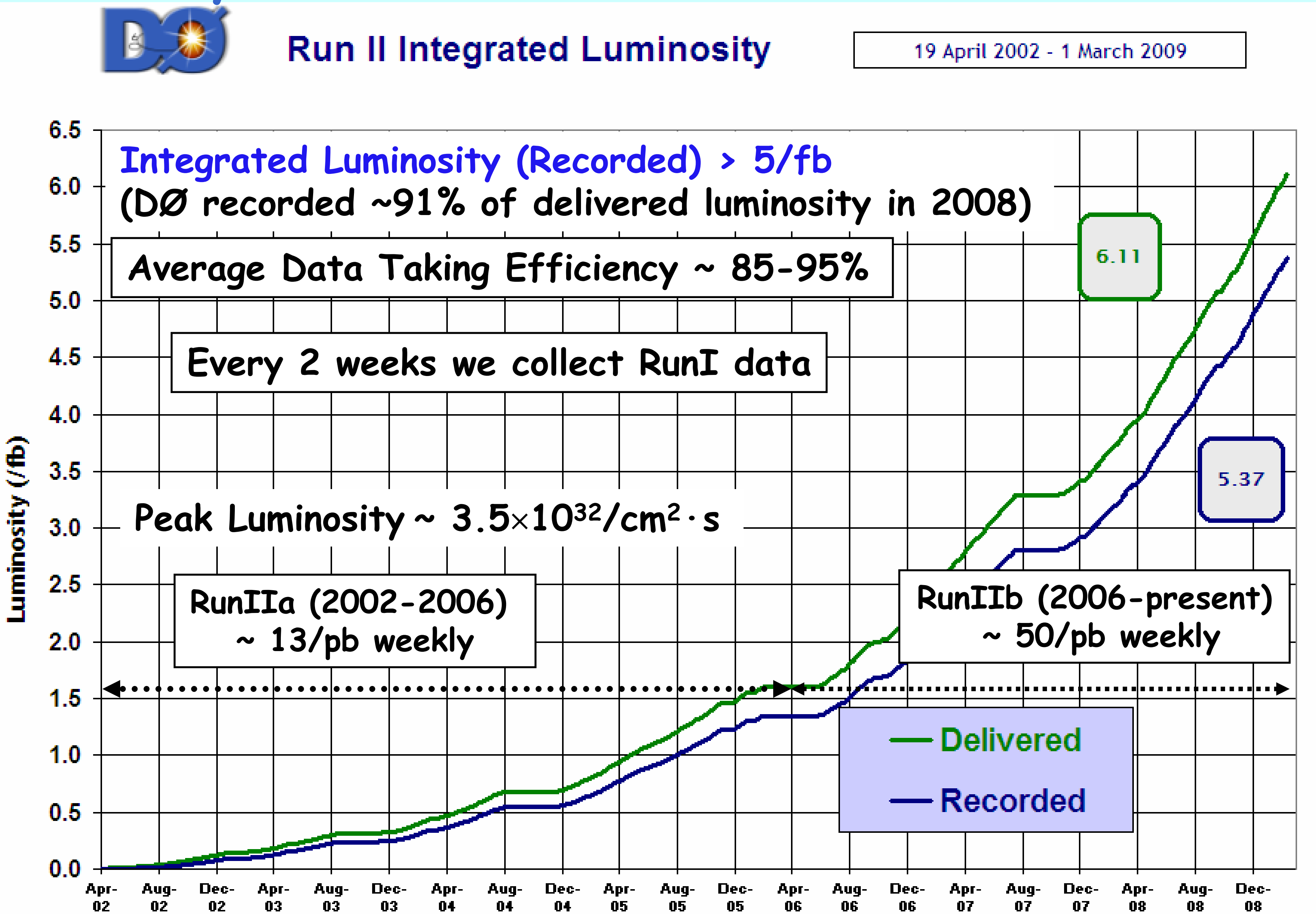
Physics Seminar 2009, March 25th, DESY-Zeuthen, Germany

Outline:

- Tevatron Introduction
- Motivation for Diboson studies
- Overview of the WW/WZ results at the Tevatron
- Studies related to WW+WZ → $l\nu jj$ at DØ
- Cross section measurement and statistical significance


CDF
Fermilab
Central Tracking
Calorimeter
Muon Tracking
DØ
$p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV
Booster
Main Injector & Recycler
Tevatron Ring (r = 1 km)
36×36 Bunches
396 ns Bunch Crossing

Luminosity Status




The
DØ
Collaboration
18 Countries
90 Institutions
554 Scientists

The DØ Detector

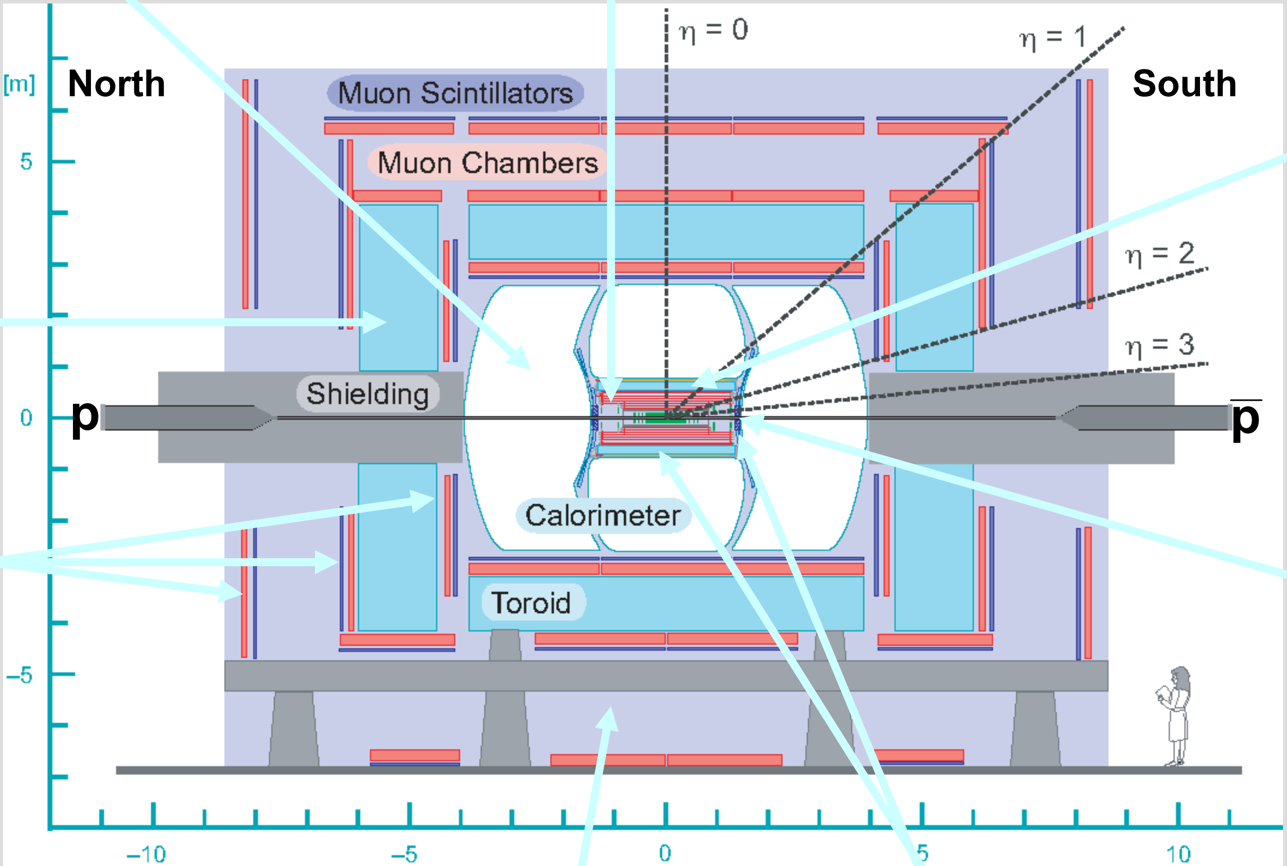
Multipurpose detector

Pseudo-rapidity: $\eta = -\ln\left[\tan\frac{\theta}{2}\right]$

Muon ID	$\lvert\eta\rvert \le 2$
Tracking	$\lvert\eta\rvert \le 3$
EM/Jet ID	$\lvert\eta\rvert \le 4$

Calorimeter (EM, FH, CH)

Tracker (SMT and CFT)



2T Solenoid Magnet

1.8T Toroid Magnet

3 Layer Muon System

Inter Cryostat Detector (ICD)

Electronics, Trigger, DAQ

Preshower detectors

EM = Electromagnetic
FH = Fine Hadronic
CH = Coarse Hadronic

SMT = Silicon Microstrip Tracker Vertex Detector
CFT = Central Fiber Tracker

Motivation for Diboson Physics

Pair production:

W/Z production in association with photon (Wγ, Zγ)
WW, WZ, ZZ

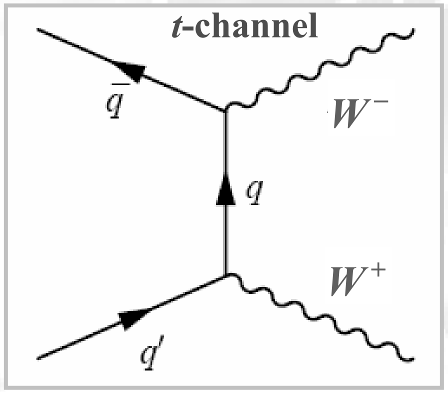
Electroweak Physics

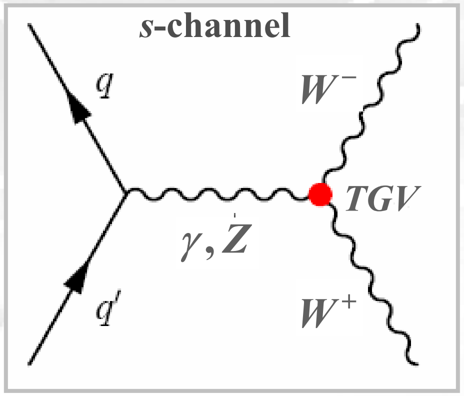
Precision measurements (tests of the SM predictions)
Search for New Physics (low energy effects; indirect searches)

Cross sections
Kinematic distributions
Gauge Boson Couplings (triple/trilinear, quartic)

Disagreement with the SM expectation would indicate the presence of New Physics

t - channel exchange diagram dominates total cross section





s - channel exchange diagram, sensitive to anomalous couplings due to the existence of trilinear gauge boson vertex (TGV)

Electroweak Physics

Trilinear Gauge boson Couplings

- **WW, WZ (Wγ) production (WWγ/WWZ vertices): charged couplings**

$$\frac{L_{WWV}}{g_{WWV}} = i g_1^V \left(W_{\mu\nu}^* W^\mu V^\nu - W_\mu^* V_\nu W^{\mu\nu}\right) + i\kappa_V W_\mu^* W_\nu V^{\mu\nu} + i\frac{\lambda_V}{M^2} W_{\lambda\mu}^* W^\mu{}_\nu V^{\nu\lambda}$$
$$- g_4^V W_\mu^* W_\nu \left(\partial^\mu V^\nu + \partial^\nu V^\mu\right) + g_5^V \varepsilon^{\mu\nu\lambda\rho}\left(W_\mu^* \partial_\lambda W_\nu - \partial_\lambda W_\mu^* W_\nu\right) V_\rho + i\tilde{\kappa}_V W_\mu^* W_\nu \tilde{V}^{\mu\nu} + i\frac{\tilde{\lambda}_V}{M_W^2} W_{\lambda\mu}^* W^\mu{}_\nu \tilde{V}^{\nu\lambda}$$

EM gauge inv. (g_1^γ = 1), C and P conserving → κ_V, λ_V, g_1^Z couplings (V = γ, Z)

C = Charge conjugation
P = Parity

SM: $g_1^Z = \kappa_V = 1, \;\; \lambda_V = 0$

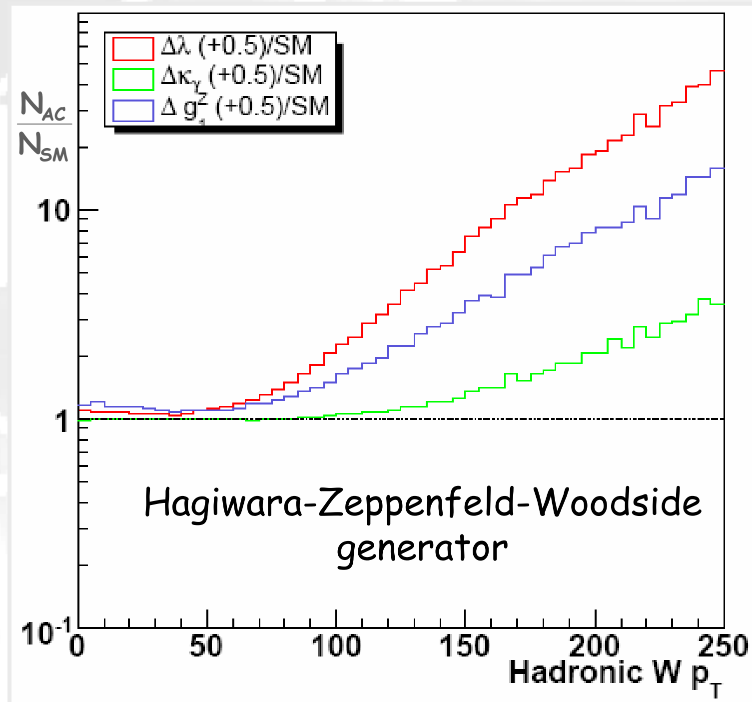
SM Deviations: $\Delta g_1^Z = g_1^Z - 1, \;\; \Delta\kappa_V = \kappa_V - 1$

$\Delta\lambda_V = \lambda_V - 0$

$\Delta \neq 0$
ANOMALOUS COUPLINGS

[J. Ellison, J. Wudka, Annu. Rev. Nucl. Part. Sci. 1998, 48:33-80]

- **Look for deviations from the SM: Cross sections, Kinematic distributions**



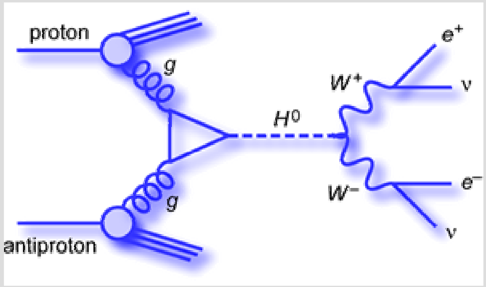
Motivation for Diboson Physics

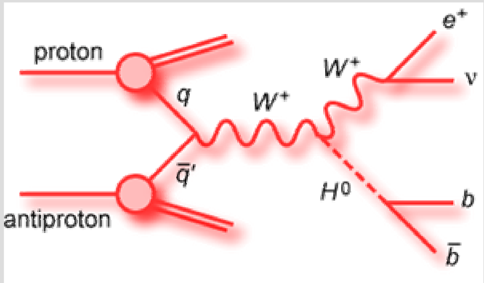
- **Higgs Physics**

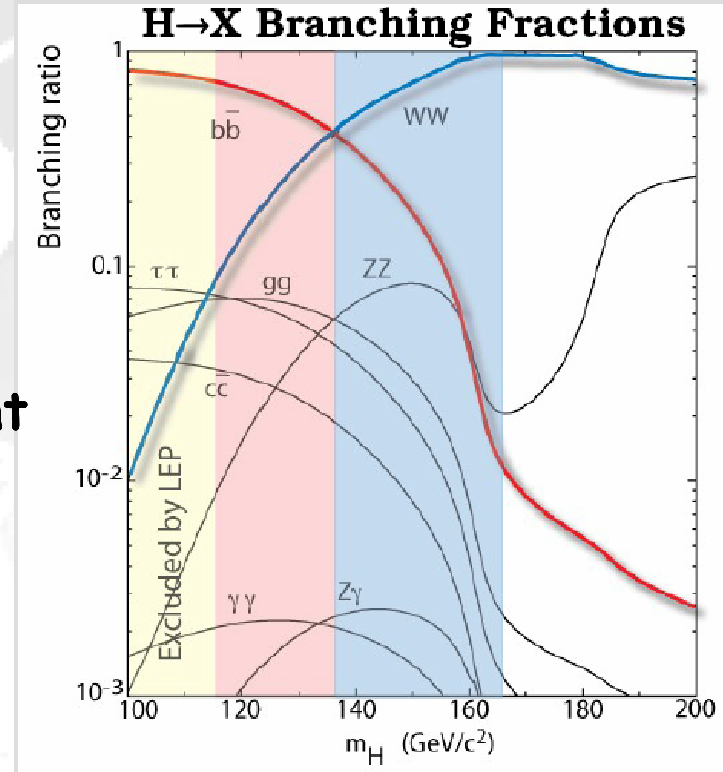
  Heavy Higgs (M_H>135 GeV):
  H → WW dominant decay mode
  Direct WW production is a significant background ⇒ essential to understand it!

  Light Higgs (M_H<135 GeV):
  WH → lvbb promising search channel
  Complementary to diboson final states: similar final state to WW/WZ → lvjj

  Similar challenges:
  Small signal in a large background!
  S/B ⇒ WH: 1.2% WW+WZ: 2.9%
  Analysis techniques, Statistical treatment







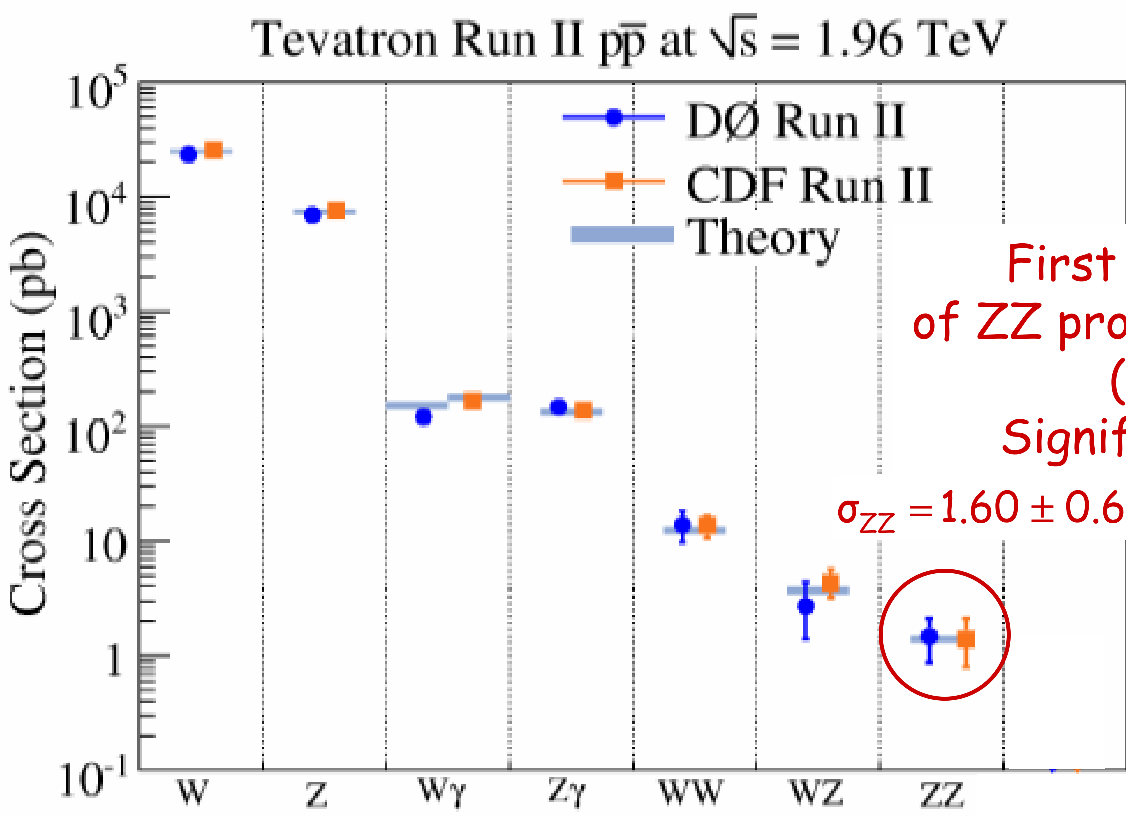
➡ WW+WZ → lvjj represents a valuable proving ground for analysis techniques used in the Tevatron Higgs search

Gauge Boson Production at the Tevatron

Vector boson factory (*L* ≈ 50/pb recorded per week)

~1.2·10^6 Ws, ~3.5·10^5 Zs, ~620 WW, ~190 WZ, ~70 ZZ, ~800 Wγ

Dibosons: Until recently, only fully leptonic final states were analyzed; small branching ratio but clean signal (low background)



First observation of ZZ production (by DØ) (llll+llνν)
Significance 5.7σ

$\sigma_{ZZ} = 1.60 \pm 0.63(\text{stat})^{+0.16}_{-0.17}(\text{syst})\,\text{pb}$

$\sigma_{ZZ}^{\text{Theory}} = 1.4 \pm 0.1\,\text{pb}$

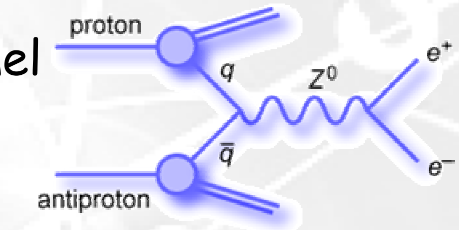
Total cross sections of Tevatron Preliminary and Published EW results

Single Boson Production (W,Z properties)

W/Z charge asymmetry

Constraints on PDFs

Most precise measurement to date by DØ in *e*-channel



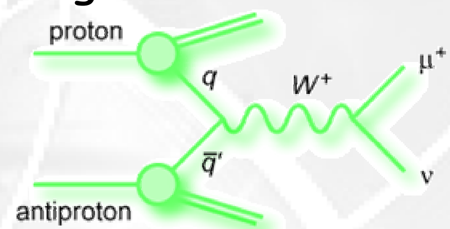
W mass

The most precise measurement by DØ (44 MeV)

The CDF+DØ combination soon:

Tevatron average uncertainty smaller than the LEP average

Forward-Backward (hadronic) Asymmetry



Measurement of $\sin^2\theta_W$

g_2 (Z/γ^* p_T) measurement

Test of the QCD predictions

Measurement of nonpertubative parameter in the Z p_T space defined in the Soft-Gluon Resummation function (BLNY formalism – Brock, Landry, Nadolsky, Yuan)

Cross Section Measurements

Good agreement with the SM prediction and other measurements

DiBoson Production (WW,WZ) ⇒

WW Production at DØ

Last published: 250/pb RunII data

- Cross section measurement
- Limits on anomalous couplings

WW→lνlν channel, l = electron/muon

Signature:

2 high p_T charged leptons

Missing Transverse Energy (MET)

Channel	Signal	Background	Candidates
e^+e^-	3.26 ± 0.05	2.30 ± 0.21	6
$e^\pm\mu^\mp$	10.8 ± 0.1	3.81 ± 0.17	15
$\mu^+\mu^-$	2.01 ± 0.05	1.94 ± 0.41	4

- **Observed significance: 5.2σ**

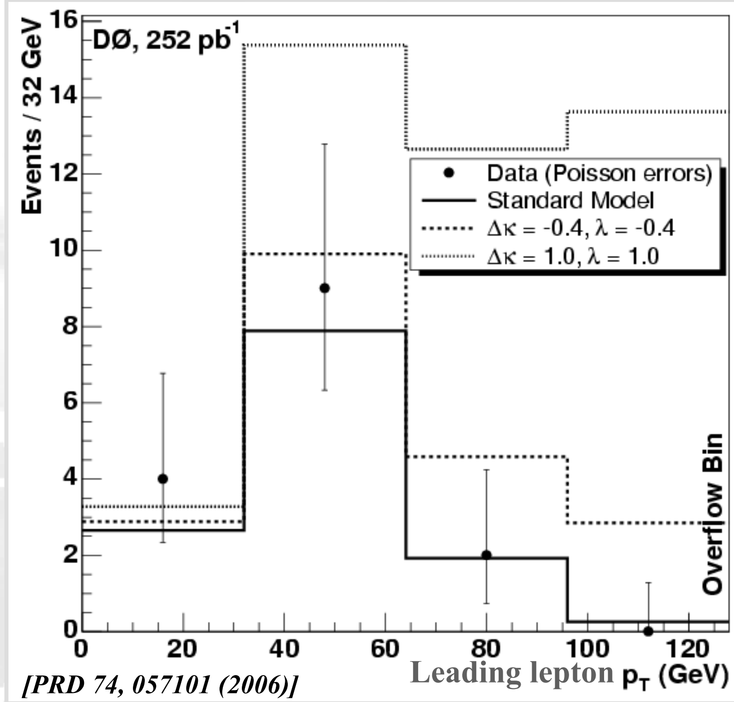
In agreement with the SM NLO:

$$\sigma_{p\bar{p}\to WW} = 13.8^{+4.3}_{-3.8}(\text{stat})^{+1.2}_{-0.9}(\text{syst}) \pm 0.9(\text{lumi})\ \text{pb}$$

$$\sigma_{WW}^{\text{Theory}} = 12.4 \pm 0.8\ \text{pb}$$

[PRL 94, 151801 (2005)]

Preliminary 1/fb: most precise measurement



[PRD 74, 057101 (2006)]

Coupling	1 – dimensional 95% C.L. Limits		Λ (TeV)
$WW\gamma = WWZ$	λ	$-0.31, 0.33$	1.5
	$\Delta\kappa$	$-0.36, 0.47$	
$WW\gamma = WWZ$	λ	$-0.29, 0.30$	2.0
	$\Delta\kappa$	$-0.32, 0.45$	
HISZ	λ	$-0.34, 0.35$	1.5
	$\Delta\kappa_\gamma$	$-0.57, 0.75$	
SM $WW\gamma$	λ_Z	$-0.39, 0.39$	2.0
	$\Delta\kappa_Z$	$-0.45, 0.55$	
SM WWZ	λ_γ	$-0.97, 1.04$	1.0
	$\Delta\kappa_\gamma$	$-1.05, 1.29$	

WW Production at CDF II

Last published: 184/pb RunII data

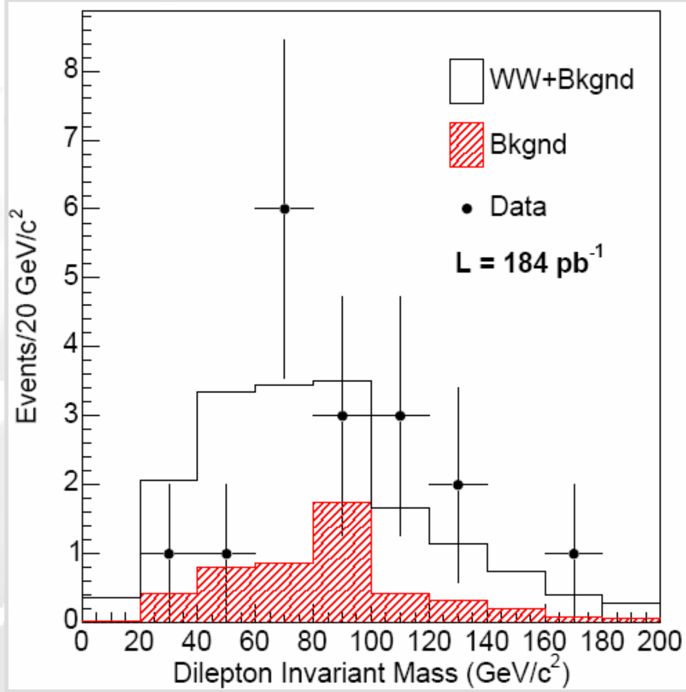
- Cross section measurement

WW→lνlν channel, l = electron/muon

	ee	$\mu\mu$	$e\mu$
Background	$1.9^{+1.3}_{-0.3}$	$1.3^{+1.6}_{-0.4}$	1.9 ± 0.4
W^+W^- Signal	2.6 ± 0.3	2.5 ± 0.3	5.1 ± 0.6
Expected	$4.5^{+1.4}_{-0.5}$	$3.8^{+1.6}_{-0.5}$	7.0 ± 0.8
Observed	6	6	5

[PRL 94, 211801 (2005)]

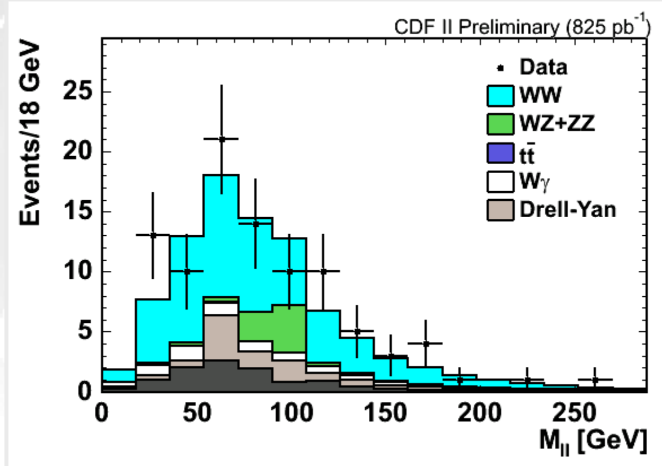
$$\sigma_{p\bar{p}\to WW} = 14.6^{+5.8}_{-5.1}(stat)^{+1.8}_{-3.0}(syst) \pm 0.9(lumi)\ pb$$



Preliminary: 825/pb RunII data

[FERMILAB-CONF-06-115-E]

Mode	ll
WW	$52.4 \pm 0.1 \pm 4.3$
Drell-Yan	$11.8 \pm 0.8 \pm 3.1$
W+jets	$11.0 \pm 0.5 \pm 3.2$
$WZ+ZZ$	$7.9 \pm 0.0 \pm 0.8$
$W\gamma$	$6.8 \pm 0.2 \pm 1.4$
$t\bar{t}$	$0.2 \pm 0.0 \pm 0.0$
Sum Bkg	$37.8 \pm 0.9 \pm 4.7$
Expected	$90.2 \pm 0.9 \pm 6.4$
Data	95



$$\sigma_{p\bar{p}\to WW} = 13.6 \pm 2.3(stat) \pm 1.6(syst) \pm 1.2(lumi)\ pb$$

WZ Production at DØ

Last published: 1/fb RunII data

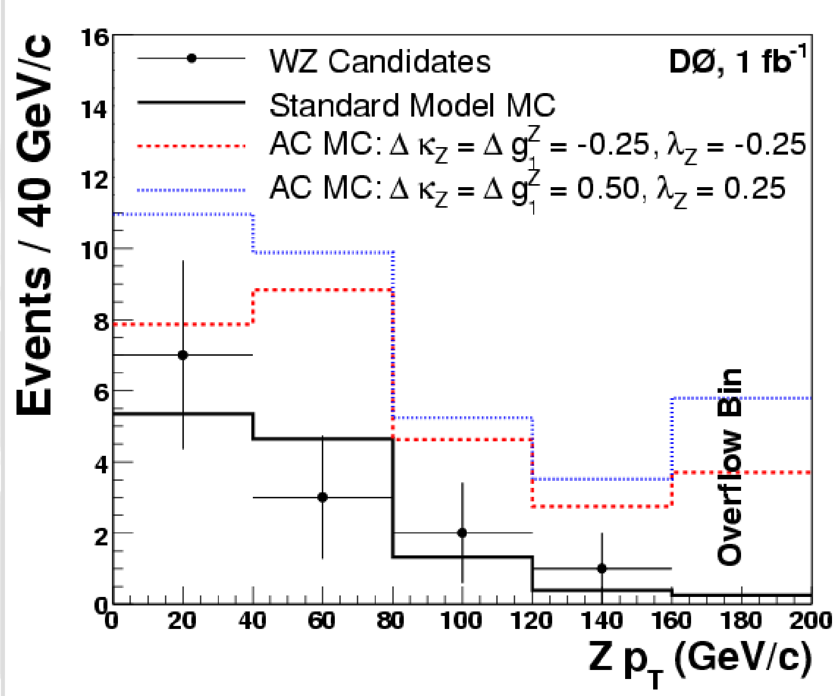
- Cross section measurement
- Limits on anomalous couplings

WZ→lνll channel, l = electron/muon

Signature:

3 high p_T charged leptons + MET

Final State	Number of Candidate Events	Expected Signal Events	Estimated Background Events	Overall Efficiency
eee	2	2.3 ± 0.2	1.2 ± 0.1	0.16 ± 0.02
$ee\mu$	1	2.2 ± 0.2	0.46 ± 0.03	0.17 ± 0.02
$\mu\mu e$	8	2.2 ± 0.3	2.0 ± 0.4	0.17 ± 0.03
$\mu\mu\mu$	2	2.5 ± 0.4	0.86 ± 0.06	0.21 ± 0.03
Total	13	9.2 ± 1.0	4.5 ± 0.6	–



1-dimensional 95% CL limits on couplings:

$\Lambda = 1.5$ TeV	$\Lambda = 2.0$ TeV
$-0.18 < \lambda_Z < 0.22$	$-0.17 < \lambda_Z < 0.21$
$-0.15 < \Delta g_1^Z < 0.35$	$-0.14 < \Delta g_1^Z < 0.34$
$-0.14 < \Delta\kappa_Z = \Delta g_1^Z < 0.31$	$-0.12 < \Delta\kappa_Z = \Delta g_1^Z < 0.29$

Observed significance: 3.0σ

In agreement with the SM NLO:

$\sigma_{p\bar{p}\to WZ} = 2.7^{+1.7}_{-1.3}(\text{stat}+\text{syst})\ \text{pb}$

$\sigma_{WZ}^{\text{Theory}} = 3.7 \pm 0.3\ \text{pb}$ *[PRD 76, 111104(R) (2007)]*

First evidence for WZ→lνll at a hadron collider

WZ Production at CDF

Last published: 1/fb RunII data

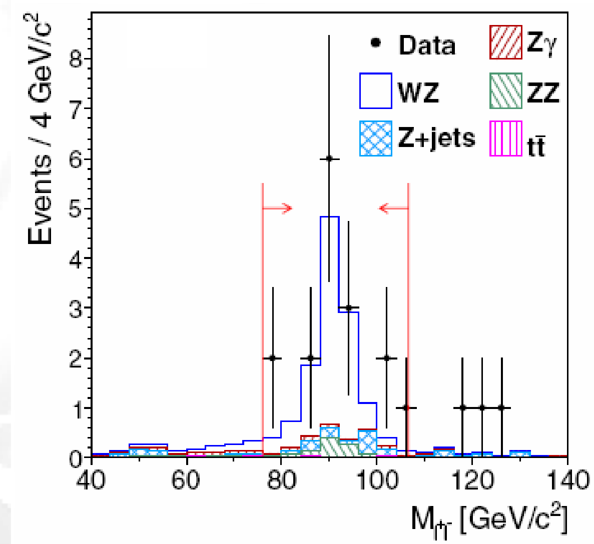
- Cross section measurement

WZ→lνll channel, l = e/μ

Z mass window (76-106) GeV

	Expectation ± stat ± syst ± Lumi
Total background	2.65 ± 0.28 ± 0.33 ± 0.09
WZ	9.75 ± 0.03 ± 0.31 ± 0.59
Total expected	12.41 ± 0.28 ± 0.45 ± 0.67
Observed	16

First observation of WZ→lνll (6σ)



$$\sigma_{p\bar{p}\to WZ} = 5.0^{+1.8}_{-1.4}(\text{stat}) \pm 0.4(\text{syst})\ \text{pb}$$

[PRL 98, 161801 (2007)]

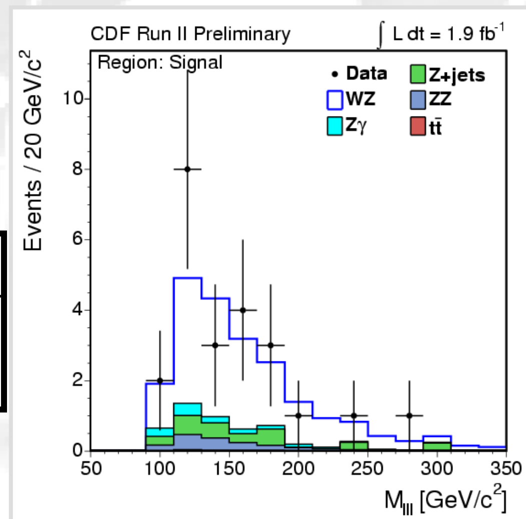
Preliminary: 1.9/fb RunII data

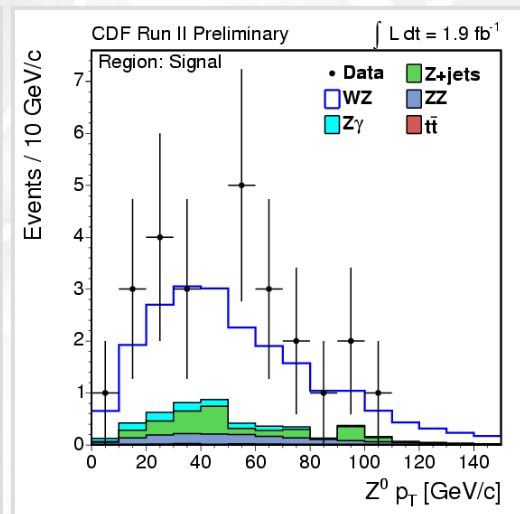
- Limits on anomalous couplings

$$\sigma_{p\bar{p}\to WZ} = 4.3^{+1.3}_{-1.0}(\text{stat}) \pm 0.4(\text{syst}) \pm 0.3(\text{lumi})\ \text{pb}$$

	Expected ± Stat ± Syst ± Lumi
Predicted	21.63 ± 0.48 ± 2.25 ± 1.15
Observed	25

1-dimensional 95% CL limits (Λ = 2 TeV)		
$-0.13<\lambda_Z<0.14$ $\Delta g_1^Z = \Delta\kappa_Z = 0$	$-0.13<\Delta g_1^Z<0.23$ $\lambda_Z = \Delta\kappa_Z = 0$	$-0.76<\Delta\kappa_Z<1.18$ $\Delta g_1^Z = \lambda_Z = 0$





Studies related to WW+WZ → lνjj at DØ

DØ Electroweak + Higgs Group Efforts

1.1/fb RunII data

- Cross section measurement (soon: Limits on anomalous couplings)

Signature: Lepton (electron or muon) + MET + 2 jets

WW/WZ → lνjj Production

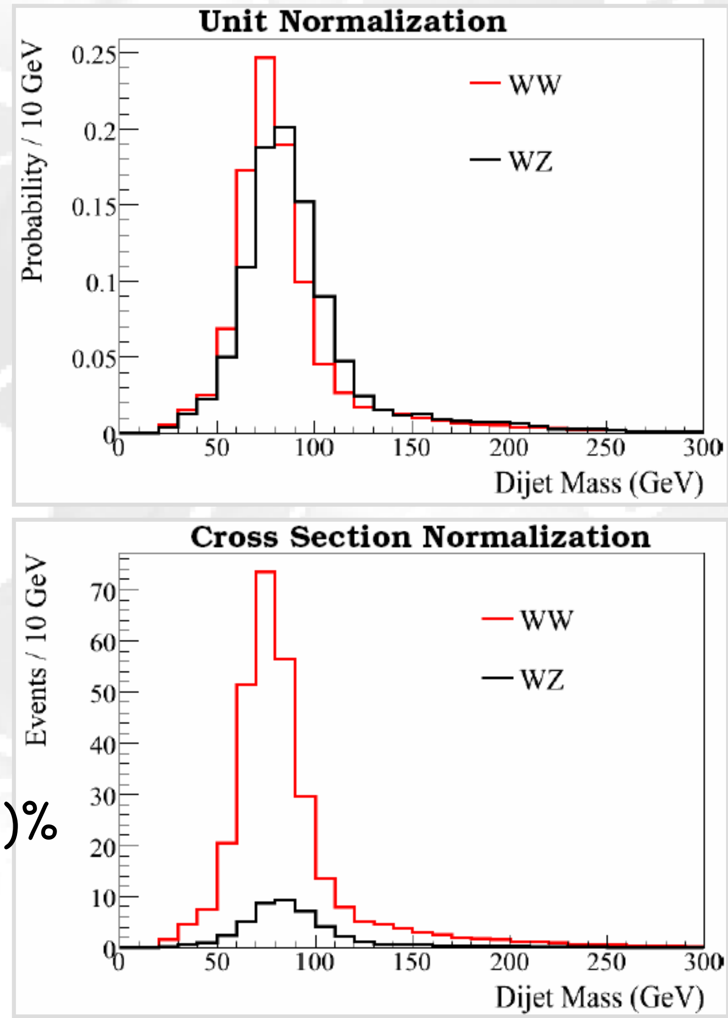
Few words on WW versus WZ:

- **WW and WZ in semileptonic final state are indistinguishable signals**

Insufficient dijet mass resolution:
- ~10 GeV difference in W/Z dijet mass peaks
- Cascade decays of heavy quarks in Z→bb contain neutrinos ⇒ reduced reconstructed dijet mass in these events

⇒ Final mass difference: ~7 GeV

Detector dijet mass resolution ~18% for dijets from W/Z decays (~15 GeV)



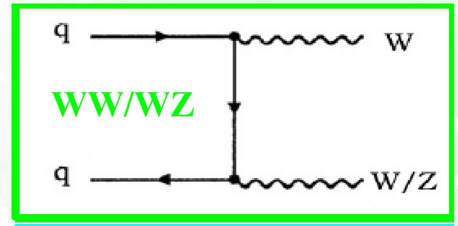
- **Branching Fractions: (larger than leptonic modes)**

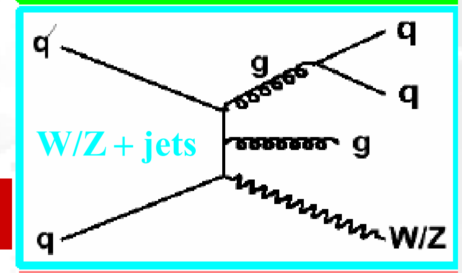
WW(WZ) → (e+μ)νjj branching ratio: ~28.5 (14.2)%
WW(WZ) → (e+μ)νjj $\sigma_{theo} \times BR$: ~3.5 (0.5) pb

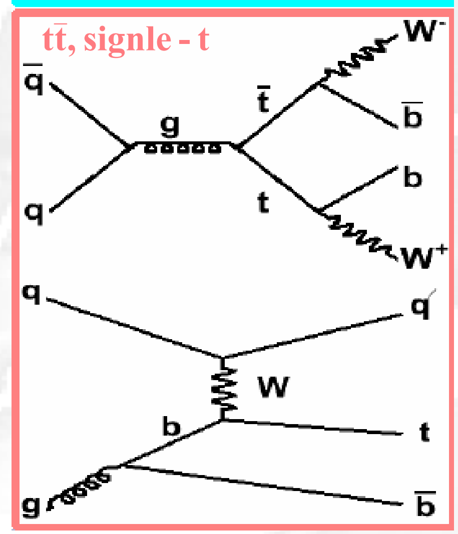
Simulated Samples

Most of event sources are generated via Monte Carlo with a full simulation of detector response

Event Source	Generator		σ(SM) / σ(WW) = 12.4 pb	
WW	Pythia	Inclusive	1.0	**(12.4 ± 0.8 pb)**
WZ	Pythia		0.3	**(3.7 ± 0.3 pb)**
ZZ	Pythia		0.1	**(1.4 ± 0.1 pb)**
W+light flavor jets	Alpgen	$W \to l\nu$	800	Challenging !
W+heavy flavor jets	Alpgen		30	
Z+light flavor jets	Alpgen	$Z \to ll$	30	
Z+heavy flavor jets	Alpgen		1	
Double-Top	Alpgen		0.6	
Single-Top	Comphep		0.2	







The rate and distributions of **Multijet** events, in which jets are misidentified as leptons, are determined from data

Huge W+jets Background with similar kinematics as WW/WZ

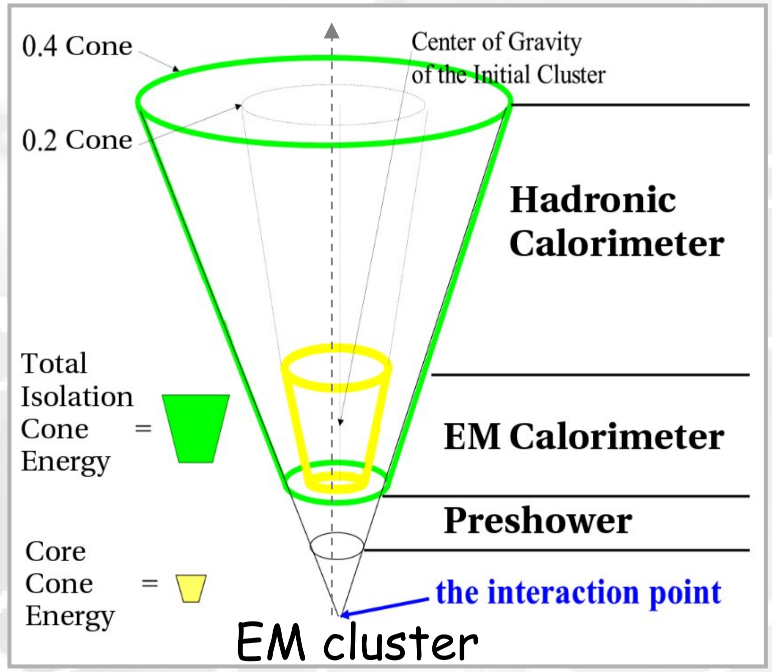
Lepton Selection

Selecting Lepton Candidates

Make sure you select what you want!

Lepton Identification
(Lepton) Track Quality
Lepton Isolation

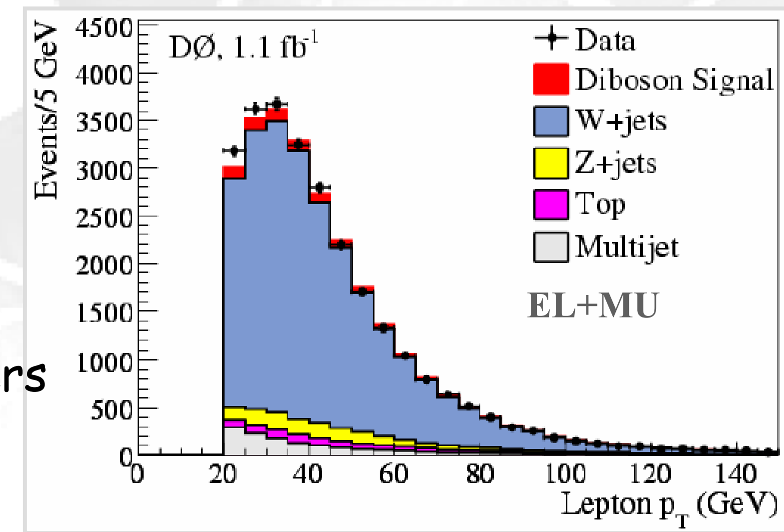
p_T>20 GeV, $|\eta|_{EL}$<1.1, $|\eta|_{MU}$<2.0
Spatial match to a central track
Veto events with multiple leptons



Electrons:

Calorimeter energy cluster in radial cone of R<0 4 $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2}$

- Calorimeter showers consistent with EL shape
- Require that 90% of energy is deposited in the EM calorimeter



Muons:

- Must have hits in at least 3 muon detector layers
- Must be isolated in both the tracker and the calorimeter

Jet Selection (W/Z jets)

Jet == Cluster Energy in cones of R<0.5

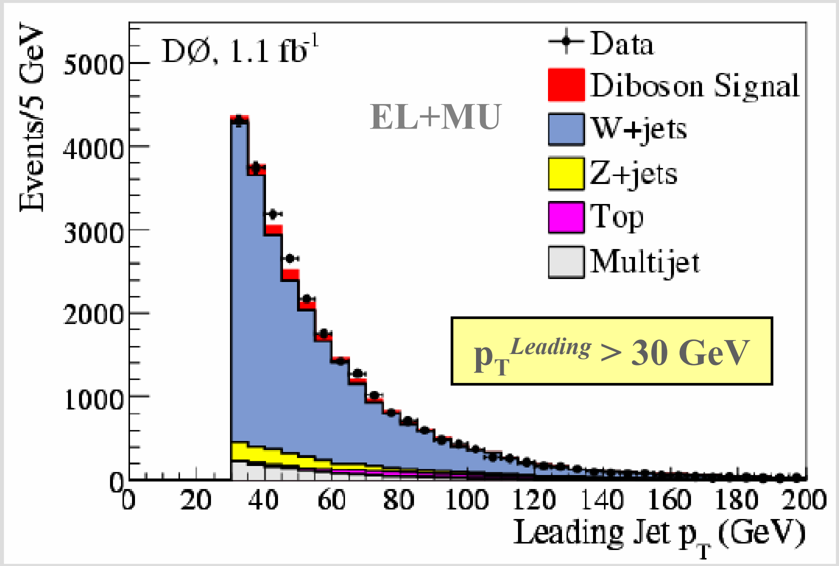
Jets from W/Z decays are highly energetic and relatively central ($|\eta|<1.1$)

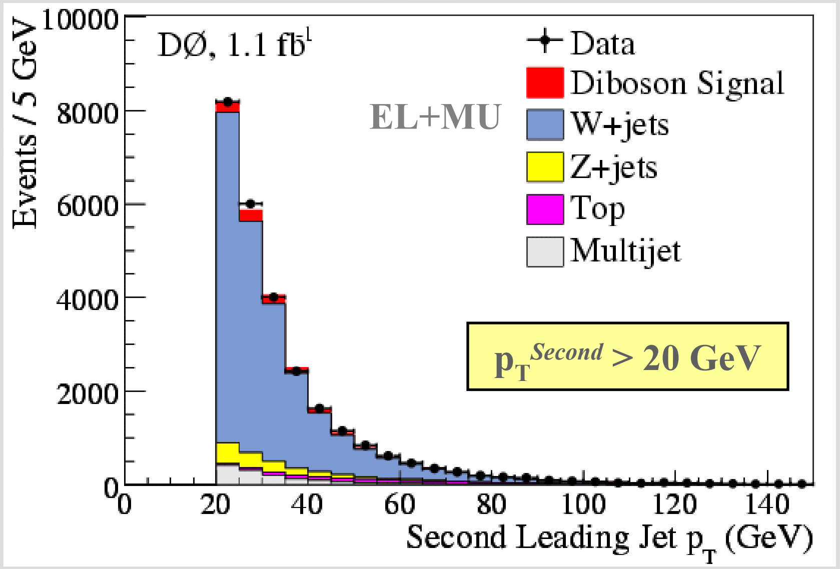
Good Jets:

Shower shape inconsistent with that of EM objects
Requirement on the EM and CH fractional energy

Minimum 2 jets in event
Leading Jet p_T>30 GeV
Second Jet p_T>20 GeV
$|\eta|_{JET}$<2.5

Jet p_T distributions well modeled





W → lν Selection

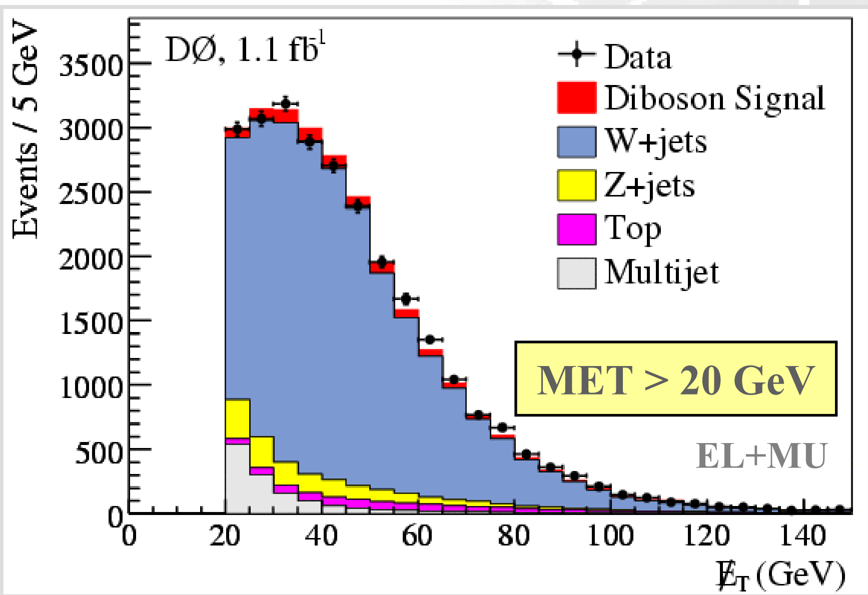
- **Neutrino manifests as an imbalance in transverse momentum (energy)**
  Events consistent with a W→lν decay have relatively high missing transverse energy (MET) due to the existing neutrino
- **Total energy in the transverse plane is conserved**
  MET = Vectorial sum of "visible" energy deposits in the calorimeter cells ($E_{x,y}$ in the *x-y* plane, plane perpendicular to the beam)

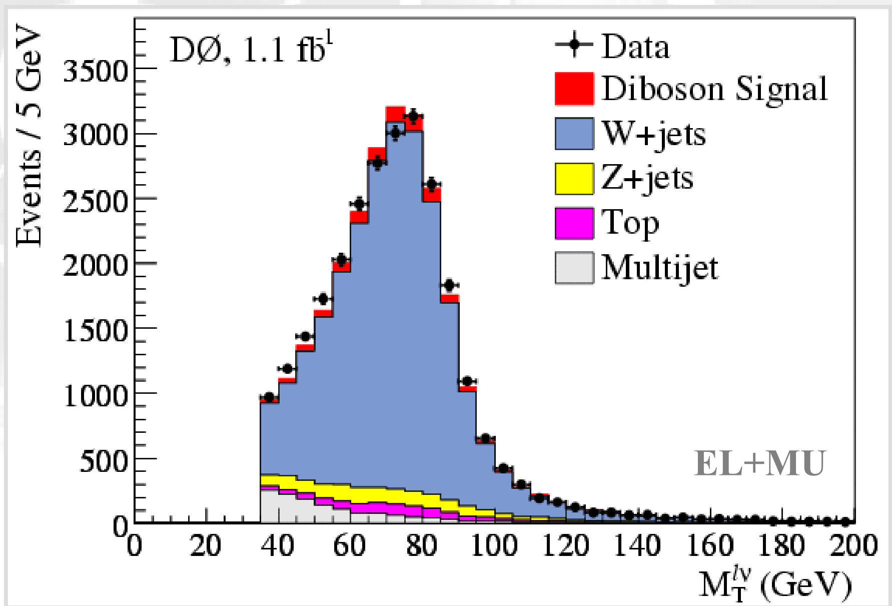
$$MET_{x,y} = -\sum_{i(E>0)} E^i_{x,y} \quad MET = \sqrt{MET_x^2 + MET_y^2}$$

$$M_T^W = \sqrt{2p_T^l \cdot MET \cdot [1 - \cos(\varphi_l - \varphi_{MET})]}$$

Corrected for: CHJets, Muons,
Jet Energy Scale, EM Scale

Reduce Multijet backgrounds: M_T^W >35GeV





Standard Monte Carlo Corrections

- **The event selection includes Efficiencies and Kinematic Corrections for known Data/Monte Carlo differences**

Z p_T :
The transverse momentum of Z bosons in Z+jets events is corrected to measurements in data.

Lepton and Jet Identification:
Percent level corrections. Often arise from changes in real detector efficiency during running period.

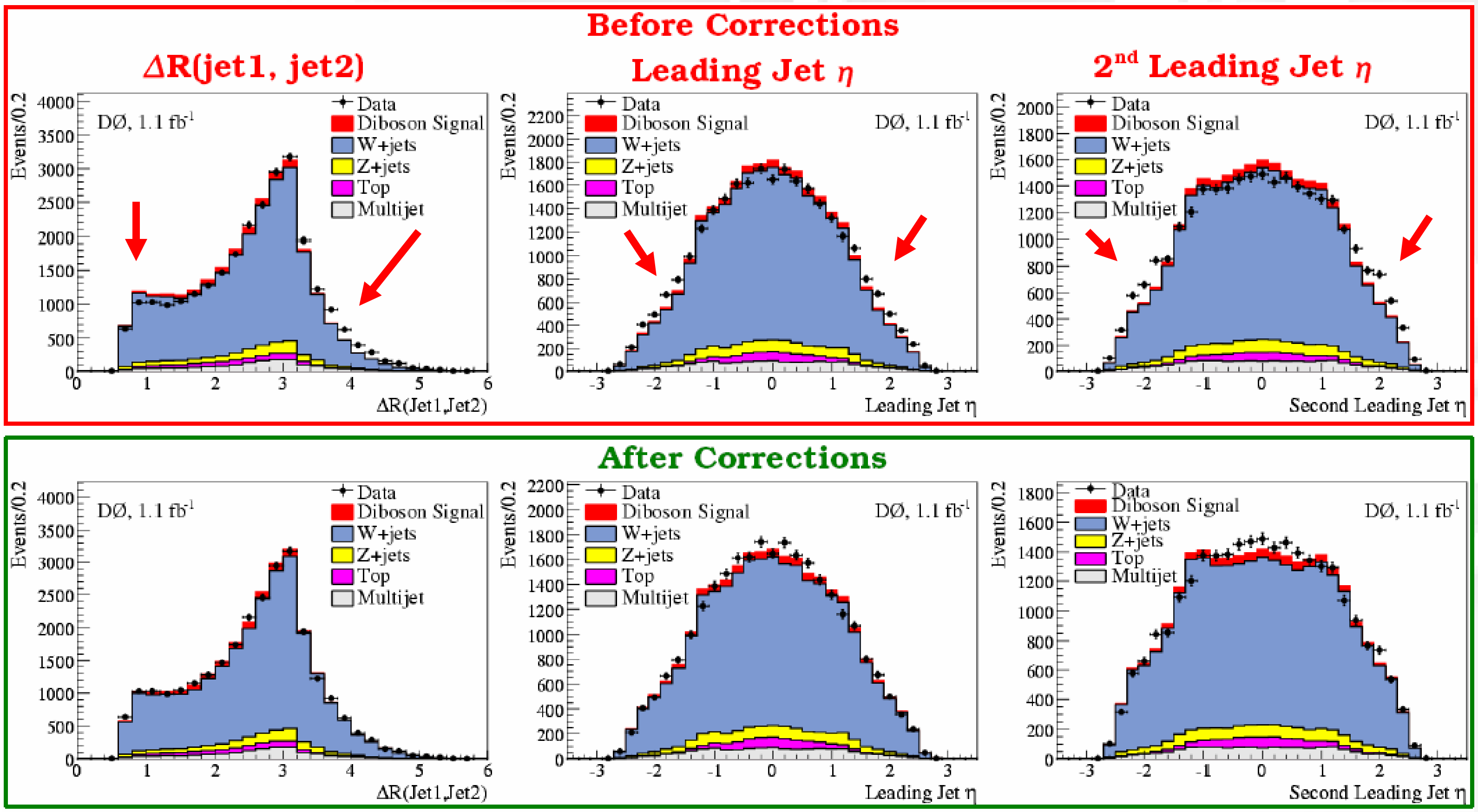
Trigger selection:
Trigger efficiencies are measured in data and propagated to simulated samples.

Luminosity profile:
The instantaneous luminosity profile of the simulation is matched to data. Helps to properly model minimum bias effects.

Beam z-position profile:
The longitudinal profile of the beam interaction region is matched to data. Impacts angular and energy calculations.

Angular Corrections

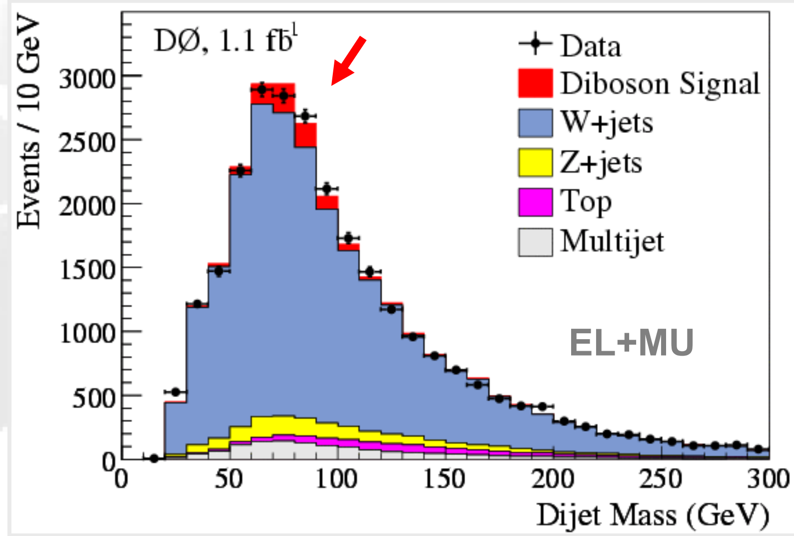
- Differences between Data and MC in jet angular distributions observed – their magnitude tells: must come from the dominant W/Z+jets background (studies ⇒ due to the relative angles of low p_T jets)
- Similar modeling effects in jet angular distributions of other generators (arXiv/hep-ph:0706.2569)
- We correct at the event level using correction functions derived from the relative Data/MC (W/Z+jets) shapes ⇒ better χ^2/ndof



Event Yields

- **Following selection and corrections ⇒ Expected and Observed number of events**

Luminosity	$e\nu jj$ channel 1067 pb^{-1}	$\mu\nu jj$ channel 1074 pb^{-1}
WV	357.5 ± 2.3	415.8 ± 2.7
W+light flavor jets	8158 ± 72	9681 ± 84
W+heavy flavor jets	2060 ± 26	2319 ± 28
Z+jets	406 ± 13	1237 ± 20
$t\bar{t}$ + single top	463.3 ± 2.2	438.0 ± 2.2
Multijet	825 ± 11	327.0 ± 9.6
ZZ	2.99 ± 0.14	11.53 ± 0.28
Total predicted	12272 ± 78	14428 ± 91
Data	12473	14392



- W+jets cross-section (i.e. k-factor): **Will be determined from data**
  Initial (scale to match $N_{\text{Data-Non-W+jets}}$): **k=1.53**

- W+jets is the dominant background (~85%) ⇒ **Signal/Background ~3%**

- **Systematic effects of order of few % on the background are important ! ⇒ needs good Alpgen modeling**

Alpgen Modeling

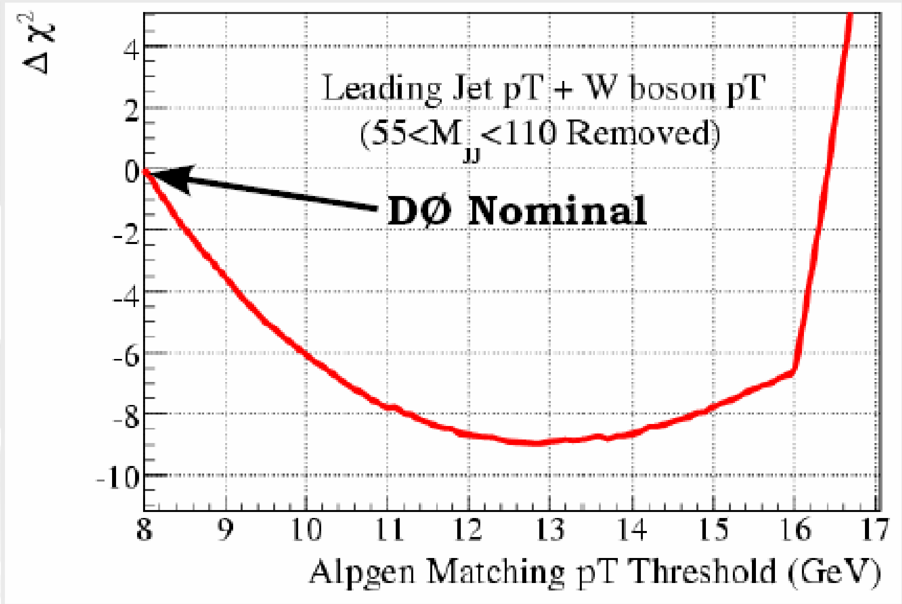
- Renormalization scale; k_T scale factor; Parton-jet matching cluster p_T threshold; Parton-jet matching clustering radius size;

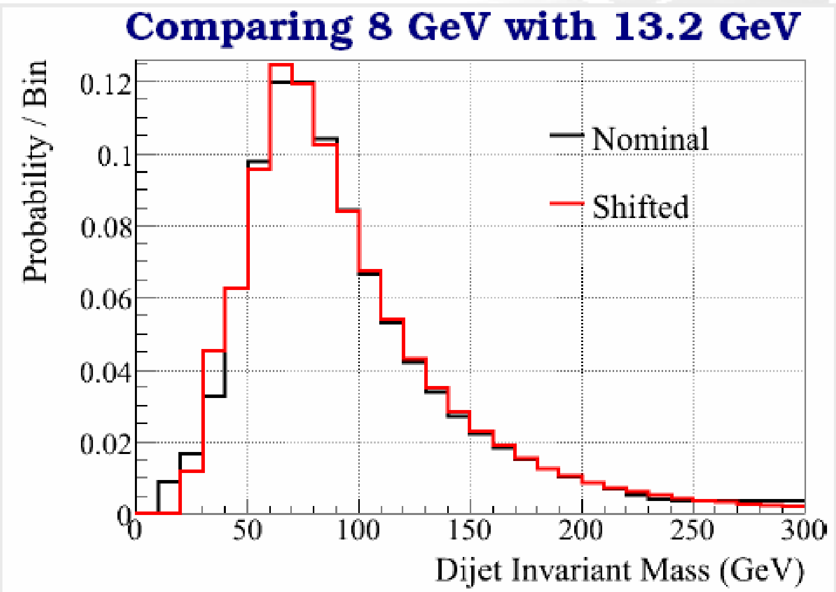
Prescription: Map χ^2 as a function of the change in each parameter we test (w/o the signal region in dijet mass)

$\Rightarrow \Delta\chi^2$ tests show no clear preference for altered parameters, aside from **Parton matching jet p_T threshold**

DØ default p_T threshold is 8 GeV. Alpgen authors' suggestion: generator level p_T cut + max(20%, 5 GeV): 13 GeV for DØ.

Propagate correction via event weights





Better Data/MC agreement at low Dijet mass (<70 GeV)

Improved Selection

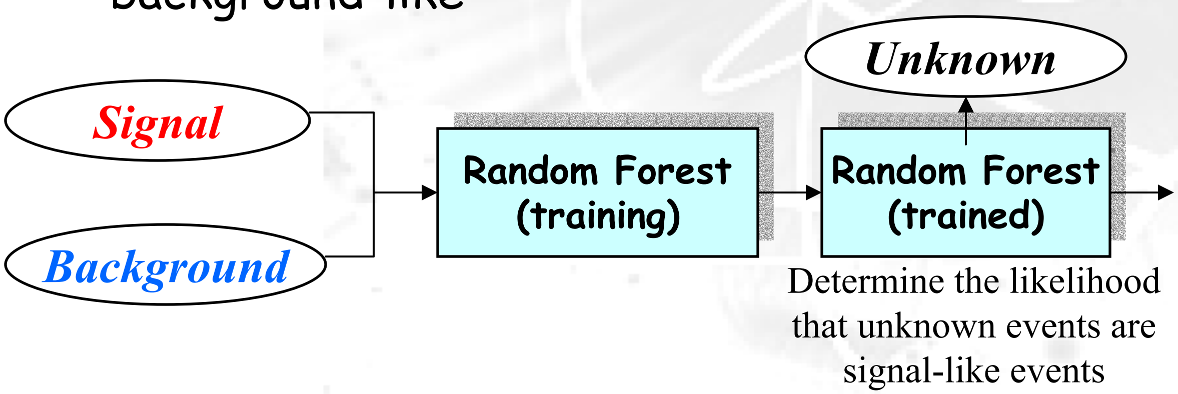
Small signal, huge background with similar kinematics …

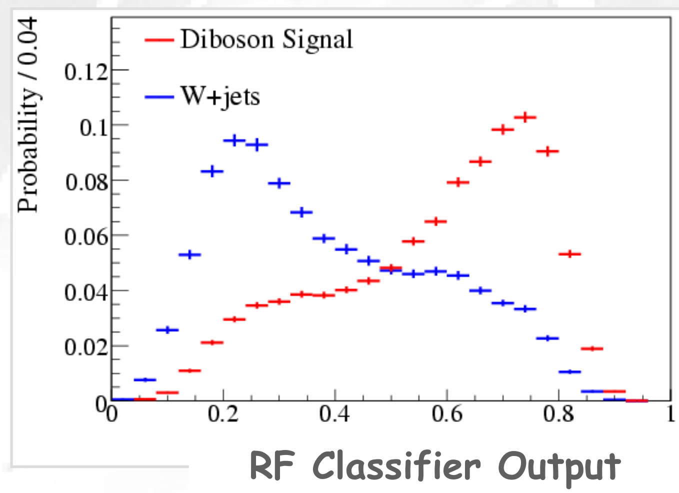
- **Multivariate (Discriminating) Techniques (Classifiers)**
  To increase the statistical power of the measurement (i.e. significance)
  (Options: Neural Networks, Matrix Elements, Likelihood ratios, Boosted Decision Trees, …)

- **Random Forest (RF) Classifier (a "Forest" of decision trees)**
  Found to be the most powerful and robust ("StatPatternRecognition" package, http://www.hep.caltech.edu/~narsky/spr.html (I. Narksy, Caltech))

Just like any other classifier …

- Trained by feeding it with events of known origin (Signal or Background)
- Trained RF classifies events of unknown origin (data) as signal-like or background-like





Random Forest

Each decision tree in the forest is independent in training and testing

- Each tree uses a random subset of the input variables ⇒ allows each tree to focus on a different subset of kinematics and correlations
- Each tree is trained using a random subset of training events ⇒ provides protection against over-training and high-weight events

The Random Forest classifier output is the average output over all trees

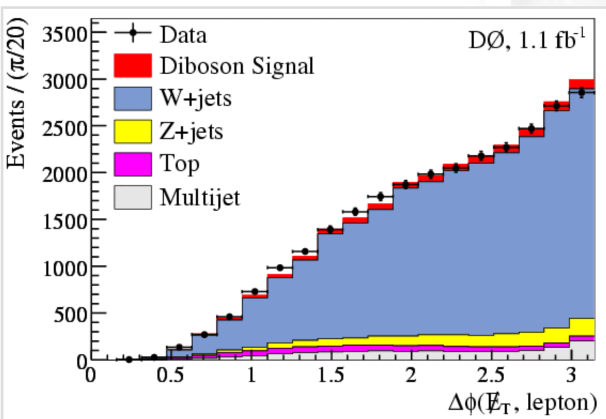
- Fluctuations and overtraining that occur for a single decision tree are reduced in the global averaging of fluctuations

RF Input Variables

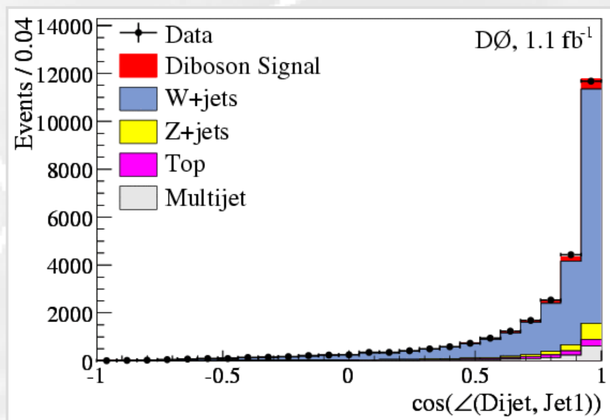
We use 13 kinematic variables as input to the Random Forest

- Each variable helps distinguish between signal and at least one background
- Ensure well modeled variables by requiring data/MC χ^2 probability outside the signal region ($55<M_{JJ}<110$ GeV) to be greater than 5%
- Variables not directly tuned to data

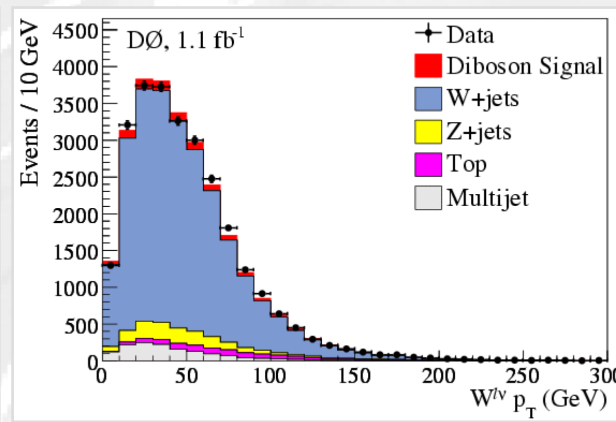
In addition to: M_{JJ}, M_T^W, p_T^{Jet2}, MET:



Δφ (l, MET)
Azimuthal separation of the charged lepton and MET



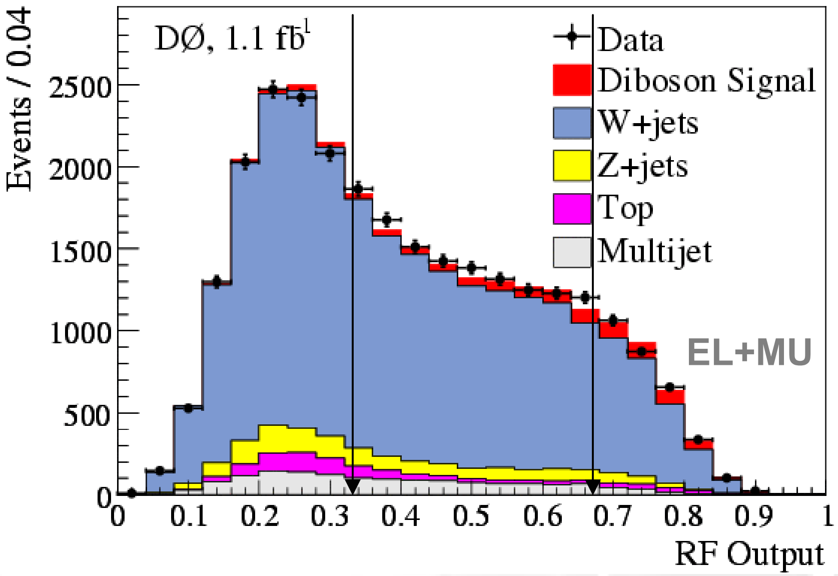
cos [∠(dijet, jet1)]
Cosine of the angle between dijet system and the leading jet in the lab frame



p_T ($W_{l\nu}$)
p_T of reconstructed W from charged lepton and MET

Random Forest Output

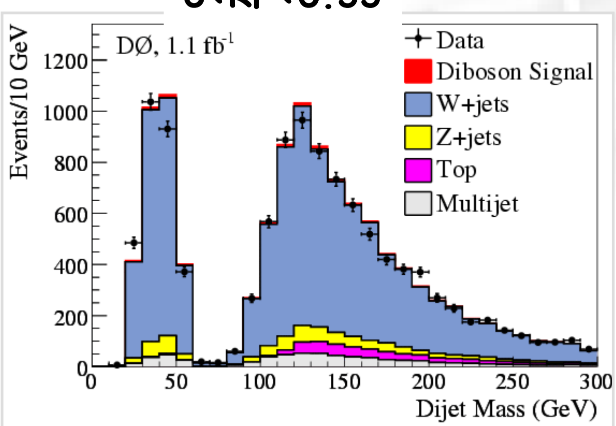
- The Random Forest output demonstrates improved separation of signal and backgrounds
- Maintains good agreement between MC and Data

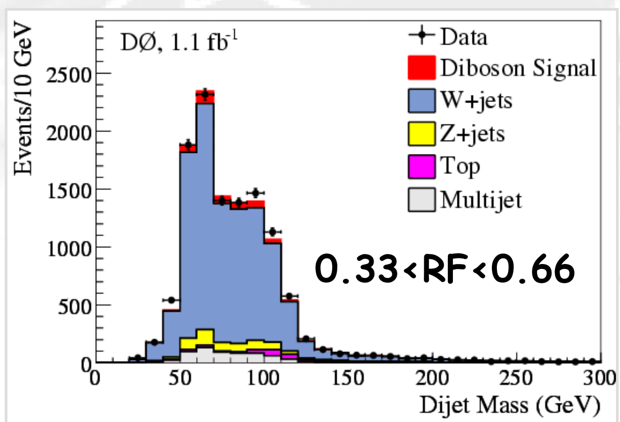


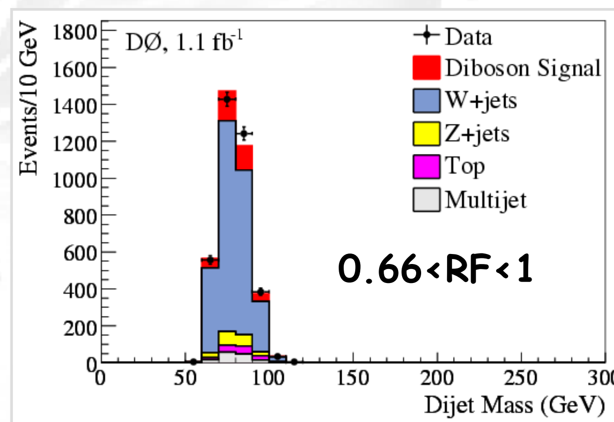
- The Dijet mass is one of the most important discriminating variables
- The RF is efficient at moving events from under the signal peak in Dijet mass (sliced RF output)

RF = 0-0.33: *Low Significance*
RF = 0.33-0.66: *Average Significance*
RF = 0.66-1.0: *High Significance*

0<RF<0.33







Systematic Uncertainties

- **The nature of systematics:**
  Normalization (Flat) or Differential (Shape) in RF distribution

For D: approximate maximal amplitude of the fluctuations in the RF output after $\pm1\sigma$ parameter changes

Source of systematic uncertainty	Diboson signal	W+jets	Z+jets	Top	Multijet	Nature
Trigger efficiency, $e\nu q\bar{q}$ channel	$+2/-3$	$+2/-3$	$+2/-3$	$+2/-3$		N
Trigger efficiency, $\mu\nu q\bar{q}$ channel	$+0/-5$	$+0/-5$	$+0/-5$	$+0/-5$		D
Lepton identification	±4	±4	±4	±4		N
Jet identification	±1	±1	±1	$\pm<1$		D
Jet energy scale	±4	±9	±9	±4		D
Jet energy resolution	±3	±4	±4	±4		N
Cross section		$\pm20^a$	±6	±10		N
Multijet normalization, $e\nu q\bar{q}$ channel					±20	N
Multijet normalization, $\mu\nu q\bar{q}$ channel					±30	N
Multijet shape, $e\nu q\bar{q}$ channel					±6	D
Multijet shape, $\mu\nu q\bar{q}$ channel					±10	D
Diboson signal NLO/LO shape	±10					D
Parton distribution function	±1	±1	±1	±1		D
ALPGEN η and ΔR corrections		±1	±1			D
Renormalization and factorization scale		±3	±3			D
ALPGEN parton-jet matching parameters		±4	±4			D

Uncertainties given in [%]

- Uncertainty on luminosity measurement 6.1%
- 100% correlated amongst signals and backgrounds
- Uncertainty sources uncorrelated among themselves

Cross Section Measurement

Input: Data, MC predictions, Statistical + Systematic Uncertainties

- **The "Best Fit" of the Signal and Backgrounds to Data using the RF output distribution:** minimizing Poisson (modified) χ^2 (ratio of Poisson Likelihoods + systematics):

Fit with respect to variations in the systematic uncertainties

$$\chi^2 \cong -2\ln\left[\frac{L^P}{L_0^P}\times\frac{L^G}{L_0^G}\right] = 2\sum_{i=0}^{N_{bins}}\left(B_i' + S_i' - d_i\right) - d_i \cdot \ln\left[\frac{B_i' + S_i'}{d_i}\right] + \sum_{k=0}^{N_{syst}} R_k^2$$

Gaussian constraint on systematics

[W. Fisher, FERMILAB-TM-2386-E]

$$B_i'(S_i') \to B_i(S_i)\prod_{k=0}^{N_{syst}}\left(1+\sigma_i^k R_k\right)$$

$B_i'(R_k), S_i'(R_k)$ - predicted number of events per bin "transformed" by the systematics

R_k - deviation from the central value of syst. uncertainty in units of σ (s.d.)

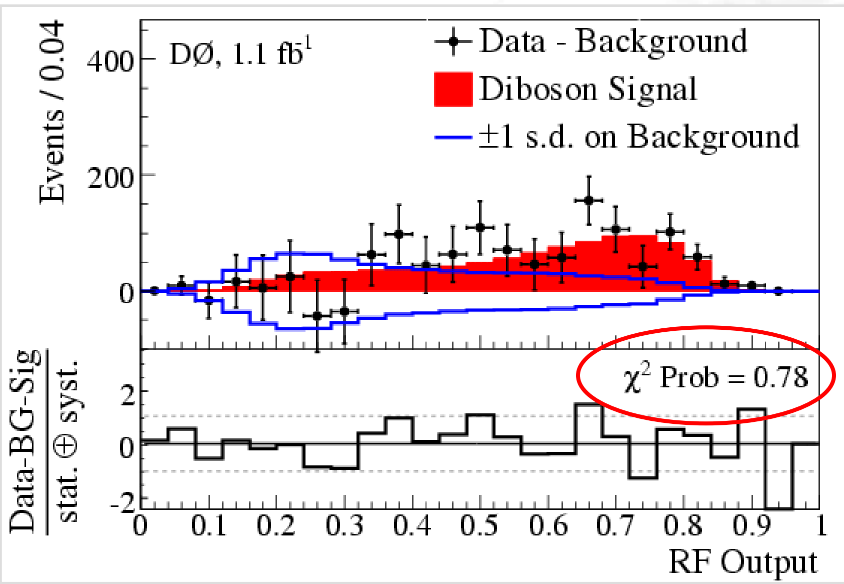
<u>Systematic uncertainties:</u>
Gaussian distributed uncertainties on the expected number of MC events

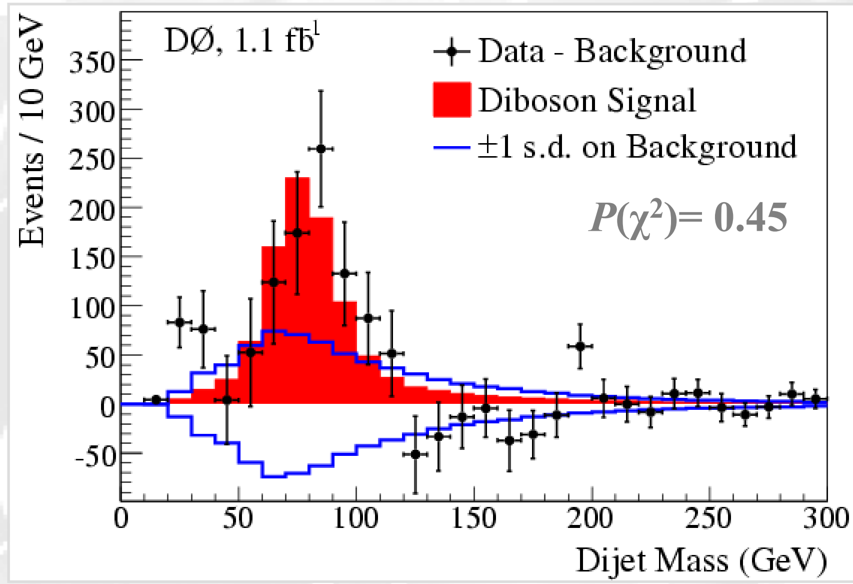
- Signal and Background distributions fluctuate within their uncertainties (with Gaussian constraint)
- Signal and W+jets normalizations (cross-section) are free parameters (Gaussian constraint removed from the sum)

Cross Section Measurement I

- Fit of the **Random Forest** outputs
- Theory prediction (NLO): 16.1±0.9 pb
- k-factor (from fit): 1.53±0.13

Observed number of signal events (WW+WZ: 963±56)





Channel	Fitted signal σ (pb)
$e\nu q\bar{q}$ RF Output	18.0±3.7(stat)±5.2(sys)±1.1(lum)
$\mu\nu q\bar{q}$ RF Output	22.8±3.3(stat)±4.9(sys)±1.4(lum)

Combined $\sigma_{WW/WZ}$:

20.2 ± 2.5(stat) ± 3.6(syst) ± 1.2(lumi) pb

Significance I

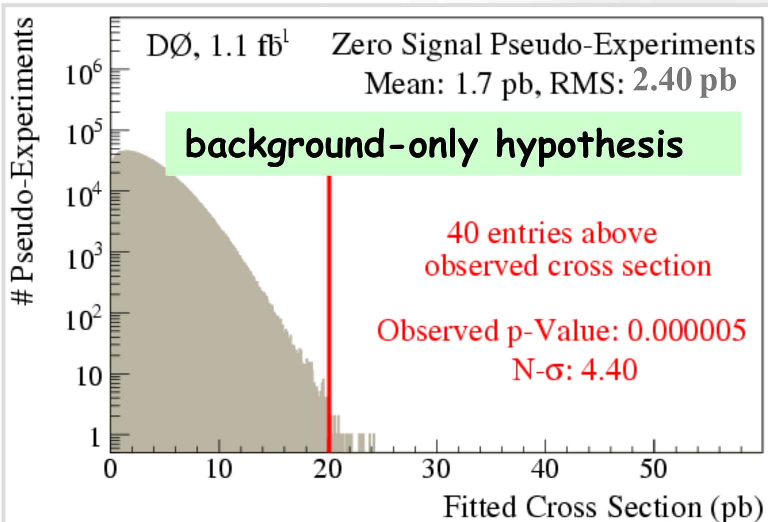
- **Significance is obtained by fitting MC templates (S+B) to pseudo-data (same procedure as for the cross section measurement)**

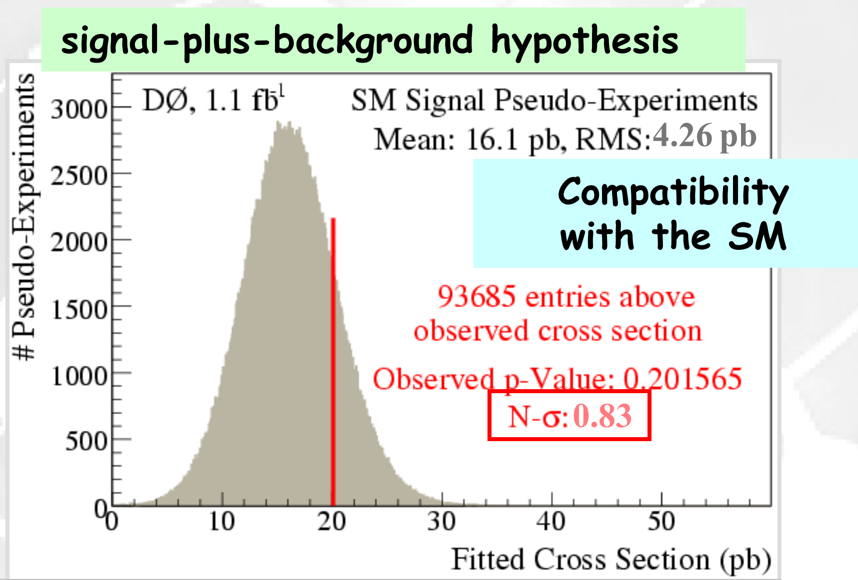
  pseudo-data == background-only hypothesis (zero-signal)

  - Randomly sample systematics (Gaussian) from their assumptions
  - Drawn Poisson trials for each bin

- **Count the number of outcomes above the Expected (Observed) cross section**

  → Expected (Observed) significance





Channel	Expected p-value (significance)	Observed p-value (significance)
$e\nu q\bar{q}$ RF Output	6.8×10^{-3} (2.5 s.d.)	3.2×10^{-3} (2.7 s.d.)
$\mu\nu q\bar{q}$ RF Output	1.8×10^{-3} (2.9 s.d.)	5.2×10^{-5} (3.9 s.d.)
Combined RF Output	1.5×10^{-4} (3.6 s.d.)	5.4×10^{-6} (4.4 s.d.)

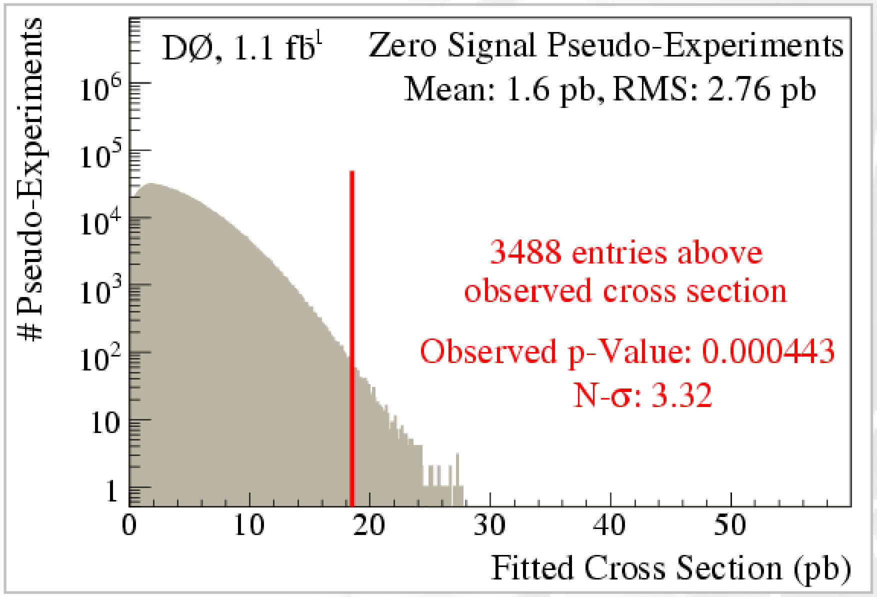
Larger acceptance and slightly smaller systematics in muon channel lead to higher expected significance.

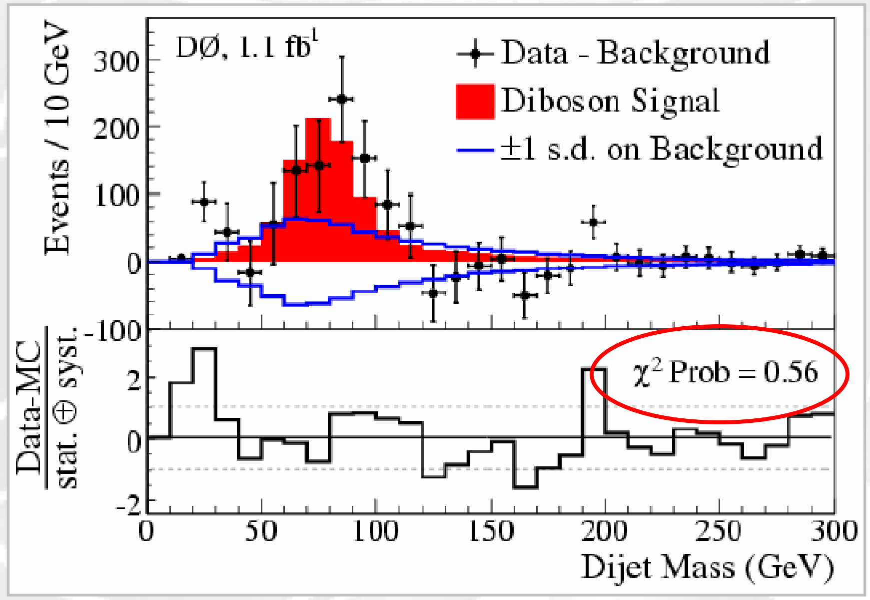
Cross Section Measurement and Significance II

- Fit of the **Dijet Mass** instead of RF outputs
  Systematic Uncertainties evaluated in the dijet mass distribution

Larger Uncertainties
Smaller Significance

Combined: $\sigma_{WW/WZ} = 18.5 \pm 2.8(stat) \pm 4.9(syst) \pm 1.1(lumi)$ pb





- **Expected p-value ~ 1.7×10^{-3}**
  $\Rightarrow$ **Significance 2.9σ**
- **Observed *p*-value ~ 4.4×10^{-4}**
  $\Rightarrow$ **Significance 3.3σ**

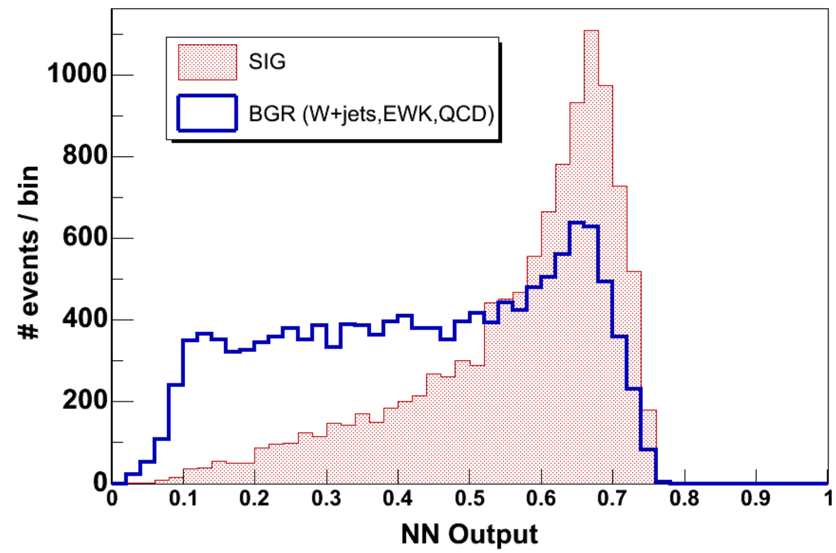
RF Fit vs. Dijet Mass Fit:
Differences indicate
different biases/sensitivities
in RF vs. Dijet mass

WW+WZ→lνjj Production at CDF II

Preliminary: 1.2/fb RunIIa data

• 95% CL limit on the cross section

WW+WZ→lνjj channel, l=el/muon



Signature:

1 high p_T charged lepton (>20 GeV)

MET>25 GeV

Dijet mass region (45-160) GeV

Improved Selection: Neural Network

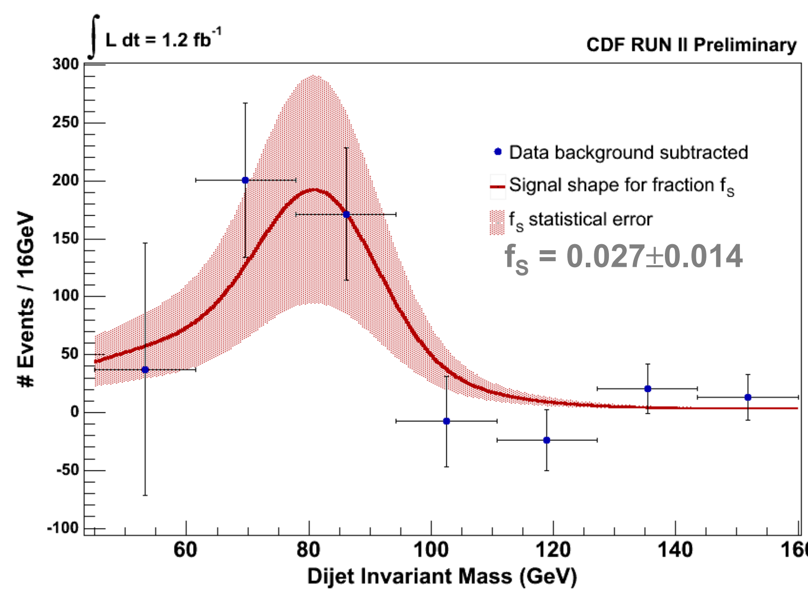
Candidate Events:	**15016**
Expected WW+WZ:	**554 ± 24**
Measured WW+WZ:	**410 ± 212 (1.7σ)**

➡ **σ×BR(W→e/μν;W/Z→jj):**

$$\sigma_{WW+WZ} \times BR < 2.88 \text{ pb at } 95\% \text{ CL}$$

$$\sigma^{Theory}_{WW/WZ} \times BR = 2.09 \pm 0.14 \text{ pb}$$

[Submitted to PRD RC (2009), hep-ex/0903.0814v1]



Summary

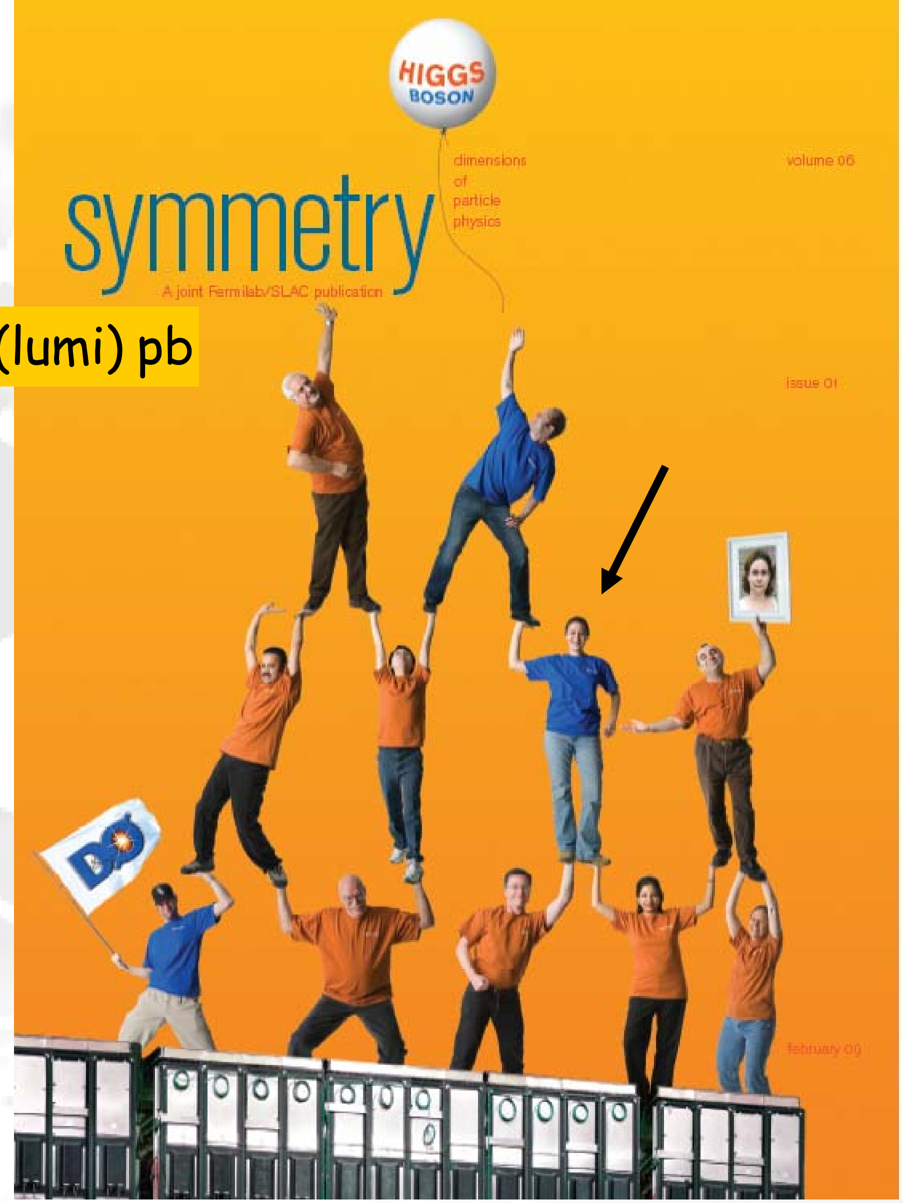
- First Evidence of WW+WZ production in the 2-jet final state at a hadron collider with observed significance 4.4σ

$$\sigma^{Measured}_{WW+WZ} = 20.2 \pm 2.5(stat) \pm 3.6(syst) \pm 1.2(lumi)\ pb$$
$$\sigma^{Predicted}_{WW+WZ} = (12.4 + 3.7) \pm 0.9 = 16.1 \pm 0.9 pb$$

- Result is consistent with previous Tevatron diboson measurements
- We boost the significance by using Random Forest Classifier (equivalent to the 35% increase of luminosity)
- DØ ability to extract small signal (like Higgs) in a large background !

Submitted to PRL
FermilabPub08/457E
arXiv:/0810.3873 [hep-ex]

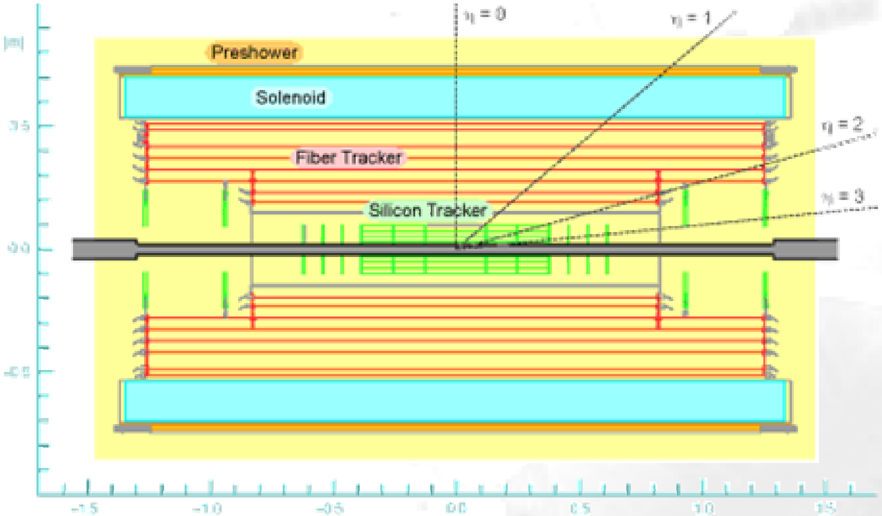
The DØ "Cheerleaders"



Backup Slides

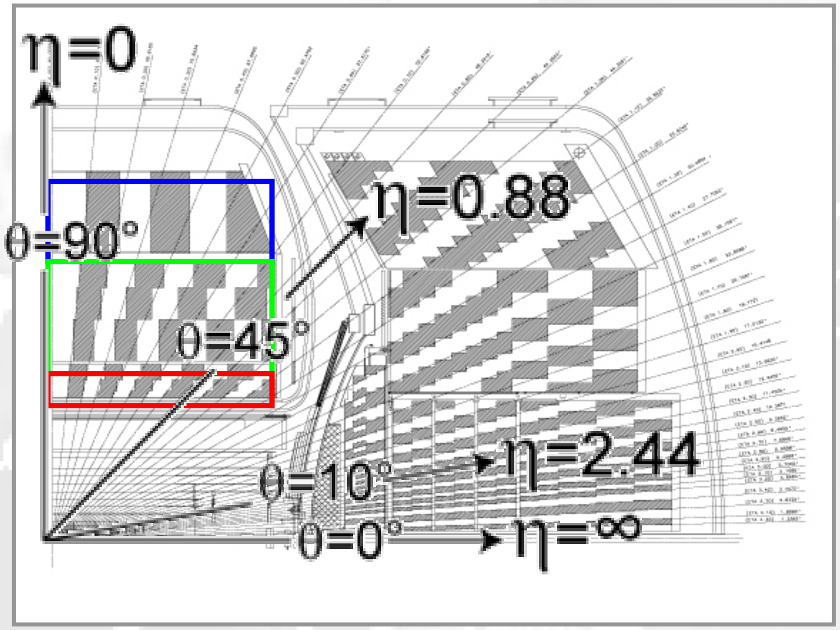
DØ Assembly Building & DØ Control Room (Tornado shelter !)

Central Fiber Tracker

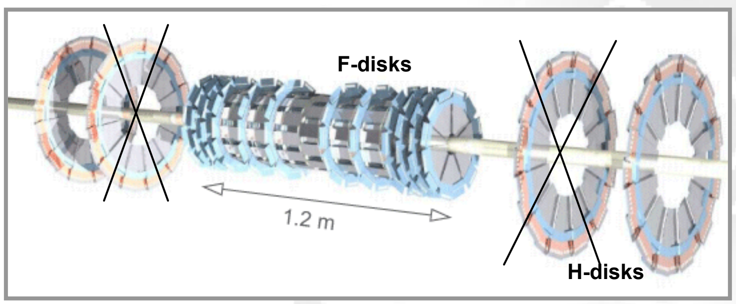


- Scintillating fibers (8 conc. cylinder)
- ~ 72k channels

Calorimeter



Silicon Microstrip Tracker Vertex Detector



4L+L0

- Six 12cm long barrels (5 detector layers) with interspersed disks (F-disks) for forward tracking
- Large area disks (2 H-disks) for forward tracking $|\eta| > 2$

- Uranium/Liquid Argon
  Electromagnetic (EM)
  Fine Hadronic (FH)
  Coarse Hadronic (CH)
- Central (CC) $|\eta|<1.1$ +
  2 Endcaps (EC) $1.4<|\eta|<4.0$
- ICD $1.1<|\eta|<1.4$
- Fine segmentation:
  $\Delta\eta\times\Delta\varphi = 0.1\times0.1$ (tower)
  (maximal shower) 0.05×0.05
- 55k readout channels

Alpgen Modeling

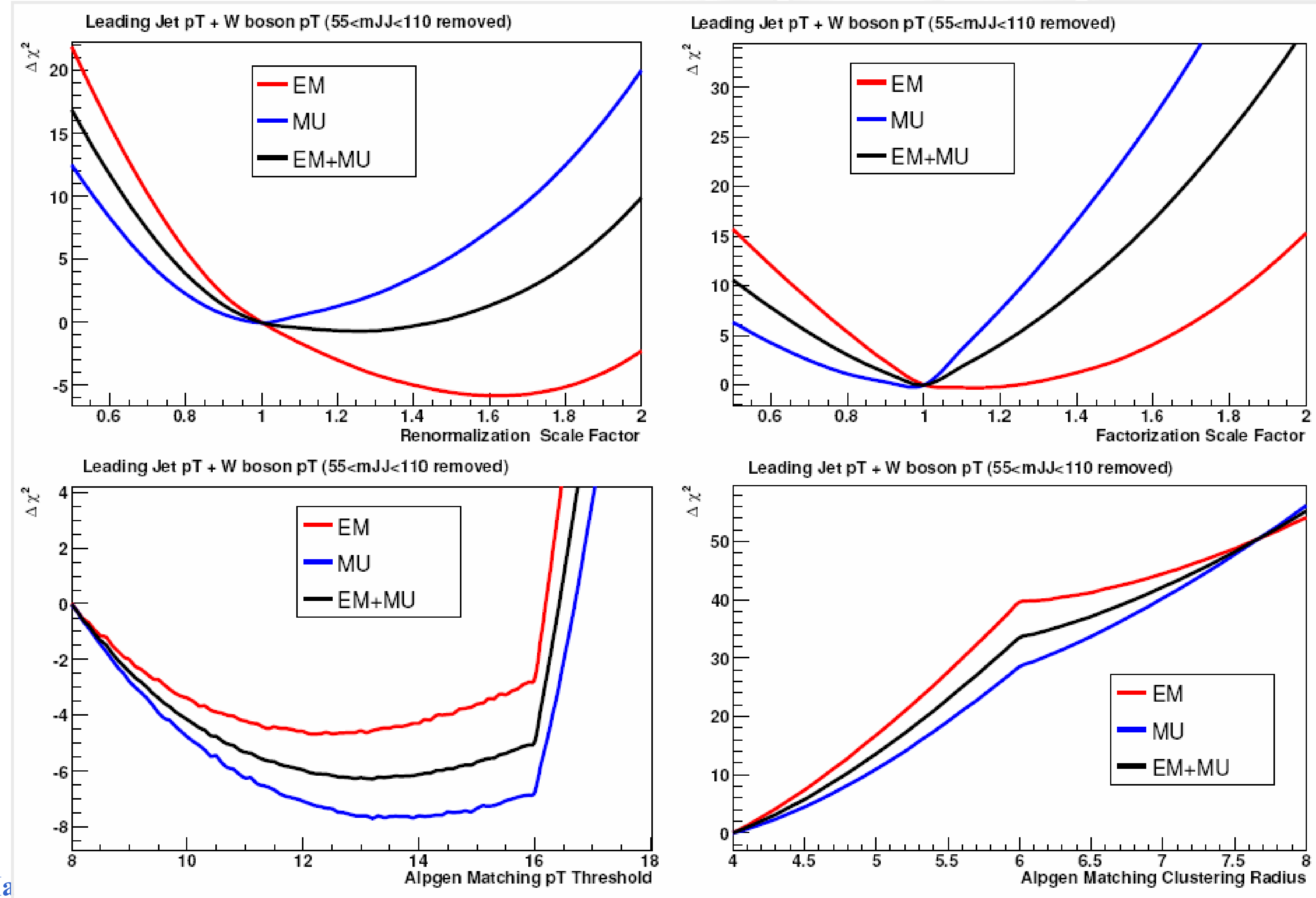
- ALPGEN: accurate description of the hard process, parton level, needed for N-jet simulation
- PYTHIA: parton showers, needed for realistic description of the final state

PROBLEM: Double counting of final states due to jets from showering

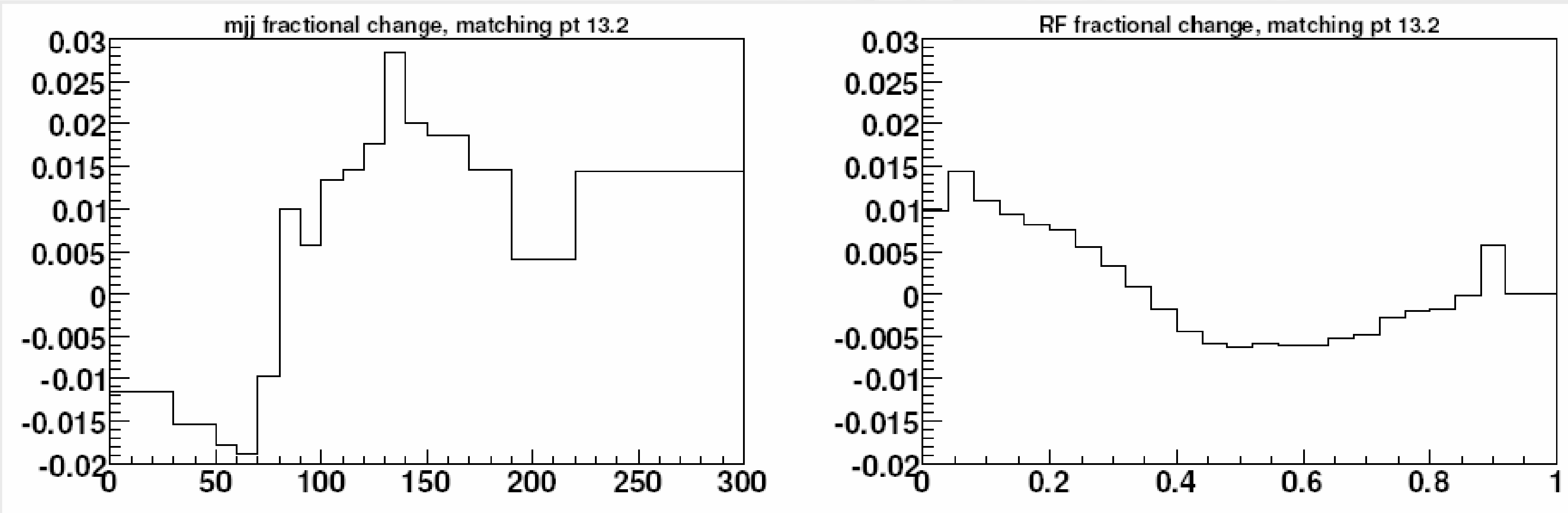
MLM parton-jet matching algorithm (ALPGEN): Cluster the showered partons into cone jets. Keep events only if each jet is matched to just one parton.

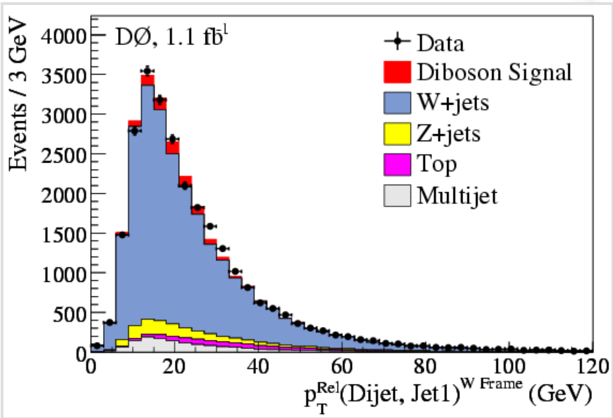
(MLM = Michelangelo Mangano)

Change in χ^2 between data and MC when varying each Alpgen parameter and simultaneously comparing the leading jet and W boson p_T distributions only using events outside of the dijet mass region 55 GeV < M_{jj} < 100 GeV



Fractional shape change between the DØ default parton-matching p_T of 8 GeV and preferred value of 13.2 GeV for dijet mass (left) and RF output (right) distributions.

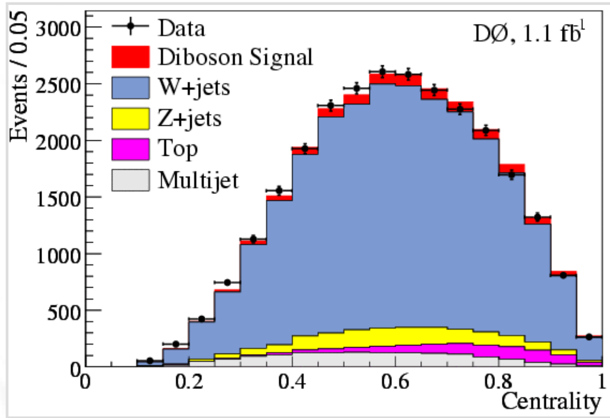




p_T^{Rel} (dijet, jet1)$_{WF}$

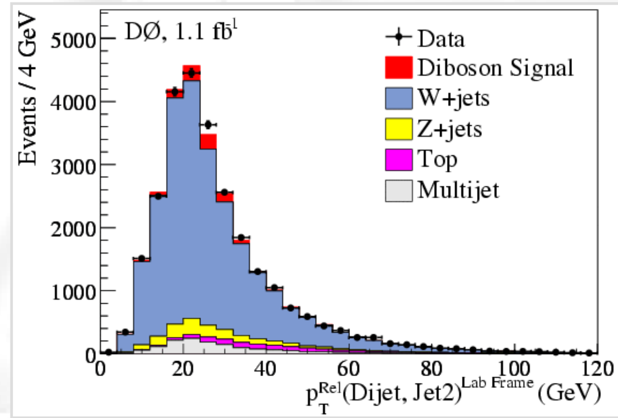
Magnitude of $p_T^{(jet1)} \perp$ to dijet system calculated in the rest frame of the reconstructed leptonic W boson

$$\left|\vec{p}_T(jet_1 + jet_2) \times \vec{p}_T(jet_1)\right| / \left|\vec{p}_T(jet_1 + jet_2)\right|$$



Centrality

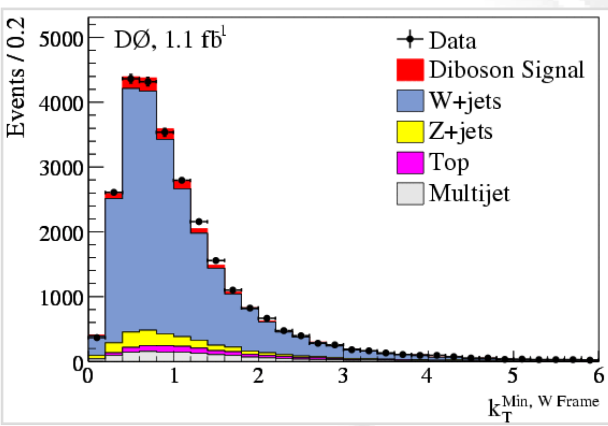
Sum of scalar p_T divided by sum of energy for the lepton+jets in the event



p_T^{Rel} (dijet, jet2)$_{WF}$

Magnitude of $p_T^{(jet2)}$ to di-jet system $\perp$ to dijet system calculated in the lab frame

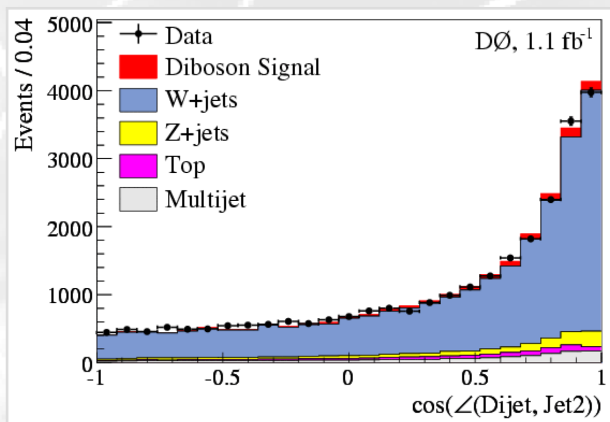
$$\left|\vec{p}_T(jet_1 + jet_2) \times \vec{p}_T(jet_2)\right| / \left|\vec{p}_T(jet_1 + jet_2)\right|$$



k_T^{Min} (W frame)

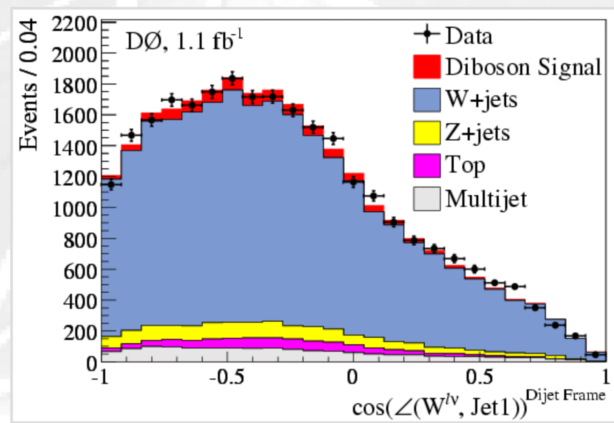
Calculated in the rest frame of the reconstructed leptonic W boson

$$\Delta R(jet_1, jet_2) \cdot E_T(jet_2) / \left(E_T^e + MET\right)$$



cos [∠(dijet, jet2)]

Cosine of the angle between dijet system and the secong jet in the lab frame



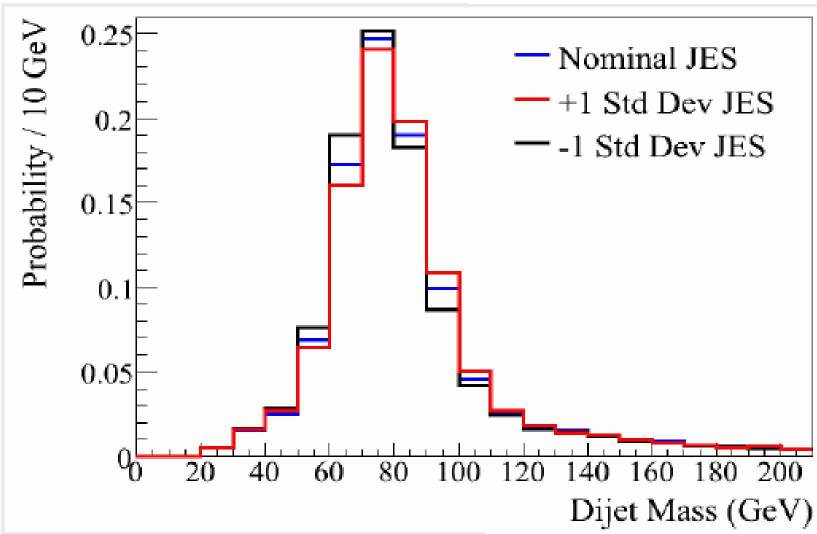
cos [∠($W_{l\nu}$, jet1)]$_{DF}$

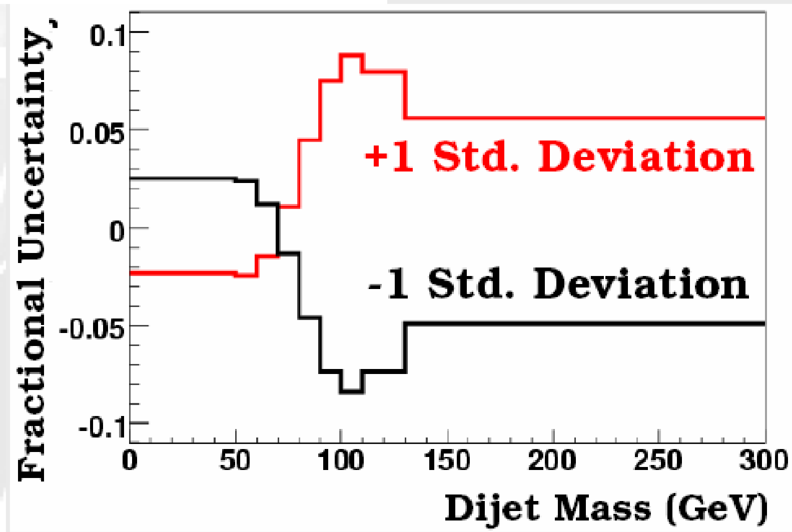
Cosine of the angle between the reconstructed leptonic W and the leading jet in the rest frame of the dijet system

Shape Uncertainty

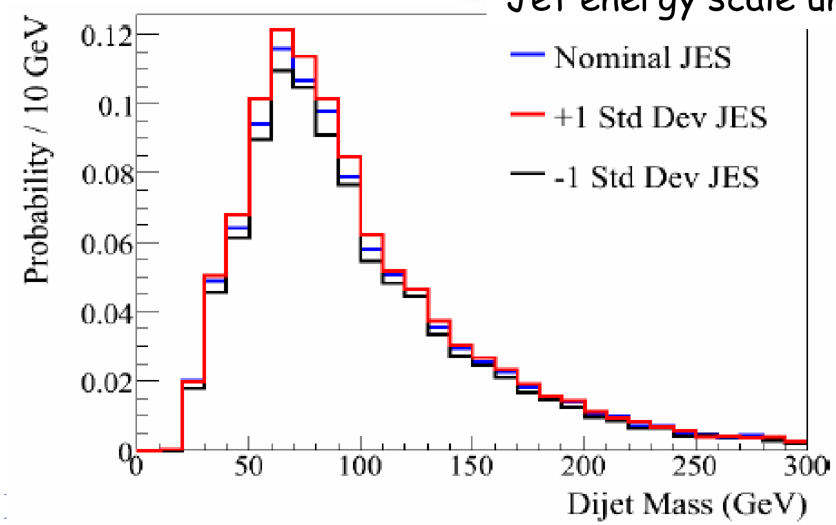
Evaluated for each uncertainty source and each sample in each channel for each (RF input) distribution separately

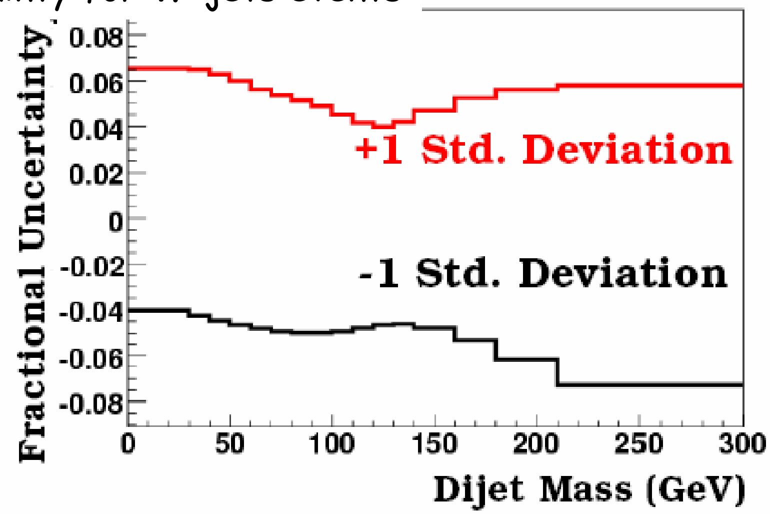
Jet energy scale uncertainty for signal events



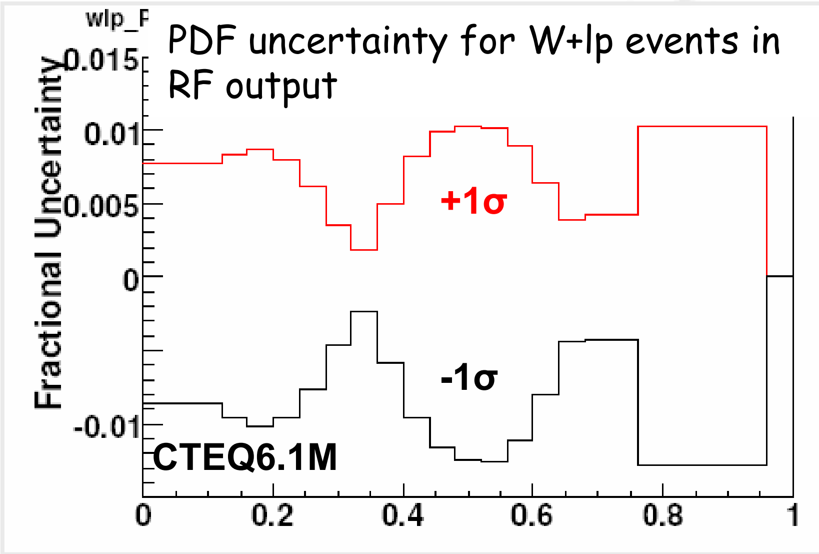


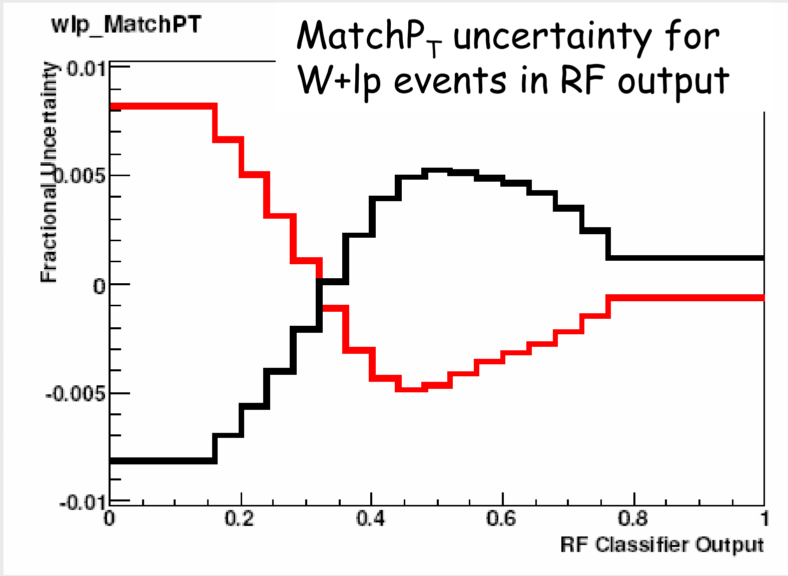
Jet energy scale uncertainty for W+jets events





Shape Uncertainty





Jet Resolutions

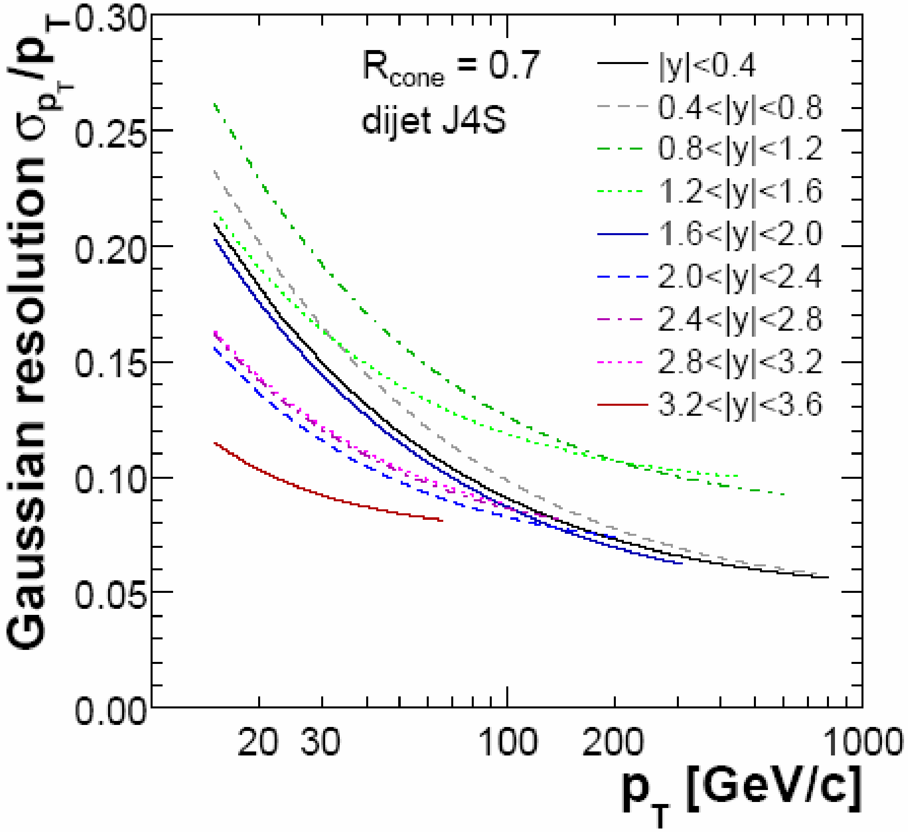


FIG. 35: Final Gaussian resolution for data as implemented in qcd_jet.caf v01-00-03.

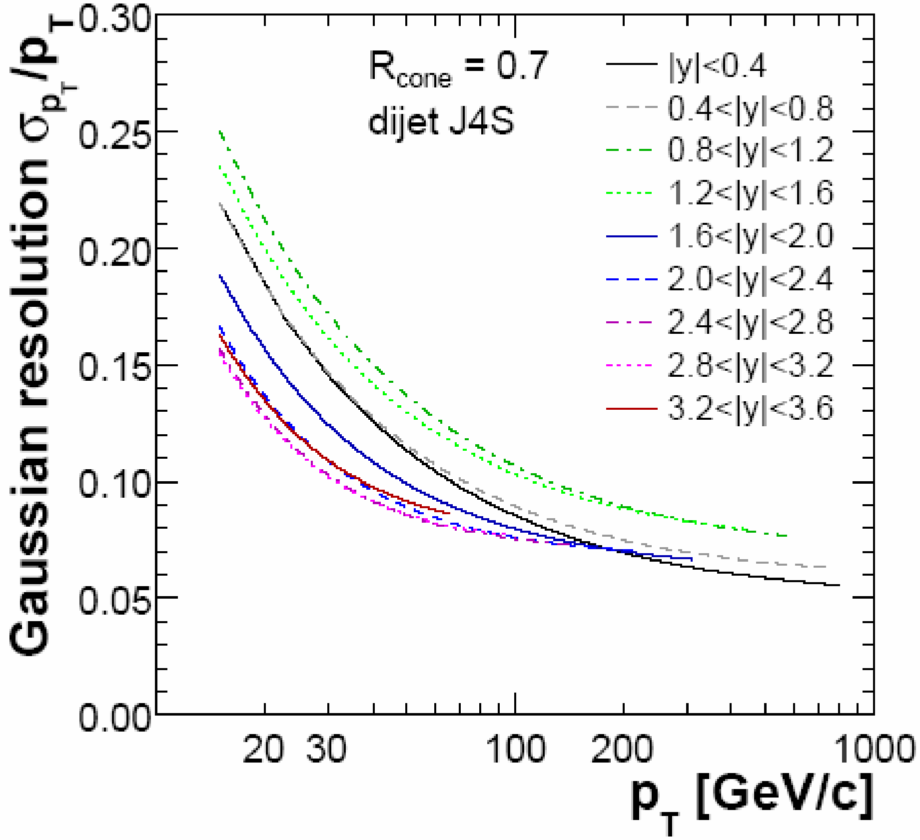


FIG. 36: Final RMS (GPT) resolution for MC (truth) as implemented in qcd_jet.caf v01-00-03.

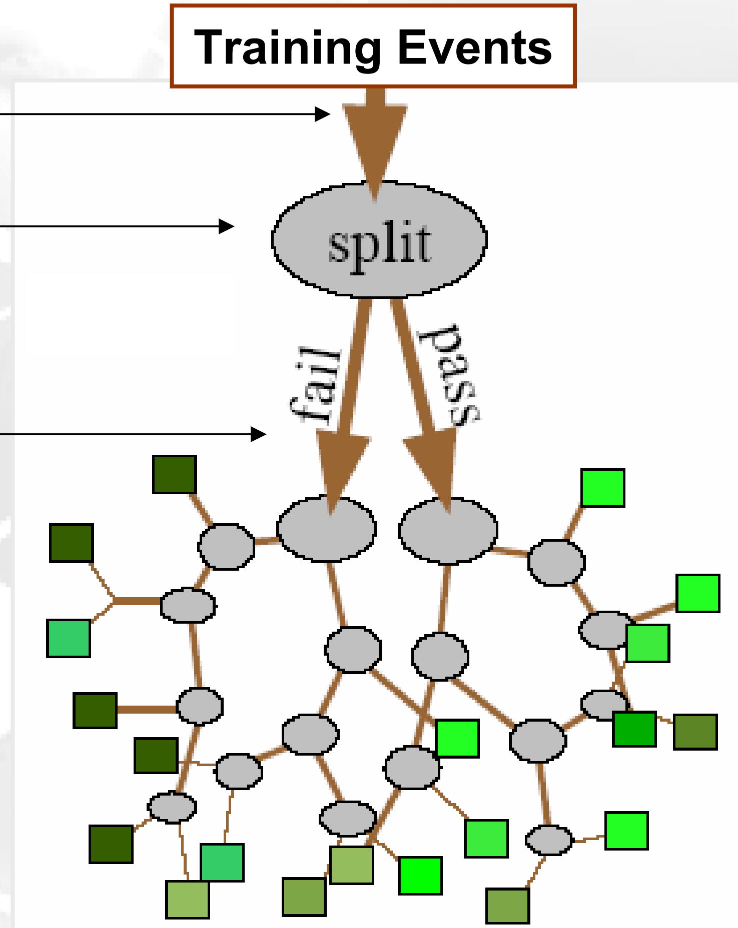
Random Forest

Decision tree is trained/grown using a set of known signal & background training events
⇒ These events go into the root node

Algorithm looks at all possible splits on all input variables and applies split giving best separation between signal and background

Events pass into one of two child nodes depending on whether they pass or fail

This process is repeated until:
- A node contains all or no signal events
- Number of events per node is less than a prespecified amount (optimized for each application)



Output for an unknown event is determined by the signal purity of the terminal node that the event ends up in