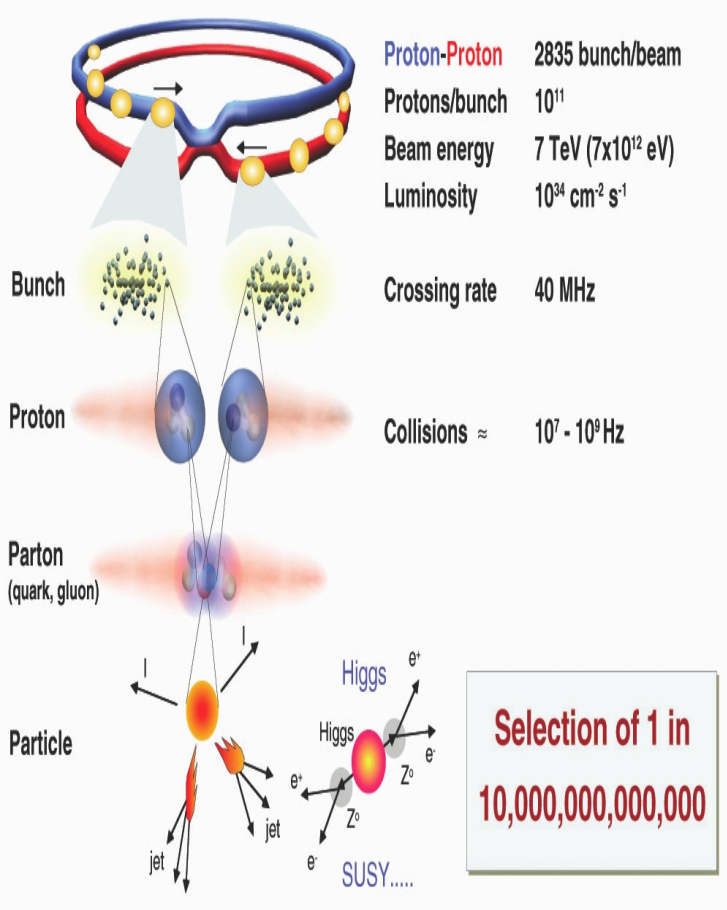
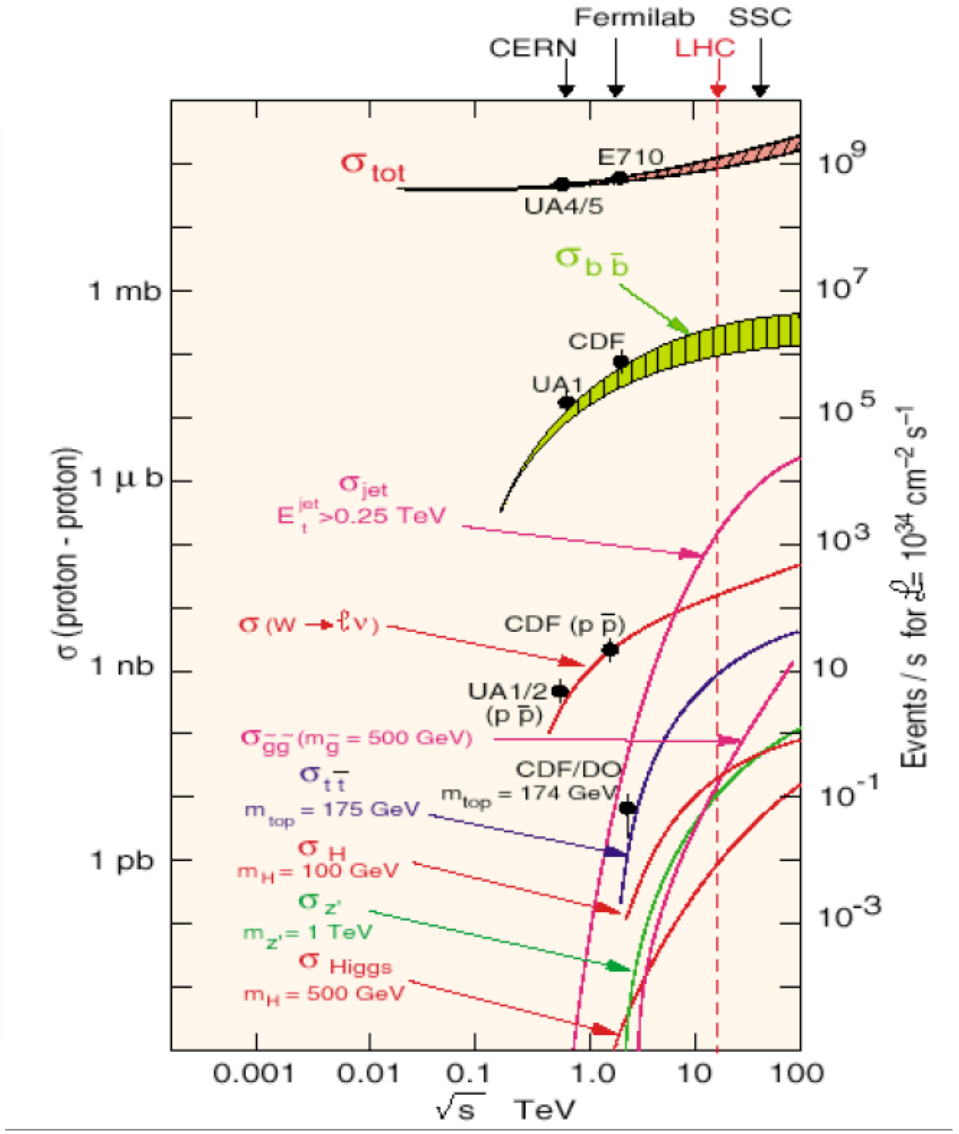
Status and Prospect of Vector Boson + Jets Study at CMS

Haifeng Pi
hpi@physics.ucsd.edu

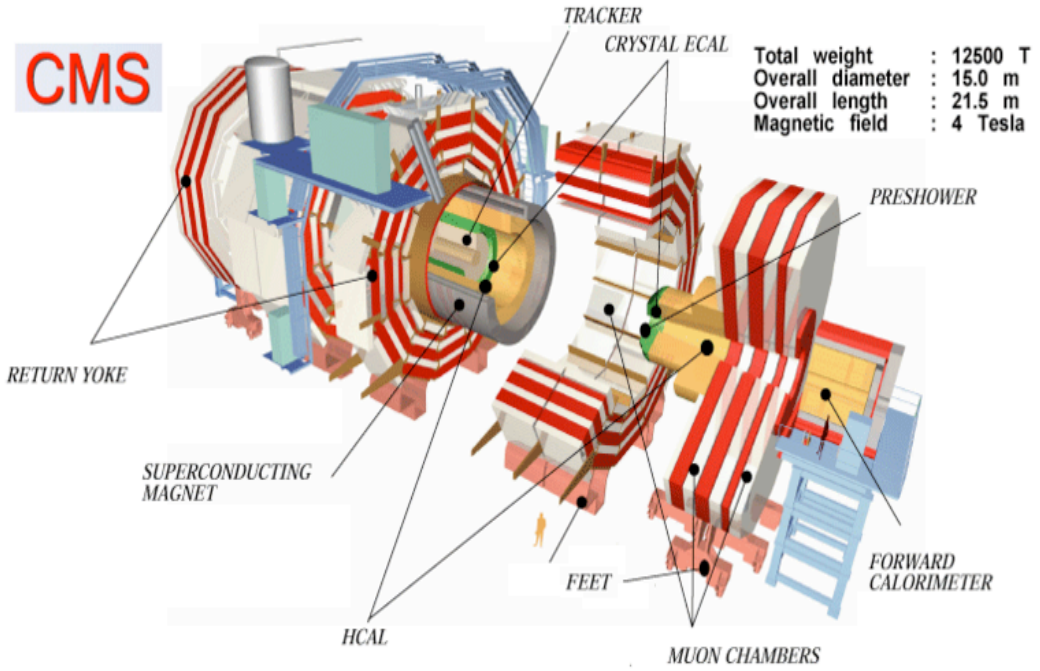
University of California, San Diego

LHC Collision and Cross Section





CMS Experiment and Detector

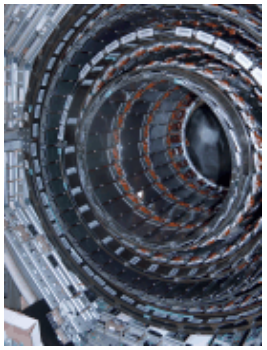


LHC

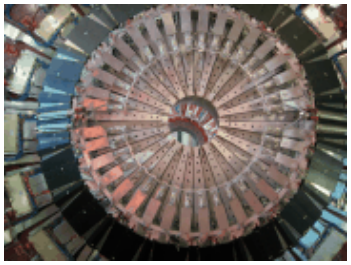
- *CME 14 TeV*
- *Luminosity 10^{33}-$10^{34}CM^{-2}S^{-1}$*
- *Crossing rate 40 MHz*
- *P-P 2835 bunch*
- *Proton/bunch 10^{11}*
- *Inelastic CrossSection 10^{8} nb*

CMS Detector

- *Tracker $|\eta| < 2.4$*
- *Muon Detector $|\eta| < 2.4$*
- *ECAL $|\eta| < 3.0$*
- *HCAL $|\eta| < 5.0$*
- *B = 4T*



Tracker (TIB, TID) (Oct. 2006)



HCAL Endcap (Feb. 2006)

Solenoid (Feb. 2006)

Muon Barrel DT + RPC (Oct. 2007)

CSC in Muon Endcap (Feb. 2006)

Inclusive W→lν and Z→ll Studies at CMS

The inclusive W and Z reconstruction and cross section measurement using in leptonic channels are important to the VB + jets study

- **Measurement of basic efficiency**
  lepton ID, selection, isolation, trigger, Z ID, W ID, Missing Et, M_T
- **Measurement of properties of inclusive W and Z production**
  cross section, distribution
- **Tuning of MC tools and study of theoretical issues**
  NLO and NNLO (related easy for inclusive analysis)
- **Study of detector systematic uncertainty**
  lepton momentum resolution, scale, misalignment
- **Measurement of QCD and other Background**

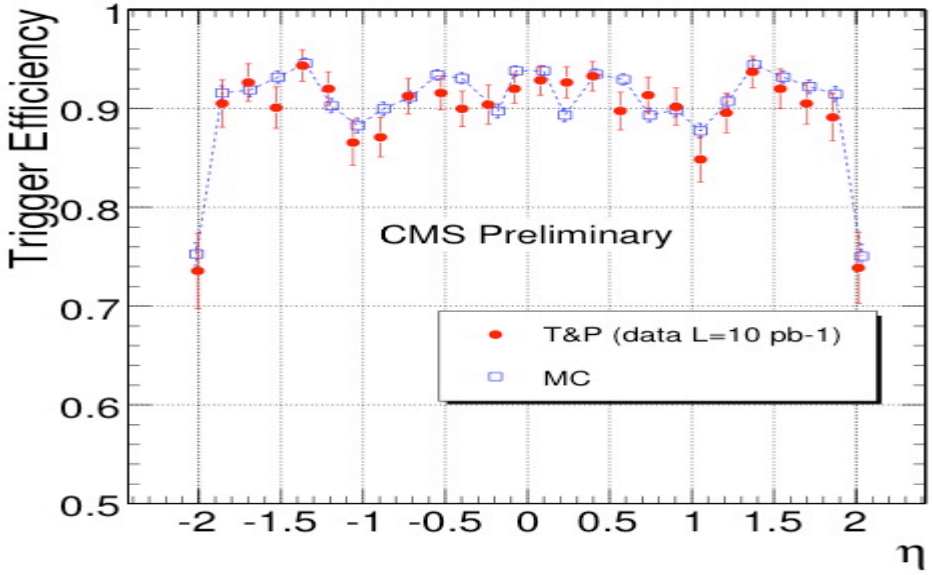
Two New Studies Recently Approved by CMS

- **Towards a measurement of the inclusive W→μν and Z→μ+μ- cross section in pp collisions at sqrt(s)=14 TeV**
  - *Available at http://cms-physics.web.cern.ch/cms-physics/public/EWK-07-002-pas-v8.pdf*
- **Towards a measurement of the inclusive W→eν and Z/γ→ e+e- cross section in pp collisions at sqrt(s) = 14 TeV**
  - *Available at http://cms-physics.web.cern.ch/cms-physics/public/EWK-07-001-pas-v5.pdf*

Many important measurements can be done with 10 pb^{-1} and extracted from dedicated studies on data

Next a few slides summarize some important results of inclusive W/Z analysis to the VB+jets study

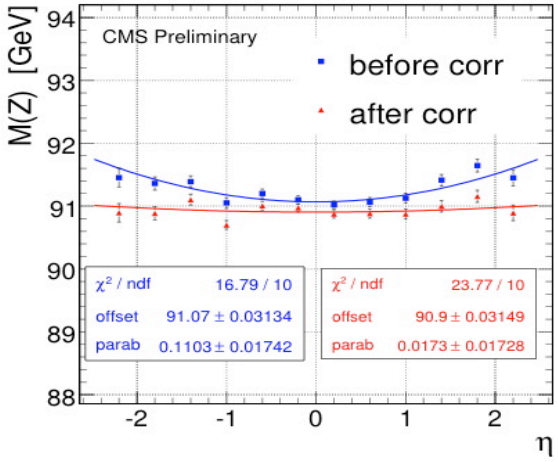
Lepton Efficiency and Systematics

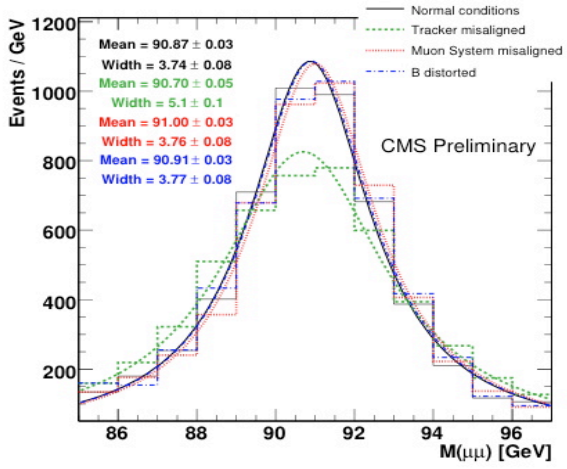


Muon Trigger Efficiency

Tag-and-probe method used to determine the muon/electron efficiency from data

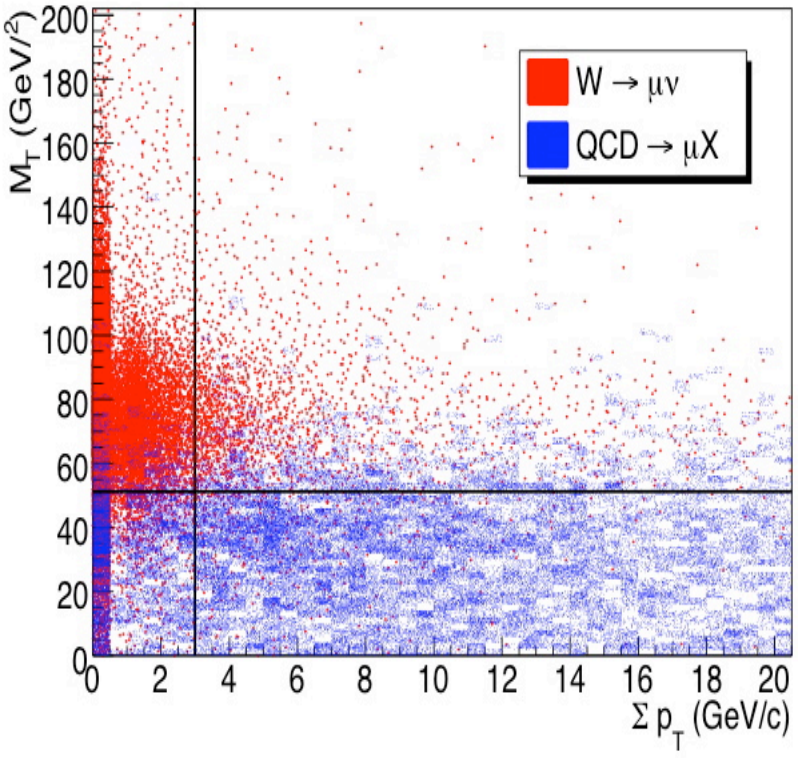
Missing Et related efficiency in W sample can be measured from Z sample



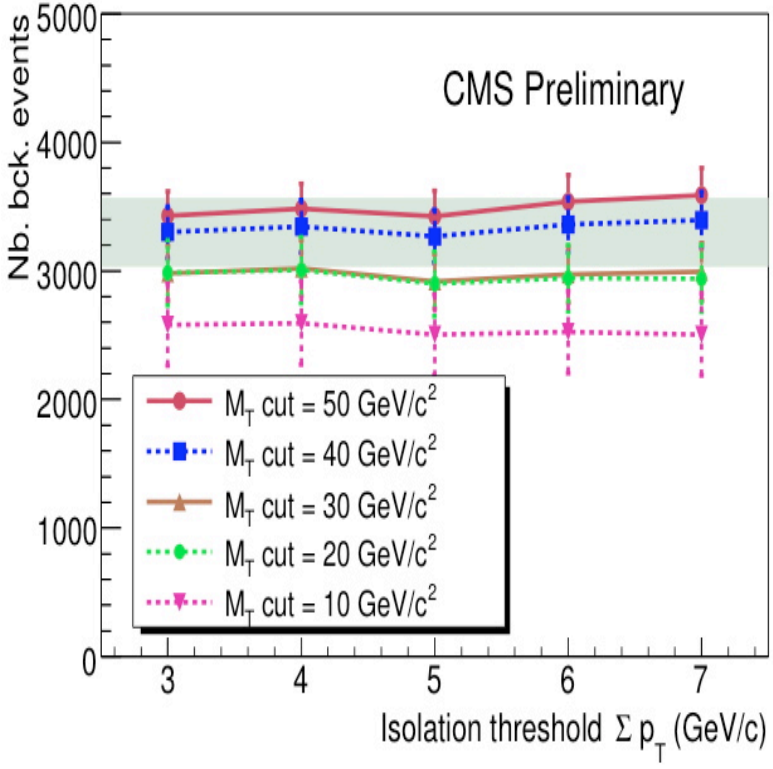


Z mass distribution vs Muon systematics

QCD Background

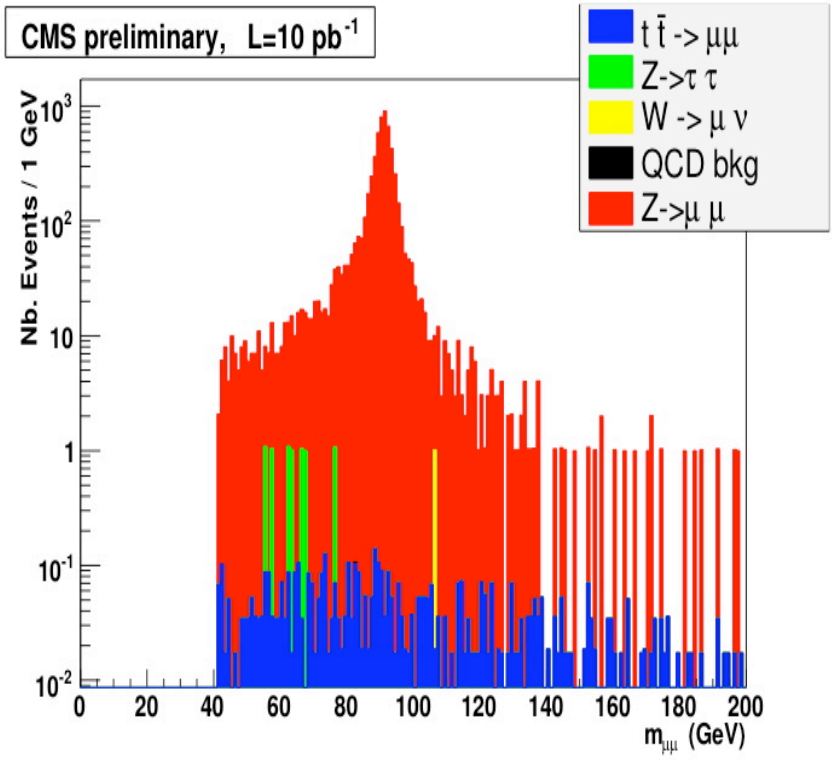


Matrix method for determining QCD faked muon rate

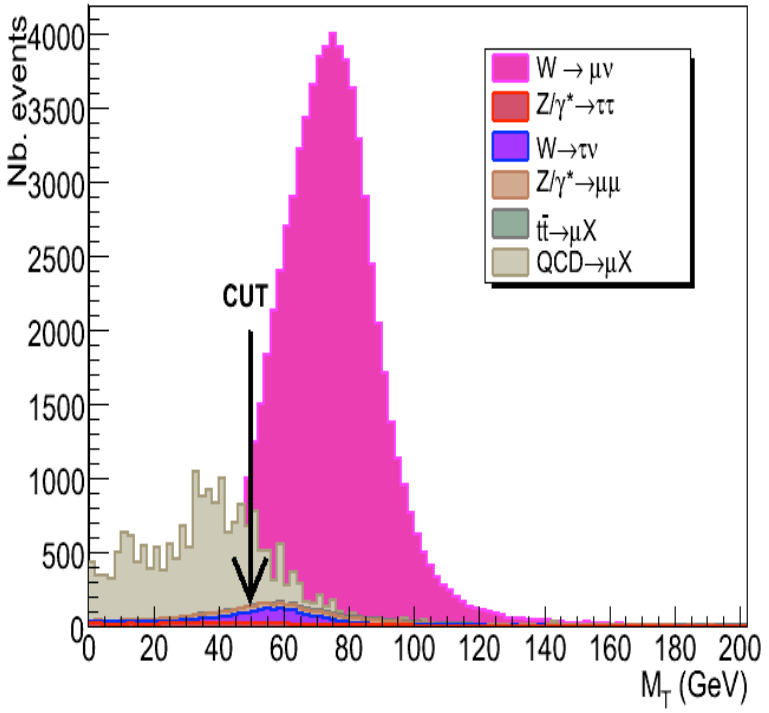


QCD faked muon rate vs isolation

Reconstructed Inclusive Z and W in Muon Channel

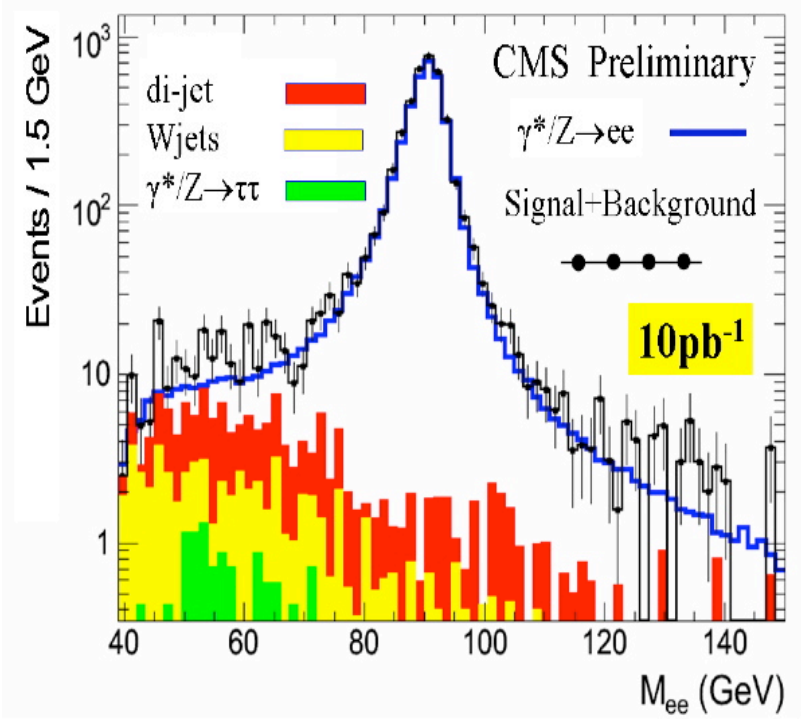


Z mass distribution using Z→μμ channel

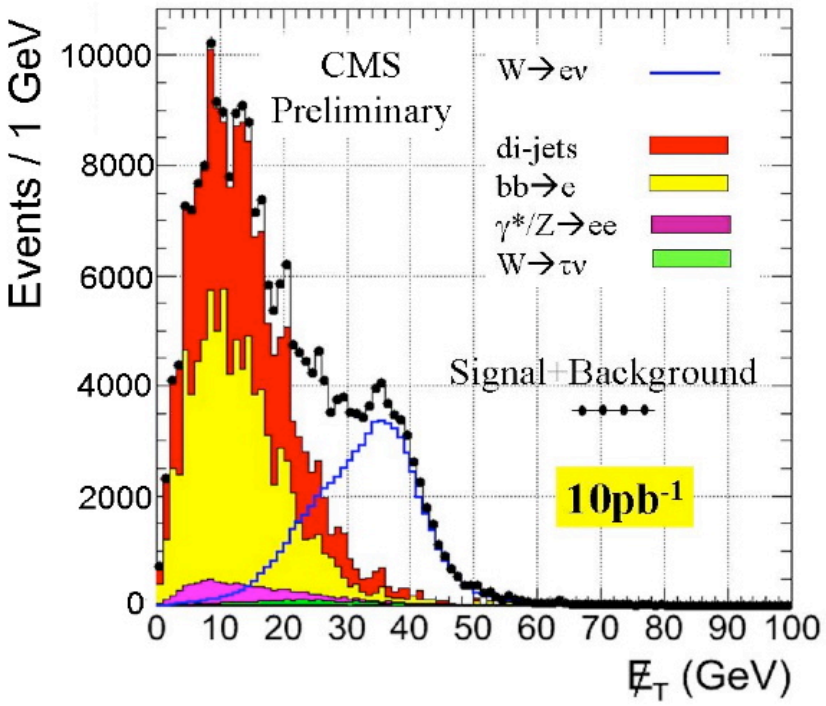


M_T distribution using W→μν channel

Reconstructed Inclusive Z and W in Electron Channel



Z mass distribution using Z→ee channel



Missing Et distribution using W→eν channel

QCD background is a severe problem for W→eν final states. Tighter electron isolation cuts are necessary.

W/Z+jets Study in 2006 and After

In the following slides, results of W/Z+jets are from the analysis performed in 2006. The approved results can be viewed at

h ttps://twiki.cern.ch/twiki/bin/view/CMS/WZJetsAN2006

The work was concentrated on

- ***Evaluation of MC-based selection efficiency of Signal & Background***
- ***Study of W/Z identification***
- ***Study of lepton isolation, primarily justified from the MC-based evaluation of QCD background for lepton+jet+Missing E_T final states***
- ***Jet selection strategy***
- ***Evaluation of missing E_T correction***
- ***Primitive study of suppressing TTbar background for high jet multiplicity final states***

The first step towards the comprehensive reconstruction and analysis of W/Z+jets

Signal and Background of W/Z+jets

- **Signal Samples: W+0/1/2/3/4 jets, Z+0/1/2/3/4 jets**
  - *Generated by latest ALPGEN (version 2.05) with matching between matrix element and parton shower provided by PYTHIA*
  - *W+0jet, W+1jet, Z+0jet and Z+1jet are not covered by the full simulated datasets, so CMS Fast Simulation Software is used for W+Njets (N=0,1,2,3,4) and Z+Njets (N=0,1,2,3,4) simulation for consistency of analysis*
  - *The partonic level W+ Njets (N=3,4) and Z+Njets (N=3,4) shared the same files of "official" CMS production. W+ Njets(N=0,1,2) and Z+Njets(N=0,1,2) were generated based on the same configuration and event generation process ...*
- **Background Samples: TTbar, WW+jets, WZ+jets, ZZ+jets, QCD events**
  - *Generated by PYTHIA*
  - *Full detector simulation*
  - *High level trigger*
  - *Detector reconstruction*

Quick Review of Selection Cuts

Table 2: Summary of Event Selection Cuts

Selection	Configuration
Pre-Selection (W/Z+jets Only)	Quark $p_T > 20$ GeV/c $\Delta R_{\rm qq} > 0.7$ $\lvert\eta_{\rm q}\rvert < 5$ $\lvert\eta_{\ell}\rvert < 10$
Trigger	Standard single muon and electron for W+jets analysis Standard di-muon and di-electron for Z+jets analysis
Lepton Calo-Isolation	Electron: $E_T^{\rm Ecal}/(E_T^{\rm Ecal} + E_T^{\rm Ecal}) > 0.95$ $0.96 < E/p < 1.25$ $\lvert E_T^{0.2} - E_T^{\rm e}\rvert < 6.0$ GeV $\lvert(E_T^{0.2} - E_T^{\rm e})/E_T^{\rm e}\rvert < 0.1$ $E_T^{0.4} < 8.0$ GeV $E_T^{0.4}/E_T^{\rm e} < 0.12$ Muon : $E_T^{0.2} < 6.0$ GeV $E_T^{0.2}/p_T^{\mu} < 0.1$ $E_T^{0.4} < 8.0$ GeV $E_T^{0.4}/p_T^{\mu} < 0.12$
Lepton Track-Isolation	$N_{\rm trk}^{0.15} < 2$, $N_{\rm trk}^{0.30} < 3$
Faked Jet Removal	$\Delta R_{\ell_{\rm j}} > 0.5$
Lepton p_T Cut	W+jets: $p_T > 30$ GeV/c Z+jets: $p_T > 20$ GeV/c
E_T^{miss} Cut	W+jets: $E_T^{miss} > 45$ GeV Z+jets: $E_T^{miss} < 40$ GeV
W/Z Identification	W+jets: $\lvert M_T - 80.4\rvert < 25$ GeV Z+jets: $\lvert M_Z - 91.2\rvert < 4$ GeV
Jet Selection	$E_T > 50$ GeV $\lvert\eta\rvert < 4.5$

Selection Strategy

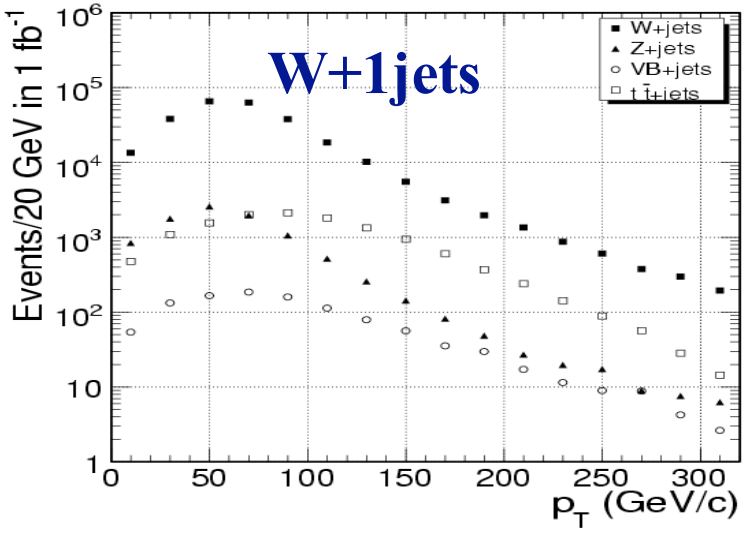
- ***Highly suppress QCD background***
- ***Reduce TTbar background***
- ***Jet counting strategy to be consistent with MC data***
- ***W and Z ID***
- ***Consideration systematics***
- ***Use full detector coverage***

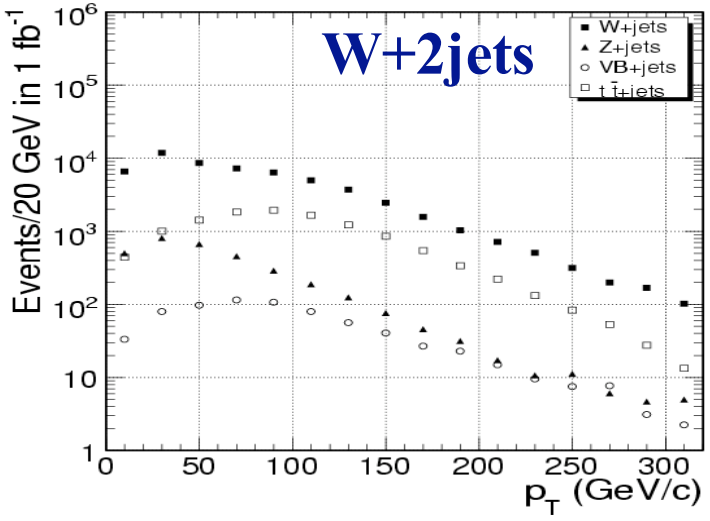
Combined Distribution in W/Z+jets Candidate Events

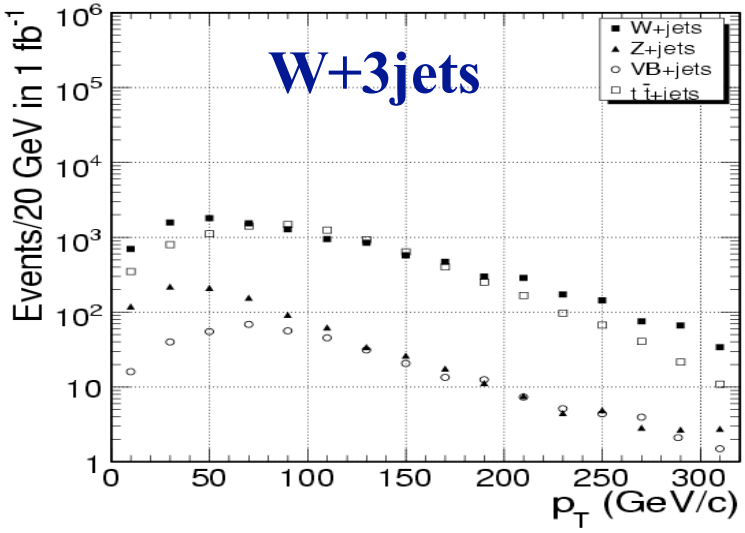
- **Combined experimental observable (signal and background)**
  - *W+≥Njets (N=0,1,2,3,4) and Z+≥Njets (N=0,1,2,3,4) candidates*
- **Using one selection schemes**
  - *Inclusive*
- **Number of events scaled to 1 fb^{-1}**

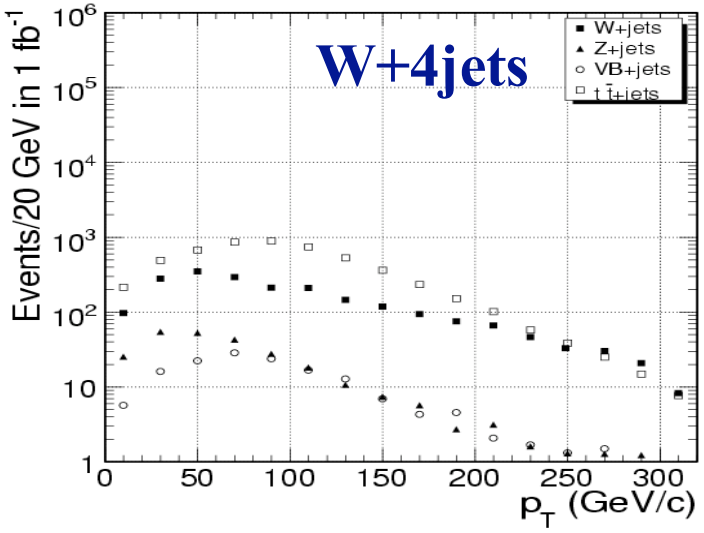
- **Basic observables of W/Z+jet properties**
  - *W/Z p_T distribution*
  - *W/Z η distribution*
  - *Jet E_T distribution*

W p_T in W+jets Selection (S & B)

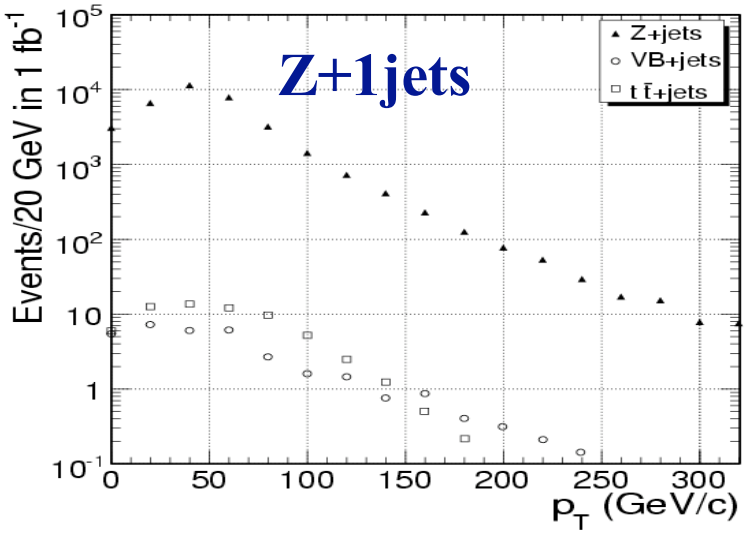


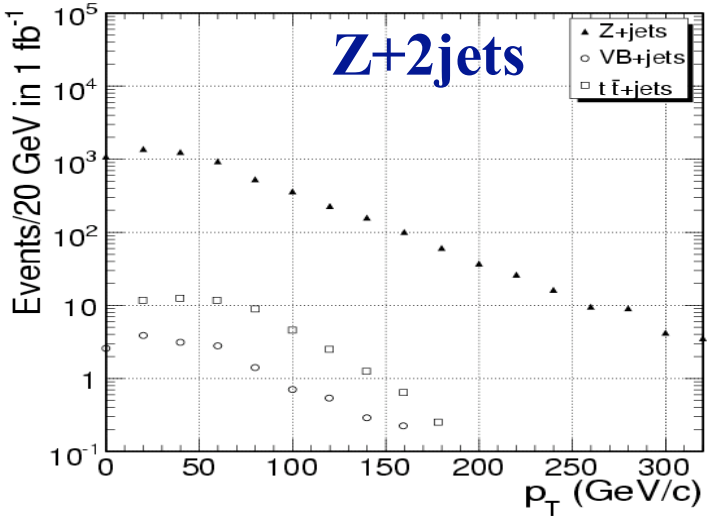


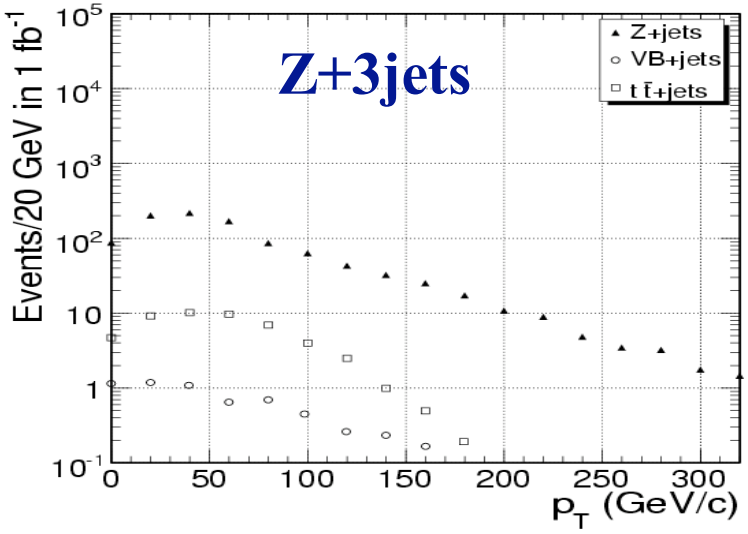


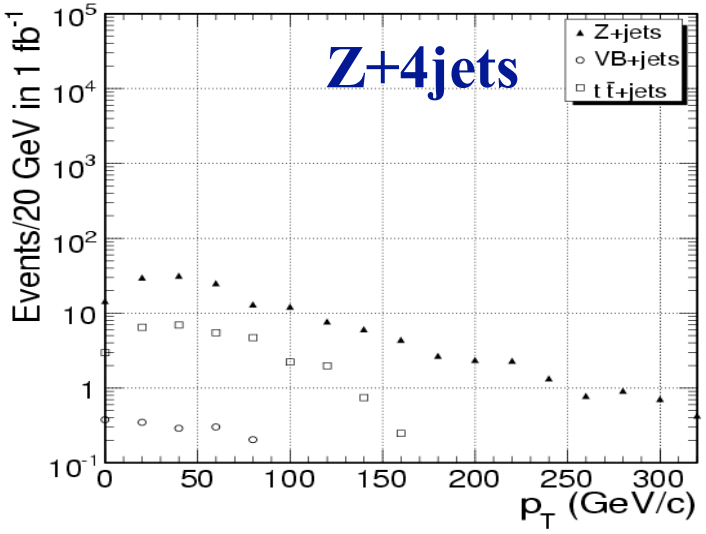


Z p_T in Z+jets Selection (S & B)



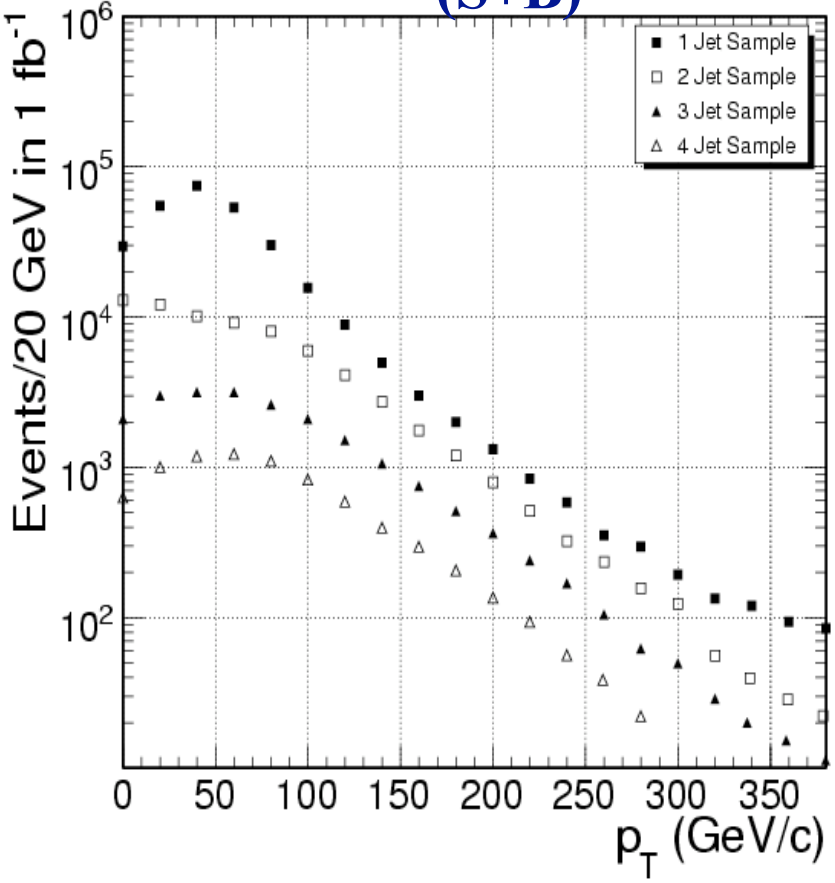




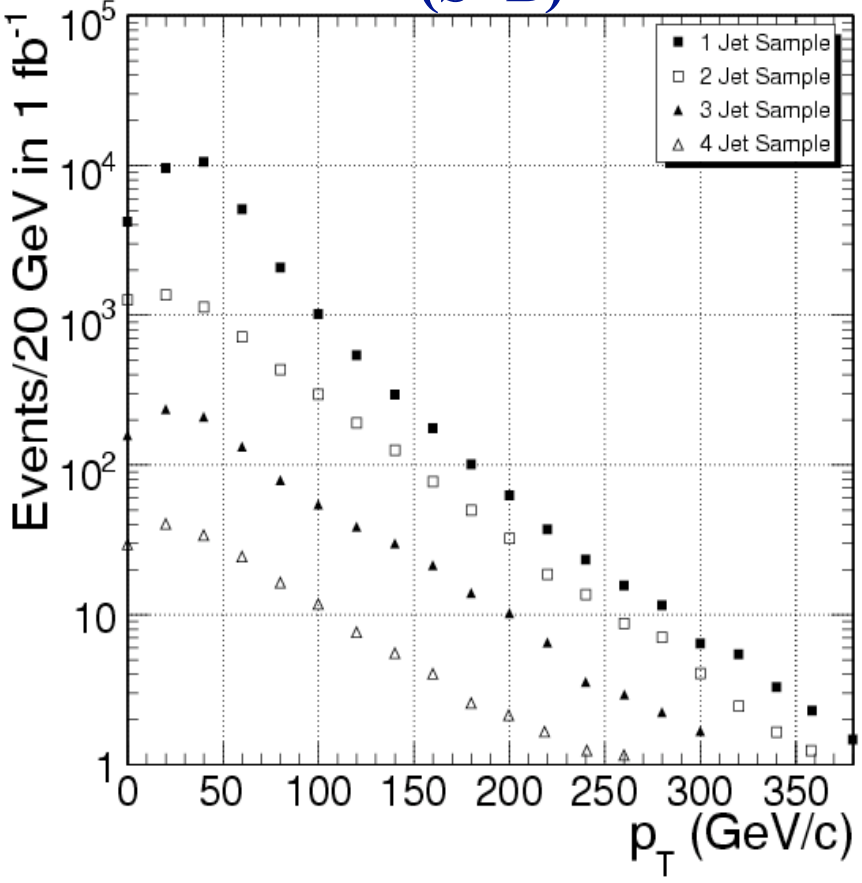


W/Z p_T in W/Z+≥Njets(N=1,2,3,4) as Observed Spectrum

W+jets (S+B)



Z+jets (S+B)



Expected Systematic Uncertainty at 1-10 fb^{-1}

Detector Systematic Uncertainty (Conservative Estimation)

Luminosity	*5%*
Jet energy scale (3%)	*8%*
Missing E_T	*5%*
Trigger & Lepton ID	*3.5%*
Lepton Isolation	*5%*

Background Subtraction ~10%

(to be evaluated and subject to S/B w.r.t. Jet multiplicity)

- **Overall 12.3% det-systematics without background subtraction**
- **Largely dependent on how jet energy scale reach w.r.t. Luminosity**
- **Missing E_T uncertainty is "over-estimated", selection cuts can reduce the systematic uncertainty**
- **Lepton ID and trigger uncertainty is much smaller**
- **Strategy of lepton isolation needs some thinking, but not a serious issue**

W/Z+jets Event Rate at 1 fb^{-1}

Table 6: Event Rate in 1 fb^{-1} with W+≥Njets(N=1,2,3,4) Selection

Channels	W+≥1jet	W+≥2jet	W+≥3jet	W+≥4jet
W+jets	260652 ± 828	56702 ± 390	10964 ± 178	2164 ± 81
Z+jets	9340 ± 96.6	3237 ± 56.9	972 ± 31.2	259 ± 16.1
$t\bar{t}$+jets	12897 ± 113.6	11842 ± 108.8	9052 ± 95.2	5420 ± 73.6
WW/WZ/ZZ+jets	1077 ± 32.8	714 ± 26.7	386 ± 19.6	151 ± 12.3
total	283966 ± 842	72495 ± 409	21374 ± 205	7994 ± 111

Table 7: Event Rate in 1 fb^{-1} with Z+≥Njets(N=1,2,3,4) Selection

Channels	Z+≥1jet	Z+≥2jet	Z+≥3jet	Z+4≥jet
Z+jets	35109 ± 187	6185 ± 78.6	977 ± 31.3	156 ± 12.5
$t\bar{t}$+jets	64 ± 8.0	58 ± 7.6	49 ± 7.0	32 ± 5.6
WW/WZ/ZZ+jets	33 ± 5.8	17 ± 4.2	5 ± 2.3	2 ± 1.4
total	35206 ± 188	6260 ± 79.1	1031 ± 32.2	190 ± 13.8

For W+jets

very serious contamination from TTbar background for W+≥Njets (N=2,3,4)

"flat" TTbar contribution for various jet multiplicities

Large uncertainty in background subtraction for W+≥Njets (N=3,4) because S/B~ 1

for low jet multiplicity, systematics dominated by detector effect

for high jet multiplicity, systematics dominated by background uncertainty

For Z+jets

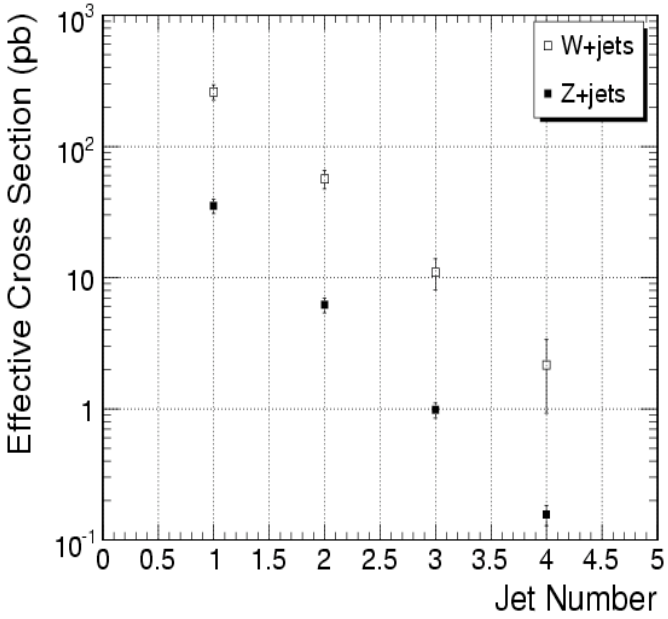
"flat" TTbar contribution for various jet multiplicities

high S/B for Z+≥Njets(N=1,2,3)

systematics dominated by detector effect

Measured Cross Section for Signal W/Z+jets using 1 fb^{-1}

$$\sigma(\mathrm{W}+\geq 1\mathrm{jet}) = 260.6 \pm 0.84^{\mathrm{sta}} \pm 34.9^{\mathrm{sys}}(\mathrm{pb})$$
$$\sigma(\mathrm{W}+\geq 2\mathrm{jet}) = 56.7 \pm 0.41^{\mathrm{sta}} \pm 9.1^{\mathrm{sys}}(\mathrm{pb})$$
$$\sigma(\mathrm{W}+\geq 3\mathrm{jet}) = 10.9 \pm 0.21^{\mathrm{sta}} \pm 2.9^{\mathrm{sys}}(\mathrm{pb})$$
$$\sigma(\mathrm{W}+\geq 4\mathrm{jet}) = 2.2 \pm 0.11^{\mathrm{sta}} \pm 1.2^{\mathrm{sys}}(\mathrm{pb})$$
$$\sigma(\mathrm{Z}+\geq 1\mathrm{jet}) = 35.1 \pm 0.19^{\mathrm{sta}} \pm 4.3^{\mathrm{sys}}(\mathrm{pb})$$
$$\sigma(\mathrm{Z}+\geq 2\mathrm{jet}) = 6.18 \pm 0.079^{\mathrm{sta}} \pm 0.77^{\mathrm{sys}}(\mathrm{pb})$$
$$\sigma(\mathrm{Z}+\geq 3\mathrm{jet}) = 0.97 \pm 0.032^{\mathrm{sta}} \pm 0.13^{\mathrm{sys}}(\mathrm{pb})$$
$$\sigma(\mathrm{Z}+\geq 4\mathrm{jet}) = 0.16 \pm 0.014^{\mathrm{sta}} \pm 0.024^{\mathrm{sys}}(\mathrm{pb})$$



- *Det-systematics dominates the uncertainty of the measurement of cross section at 1 fb^{-1}*
- *High jet multiplicities of W+jets have very large uncertainty*
- *Based on systematics, W+≥1/2jet and Z+≥1/2/3/4jets provides very good sensitivity for cross section measurement and theoretical model-dependent analysis*
- *Improving S/B for W+≥3/4jets is important to extend the sensitivity for high jet multiplicities of W+jets and comparison study with Z+jets*

The effective cross section is what we observed for a given selection. It is detector and selection dependent. Unfolding of jet multiplicity wasn't done.

Theoretical uncertainty related to NLO

The calculation of NLO cross section for high jet multiplicity events and incorporation with the matching mechanism between ME and PS remain a question for the theoretical predication for W+jets and Z+jets.

- ***50 GeV jet E_T cut highly suppresses the soft gluon radiation, which is the main effect of NLO correction.***
- ***Tevatron illustrated the basic consistency between the data and NLO-based prediction***
- ***LO-based prediction using 20-50 GeV jet E_T threshold will be tested with LHC data. We expect the physics properties (e.g. shape of the jet E_T distribution) of W+jets and Z+jets can be well predicted via a LO-based event generator when a high jet E_T threshold is chosen.***

- ***We expect the upper limit of NLO correction of WW/WZ/ZZ for W+1jets, W+2jets, W+3jets and W+4jets selection is 0.1%, 0.5%, 0.9% and 0.9% respectively. Similar expectation applies to their influence on Z+jets.***
- ***TTbar is a very serious background. The high E_T jets in TTbar are dominated by b jets from t->Wb and jets from W hadronic decay, we expect the physics properties of TTbar can be well predicted via a LO-based event generator.***

Summary

CMS will carry out comprehensive studies of W/Z+jets

- *Understand detector performance and tune the event generation and simulation*
- *Characterize and optimize the physics objects with their reconstruction*
- *Evaluate important trigger streams*
- *Strongly affect the new physics search*

Work towards the solutions for many important issues under the condition of real data analysis (especially data driven method ...)

- *Reduce the systematic uncertainty*
- *Measure the fake lepton rate from QCD and optimize the isolation*
- *Measure the lepton-related reconstruction and selection efficiency*
- *Improve the reconstruction of jet and missing Et*
- *Explore a variety of data driven analysis strategies*

With 1 fb^{-1} or less of LHC data

- *Most measurement w.r.t. W/Z+jets can be "done" with significant statistics*
- *Many optimizations based on real data analysis just start*
- *"Discovery" of problems in the system*
- *Work towards significant improvement in those difficult topics (?): high jet multiplicity final states, VB + heavy flavor ...*

Outlook: Detector Simulation and Reconstruction in Real Data Analysis

Any comparison between theoretical prediction and real data depends on a decent simulation which is reflected in many small details : *jet shape, low pt particles, underlying event structure, pileup, soft track in the lepton isolation cone, fake track rate ...*

- *To tune the simulation based on the real data is not an easy job*
- *To make the tuning/comparison possible, dedicated selection strategy needs to be studied (the comparison are better to be done via "clean" samples)*
- *The inconsistency will affect the analysis selection strategy. Need to minimize the impact of those inconsistency*
- *The inconsistency between the simulation results and data might also be our incomplete knowledge in the event generation level*

Differentiating the generation and simulation issues is important

Comprehensive investigation of reconstruction performance vs jet multiplicity

- *The dependency of jet and missing E_T scale/resolution on the jet multiplicity is potentially a serious issue*
- *The data driven jet energy scale is primarily derived from lepton(s)/γ + 1 jet event. The systematic uncertainty from 3rd or 4th jet needs to be studied*
- *Lepton isolation using calorimeter information needs to be evaluated*

Do reconstruction algorithm and calibration affect the results of comparison (other than affecting the S/B)? If so, in what level?

Outlook: Physics Analysis

Assuming the reconstruction and basic selection efficiency are well understood, physics analysis can be more concentrated on

- ***Optimizing the offline algorithm to improve the resolution of high level physics objects***
  - *Jet Et, Missing Et, lepton Pt*
- ***Increasing signal to background ratio***
  - *W/Z ID, lepton isolation, correlation between Missing Et and lepton/jet*
- ***Looking for best observables to resolve the theoretical uncertainty***
  - *Need to develop a comprehensive set of observable system sensitive to physics prediction, reducing the background, and canceling the systematic uncertainty*

Studies of high jet multiplicity final states is important but challenging due to large background contamination

- ***Complicated selection techniques introduces more systematic uncertainties in the selection chain***
  - *But reducing background significantly reduces systematic uncertainty due to background subtraction.*
- ***Jet performance plays an important role***
  - *Does jet merging-splitting help increase the sensitivity to physics predictions?*
  - *Sensitive to jet counting strategy*