Search for ZZ production in the 4 lepton channel with 4.8 fb^{-1} of data

M. Bauce

INFN and University of Padova

CDF Italia
1-2 September 2009

Summary

Analysis Goal:

Observation of ZZ and measurements of cross section in 4lepton decay in $p\bar{p}$ collision at $\sqrt{s} = 1.96$ TeV at CDF.

Summary

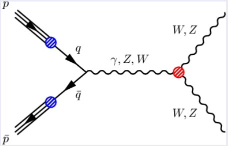
- Analysis technique
- Background modelling
- Analysis significance
- Cross-section measurement

For reference:

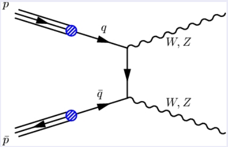
- More details can be found in
  - CDF Note 9852
  - `http://www-cdf.fnal.gov/physics/ewk/2009/ZZllll/ZZWeb/index.html`
- This is an extension of a previous analysis (CDF Note 9137 on 1.9 fb^{-1}):
  - new muon categories, adds new trigger paths (`PHIGAP` and `CMX-MsKs`)
  - extended to 4.8 fb^{-1}

Importance of diboson analyses

- Diboson production at CDF:
  - s-channel:



  - t-channel:



- Foundamental checks of the Standard Model.
- Important channel in the search of new physics:
  - Higgs decay channel in the high mass region:
    - $H \to WW$
    - $H \to ZZ$
  - Anomalous Trilinear Gauge Couplings (TGCs).

ZZ→lll'l' analysis

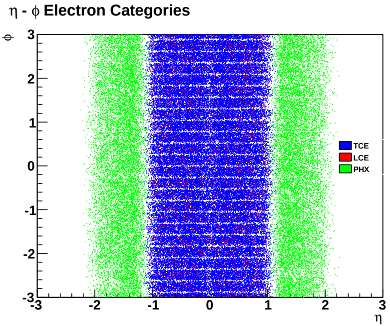
- At $\sqrt{s}$=1.96 TeV the NLO theoretical production cross section is: $\sigma_{p\bar{p}\to ZZ} = 1.4 \pm 0.1$ pb

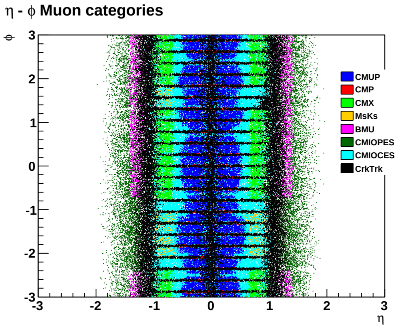
The decay channel considered (4 leptons) has two main features:

- Very low branching ratio: ~ 0.6 %
- Very low background (no physical processes).

Leptons considered

- Electrons:
  - Central Electrons, $|\eta| < 1$, TCE (*triggerable*)
  - Forward Electrons, $1 < |\eta| < 2$, PHX
- Muons:
  - Central Muons, $|\eta| < 0.6$, CMUP, CMP (*triggerable*)
  - Intermediate Muons, $0.6 < |\eta| < 1$, CMX arches, MiniSkirt-KeyStone (*triggerable*)
  - Forward Muons, $1 < |\eta| < 1.5$, BMU
  - Stubless Muons (m.i.p.), CMIOCES $|\eta| < 1$, CMIOPES $1 < |\eta| < 1.5$
- Track:
  - tracking coverage $|\eta| < 1.2$, CrkTrk





Data samples

Trigger paths

- MUON_CMUP18
- MUON_CMX18
- MUON_CMP18_PHIGAP
- ELECTRON_CENTRAL18

Good run lists - v29

Good run list	$\mathcal{L}$ (pb^{-1})
EM_NOSI	4828.8
EM_CMUP_NOSI	4772.5
EM_MU_NOSI_CMXIGNORED	4661.5
EM_SI	4549.4
EM_CUP_SI	4499.5
EM_MU_SI_CMXIGNORED	4394.8

Table: These luminosity have been scaled by the factor of 1.019.

Event selection

Signal region

- Exactly 4 leptons.
- Trigger Lepton ($p_T \geq 20 GeV$)
- Other leptons $p_T \geq$ 10 GeV
- No Track Iso (gain in acceptance)
- min $dR_{Leptons} > 0.1$
- M_{ll1} in (76-106) GeV (closest to the nominal Z mass, X-axis),
- M_{ll2} in (40-140) GeV (Y-axis).
- ll must have Same Flavor Opposite Charge

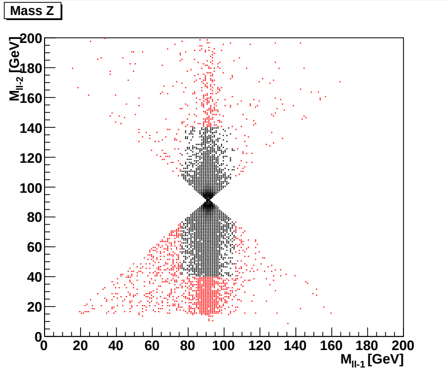


Figure: M_{ll-2} vs. M_{ll-1} for expected signal from MC.

Multi step efficiencies from MC ZZ sample.

	Efficiencies	
Triggered	one triggerable letpon scaled with trigger efficiency	**67 %**
4-Recons	4 leptons fully reconstructed in the detector	**20 %**
Pt min	min Pt request (20 GeV, 10 GeV, 10 GeV, 10 GeV)	**93 %**
dR min	$\min(dR_{all-leptons}) > 0.1$	**100 %**
Mass cut	76 GeV $< M_{ll-1} <$ 106 GeV , 40 GeV $< M_{ll-2} <$ 140 GeV	**80 %**
Total		**9.7 %**

Each efficiency is referred to the previous cut.

Expected number of signal events

- From ZZ MC sample we obtain: **4.68** $\pm$ **0.02(stat)** $\pm$ **0.76(syst)** in 4.8 fb^{-1}

Background modeling

Background

- Z+jets and $Z\gamma$+jets are the dominant background
- jet(s) mis-identified as lepton(s) $\rightarrow$ *fake* lepton(s).
- count the number of 3leptons + 1fake & 2leptons+2fakes events.

Two methods:

- Data driven estimated on Data using *fake-rates*.
  - Evaluate in jet data the probability for a lepton candidate (*denominator*) to be identified as a lepton (*fake rates*).
  - Apply this fake rates to our data sample for events with denominator object(s) to derive the expected background contribution.
- MC driven estimated the contribution of several process (WZ, $t\bar{t}$, Z γ) to our analysis
  - Just for cross check. Reported in backup slides.

Data Driven

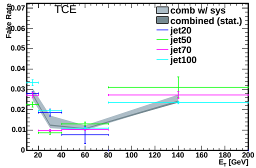
Fake Rates

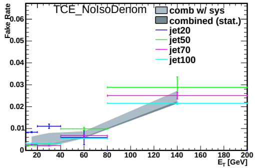
Calculate *Fake Rates* on a sample enriched in jets subtracting real EWK contribution:

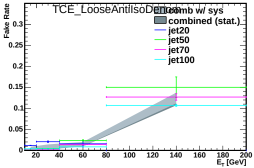
- Denominator: candidate lepton (almost a good track)
- Numerator: Fully identified lepton

$$f_i(p_T) = \frac{N_i(identified\ leptons)}{N_i(denominator\ objects)}$$

- Subtract expected contribution from real leptons.
- Fake Rates are calculated as function of p_T for each lepton type.
- Averaging measurements in JET20, JET50, JET70, JET100.
- Take the maximum spread as systematic uncertainty.







Background estimation

- To evaluate $Z(\gamma)$+jets background events we estimate the number of events with 3leptons+1fake and 2leptons+2fakes.

All numbers are expressed in 100^{ths} of event		
$\mathcal{L} = 4.8\ \mathrm{fb}^{-1}$	**Total Fake Estimation**	**ZZ contamination**
Standard (CalIso<0.3)	$19.31 \pm 6.57(\mathrm{stat.}) \pm 4.85(\mathrm{syst.})$	12.35
NoIso	$8.41 \pm 2.35(\mathrm{stat.}) \pm 3.60(\mathrm{syst.})$	2.67
AntiIso (CalIso>0.2)	$4.15 \pm 1.62(\mathrm{stat.}) \pm 2.87(\mathrm{syst.})$	0.96

- We chose the Data driven method with AntiIso denominators because it a priori has the better statistical precisions combined with a smaller ZZ contamination:
  - $4.15 \pm 1.62(stat.) \pm 2.87(syst.)$ 100^{ths} of event

Systematics

Systematics considered in this analysis

Fractional Uncertainty (%)		
NLO Acceptance	10.0 %	same as previous analysis
Cross -section	10.0 %	theoretical uncertainty
PDF uncertainty	2.7 %	same as previous analysis
Luminosity	6.0 %	from joint physics group
LeptonID $\pm 1\ \sigma$	3.6 %	varying ID SF in MC signal sample
Trigger Efficiency	2.1 %	varying ϵ^{trig} in MC signal sample
Fake rates: 3l+1d	42 %	from spread in fake rate calculation
Fake rates: 2l+2d	113 %	from spread in fake rate calculation

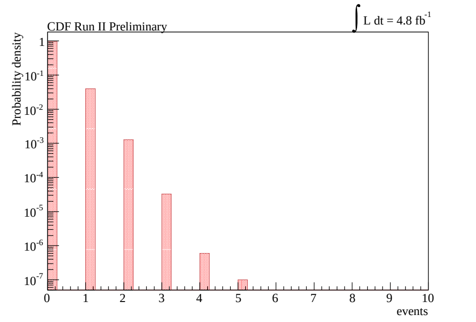
- The main systematic contribution is from Fake Rates.

Significance

- Calculate the probability that the background fluctuate to give a number of events greater or equal to the expected ones.
- With pseudo-experiments:
  - We use a poissonian distribution of the background b
  - Add a Gaussian smearing of the probability distribution for systematic

$$b \longrightarrow b \cdot \prod_{i=1}^{\#sys} (1 + s_i \cdot x_i)$$

  - x_i is a normal distributed variable; s_i is the systematic considered.
- Evaluate the probability to have $n_{obs} \geq n_{exp}$.

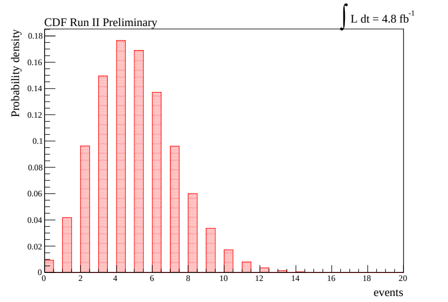


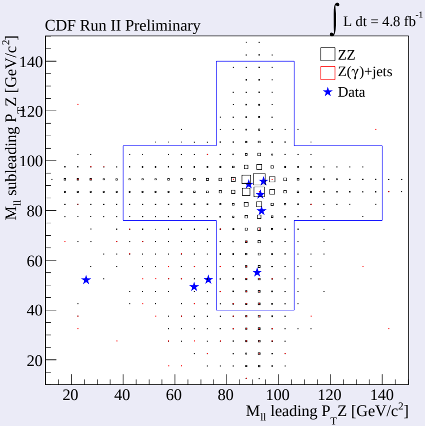
Expected Results

Discovery Probability

Events in $\mathcal{L} = 4.8$ fb^{-1}	
Signal	4.68 ± 0.02(stat.) ± 0.76(syst.)
Z(γ)+jets	0.041 ± 0.016(stat.) ± 0.029(syst.)
Total expected	4.72 ± 0.03(stat.) ± 0.76(syst.)

Probability of Observing a Signal	
Significance	Probability
2σ	0.99
3σ	0.95
5σ	0.70





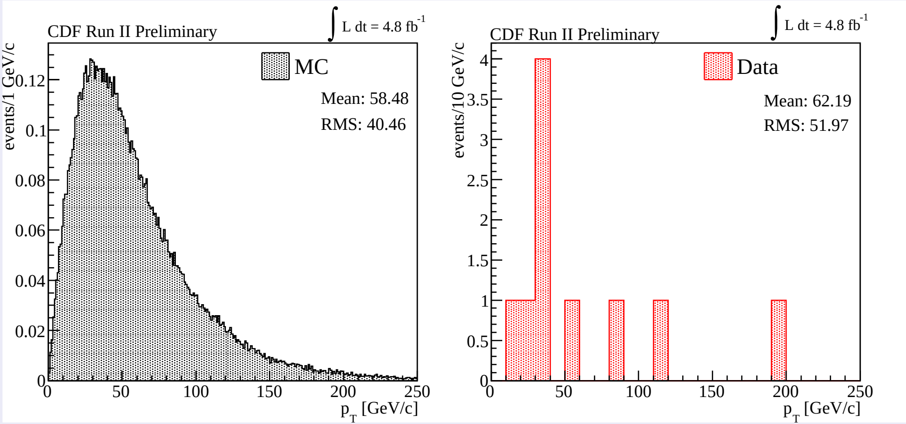
RESULTS:

FIRST CDF OBSERVATION

- 5 events found in data
- p-value $= 1.2 \pm 0.3 \cdot 10^{-8}$
- significance $= 5.7\ \sigma$

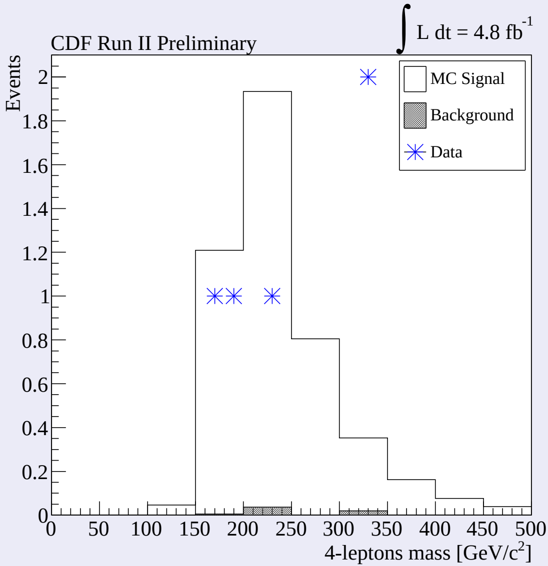
Candidate	leptons	M_{ll-1}	M_{ll-2}	4 lepton invariant mass
1	$trk\mu/\mu\mu$	90.5 GeV/c^2	88.5 GeV/c^2	324.8 GeV/c^2
2	$trk\mu/\mu\mu$	91.6 GeV/c^2	94.2 GeV/c^2	169.4 GeV/c^2
3	$ee/\mu\mu$	93.0 GeV/c^2	86.4 GeV/c^2	191.9 GeV/c^2
4	$ee/\mu\mu$	93.3 GeV/c^2	79.7 GeV/c^2	229.2 GeV/c^2
5	$\mu\mu/\mu\mu$	91.7 GeV/c^2	55.1 GeV/c^2	325.0 GeV/c^2

p_T distribution for the two Zs in MC and Data



Small statistics but no anomalous behaviour.

4 leptons invariant mass for events in data



Cross-section

Cross section calculated with the standard formula:

$$\sigma_{ZZ} = \frac{N_{obs} - N_{bck}}{\mathcal{L} \cdot \epsilon}$$

- $N_{obs} = 5$
- $\epsilon = 0.065\ \%$
- $N_{bck} = 4.15 \pm 1.62(stat) \pm 2.87(syst)\ \cdot 10^{-2}$
- Considered: $\sigma_{\mathcal{L}} \rightarrow 6\%$ $\sigma_{\epsilon_{Trig}} = 2.1\ \%$ $\sigma_{\epsilon_{lepID}} = 3.6\ \%$ $\sigma_{X_{th}} = 10\ \%$

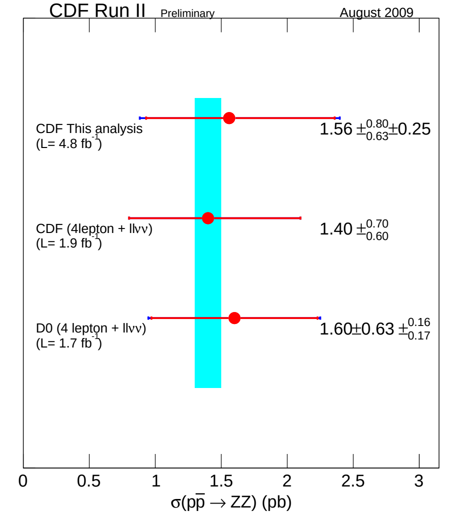
$$\sigma_{ZZ} = 1.56^{+0.80}_{-0.63}(stat) \pm 0.25(syst)\ pb$$

$$(expected\ at\ NLO\ \sigma_{ZZ} =\sim 1.4\ \pm\ 0.1\ pb)$$

previous CDF analysis

$$\sigma_{ZZ} = 1.4^{+0.7}_{-0.6}(stat + syst)\mathrm{pb}$$

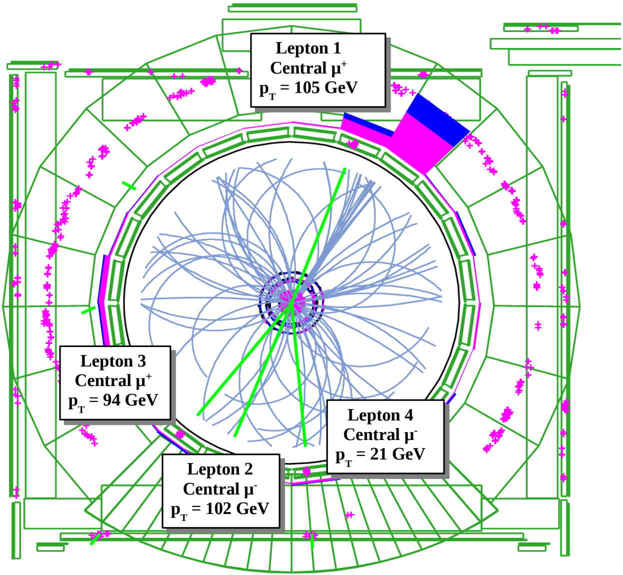
Different cross σ_{ZZ} measurements

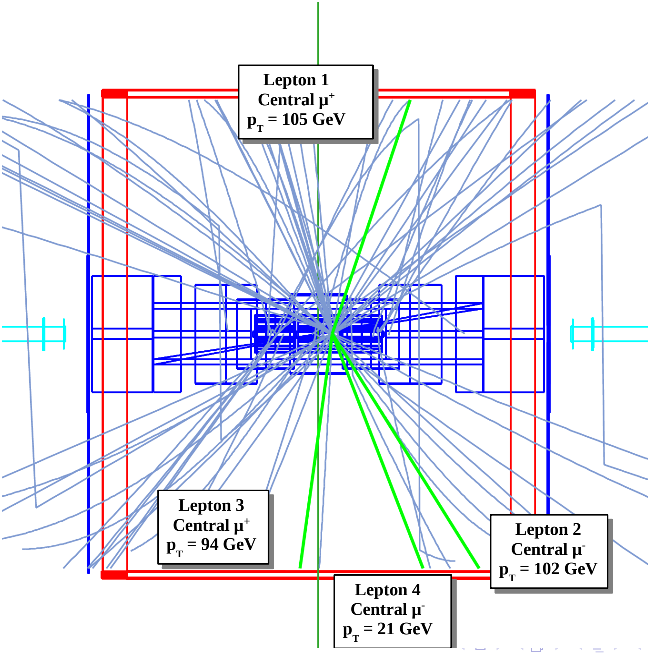


Conclusions

- First CDF observation of ZZ production: 5.7 σ
- Result blessed on 20^{th} of August 2009.
- Cross section measurement:

$$\boxed{\sigma_{ZZ} = 1.56^{+0.80}_{-0.63}(stat) \pm 0.25(syst)\ pb}$$


Lepton 1
Central μ+
pT = 105 GeV
Lepton 3
Central μ+
pT = 94 GeV
Lepton 4
Central μ-
pT = 21 GeV
Lepton 2
Central μ-
pT = 102 GeV


Lepton 1
Central μ+
pT = 105 GeV
Lepton 3
Central μ+
pT = 94 GeV
Lepton 2
Central μ-
pT = 102 GeV
Lepton 4
Central μ-
pT = 21 GeV

backup

MC samples

Signal samples

- ZZ $\rightarrow$ ll+X - Pythia
  - 22 ev/nb
  - NLO cross section 1.511 pb
  - Filter on two leptons with M_{ll}>15 GeV
  - samples: we0s7d, we0sdd, we0sid, we0smd, we0spd, we0scf, we0sif.

Background samples

- WZ $\rightarrow$ ll - Pythia
  - 7 ev/nb
  - NLO cross section 0.365 pb
  - samples: we0s6d, we0scd, we0shd, we0sld, we0sod, we0sbf, we0shf.
- $t\bar{t}$ - Pythia
  - 10 ev/nb
  - NLO cross section 0.68778 pb
  - te0s2z
- Zγ $\rightarrow$ ll - Pythia
  - 5 ev/nb
  - NNLO cross section 10.33 pb - k-factor 1.36
  - Filter on two leptons with M_{ll}>15 GeV
  - samples: re0s33(ee), re0s34($\mu\mu$), re0s37($\tau\tau$)

Trigger Efficiencies

- Calculated directly from $Z \to ll$ data using *Tag & Probe* method for muons with $p_T > 20$ GeV.
- Treat 2D/3D triggers (XFT stereo upgrade) separately.
- Apply livetime corrections when necessary for different DPS.
- CMUP/CMX efficiencies calculated with CMX-CMUP events.

e.g.

$$\epsilon(CMUP, 2D) = \frac{\#(CMX - 2D \;\&\&\; CMUP - 2D)}{\#CMX - 2D}$$

$$\epsilon(CMUP, 3D)[I] = \frac{\#(CMX - 2D \;\&\&\; CMUP - 2D \;\&\&\; CMUP - 3D)}{\#(CMX - 2D \;\&\&\; CMUP - 2D)}$$

$$\epsilon(CMUP, 3D)[II] = \frac{\#(CMX - 3D \;\&\&\; CMUP - 3D)}{\#CMX - 3D}$$

- CMP efficiencies calculated with CMP-CMX events

$$\epsilon(CMP) = \frac{\#(CMX - 2D \;\&\&\; CMP)}{\#CMX - 2D}$$

Run Period	P14-P19	P20	P21	P22	P23	P14-P23
Trigger ϵ						
CMUP 2-D	0.917	0.916	0.902	0.905	0.912	0.913
CMUP 3-D (I)	0.826	0.867	0.851	0.857	0.869	0.842
CMUP 3-D (II)	0.828	0.871	0.853	0.852	0.868	0.843
CMX ARCH 2-D	0.944	0.921	0.954	0.967	0.956	0.948
CMX ARCH 3-D (I)	0.916	0.898	0.919	0.937	0.899	0.916
CMX ARCH 3-D (II)	0.901	0.883	0.896	0.923	0.885	0.899
CMX MS/KS 2-D	0.828	0.802	0.826	0.774	0.863	0.822
CMX MS/KS 3-D (I)	0.810	0.783	0.813	0.752	0.811	0.801
CMX MS/KS 3-D (II)	0.794	0.753	0.787	0.775	0.808	0.787
CMP PHI-GAP	0.769	0.794	0.802	0.797	0.838	0.787
CMP PHI-GAP (II)	0.787	0.848	0.854	0.870	0.900	0.828

LiveTimes	P14-P19	P20	P21	P22	P23	P14-P23
CMUP 2-D	0.876	0.830	0.776	0.800	0.813	0.837
CMUP 3-D	1.000	1.000	1.000	1.000	1.000	1.000
CMX 2-D	0.802	0.750	0.667	0.701	0.711	0.750
CMX 3-D	0.992	1.000	1.000	1.000	1.000	0.996
CMP PG	0.920	0.876	0.813	0.800	0.811	0.871
CMU EG	0.933	0.93	0.83	0.71	0.74	0.870

- Good agreement with Joint Physics results.
- Trigger efficiencies on CMX Ms/Ks did not look as expected. We now derive them by scaling the efficiencies for CMX arches ($\sim$ 10% difference).
- On these efficiencies we take as error $\sim$ 2% added in quadrature to the average spread between different calculation of the efficiency (*e.g. CMUP 2D (I) - CMUP 2D (II)*).
- TCE & PHX trigger efficiencies are the same calculated by Joint Physics group.

Lepton ID

Lepton considered

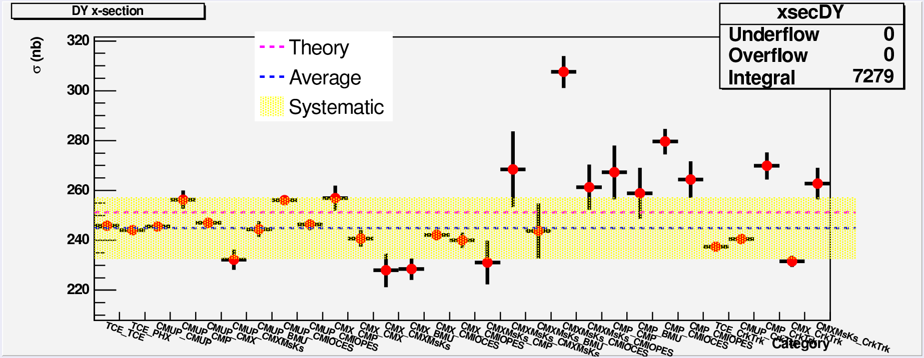
- Electron: TCE, PHX.
- Muon: CMUP, CMP, CMX, MsKs, BMU, CMIOCES, CMIOPES In blue new muon categories .
- Track: CrkTrk
- No Track isolation requirement.
- Most of our lepton definitions are the same of Joint Physics.

Lepton ID efficiencies

- Reconstruction and ID efficiencies calculated in Data and MC - divide the two results to have a scale factor for MC simulations.
- *Tag & Probe* method: Z candidates with 76 GeV $< M_{ll} <$ 106 GeV. see CDFNote 8538
- Generic probe for Central and Forward muons $\rightarrow$ Fiduciality for the various efficiencies.
- Background subtraction from Z sidebands.
- Used the same lepton ID scale factors of the blessed analysis $H \rightarrow WW$.

Drell-Yan cross section

- Checks modeling of selection efficiencies for the different lepton types.
- Measured in data (4.8 fb^{-1}) and $Z \to ee, \mu\mu$ Pythia MC.



- Most of lepton categories are in good ageement with theoretical prediction.

$Z \to \tau\tau$

Theory in $\mathcal{L} = 4.8 fb^{-1}$

$X_{ZZprod} = 1.511$ pb   7252.8 ZZ events

Different BR

- $Z \to e^+e^- = 3.363$ %
- $Z \to \mu^+\mu^- = 3.366$ %
- $Z \to \tau^+\tau^- = 3.370$ %

- $\tau \to e\nu(\gamma) = 19.6$ %
- $\tau \to \mu\nu = 17.36$ %

Actual signal region

BR($ZZ \to S.F.O.S.$) $\simeq 0.517$ % $\to$ 37.50 events (Expected: 4.8 $\to \epsilon \sim 15$ %)

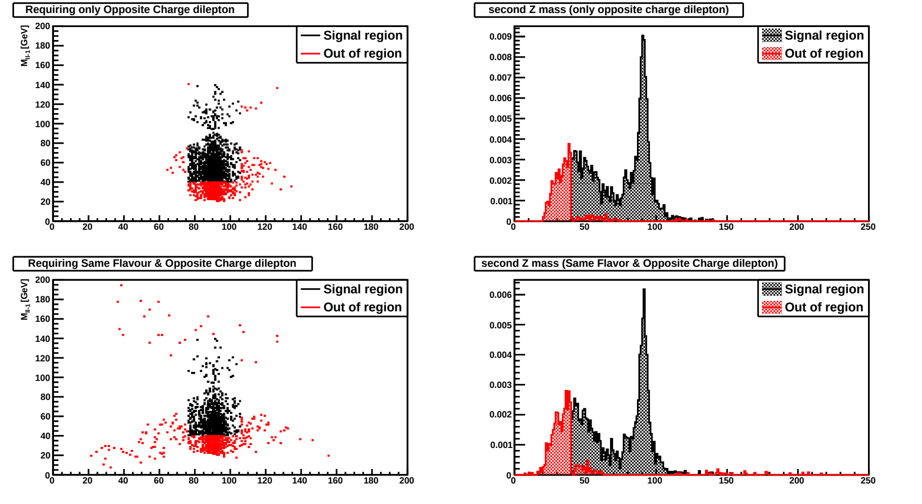
Signal region with different flavour

BR($ZZ \to D.F.O.S.$) $\simeq 0.017$ % $\to$ 1.23 events (4.8 + 0.18) $+ \leq 3.75\%$

- $\leq$ due to M_{ll} cuts: probably we need to take (20 - 140) GeV with lot of background.

ZZ with one $Z \to \tau\tau$

From ZZ MC sample selected events with $\tau^+\tau^-$ generated.



ZZ with one $Z \rightarrow \tau\tau$

Statistics

In $\mathcal{L}$=3.6 fb^{-1} 12.8 100^{ths} of event with $N_\tau \geq$ 2.

- Requiring just Opposite Charge:
  - 8.4 100^{ths} event with 40 GeV $\leq M_{ll-2} \leq$ 140 GeV
  - 12.2 100^{ths} event with 20 GeV $\leq M_{ll-2} \leq$ 140 GeV
- Requiring Same Flavor Opposite Charge (actual signal region):
  - 5.2 100^{ths} event with 40 GeV $\leq M_{ll-2} \leq$ 140 GeV
  - 8.2 100^{ths} event with 20 GeV $\leq M_{ll-2} \leq$ 140 GeV

Fake Rates

Fakeable Electron
$E_T > 10 GeV$
$E_{Had}/E_{EM} < 0.125 + 0.00045 \cdot E$
Isolation < 0.3
Has a good quality track
Is not a conversion
Fiduciality to central or plug

Fakeable Muon
$p_T > 10 GeV$
Axial SL ≥ 2 with ≥ 5 hits
Stereo SL ≥ 2 with ≥ 5 hits
Track $\lvert z_0 \rvert < 60$ cm
Track $\lvert d_0 \rvert < 0.2$ cm (< 0.02 cm with silicon)
$E/P < 1$
Isolation < 0.3
$\chi^2/dof < 4.0$ (< 3.0 if Run > 186598)
Fiduciality to different muon chambers

Data Driven

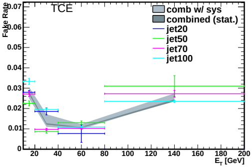
Fake Rates

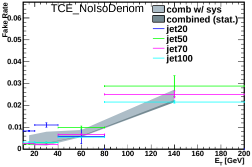
Calculate *Fake Rates* on a sample enriched in jets subtracting real EWK contribution:

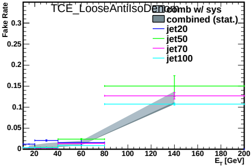
- Denominator: candidate lepton (almost a good track)
- Numerator: Fully identified lepton

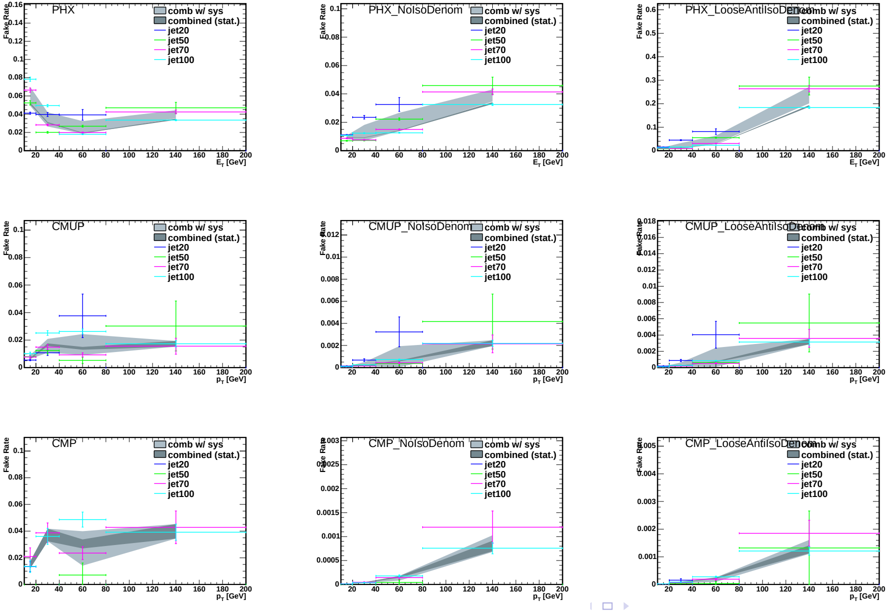
$$f_i(p_T) = \frac{N_i(identified\ leptons) - \sum_{j \in EWK} N_{ij}(Identified\ leptons)}{N_i(denominator\ objects) - \sum_{j \in EWK} N_{ij}(Denominator\ objects)}$$

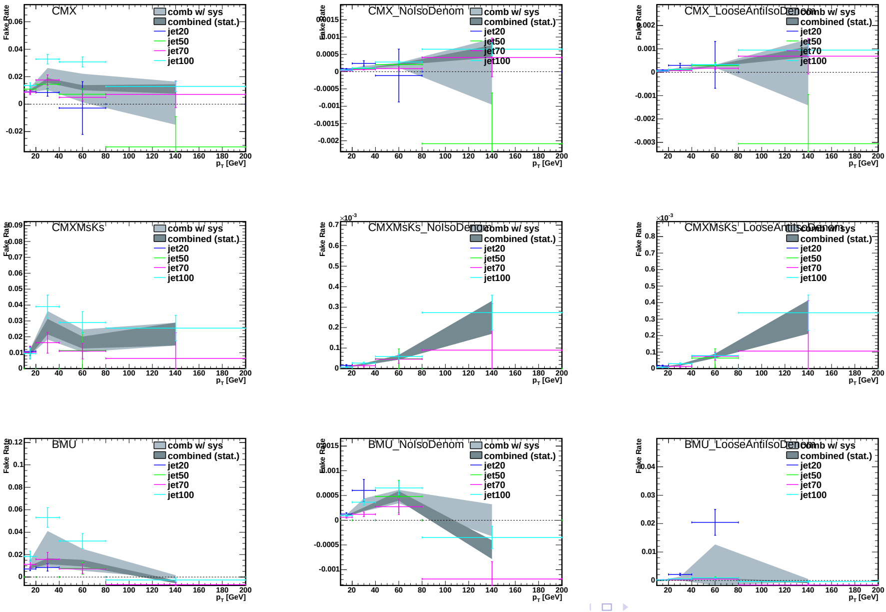
- Fake Rates are calculated as function of p_T for each lepton type.
- Averaging measurements in JET20, JET50, JET70, JET100.
- Take the maximum spread as systematic uncertainty.

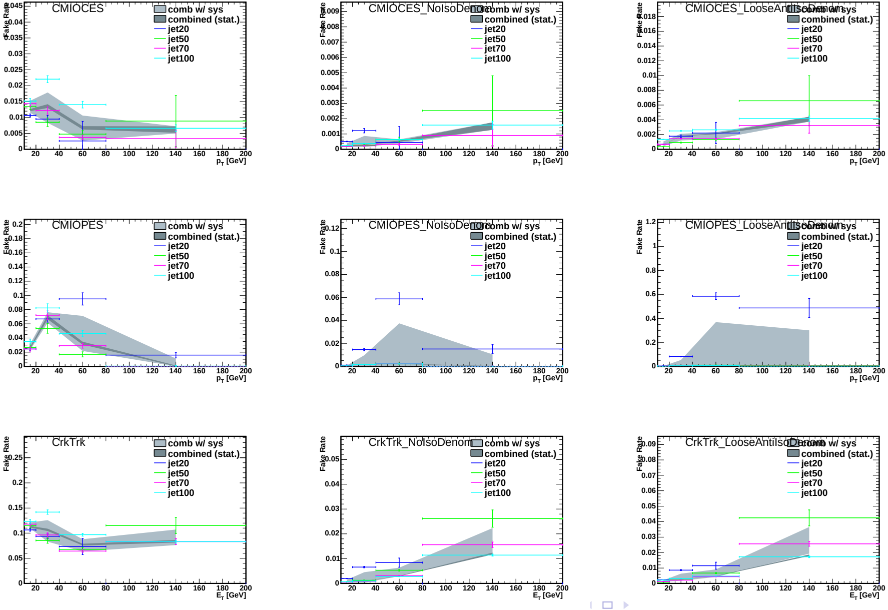







PHX
PHX_NoIsoDenom
PHX_LooseAntiIsoDenom
CMUP
CMUP_NoIsoDenom
CMUP_LooseAntiIsoDenom
CMP
CMP_NoIsoDenom
CMP_LooseAntiIsoDenom
comb w/ sys
combined (stat.)
jet20
jet50
jet70
jet100
Fake Rate
E_T [GeV]
p_T [GeV]


CMX
CMX_NoIsoDenom
CMX_LooseAntiIsoDenom
CMXMsKs
CMXMsKs_NoIsoDenom
CMXMsKs_LooseAntiIsoDenom
BMU
BMU_NoIsoDenom
BMU_LooseAntiIsoDenom
comb w/ sys
combined (stat.)
jet20
jet50
jet70
jet100
Fake Rate
p_T [GeV]
$\times 10^{-3}$


CMIOCES
CMIOCES_NoIsoDenom
CMIOCES_LooseAntiIsoDenom
CMIOPES
CMIOPES_NoIsoDenom
CMIOPES_LooseAntiIsoDenom
CrkTrk
CrkTrk_NoIsoDenom
CrkTrk_LooseAntiIsoDenom
comb w/ sys
combined (stat.)
jet20
jet50
jet70
jet100
Fake Rate
p_T [GeV]
E_T [GeV]

Background estimation

- We want to measure the number of events with:
  - 3 leptons + 1 fake
  - 2 leptons + 2 fakes
- Evaluate N_{3l+1f} and N_{2l+2f} from events with 3l+1d or 2l+2d - avoiding overlap.
- Weight those events with the fake rates of each denominator object present.
- For the 2 leptons + 2 fakes contribution:

$$N_{Z+2fakes} = \sum_{d_1,d_2} f_{d1}(p_T^{d1}) \cdot f_{d2}(p_T^{d2}) = f^2 \cdot N_{2l+2d}$$

- While for 3 leptons + 1 fake there are two contributions:

$$N_{3l+1d} = N_{Z\gamma+d} \; + \; N_{2l+1f+1d} \; = \; N_{Z\gamma+d} \; + \; 2 \cdot f \cdot N_{2l+2d}$$

- Finally the total background, due to both 3l+1f and 2l+2f is

$$N_{Z\gamma+1fake} \; + \; N_{Z+2fake} \; = \; f \cdot N_{3l+1d} - f^2 \cdot N_{2l+2d}$$

- For each event considered all the possible combinations of denominators.

A matter of sign

3l+1d and 2l+2d

Writing in short notation:

$$N_{Z+2fakes} = \sum_{d_1,d_2} f_{d1}(p_T^{d1}) \cdot f_{d2}(p_T^{d2}) = f^2 \cdot N_{2l+2d}$$

The number of events with only one denominator is due to 2 contribution:

$$N_{3l+1d} = N_{Z\gamma+d} \; + \; N_{2l+1f+1d} \; = \; N_{Z\gamma+d} \; + \; 2 \cdot f \cdot N_{2l+2d}$$

$$N_{3l+f} = f \cdot N_{3l+d} = f \cdot N_{Z\gamma+d} \; + \; 2 \cdot f^2 \cdot N_{2l+2d}$$

So the total background, due both to $3l + 1f$ and $2l + 2f$ is

$$\begin{aligned} N_{Z\gamma+1fake} \; + \; N_{Z+2fake} &= f \cdot N_{Z\gamma+d} + f^2 \cdot N_{Z+2d} \\ &= f \cdot N_{Z\gamma+d} + 2f^2 \cdot N_{2l+2d} - f^2 \cdot N_{2l+2d} \\ &= f \cdot N_{3l+1d} - f^2 \cdot N_{2l+2d} \end{aligned}$$

Events with 3l+1d and 2l+2d are weighted with opposite sign.

Background overstimation

In Data Driven Background calculation, for each denom found, all candidates are duplicated and to the new ones the denominator is added and so on. In this process are neglected all the f^2 terms in event weight.In the case of an event with 2d we have

- 4l(+2d) 1 instead of $1-(f_1+f_2-f_1f_2)$
- 3l+$d_1(+d_2)$ f_1 instead of $f_1(1-f_2)$
- 3l+$d_2(+d_1)$ f_2 instead of $f_2(1-f_1)$
- 2l+d_1+d_2 f_1f_2.

The former are correct if we consider the probability to have more than 4 leptons while the latter should be used in the case of just 4 leptons.

There should be no problems in such neglection because of the smallness of fake rates only if we have no more than 2 denominators per candidate.

In some case we really have more than 2 denominator candidate so in this way we are overstimating background.

Heavy flavour contribution

In JET sample - where we calculate *Fake Rates*

- $\boxed{\frac{\sigma(b\bar{b}-jets)}{\sigma(q\bar{q}-jets)} \sim \mathbf{10^{-2}}}$ $\qquad (p_T^j \sim 50\ GeV)$

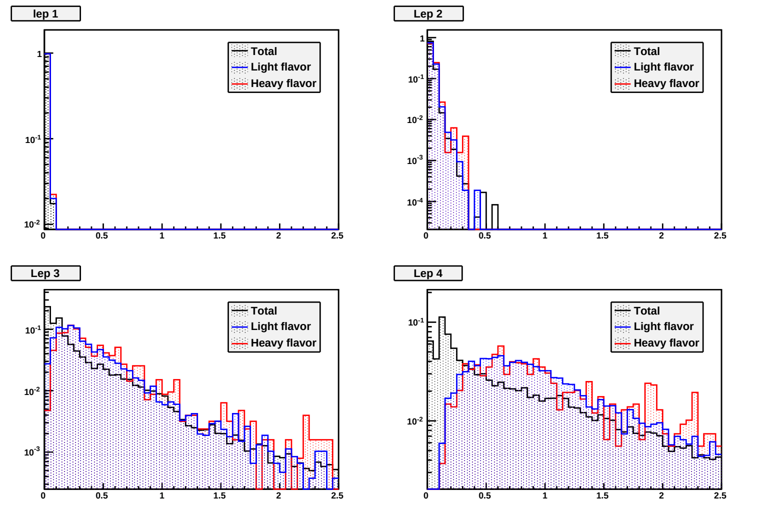
In our Data Sample: 2l+2d candidates

- Required at least 2 real leptons in [76,106] GeV dilepton mass range.
- Contribution from **Z + jets**: $\frac{Z+b-jet(s)}{Z+jet(s)} \sim 5 \cdot 10^{-3} \rightarrow$ Negligible
- Main contribution to heavy flavor from $ZZ \rightarrow l^+l^- + b\bar{b}$
  - **From MC:** 2.97% of the total ZZ contamination in 2l+2d sample is from heavy flavor.
- **2l+2d sample** Data Driven:
  - **106 candidates** corresponding to 0.082 100^{ths} of event
  - 0.014 100^{ths} ZZ contamination $\rightarrow$ 17.4 % of total 2l+2d sample
  - $\boxed{0.174_{ZZ-cont} \times 0.03_{b\bar{b}} \sim \mathbf{5 \cdot 10^{-3}}}$ $\rightarrow$ **less than 1 candidate** over 106 is from heavy flavor.

Heavy flavor contribution

MC sample:

selected $ZZ \rightarrow q\bar{q}l^+l^-$ where $q\bar{q}$ are the 2 denominators   note: DrawNormalized



Background modeling

MC driven

Evaluate the contribution of the following physics process looking for signal (4-leptons) in MC

- WZ: 0.335 100^{ths} of events.
- $t\bar{t}$: 0.023 100^{ths} only 2l+2d.
- $Z\gamma = 0.27 \pm 0.05$ 100^{ths}
  $NoIso = 0.17 \pm 0.01$ 100^{ths}.
- **Total: 0.57** $\pm$ **0.05** 100^{ths}.

- Number of background events from MC simulation seems to be lower than what measured using Data driven methods, but they agree within the errors.
- To compare these results to that obtained with Data Driven method we must remember that in data there is contamination of real ZZ events.

Significance

- Calculate the probability that the background fluctuate to give a number of events greater or equal to the expected ones.
- With pseudo-experiments:
  - Start from poissonian distribution n_{cand}^{3l+1d} and n_{cand}^{2l+2d}.
  - Weight those distribution with average fake rates.
  - Gaussian smearing of the probability distribution for systematic

$$b \longrightarrow b \cdot \prod_{i=1}^{\#sys}(1 + s_i \cdot x_i)$$

- Evaluate the probability to have $n_{obs} \geq n_{exp}$.