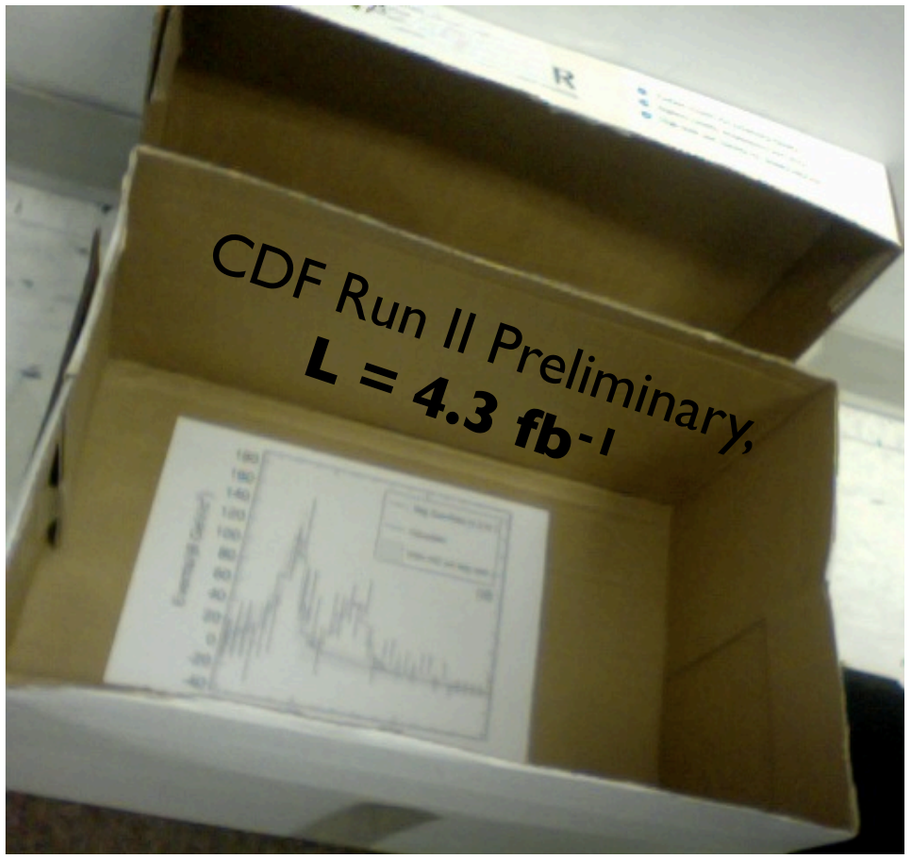
Update on dijet mass spectrum in W + 2jets events



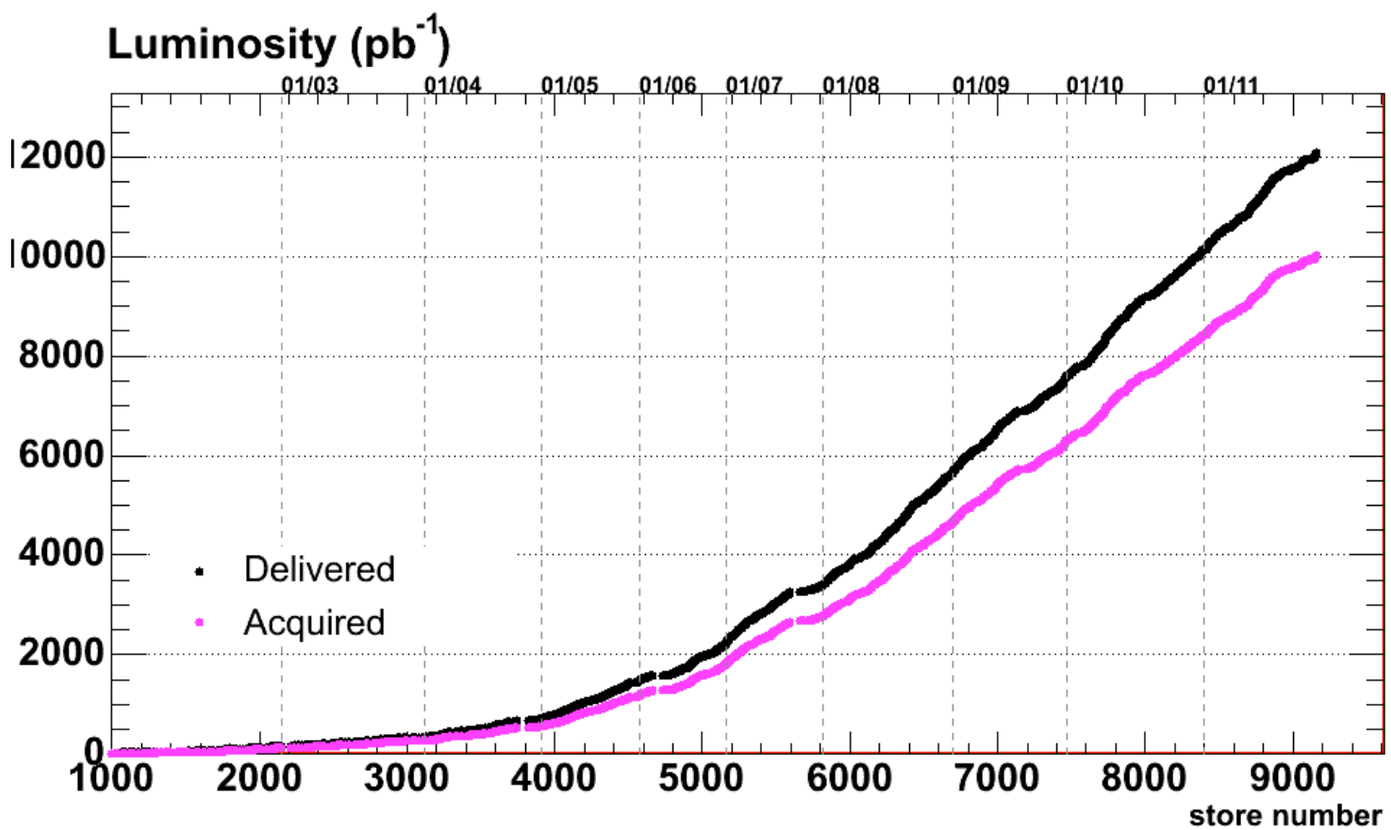
M. Trovato*
on behalf of the CDF collaboration

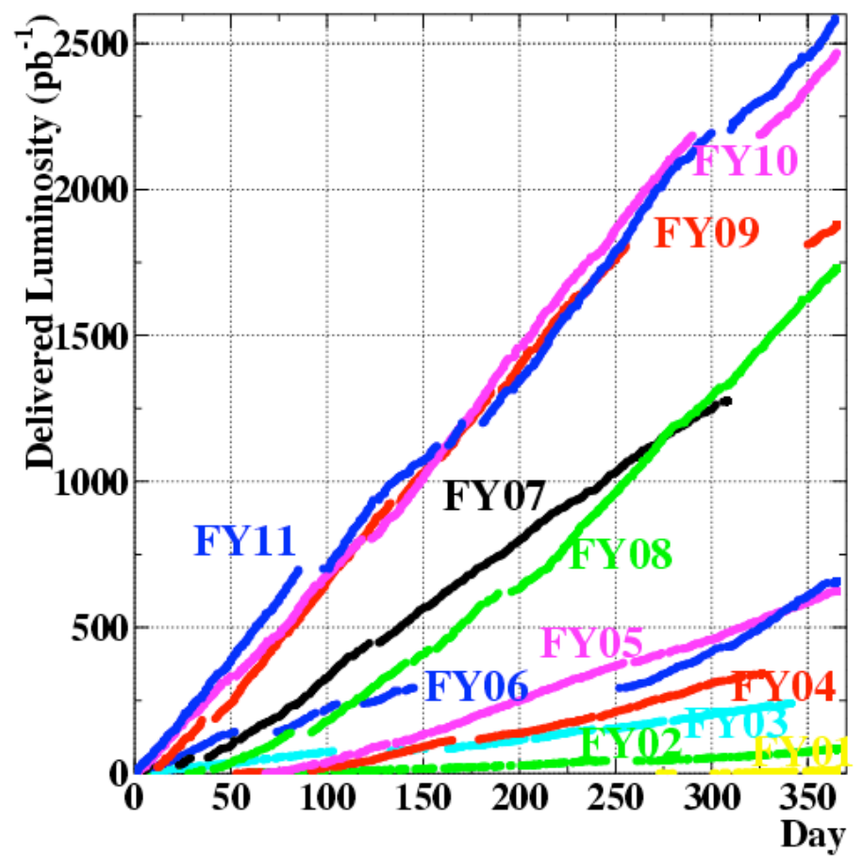
***: Scuola Normale Superiore - Pisa & Fermilab**

Tevatron Performance

- Integrated luminosity:
  - (Tevatron) delivered ~12 fb^{-1}
  - (CDF) acquired ~10 fb^{-1}

End of operations: September 2011





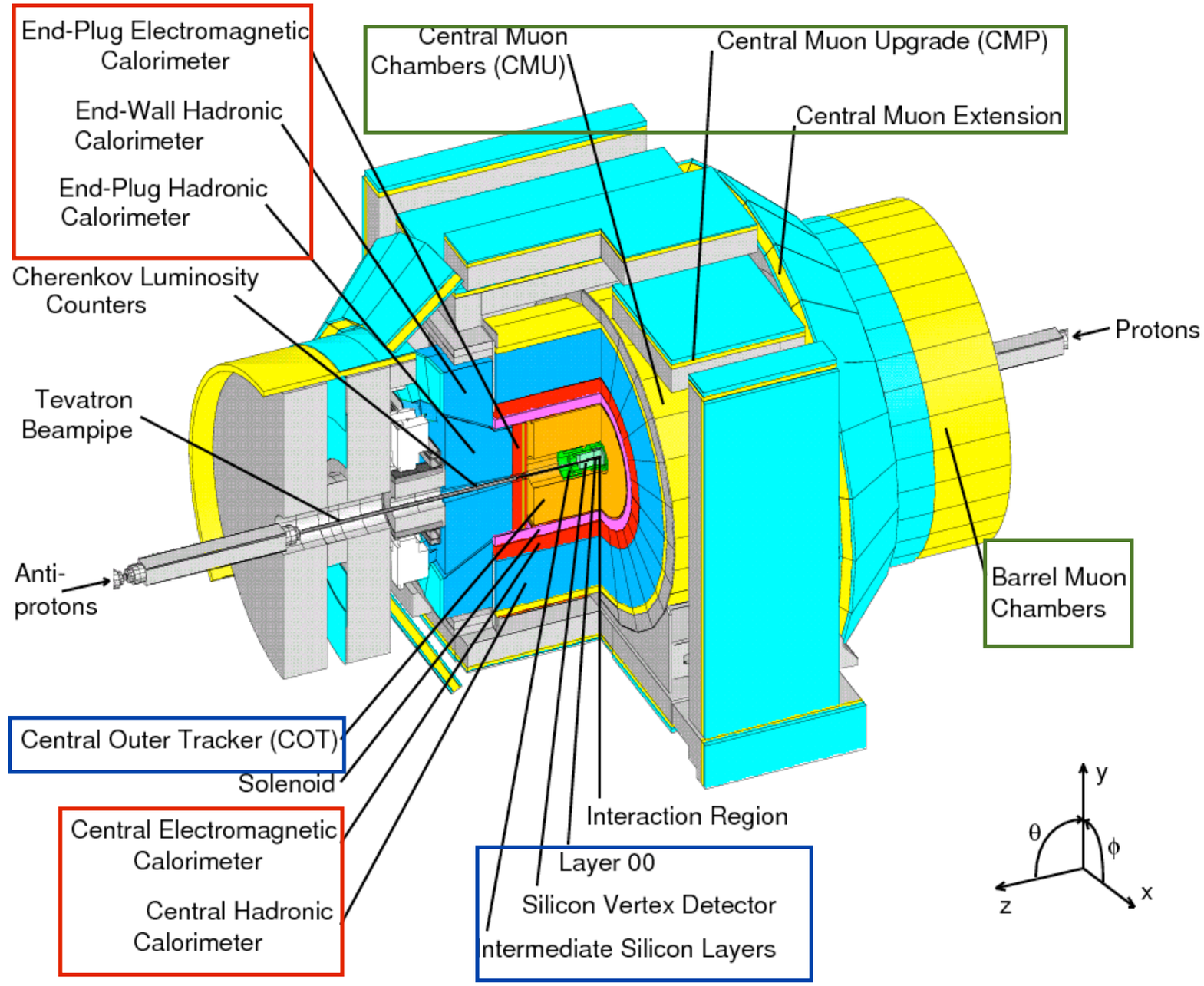
Thanks to the Accelerator Division for the amazing performance over the past ~25 years

CDF II Detector

- Tracking system in 1.4 T magnetic field
  - Silicon detector
  - Central drift chamber

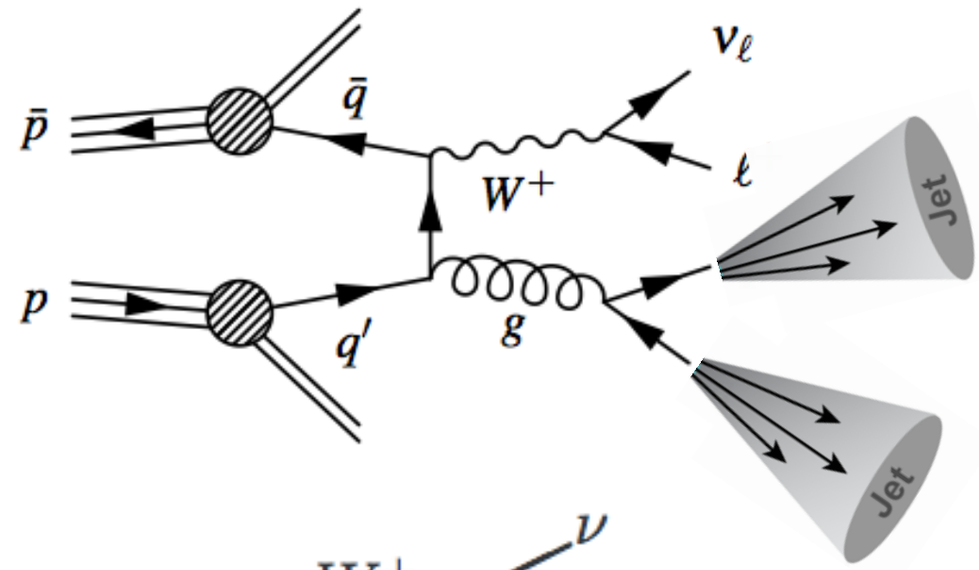
- Electromagnetic and hadronic calorimeters up to $|\eta|<3.6$

- Muon chambers up to $|\eta|<1.5$

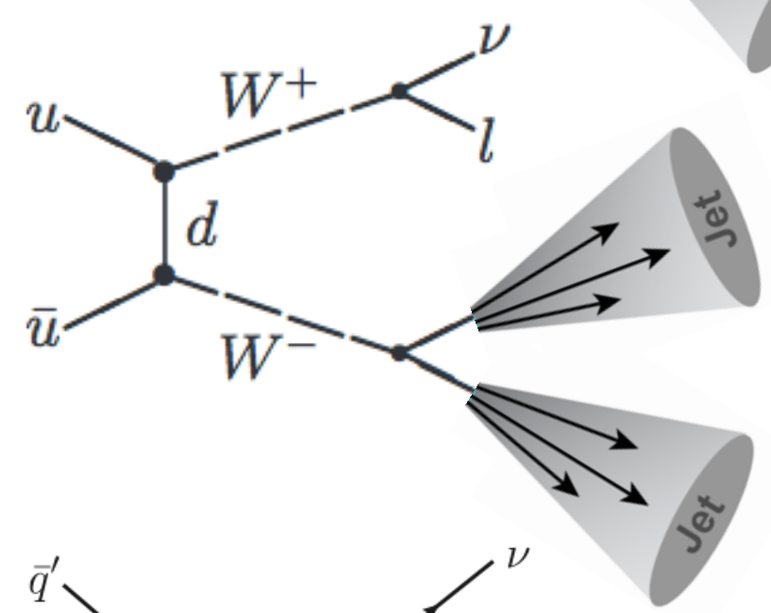


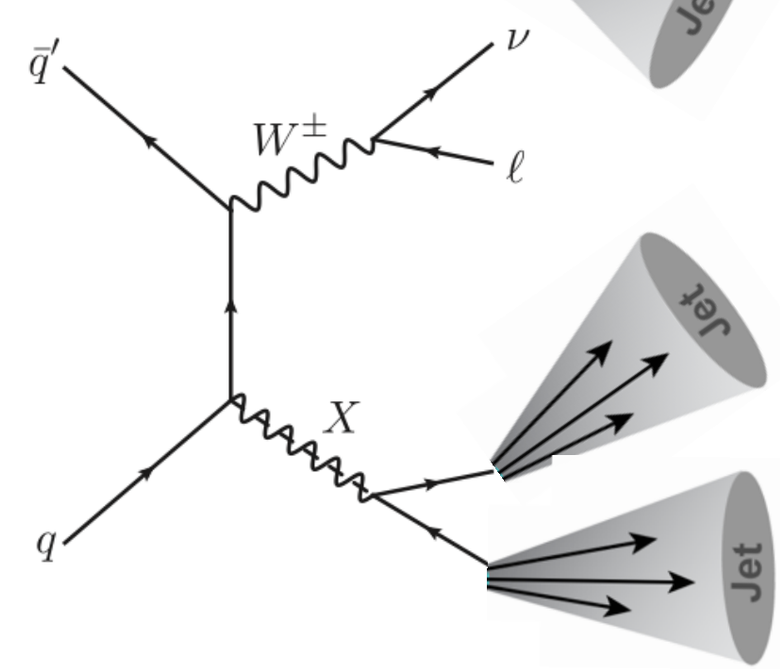
Motivation

- Measurement of associated production of W bosons and jets is a <u>fundamental test of the standard model (SM)</u>



- W+jets production is:
  - ❖ a dominant background of important processes being studied at the Tevatron
    - Diboson, single-top, Higgs, etc.
  - ❖ a background in searches for new physics beyond SM ("X")

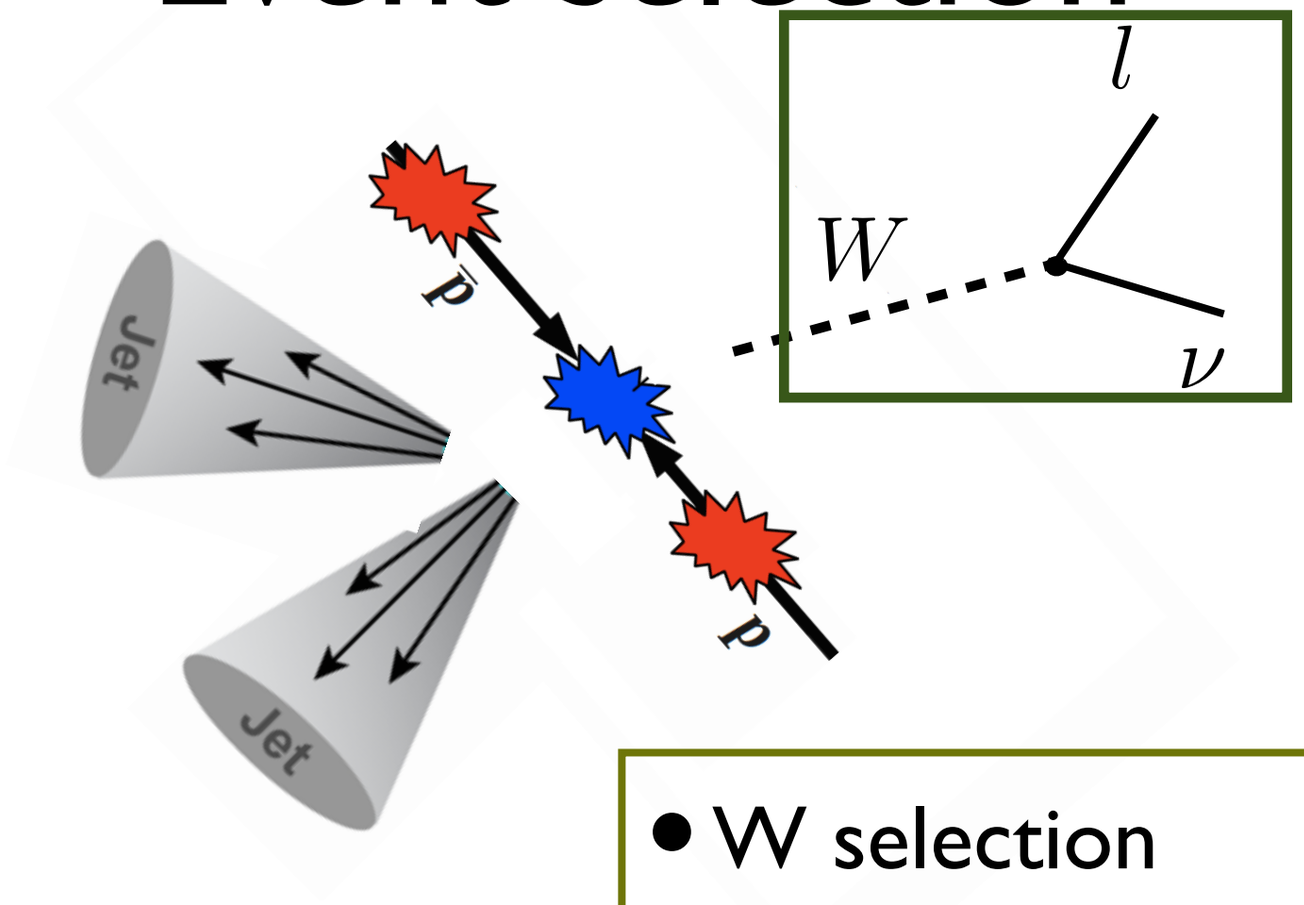




Outline & Documentation

- Analysis outline
- Previous CDF results and results from other experiments
- Update of previous results using the full CDF dataset
- Discussion of improved analyses techniques and results
- Studies in orthogonal samples
- ***CDF Public Note 10973***
  - ***A more detailed description will be published shortly***

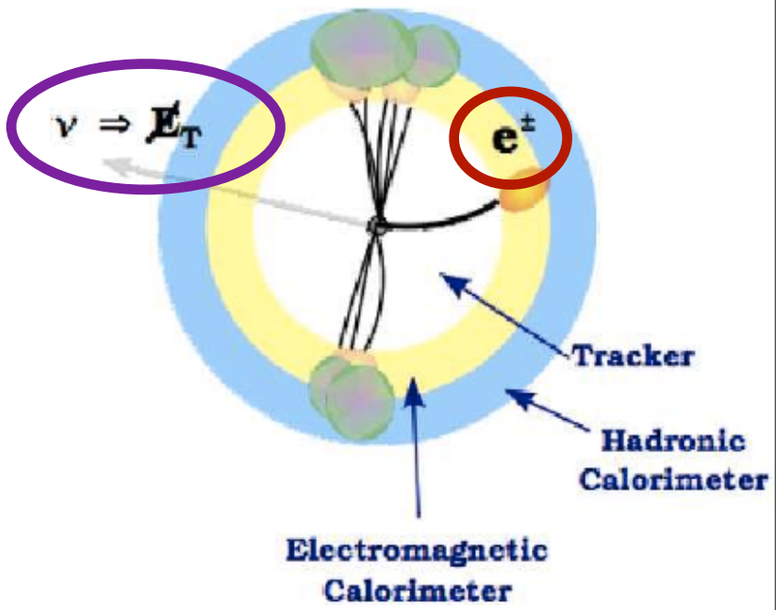
Event selection

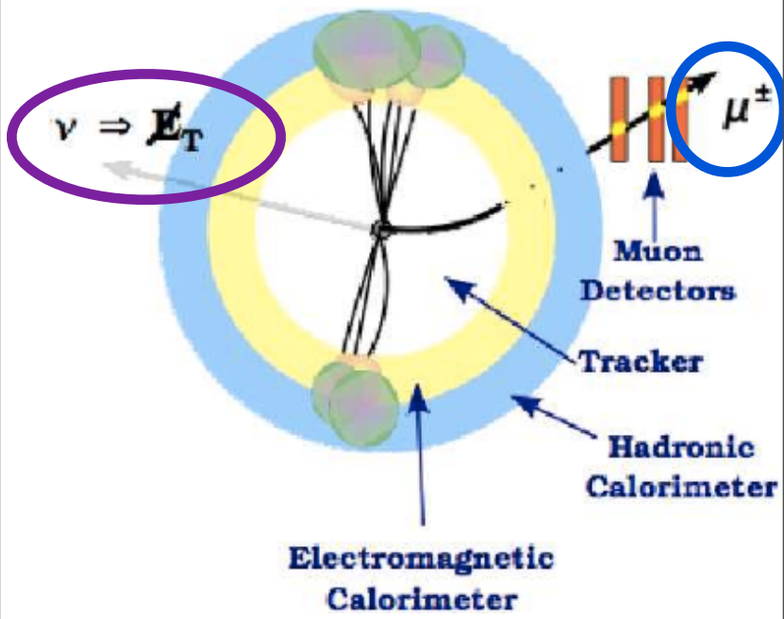


- W selection
  - lepton (electron/muon) P_T>20 GeV, |η|<1
  - $\not{E}_T$ >25 GeV
  - M_T^W > 30 GeV

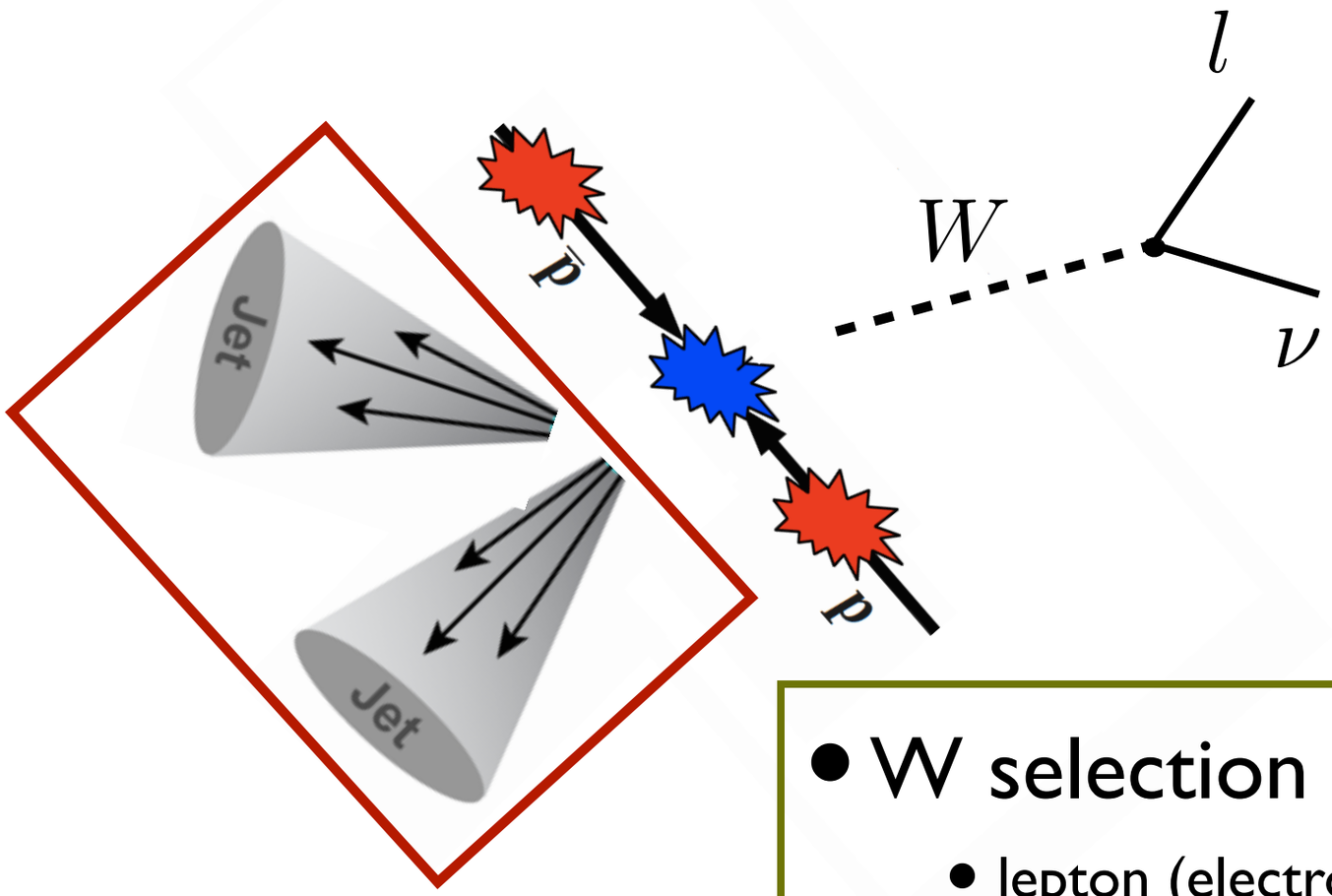
Object identifications - leptons, ν

- Want to efficiently collect Ws:
  - Trigger on charged lepton
  - Offline lepton identification:
    - Electrons:
      - Track matched to an isolated calorimetric deposit of one or two towers
        - Mostly electromagnetic (>90%)
    - Muons:
      - Isolated track matched to a small calorimetric deposit and hits in the muon chambers.
    - Neutrinos: large $\not{E}_T$
      - $\not{E}_T$ corrected for calorimeter inefficiencies and the presence of muons





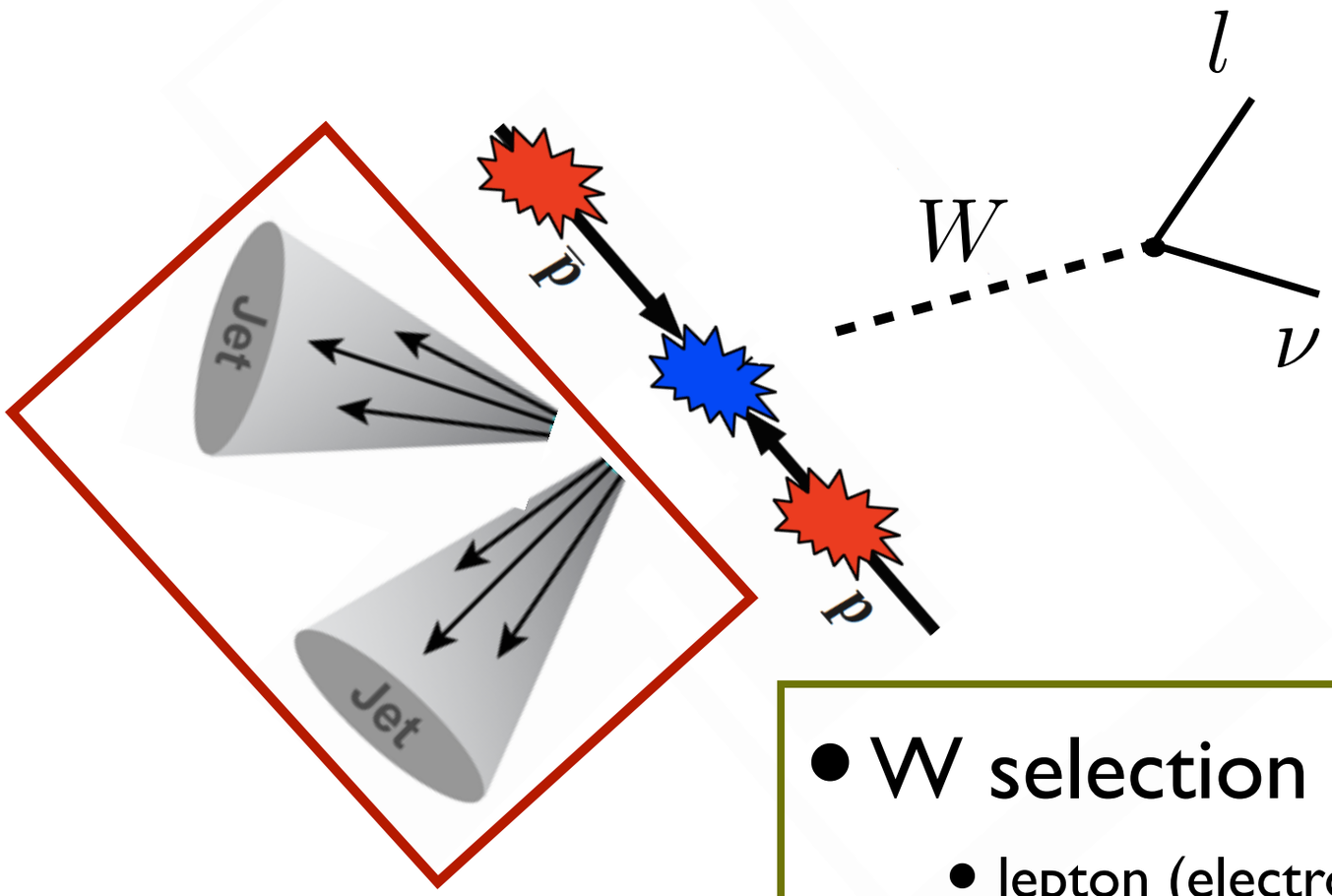
Event selection



- W selection
  - lepton (electron/muon) P_T>20 GeV, |η|<1
  - $\not{E}_T$ >25 GeV
  - M_T^W > 30 GeV

- Dijet Selection
  - Two jets ("jet1","jet2") each with:
    - E_T>30 GeV, |η|<2.4
    - No other jets passing these criteria

Event selection



- W selection
  - lepton (electron/muon) P_T>20 GeV, |η|<1
  - $\not{E}_T$ >25 GeV
  - M_T^W > 30 GeV

- Dijet Selection
  - Two jets ("jet1","jet2") each with:
    - E_T>30 GeV, |η|<2.4
    - No other jets passing these criteria

- Additional cuts
  - P_T(jet1,jet2)>40 GeV,
  - |Δη(jet1,jet2)|<2.5
  - Δφ($\not{E}_T$,jet1)>0.4

Modeling of the data sample

- Attempt to model the data in terms of known contributions

Modeling of the data sample

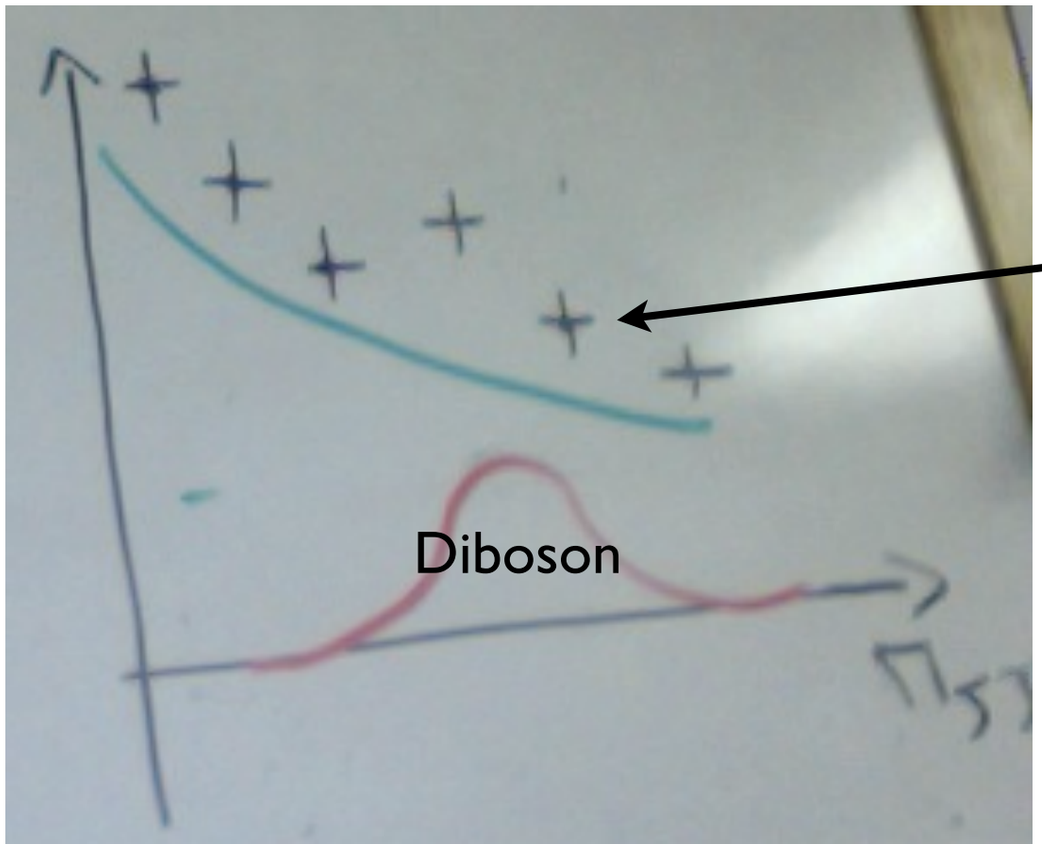
- Attempt to model the data in terms of known contributions

- *But what if those are not enough? Found new physics?*

Modeling of the data sample

How do we model the data exactly?

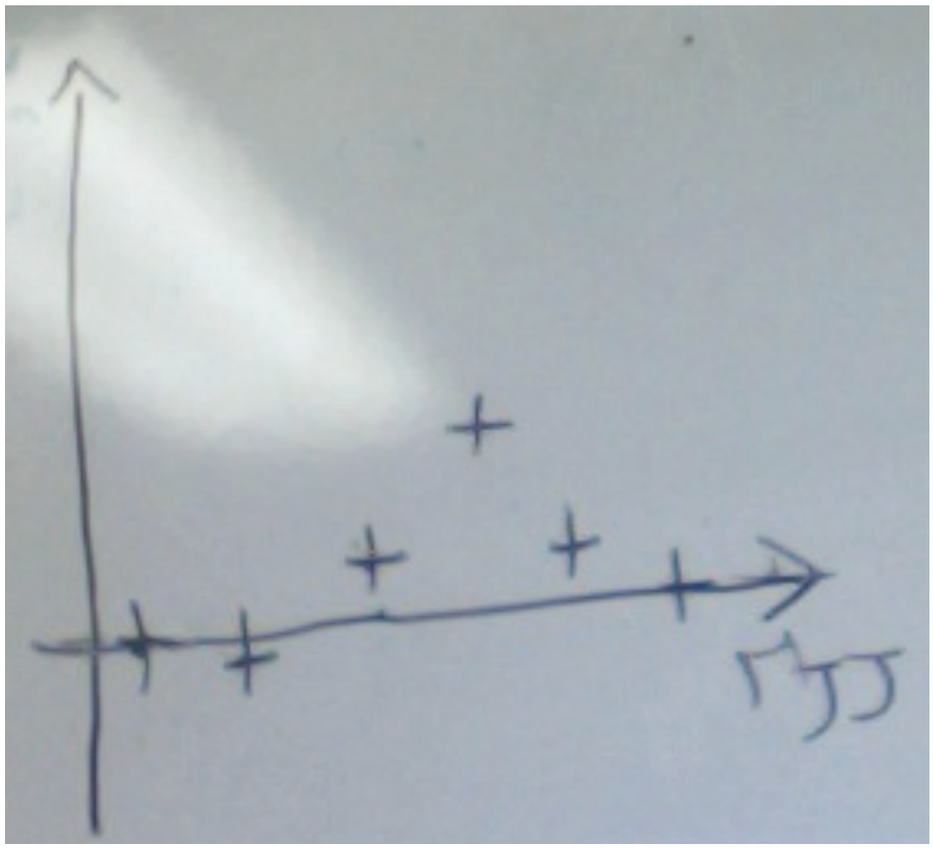
Fit in dijet-mass ("Mjj")



After the fit

Data-SM

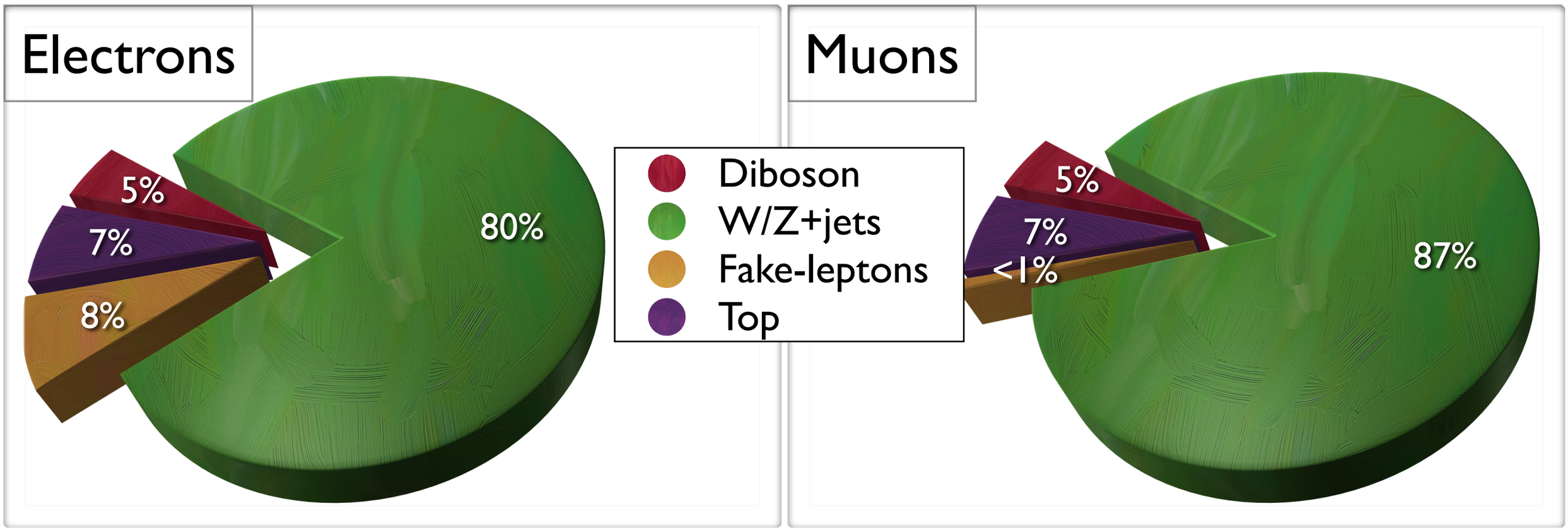
(excluding Diboson)



Modeling of the data sample

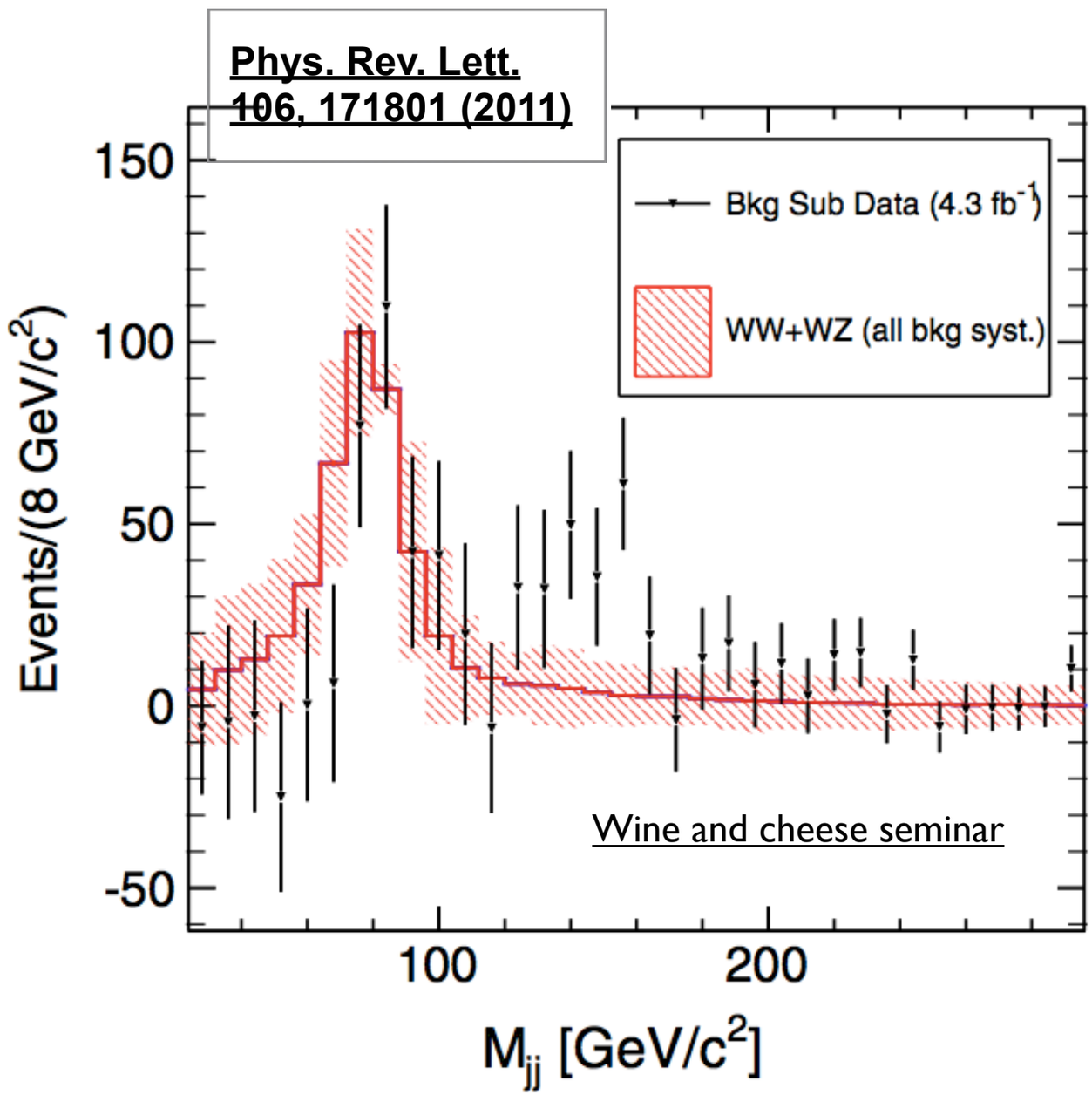
Process	Shapes	Rate constraint
$WW/WZ/ZZ$ inclusive	PYTHIA	Theory
$t\bar{t}$	PYTHIA	Measured cross-section
single-top	MadEvent+PYTHIA	Theory
$Z \to e, \mu, \tau$+jets	ALPGEN+PYTHIA	None
W+jets	ALPGEN+PYTHIA	None
Fake leptons from multi-jets	from data	from data

Composition of the data sample



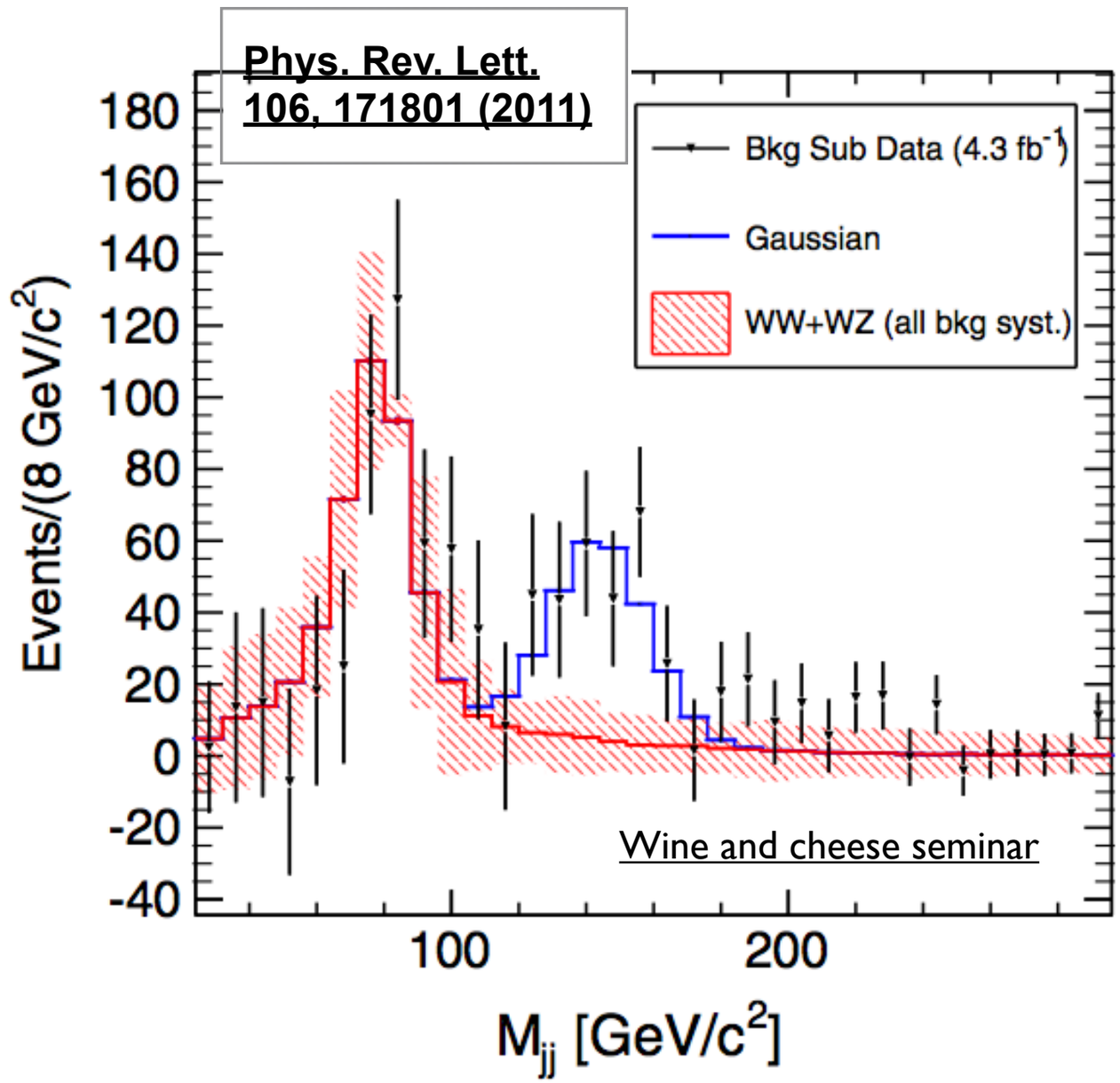
Previous results

- CDF has reported an excess of events around a dijet mass of 145 GeV using a 4.3/fb dataset. Such an excess "is not described by current theoretical predictions within the statistical and systematic uncertainties"



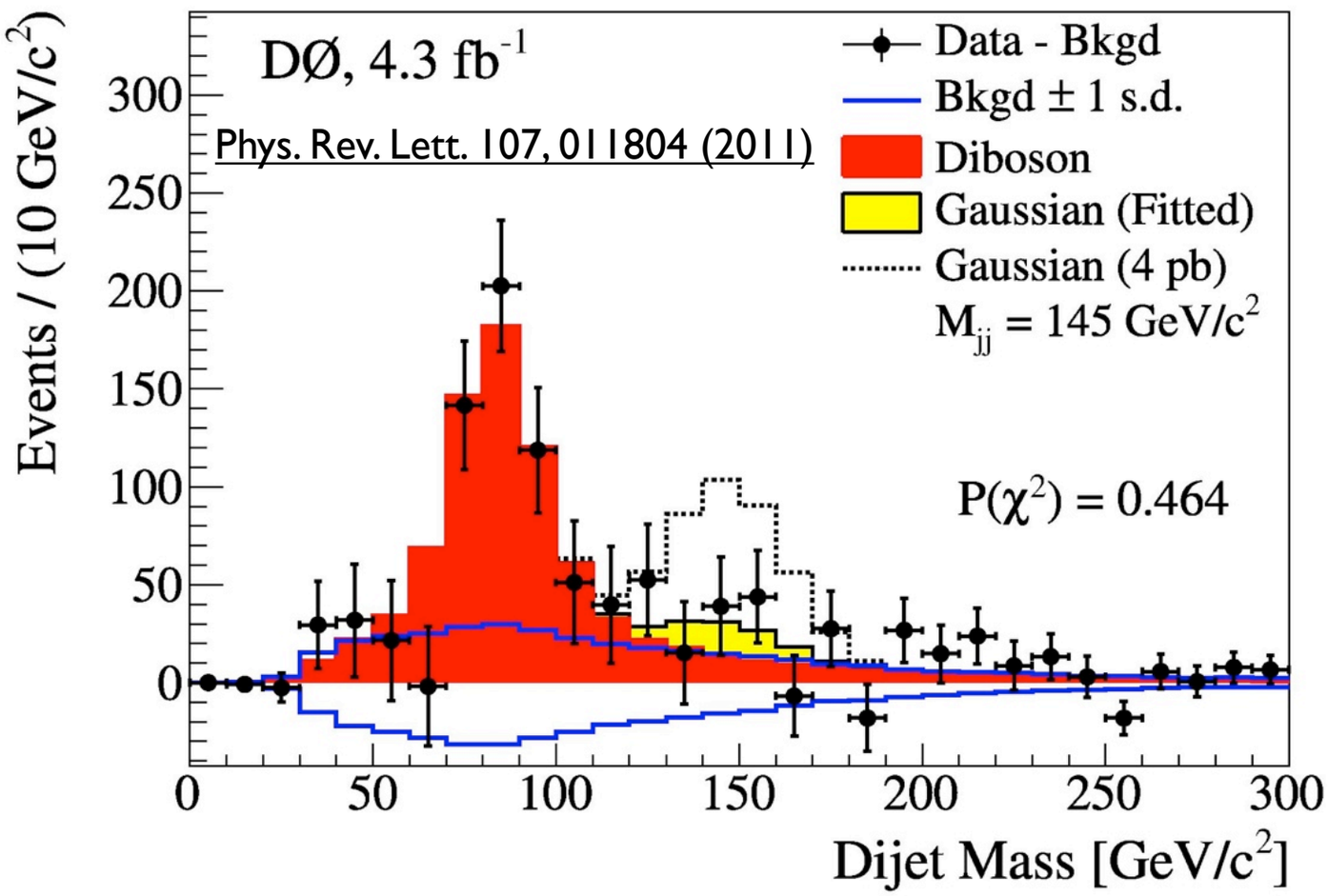
Previous results

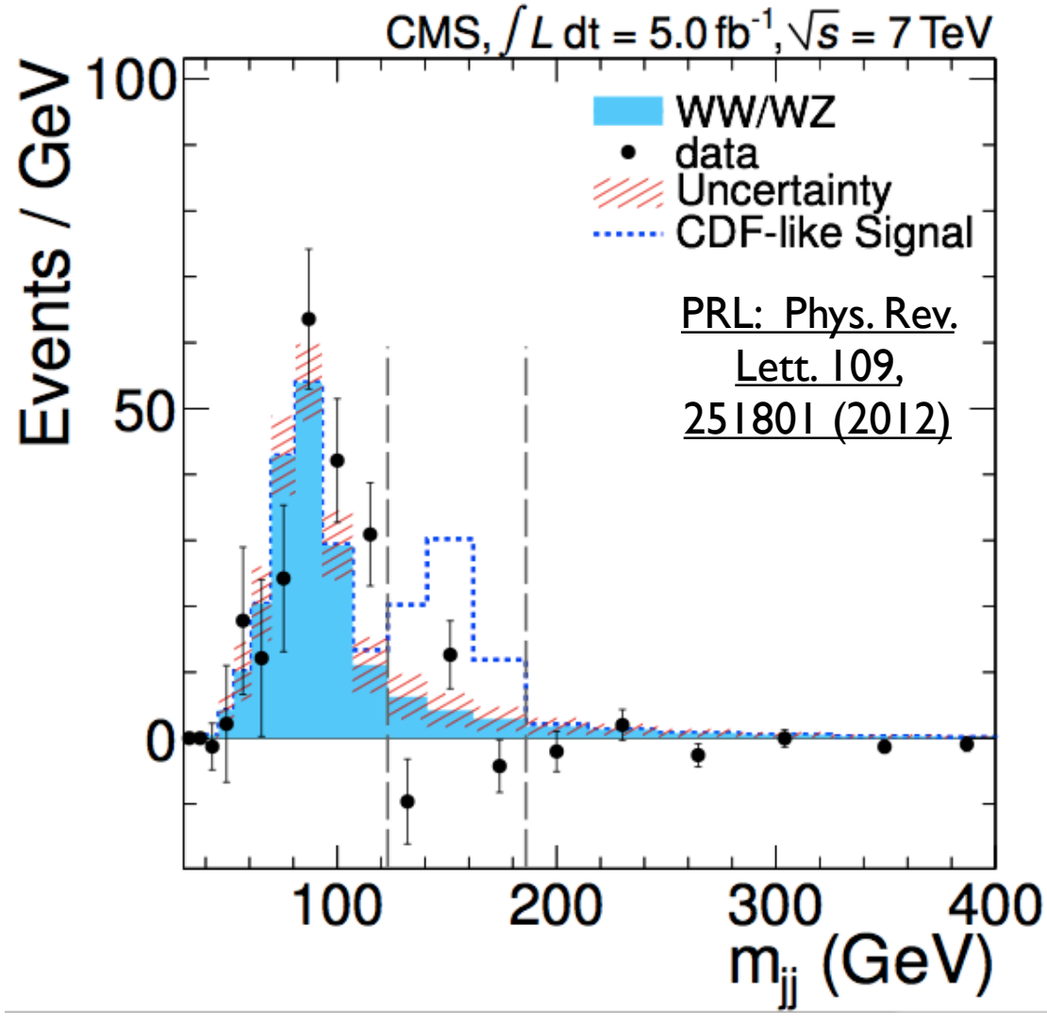
- "One possible way to interpret this disagreement is as an excess in the 120–160 GeV/c^2 mass range" with the significance of 3.2 σ



State of the art from other experiments

- Both D0 and CMS performed similar analyses

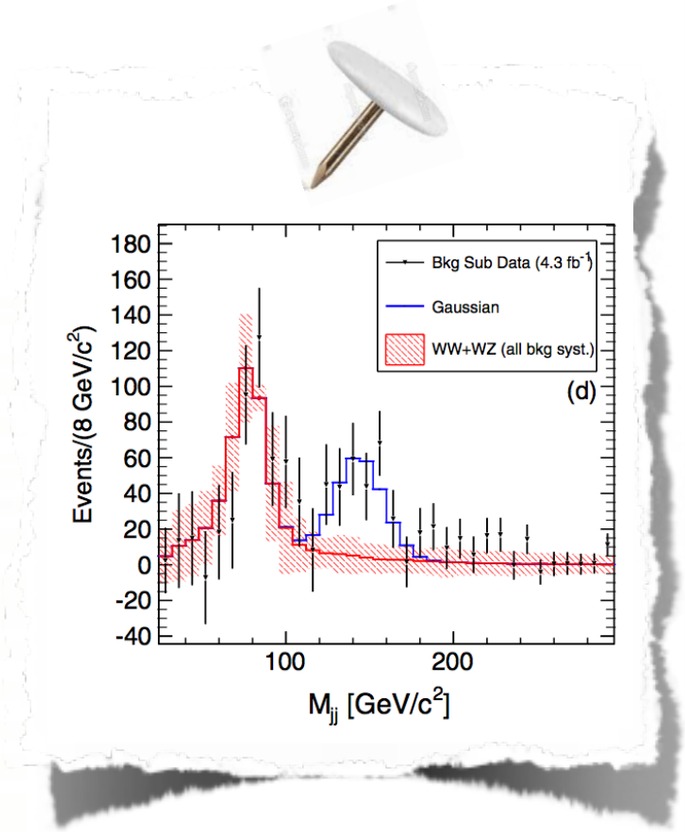




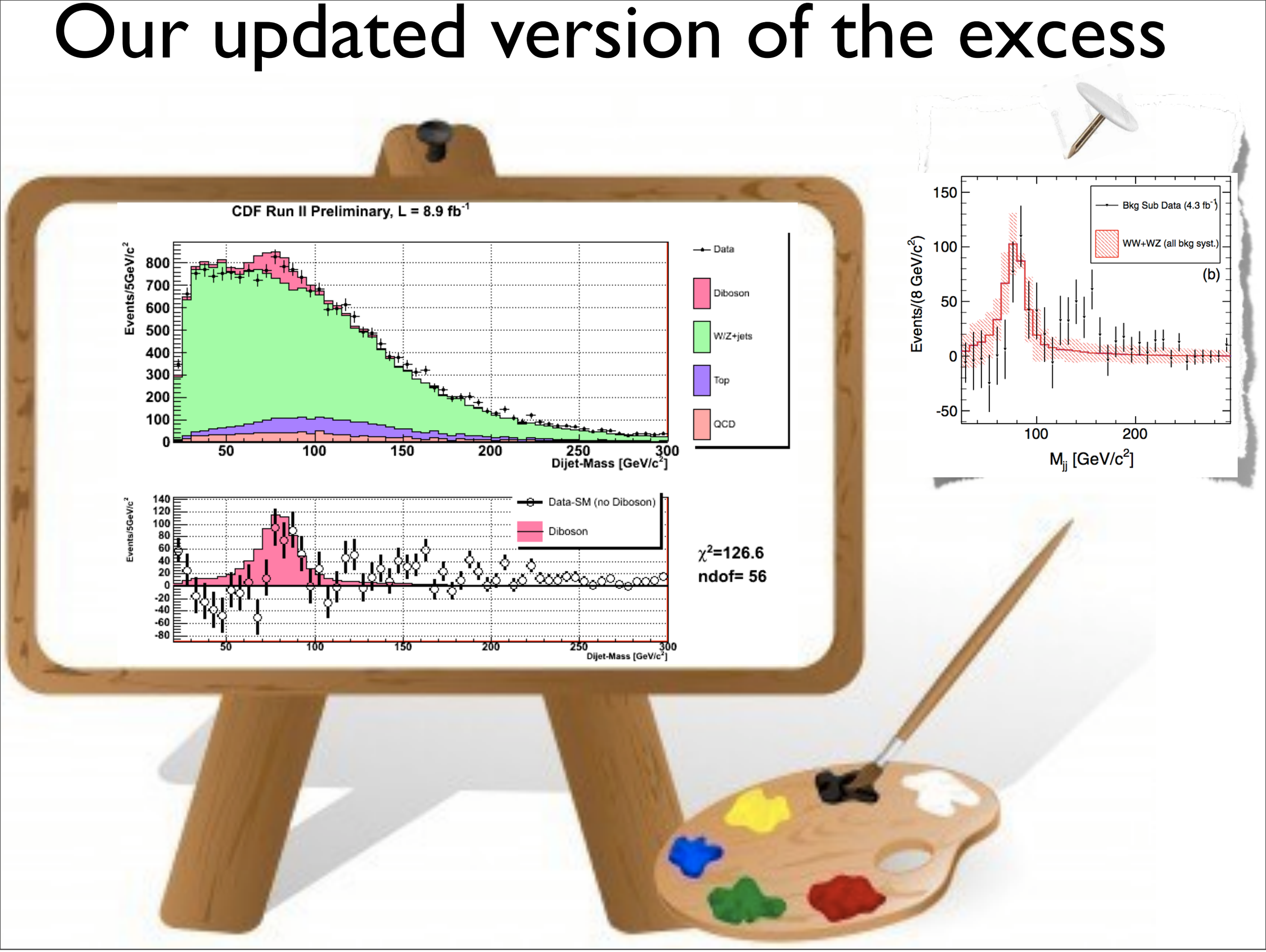
- Also, preliminary analysis from ATLAS (ATLAS-CONF-2011-097)

Similar effects are not observed

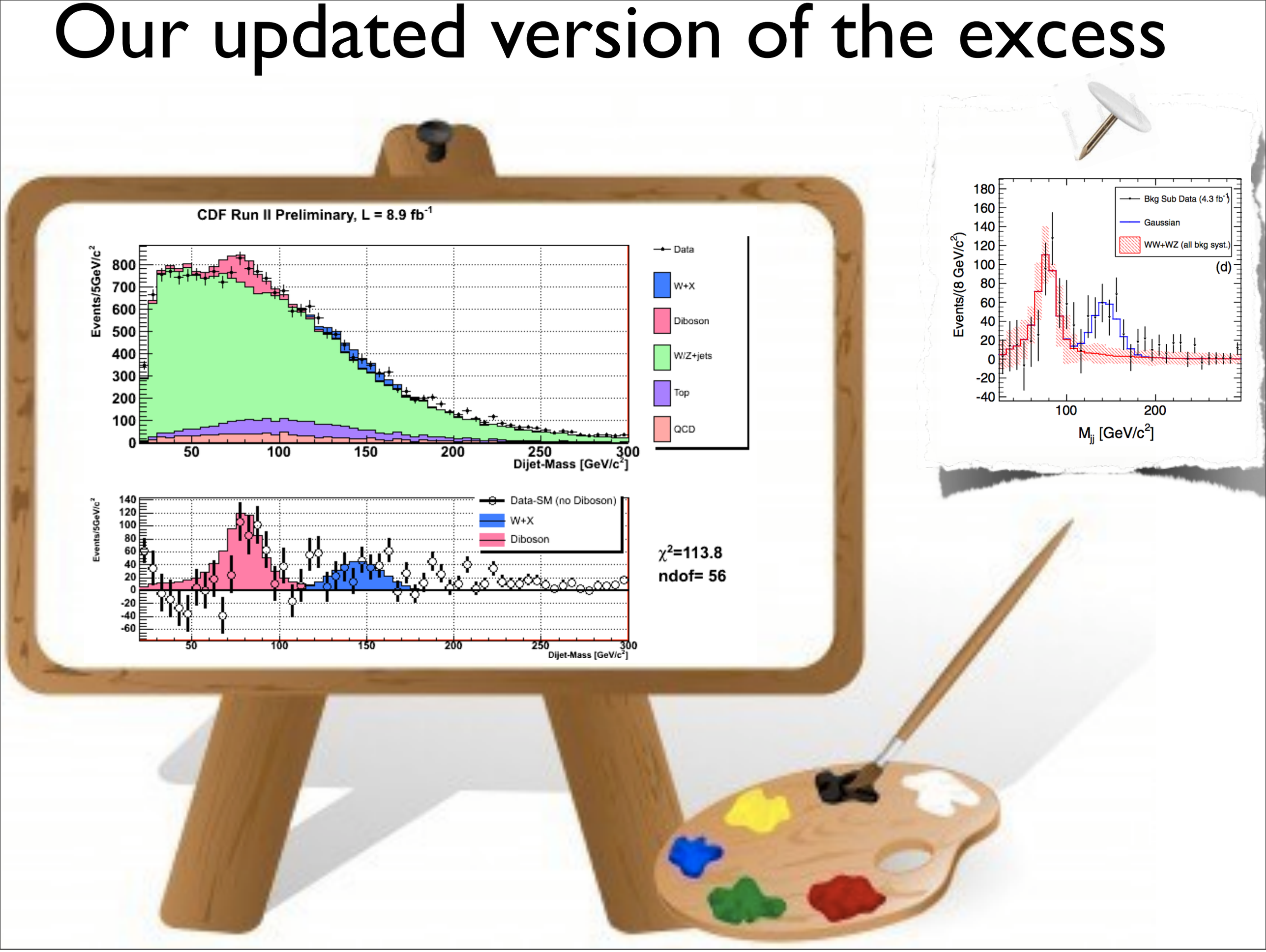
- As a first step, we will repeat the same analysis as in 2011 (4.3 fb^{-1}) with the full CDF dataset
  - Same event reconstruction and selection
  - Same modeling of the data sample



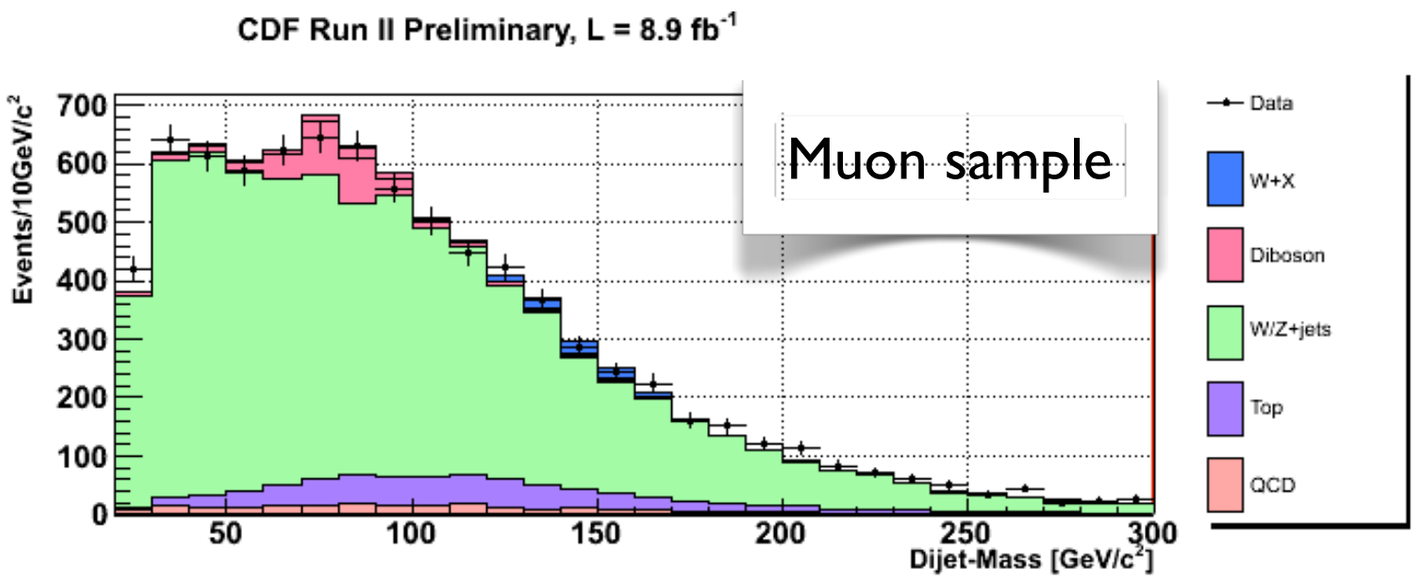
Our updated version of the excess

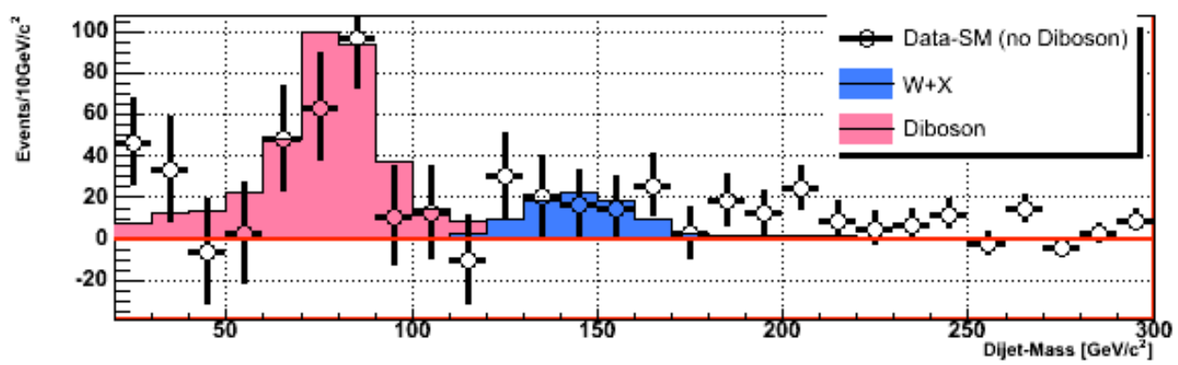


Our updated version of the excess

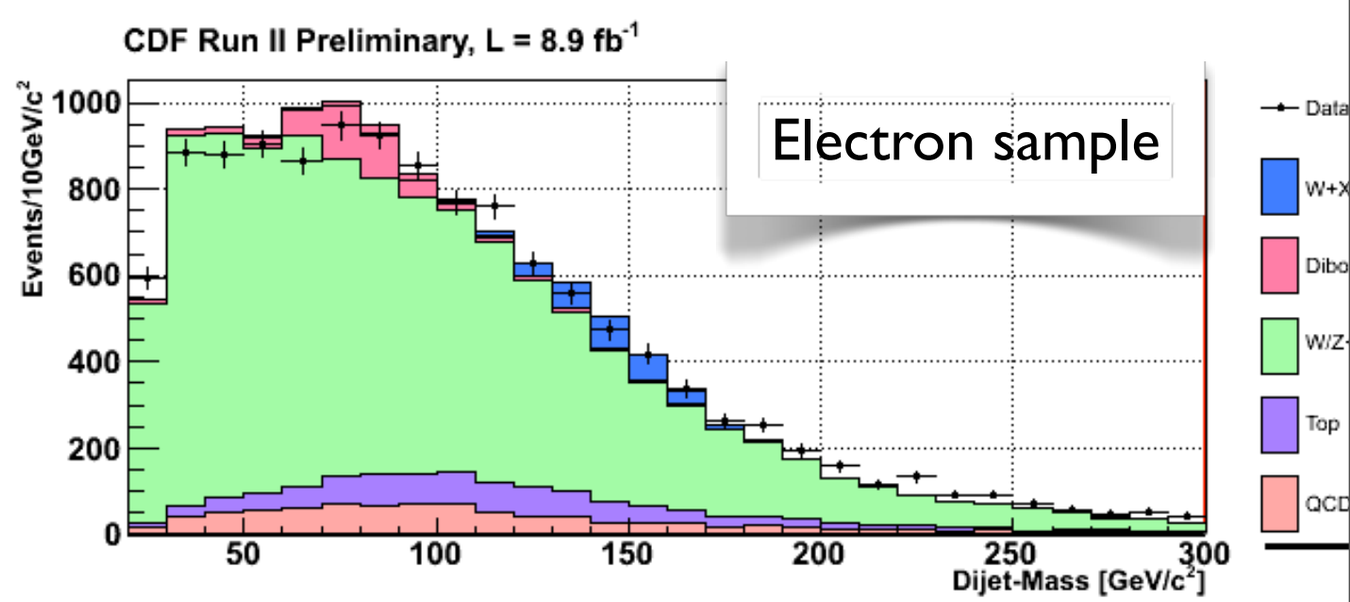


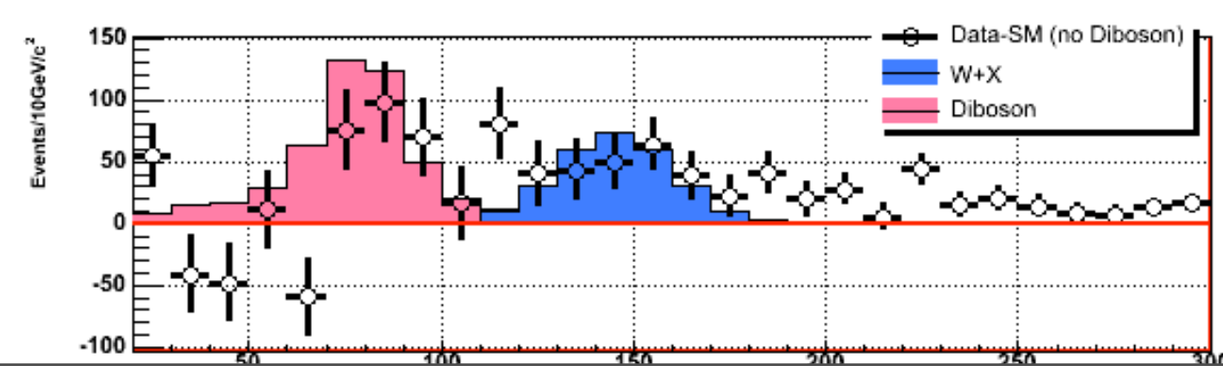
Our updated version of the excess





We also look at the <u>muon</u> and <u>electron</u> samples <u>separately</u>





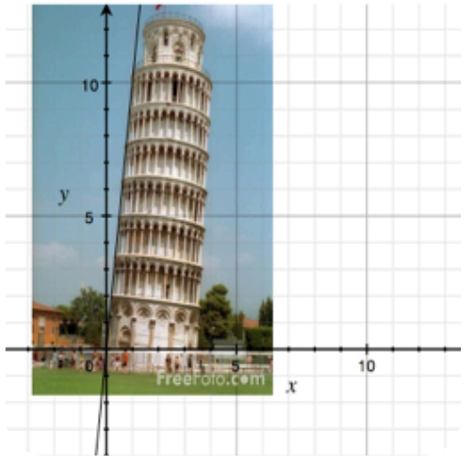
<u>Similar effects in both samples</u>

- Reproducing the same analysis with the full CDF dataset we observe a similar excess
- However, we are aware that other experiments don't observe the same excess
- What's going on?

Moving on

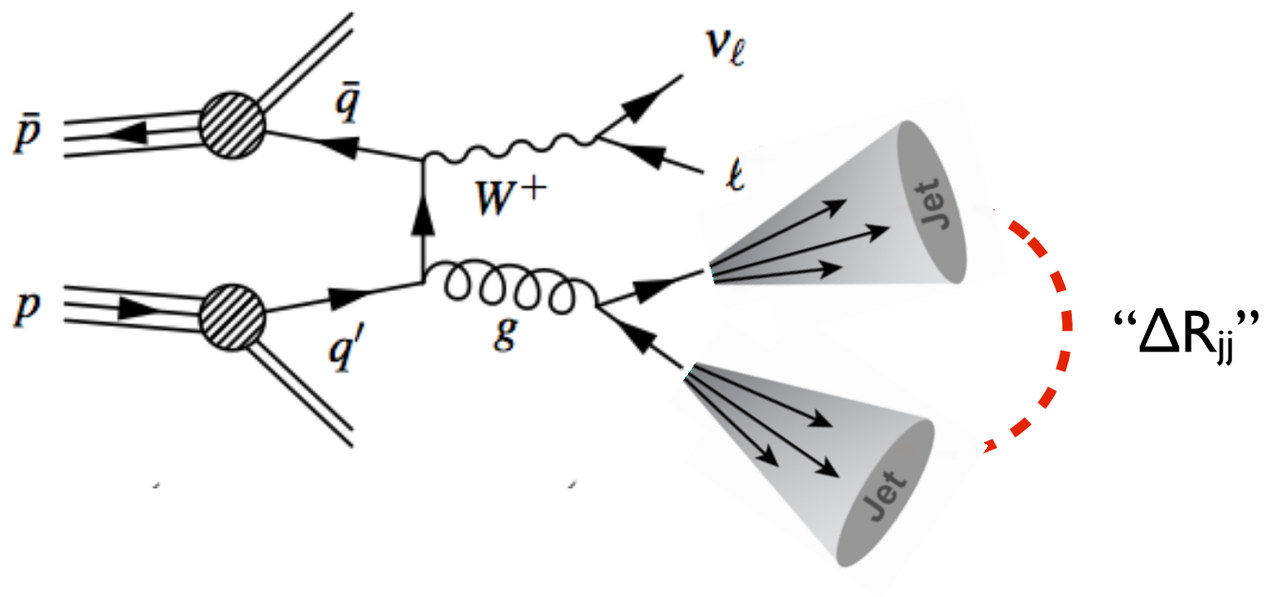
- In the past two years we studied the W+2jets sample intensively, including the possibility of **deficiencies in our modeling**
- In parallel we studied other samples as well
  - Large $\not{E}_T$ + 2jets
  - Two charged leptons + 2jets

  } more details later



- Here we focus on the three issues found to be most relevant:
  - Modeling of events with small dijet opening angle
  - Modeling of the jet response
  - Modeling of the fake-leptons
- All the details about our studies will be published soon

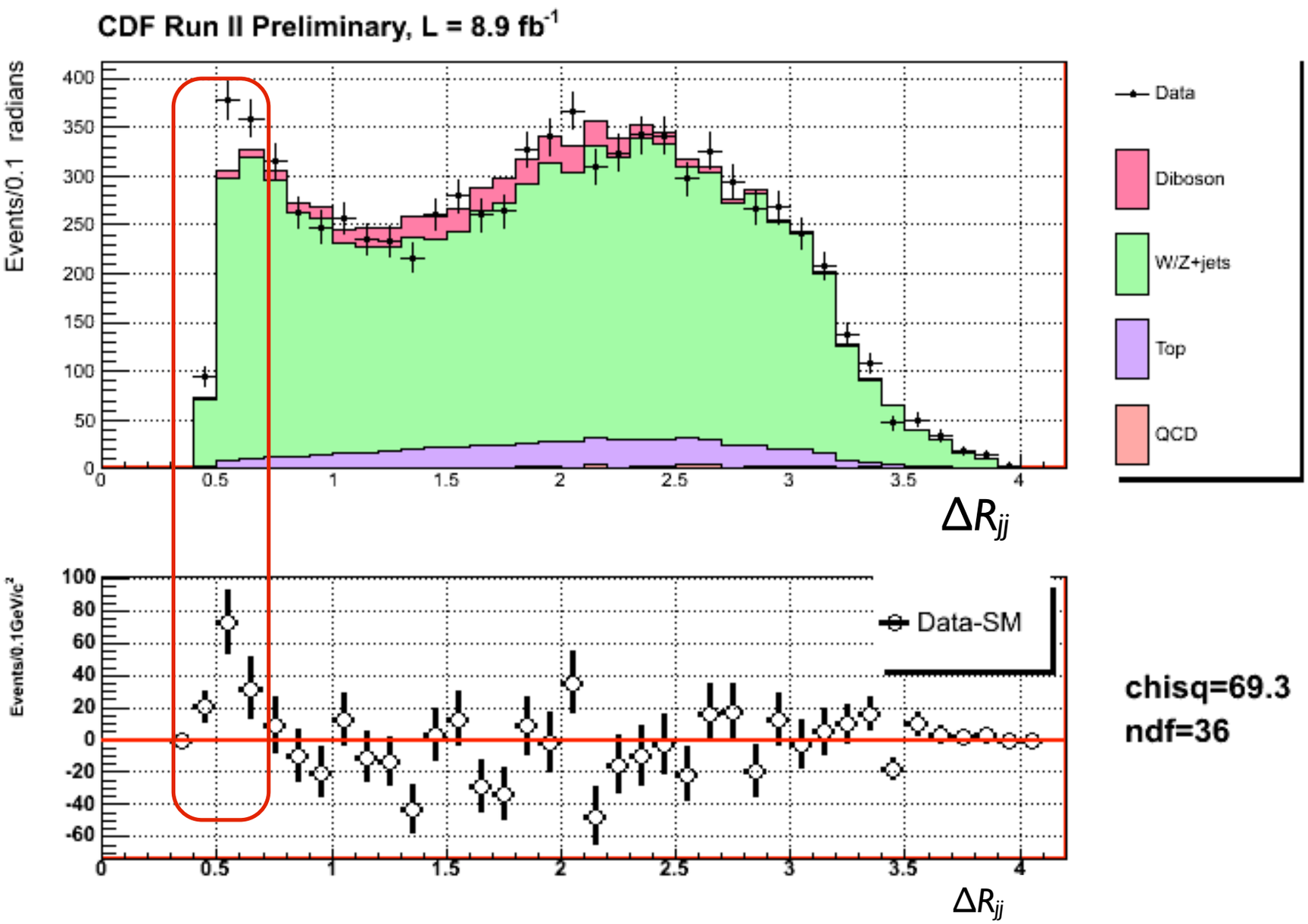
- Modeling of events with small dijet opening angle



- Modeling of the jet response
- Modeling of the fake-leptons

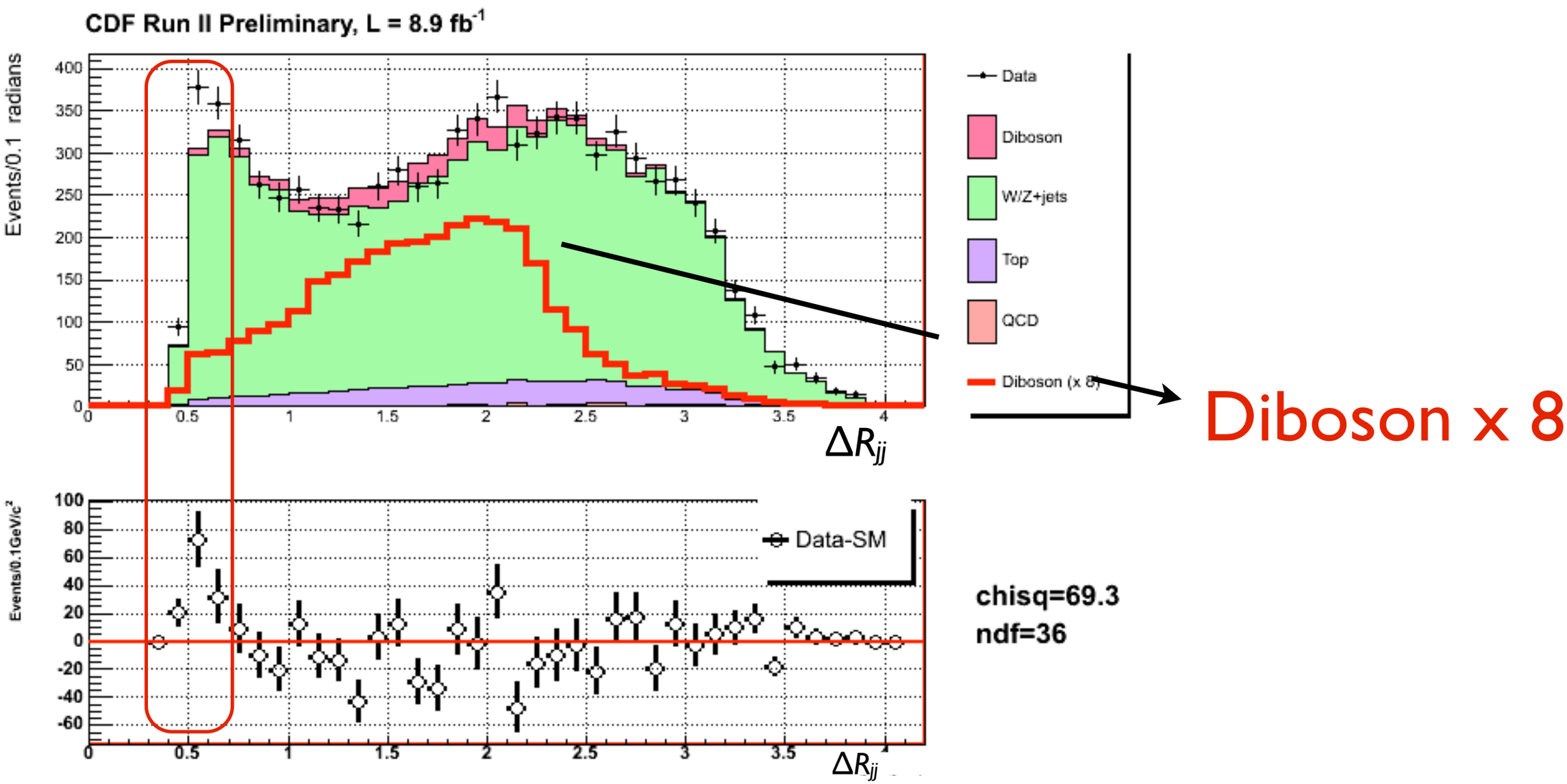
Event selection: Additional cut

- We noticed disagreements between data and predictions at low dijet opening angles
  - Clustering not properly simulated for closely spaced jets?



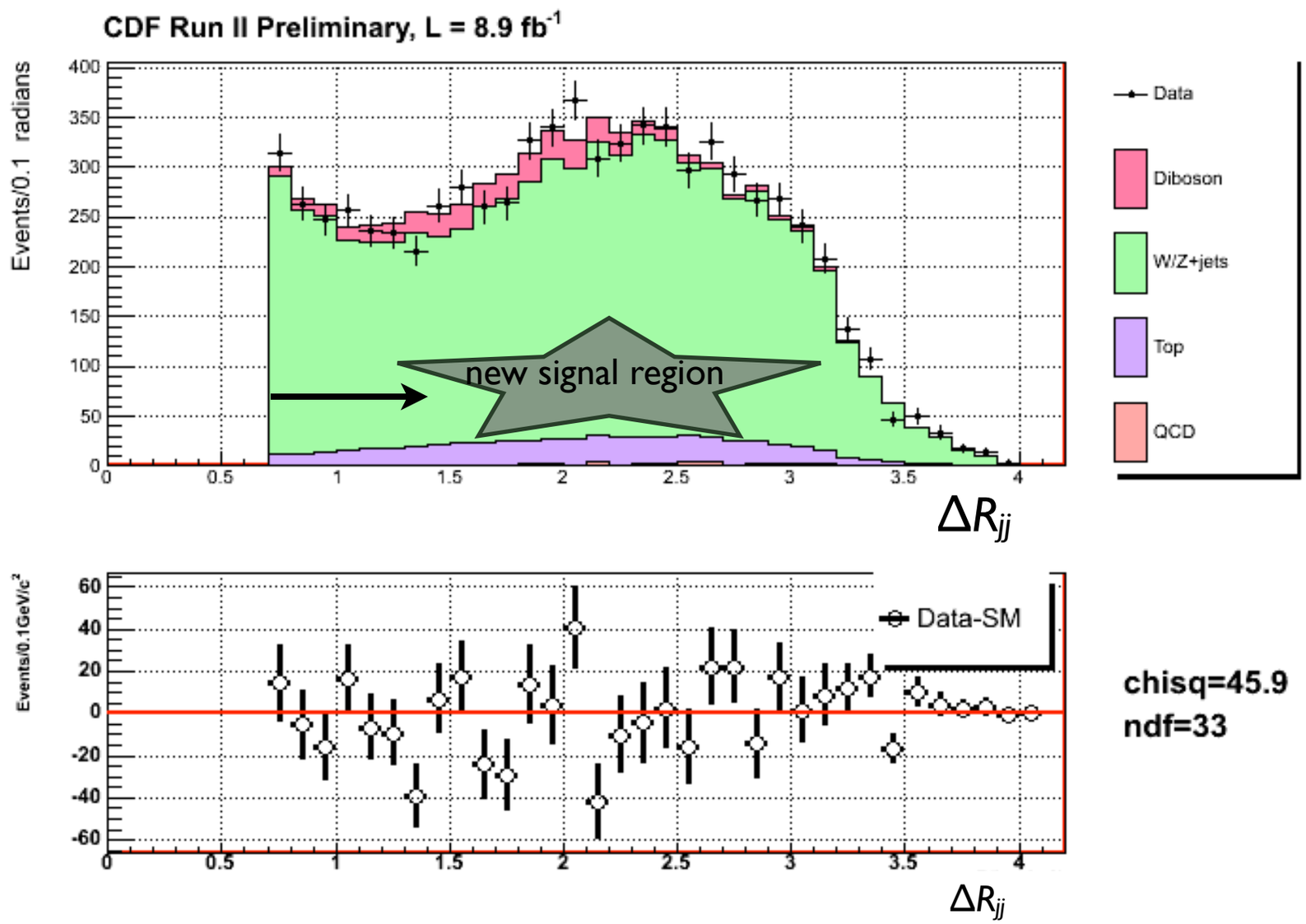
Event selection: Additional cut

- We noticed disagreements between data and predictions at low dijet opening angles
  - Clustering not properly simulated for closely spaced jets?
  - However, heavy particles decay to jets with large opening angle



Event selection: Additional cut

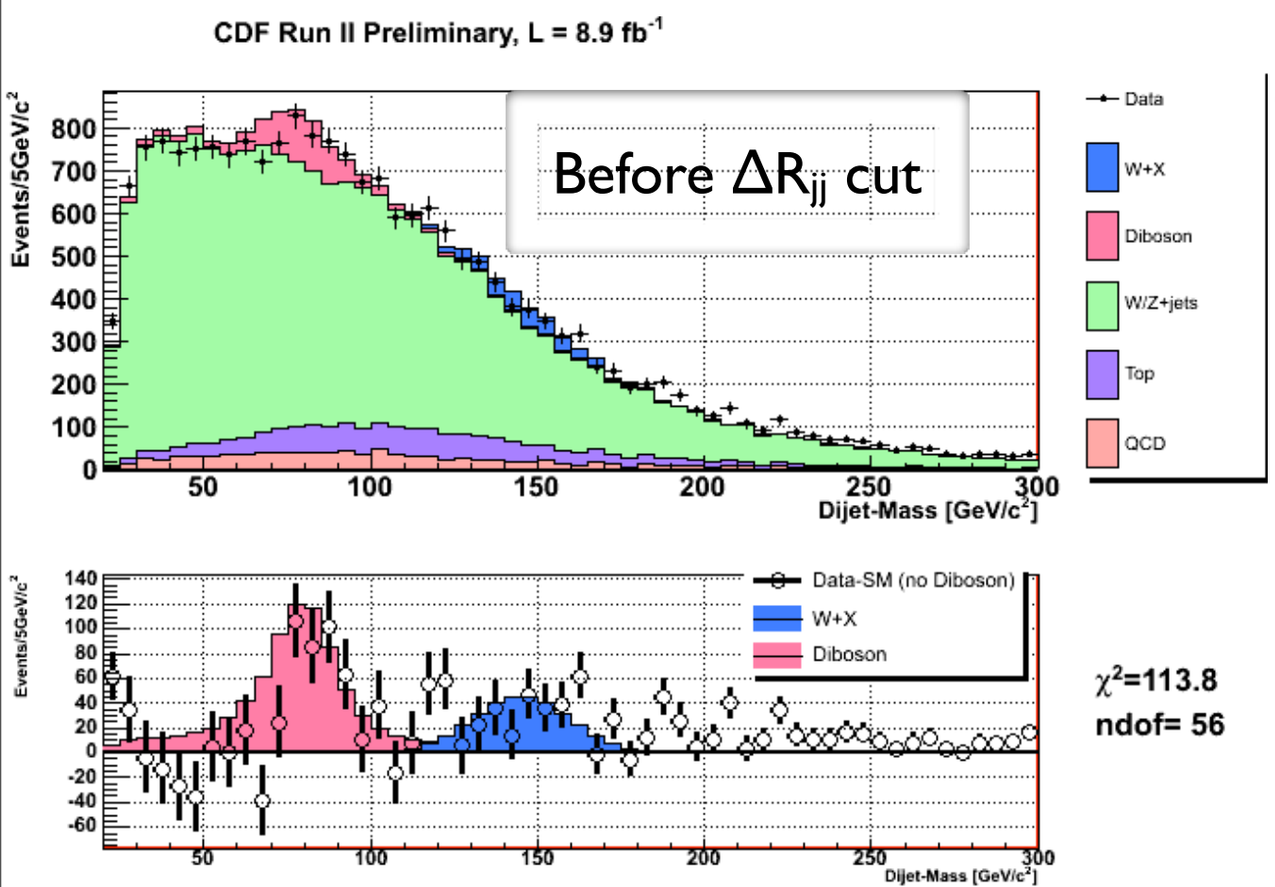
- We noticed disagreements between data and predictions at low dijet opening angles
  - Clustering not properly simulated for closely spaced jets?
  - However, heavy particles decay to jets with large opening angle
  - We therefore require $\Delta R_{jj}>0.7$

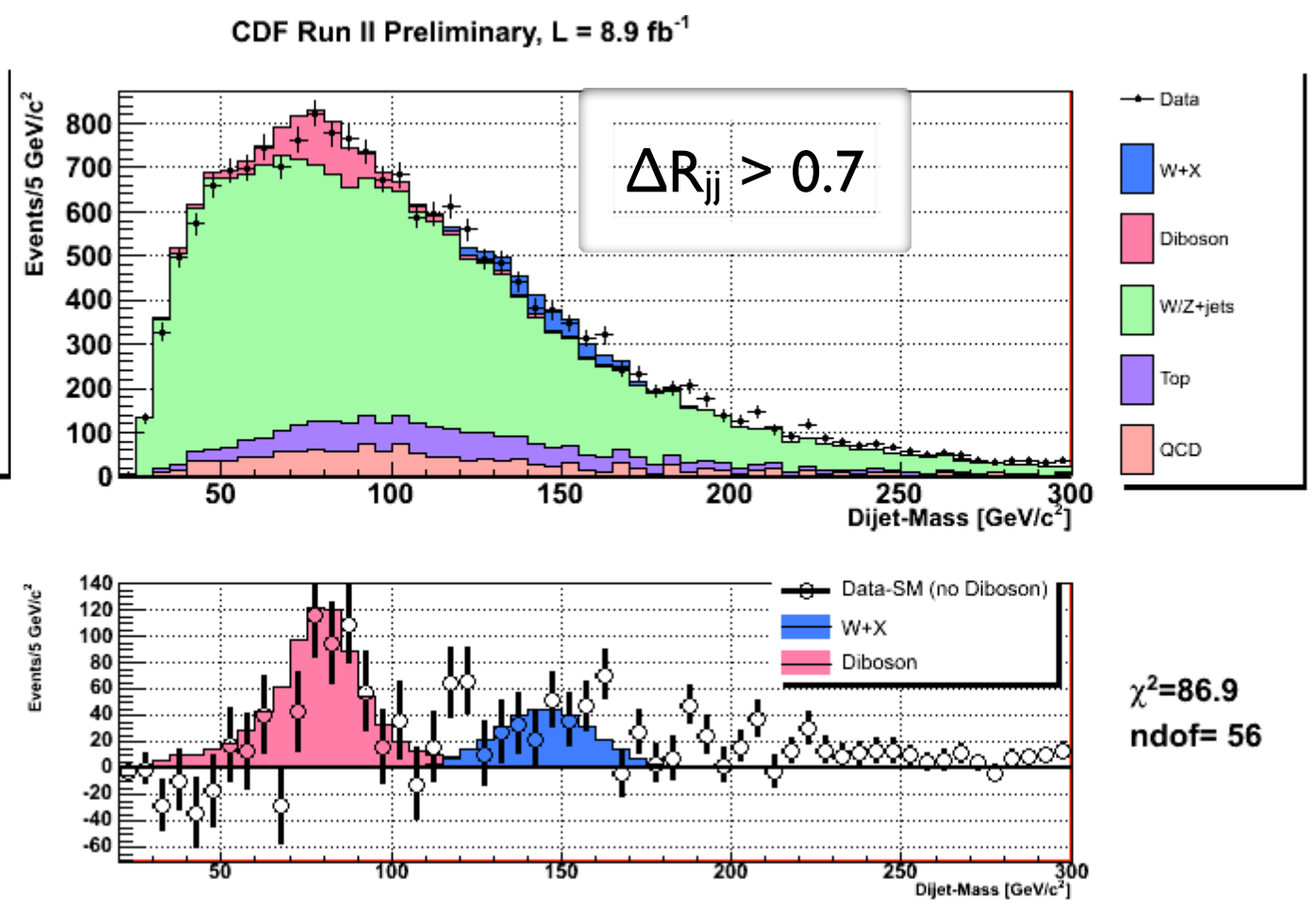


chisq=45.9
ndf=33

Event selection: Additional cut

- ΔR_{jj}>0.7 cut has a minor effect at large dijet masses
- But improved agreement with the predictions for low dijet masses

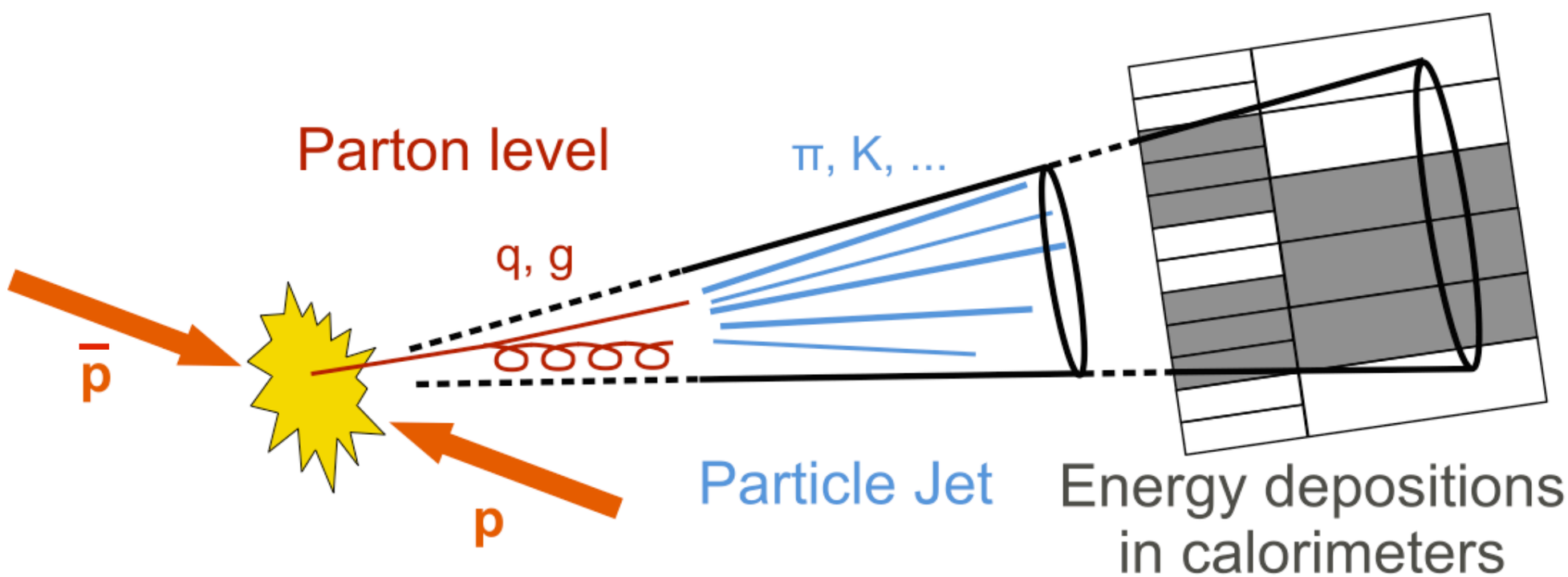




- Modeling of events with small dijet opening angle
- Modeling of the jet response
- Modeling of the fake-leptons

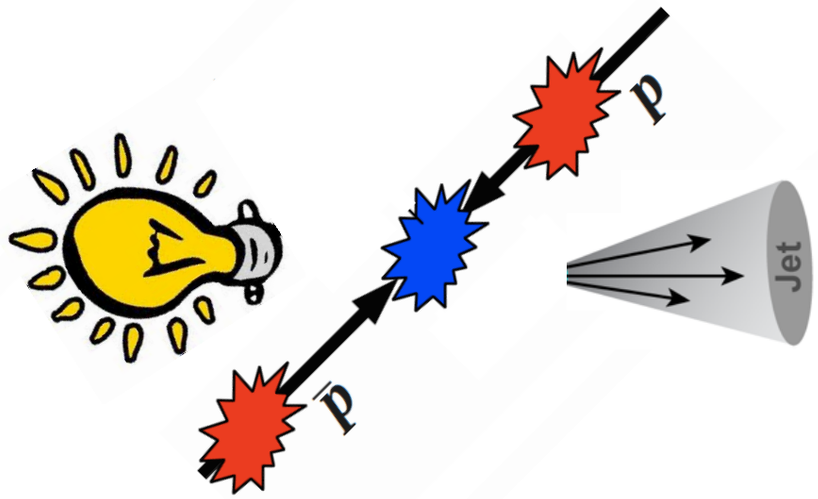
Object identifications - jets

- Jets of particles **in this analysis**
  - Collection of calorimeter towers clustered with "**JETCLU"** cone algorithm (radius **R=0.4**)
  - Electromagnetic fraction less than 0.9



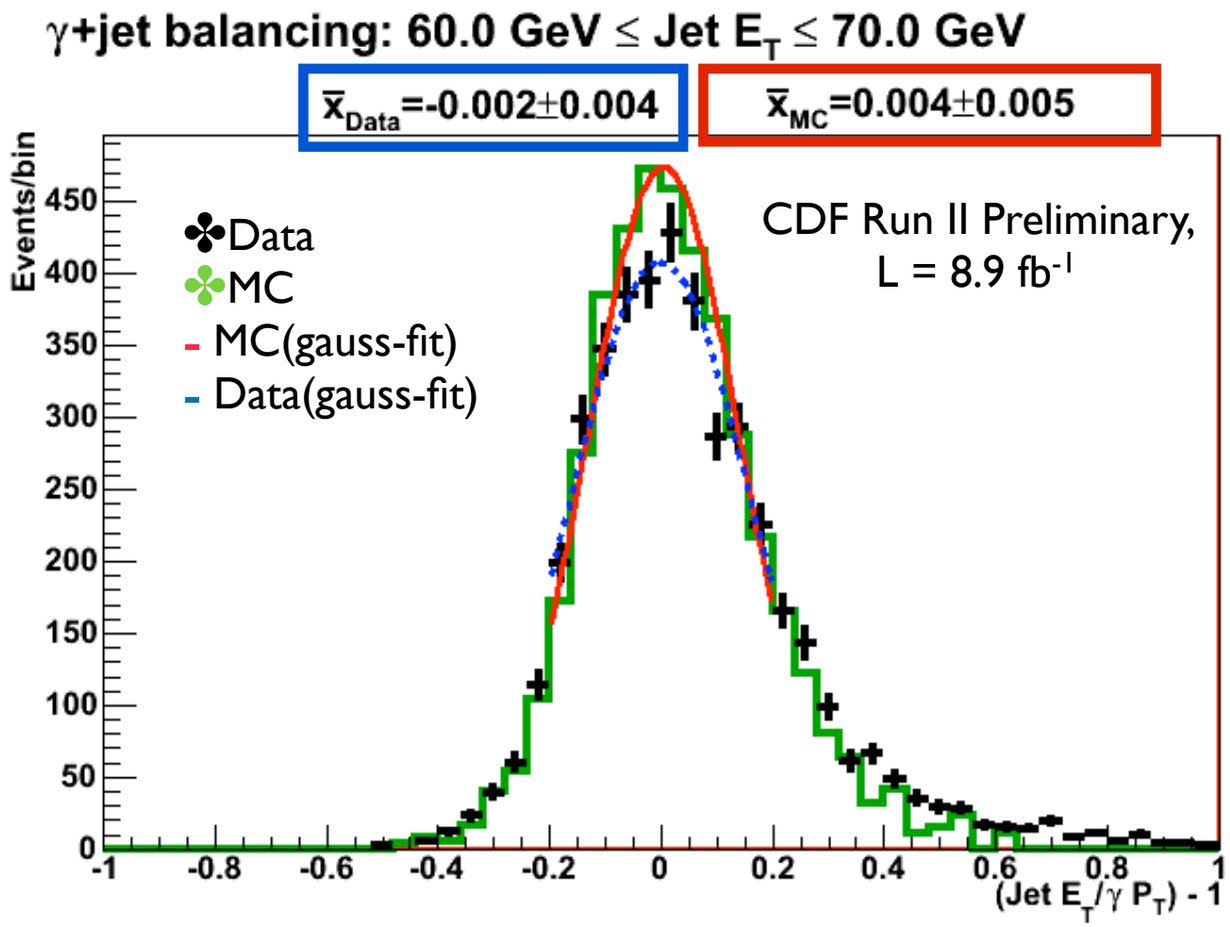
Object identifications - jets

- Jet energy scale ("JES")
  - Transfer function from the tower-cluster to the hadronized parton energy (true energy)
  - Validated in γ+jet events
    - If no extra-activity in the calorimeter $E^{jet}_T / E^{\gamma}_T \sim 1$



Object identifications - jets

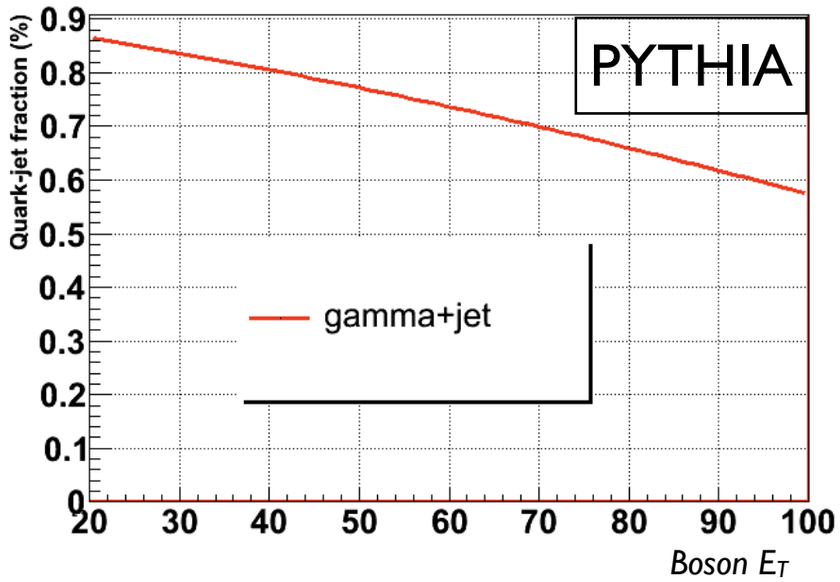
- Critical to this analysis when validating the JES is the **compatibility between Data and MC**



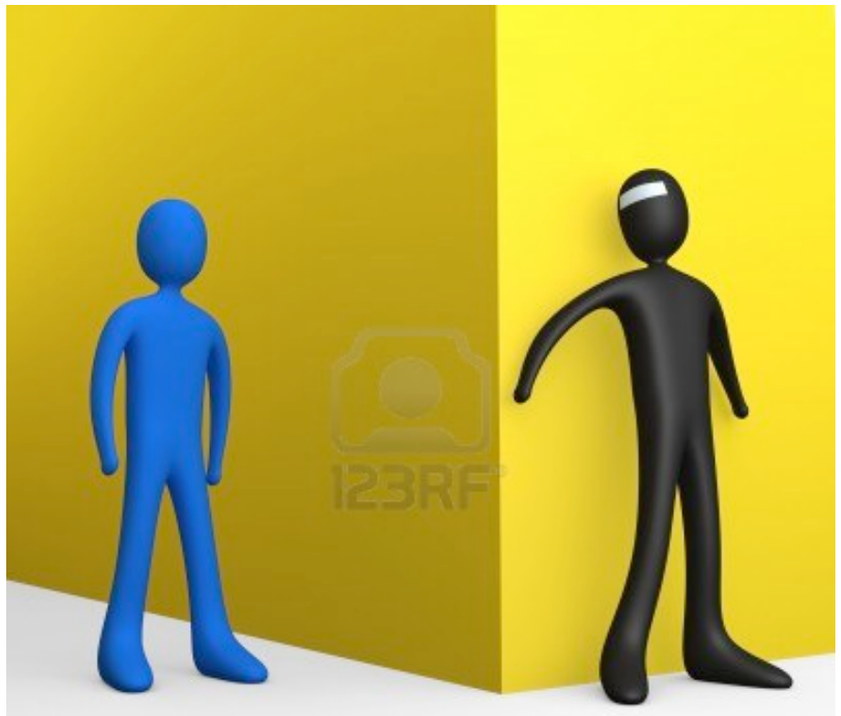
✓Jet balancing against the photon is **compatible** between PYTHIA Tune A **MC** and **Data** within statistical uncertainty

Object identifications - jets

- However, γ+jet events are quark dominated (>70% at E^{γ}_{T}<80 GeV)



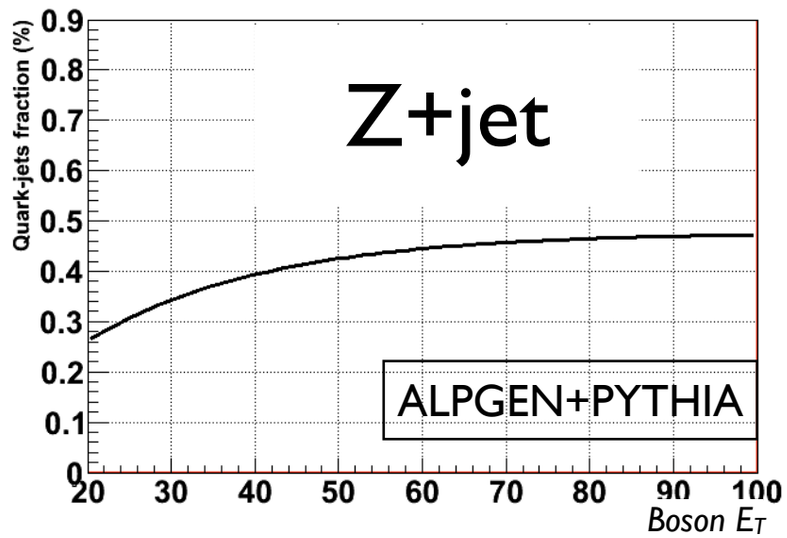
- *Question: how well does this JES work in samples with higher gluon content?*



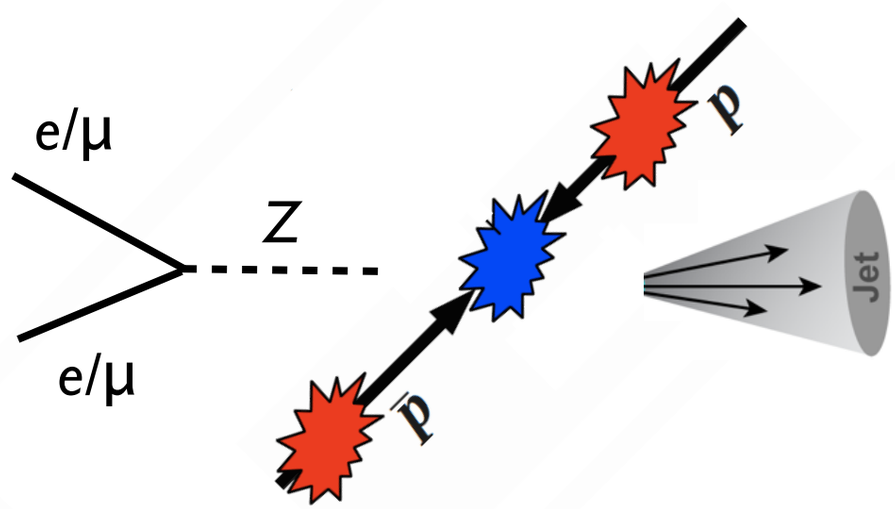
Object identifications - jets

NEW

- Back when JES was derived we didn't have the statistical power (200 pb^{-1}) to investigate other samples with larger gluon-jets content



- Now with the full dataset (8.9 fb^{-1}) we can do that

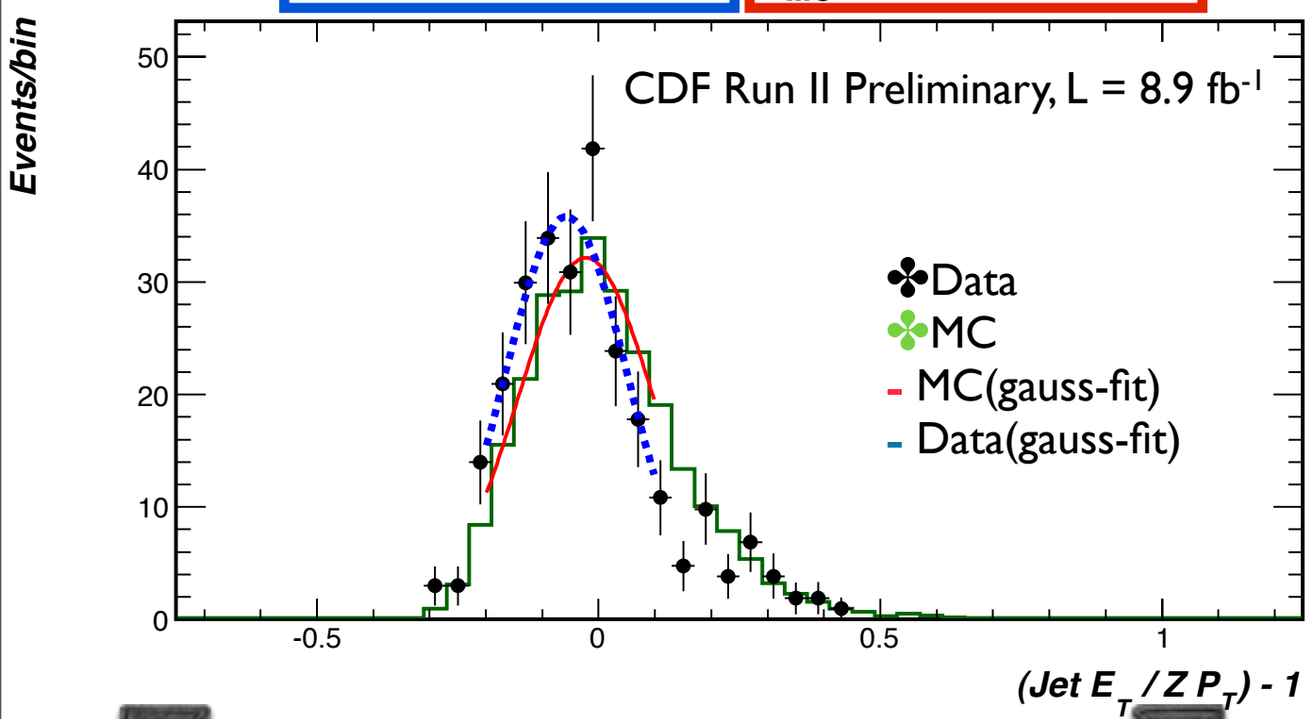


Object identifications - jets

- Critical to this analysis when validating the JES is the **compatibility between Data and MC**

Z-Jet Balancing: 60.0 GeV < Jet E_T < 70.0 GeV

$\bar{x}_{Data} = -0.057 \pm 0.012$, $\bar{x}_{MC} = -0.023 \pm 0.004$



ALPGEN+PYTHIA Tune A **MC** and **Data** are **not compatible** within statistical uncertainty

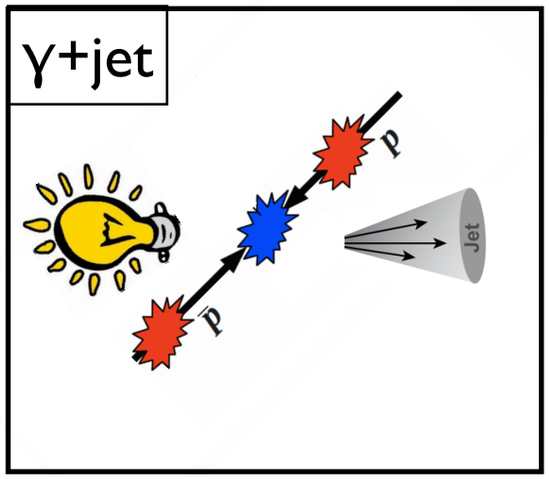
© 3poD * www.ClipartOf.com/15462

Object identifications - jets

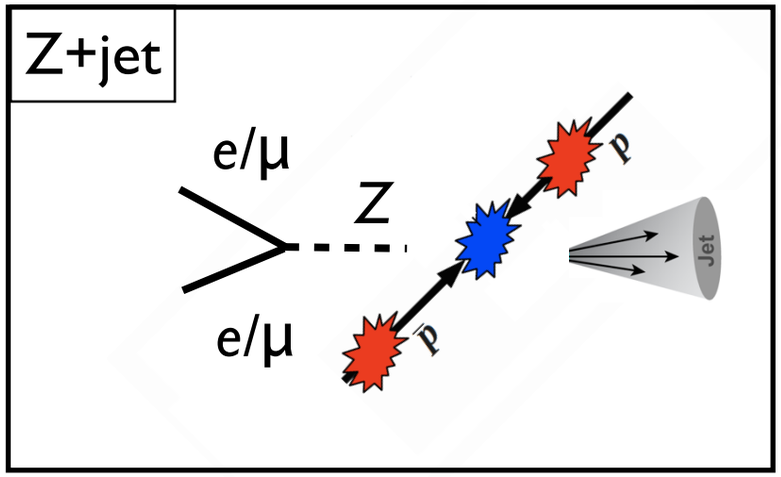
NEW

Object identifications - jets

jet balancing

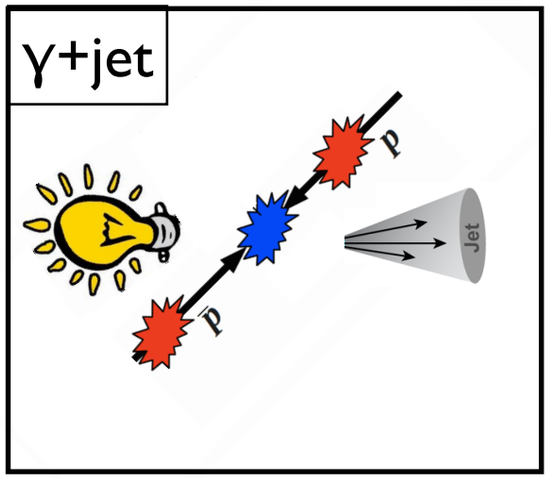


,

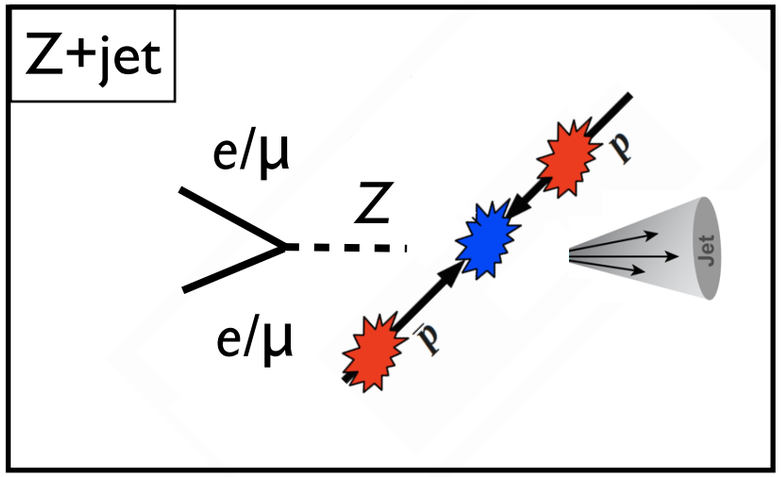


Object identifications - jets

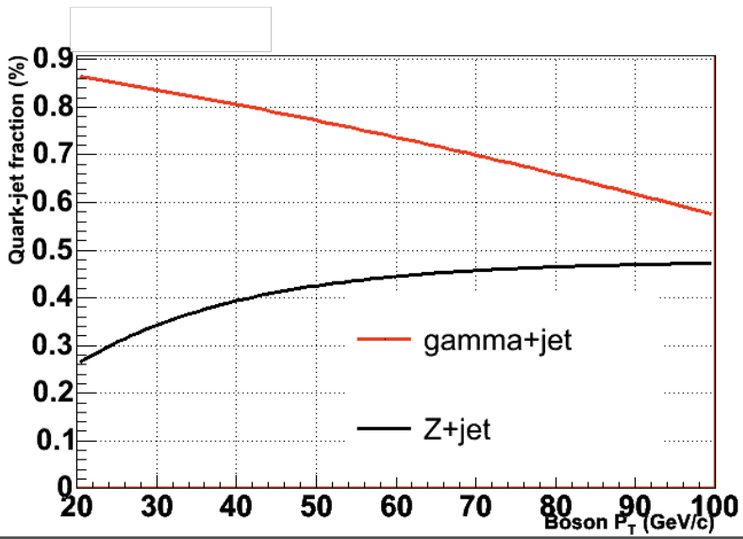
jet balancing



,



Quark-jet fraction
(from MC)



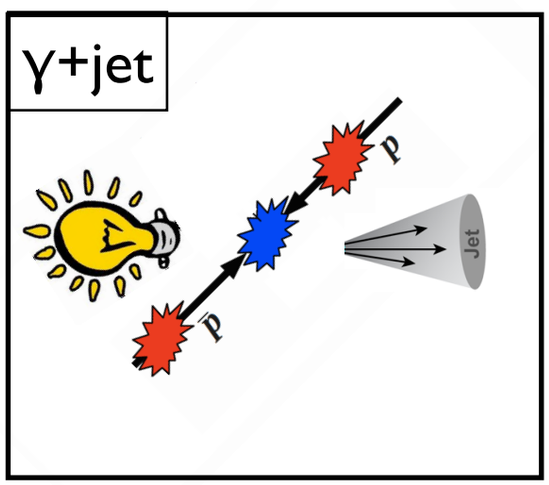
, (cross-checked with data)

Object identifications - jets

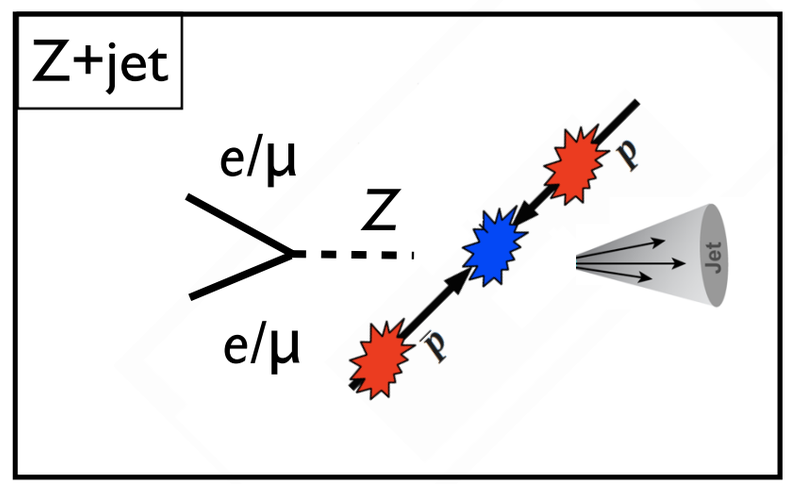
NEW

We will make the assumption that the observed discrepancies between γ+jet and Z+jet originate from differences in the modeling of the jet response for quarks and gluons

jet balancing

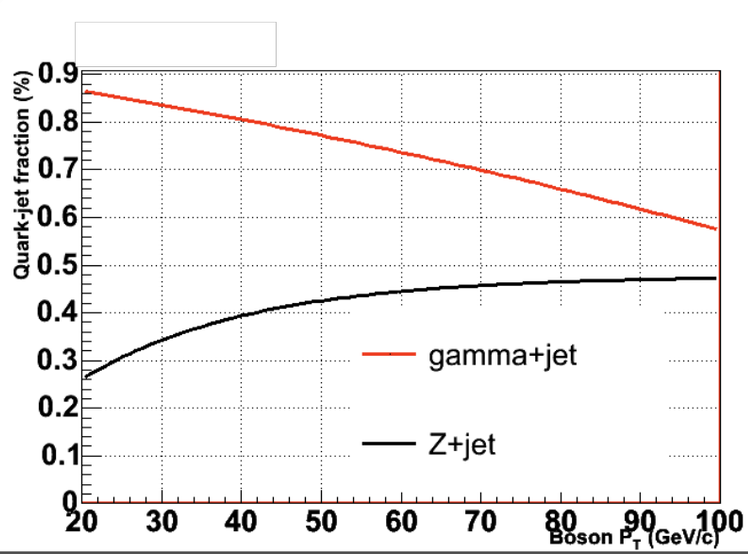


,



:

Quark-jet fraction (from MC)



, (cross-checked with data)

= quark-JES

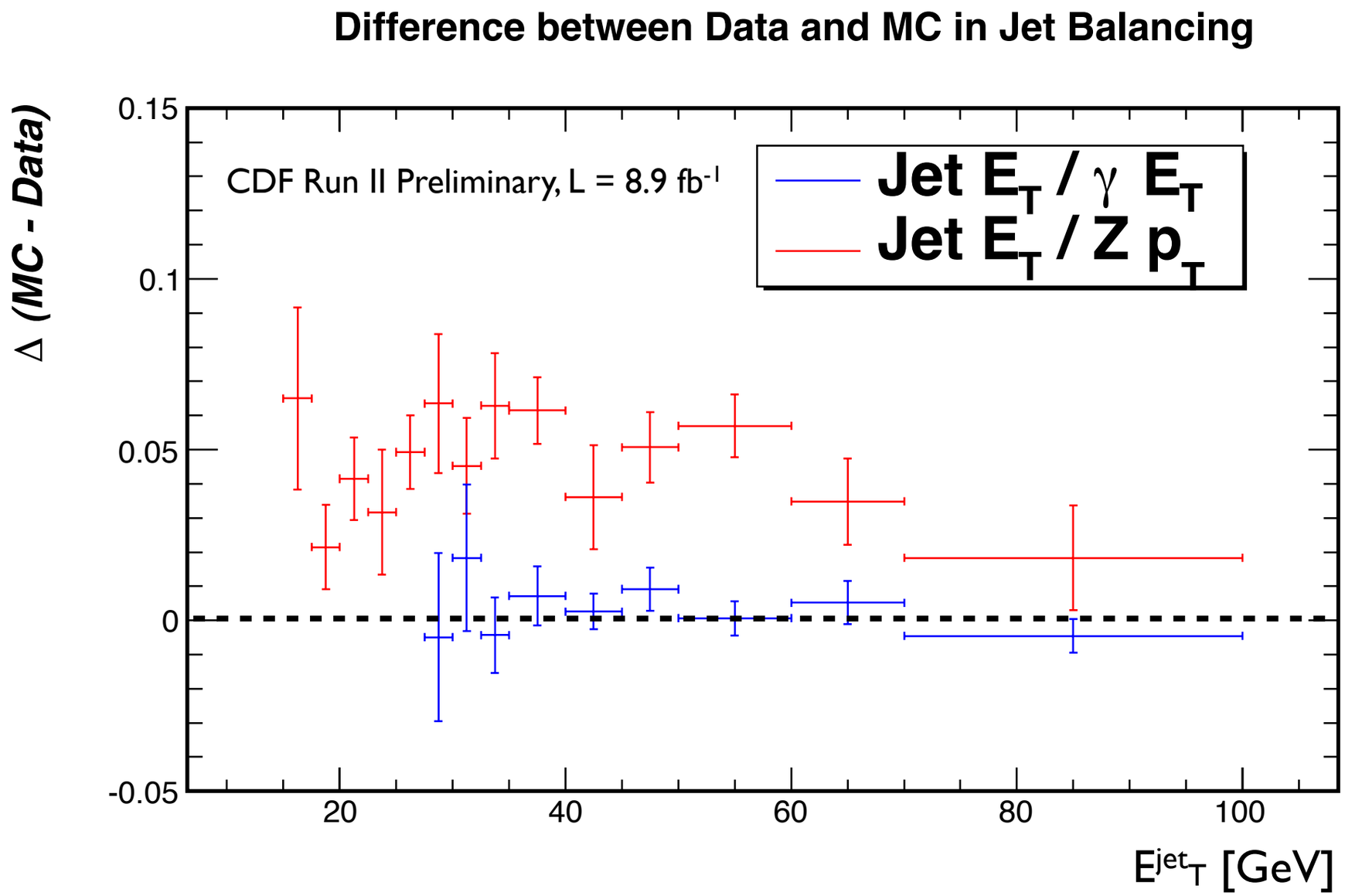
gluon-JES

Object identifications - jets

NEW

Jet balancing overview

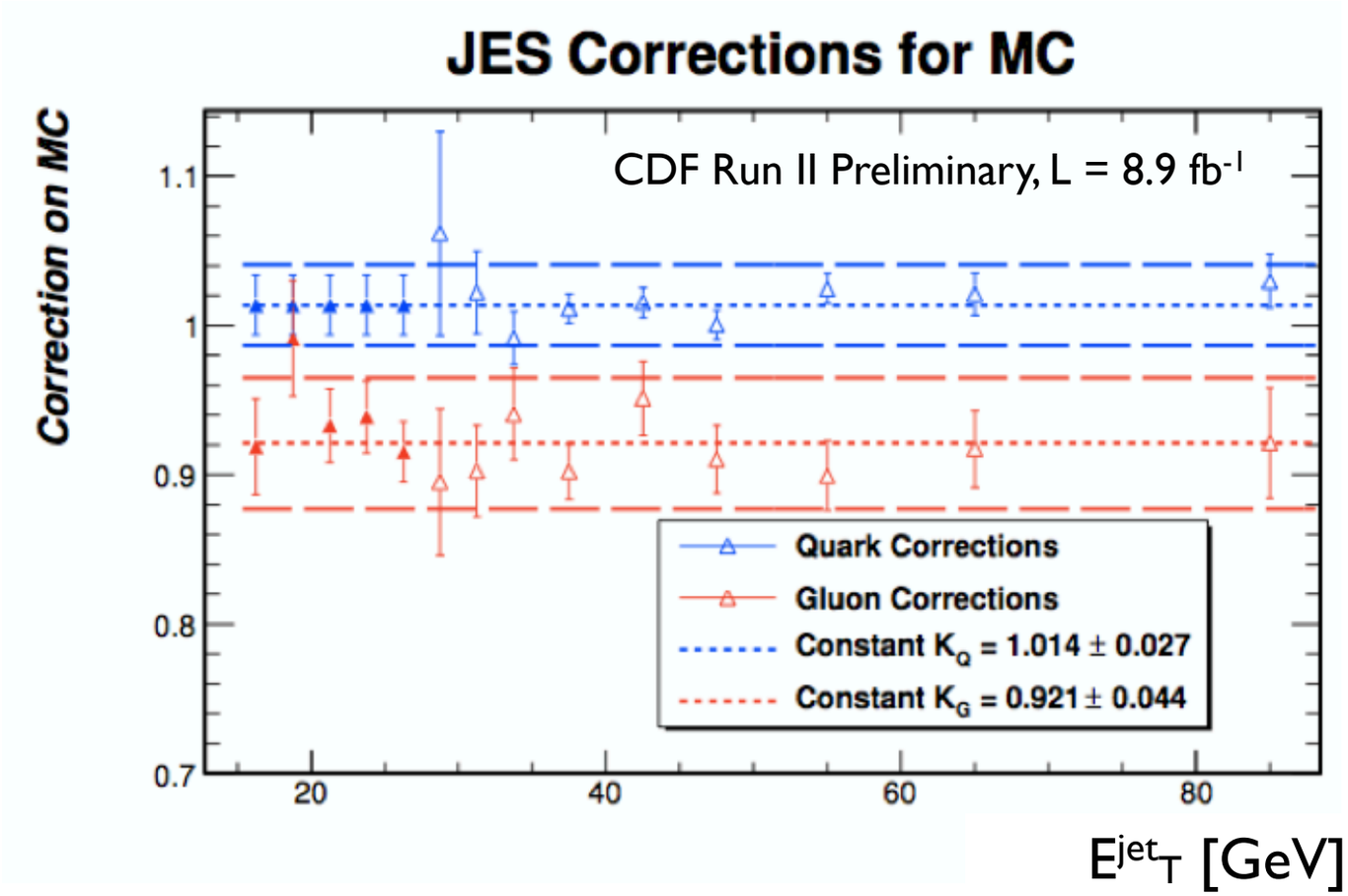
- Balancing performed in several jet E_T bins



Object identifications - jets

NEW

extraction
of quark/gluon-JES
(derived for JETCLU 0.4 cone jets
and PYTHIA showered MC)



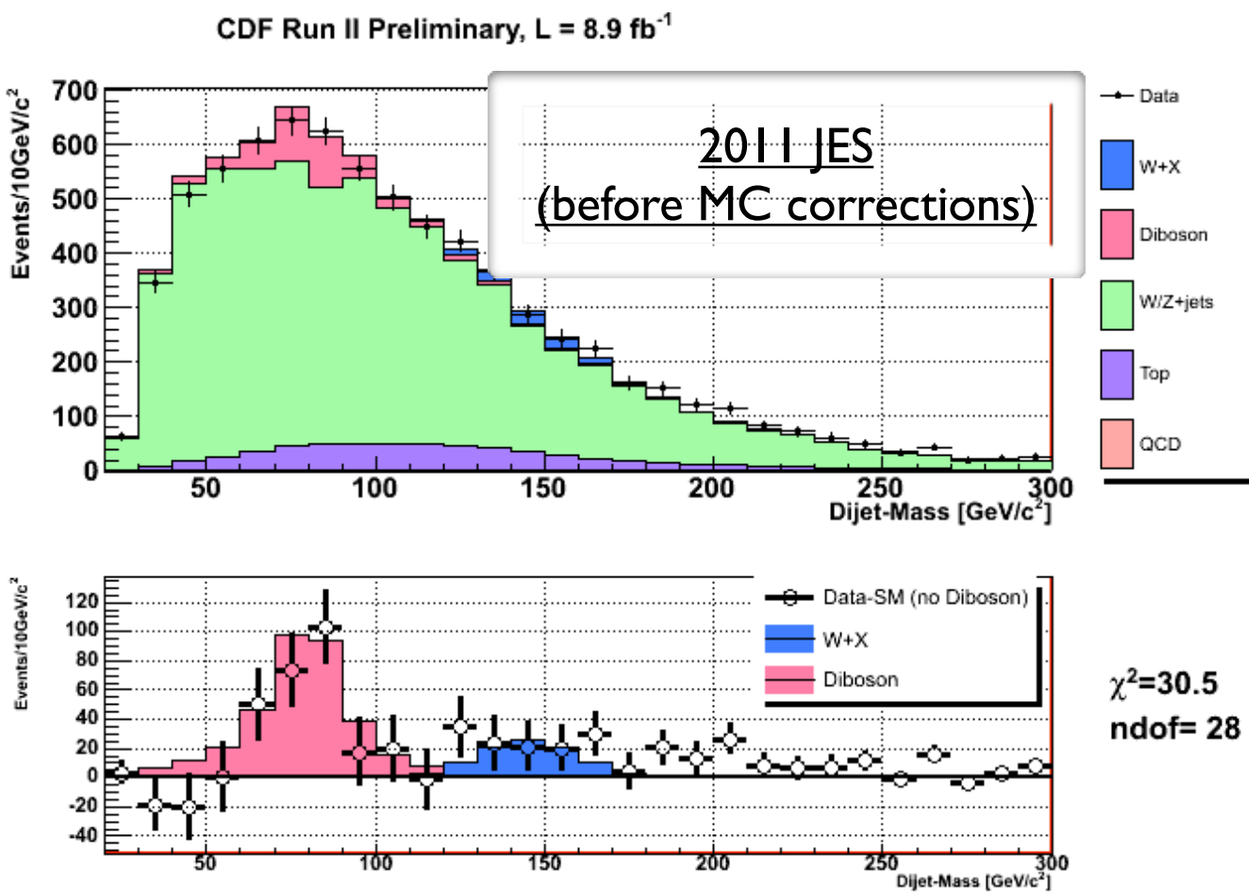
With respect to 2011 JES

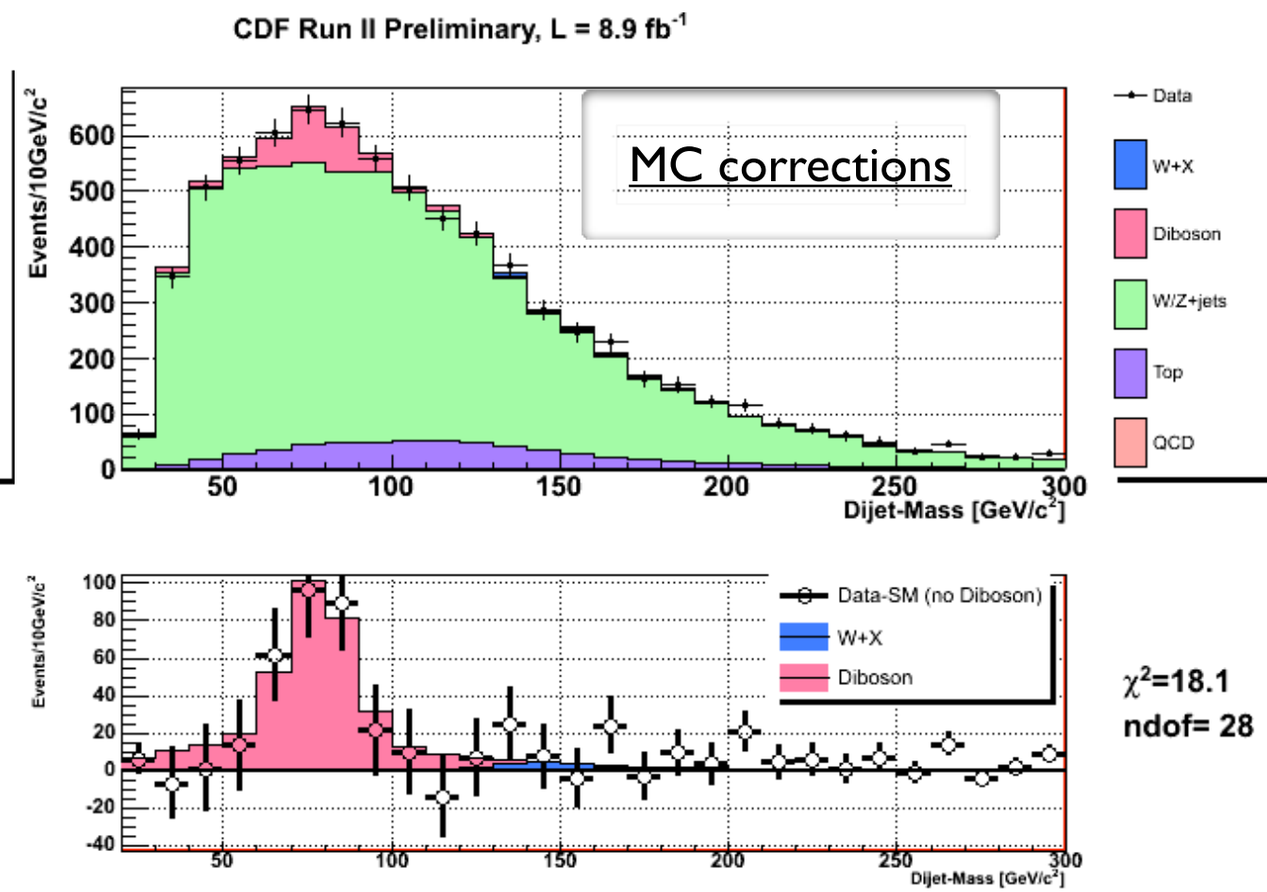
- MC:
  - quark-jets: (+1.4 +/- 2.7)%
  - gluon-jets: (-7.9 +/- 4.4)%

(No corrections applied to data)

Effect of the MC corrections
(Muon sample)

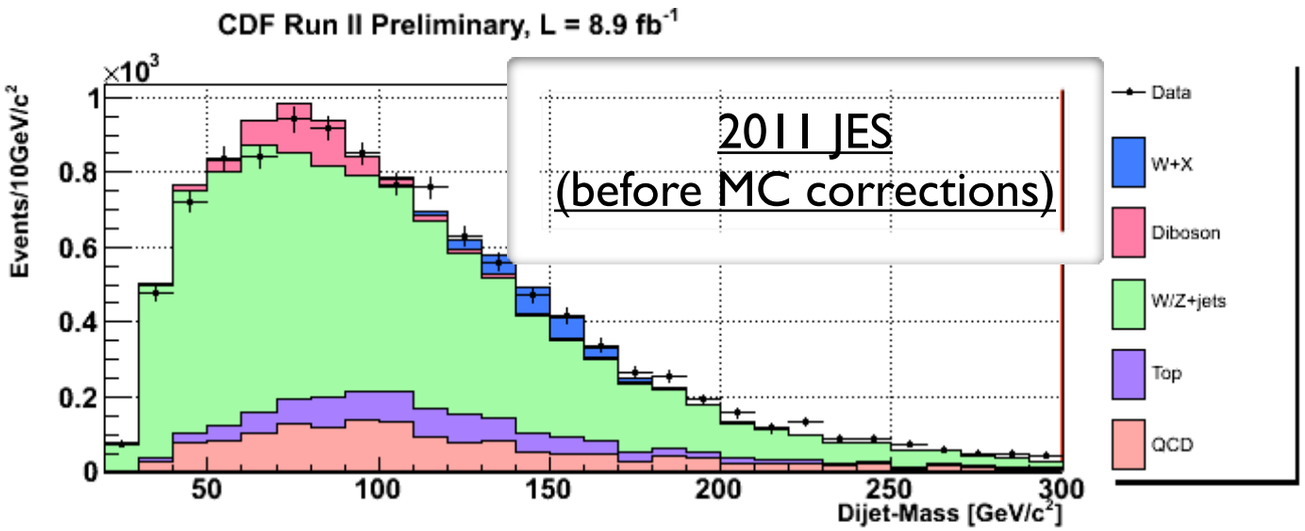
- Dijet mass in the muon sample with
  - standard JES MC (as in 2011)
  - JES MC corrections

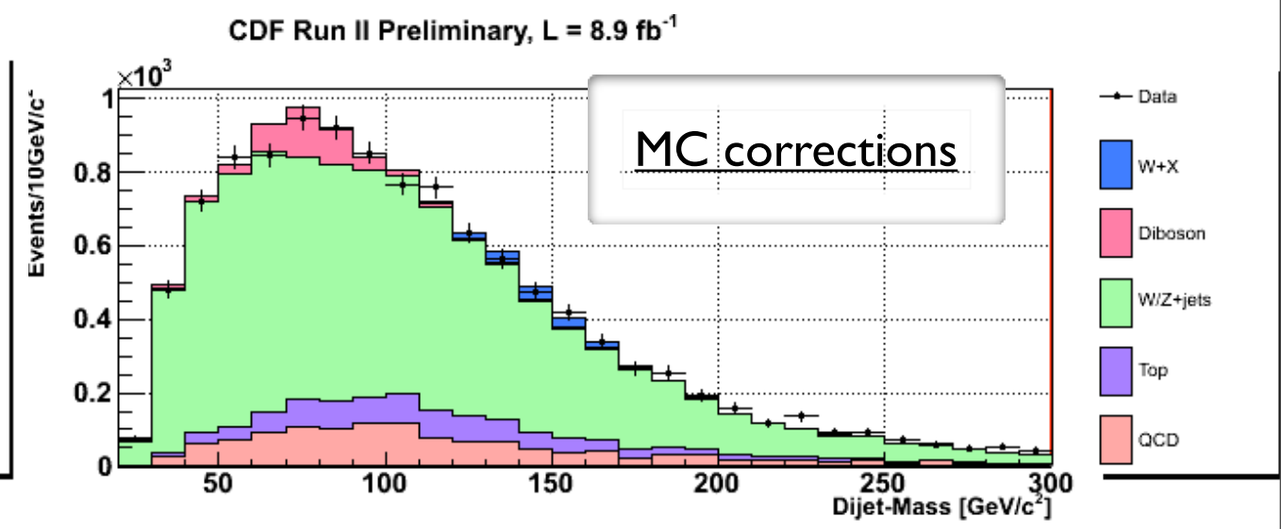


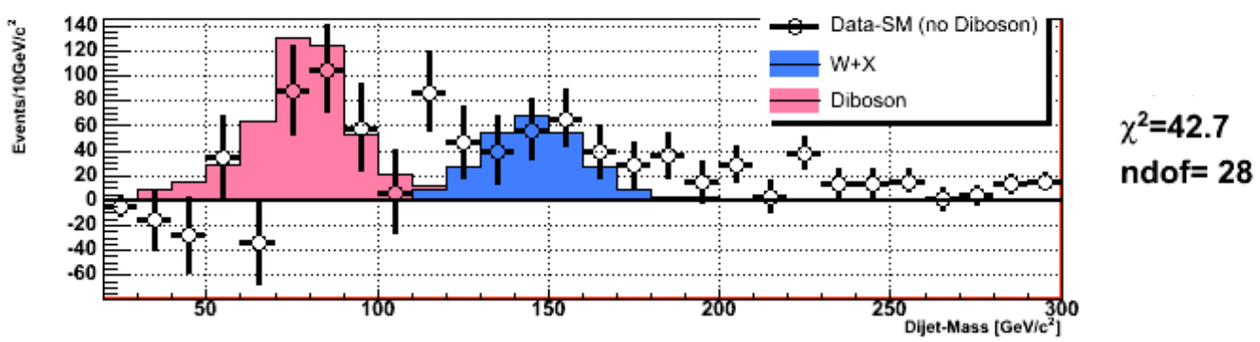


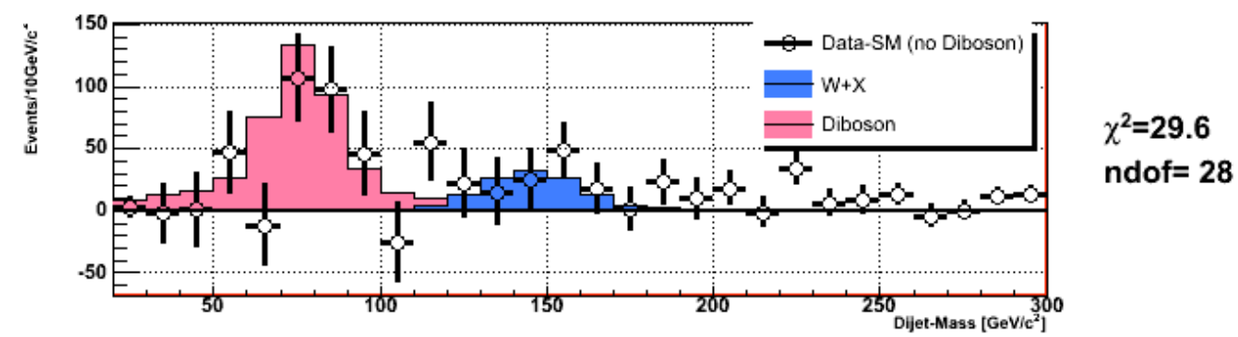
After JES MC corrections
muon sample is well described

Effect of the MC corrections
(Electron sample)









- JES MC corrections improve the description of the data
- However, the same **behavior** is not seen in the **muon** and electron **samples**

Question: what can be different between the muon and electron samples?

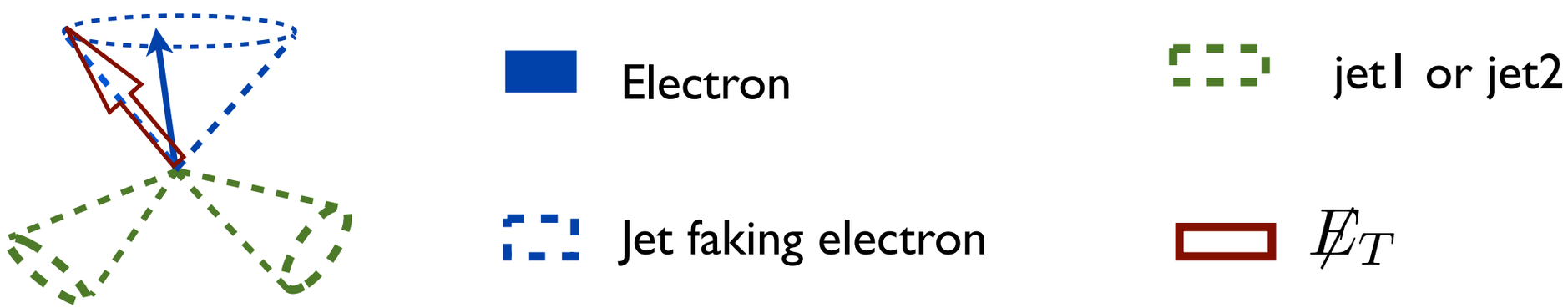
(Most obvious) answer: fake-leptons background

- Muons: neglibible
- Electrons: sizeable

- Modeling of events with small dijet opening angle
- Modeling of the jet response
- Modeling of the fake-leptons
  - the electron faking mechanism might not be accurately modeled

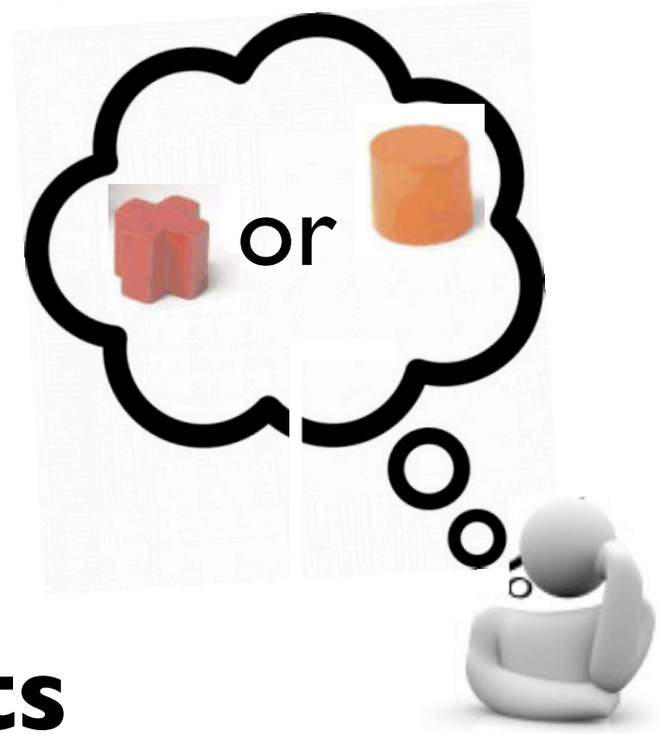
QCD multi-jets background

- Multi-jet selected events are typically 3 jets events
  - One jet faking the identified electron
  - $\not{E}_T$ generated by mis-measured calorimetric objects

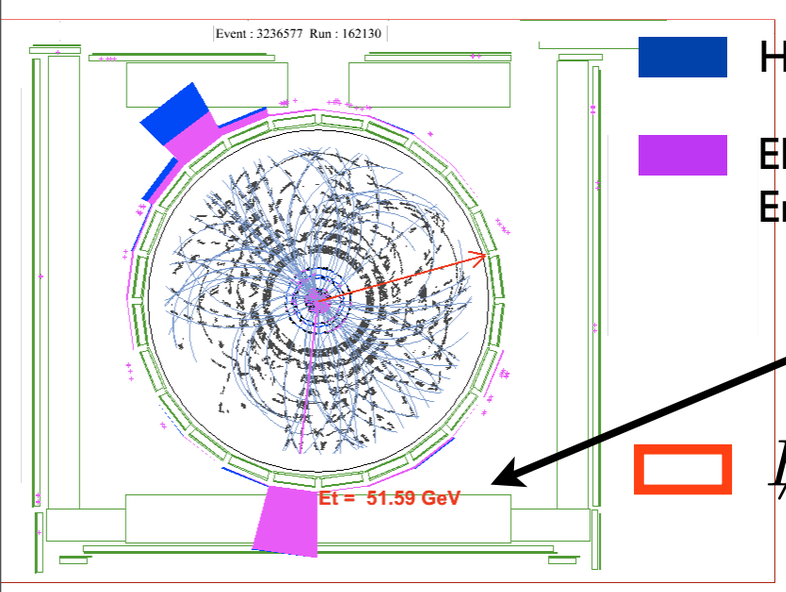


QCD multi-jets background

How do I model it?



- **Using "non-electron" events**
  - Non-electrons pass all the electron-candidate identification requirements except two of them (e.g: electromagnetic fraction)
  - Same kinematic requirements are applied to select non-electron and electron-candidate events



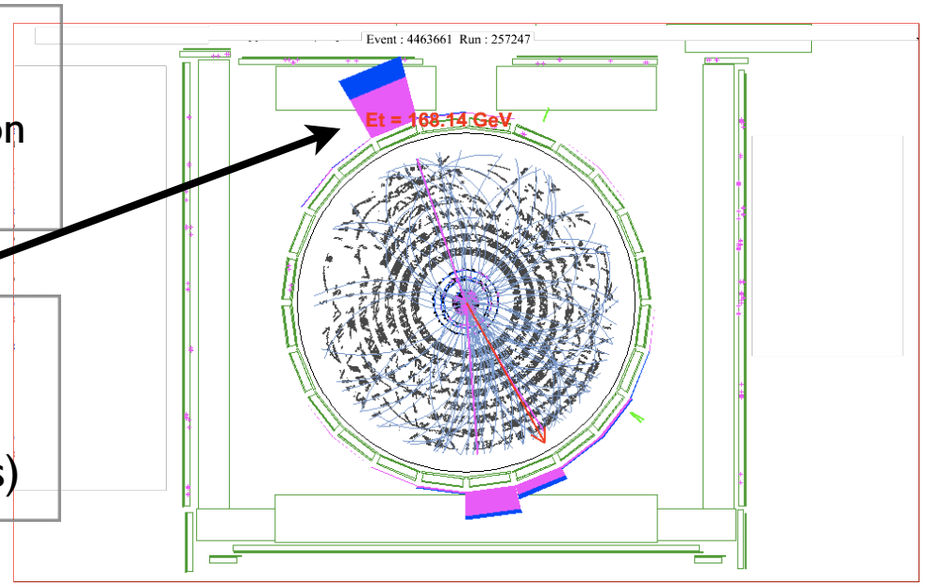
Hadronic Energy

Electromagnetic Energy

$\not{E}_T$

Electron-candidate
(passing all the identification requirements)

Non-electron
(failing two identification requirements)

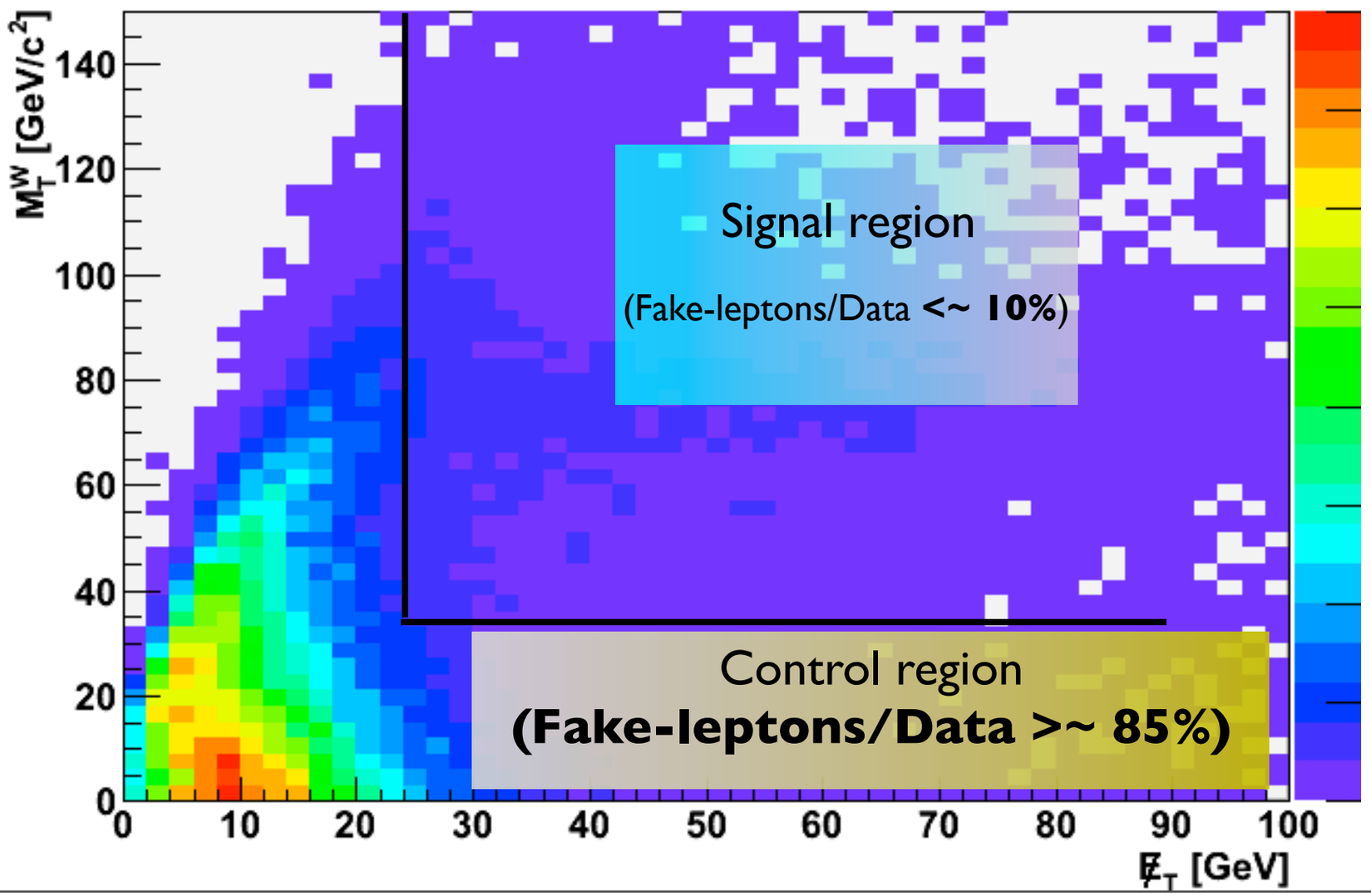


QCD multi-jets background

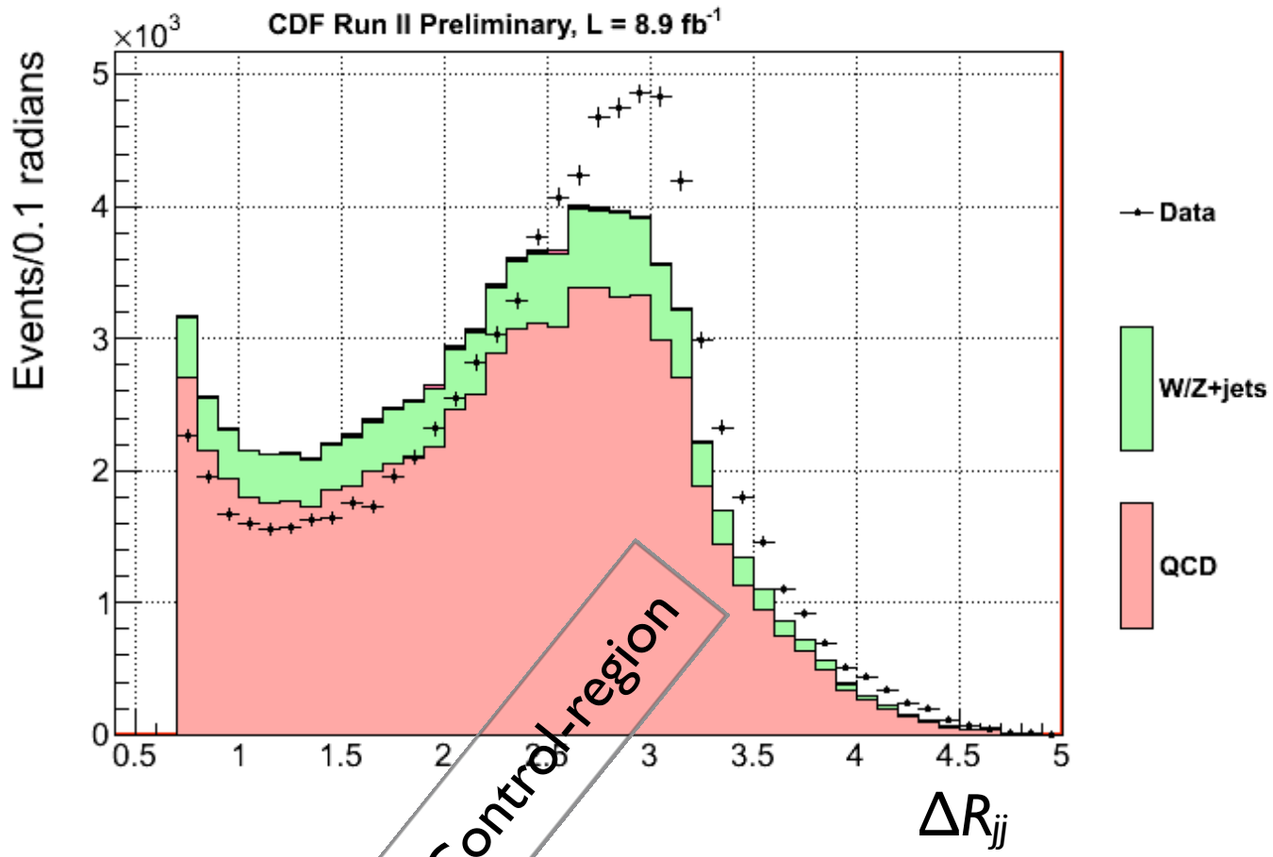
- How do we judge whether the non-electron model is appropriate or not?

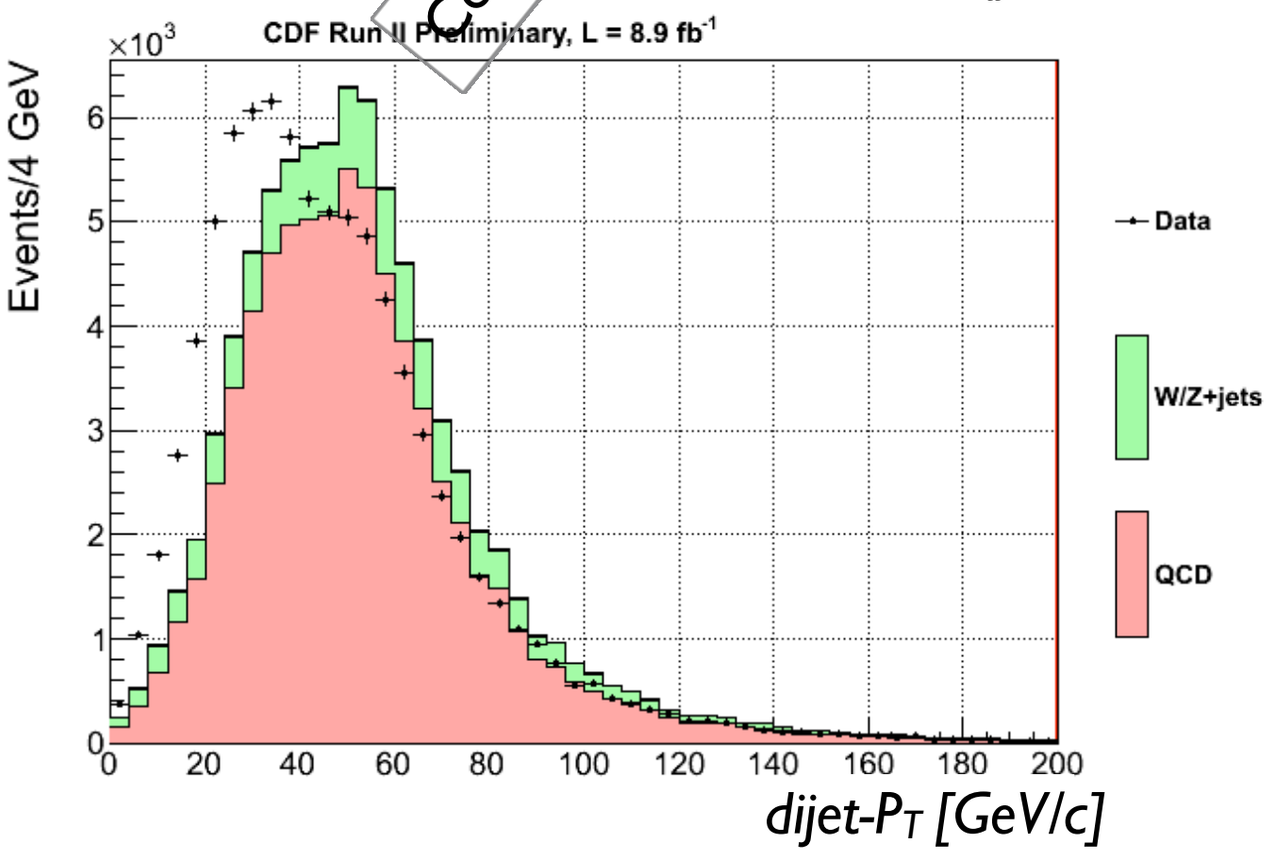
Control region

- *Defined to enhance the fake-leptons contamination in the sample*
- Need to magnify the symptoms to spot the problems!



QCD multi-jets background





WRONG
WAY
GO BACK

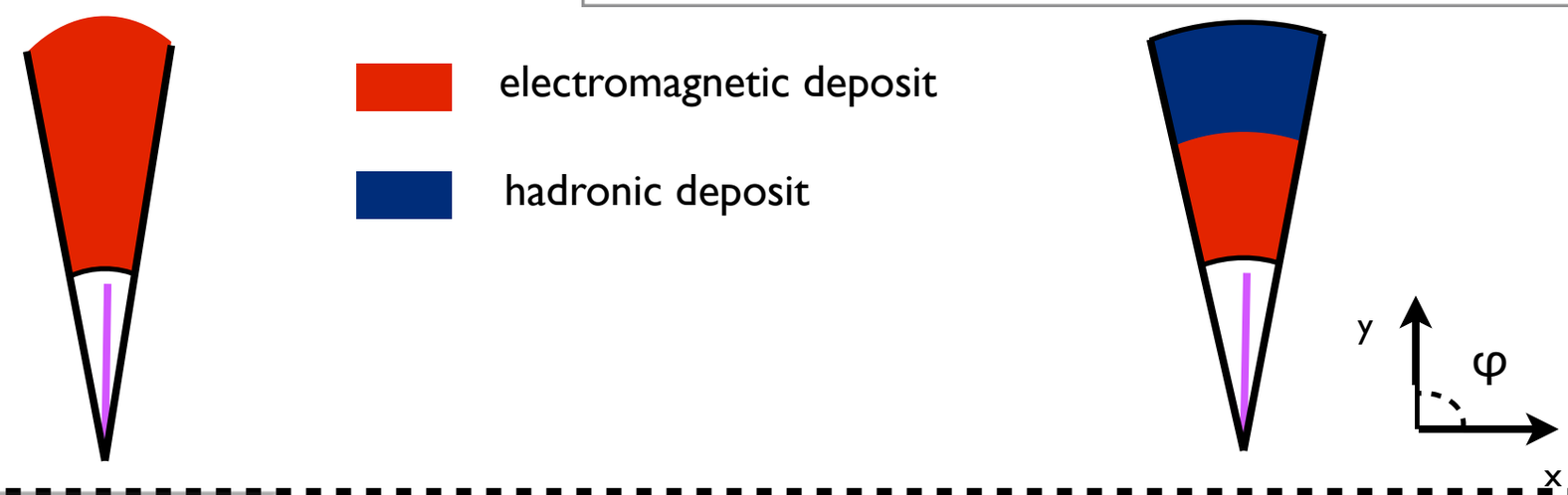
QCD multi-jets background

- *Quick Reminder:* differences between electron candidates and non-electrons

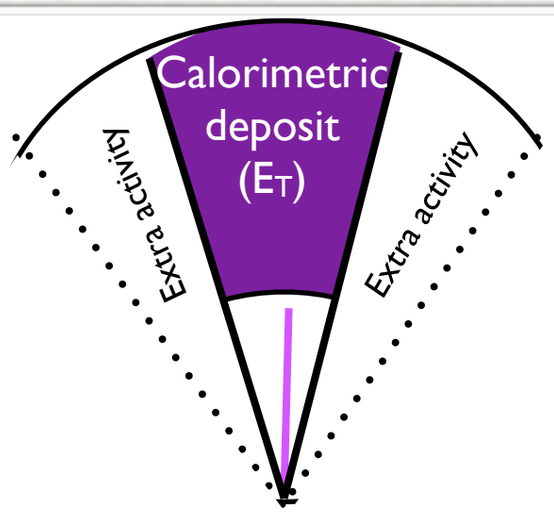
Electron candidate

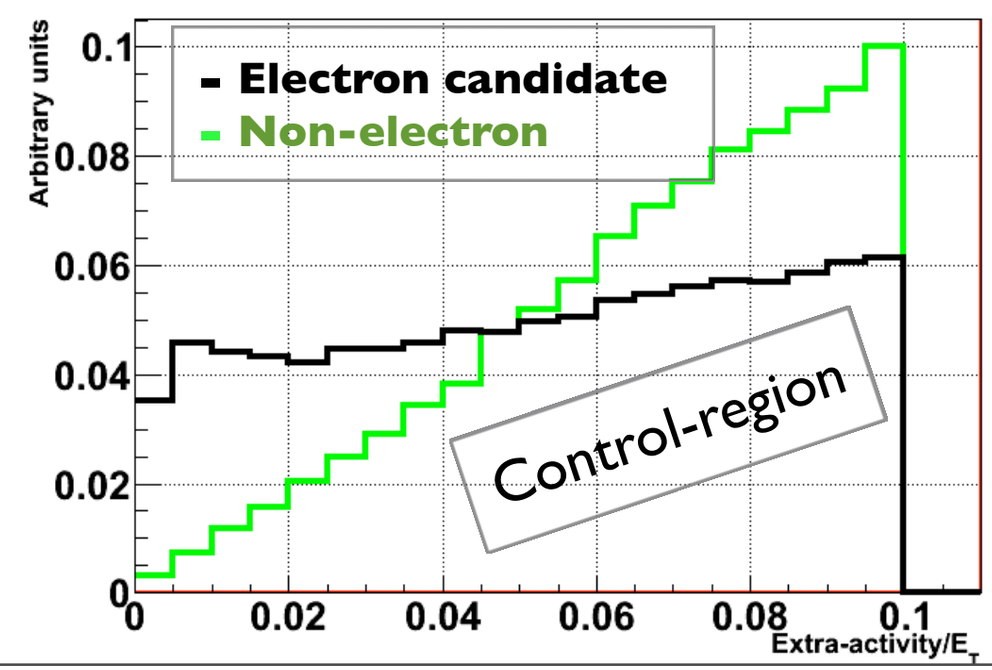
Non-electron
(failing two identification requirements)

1. Electromagnetic fraction



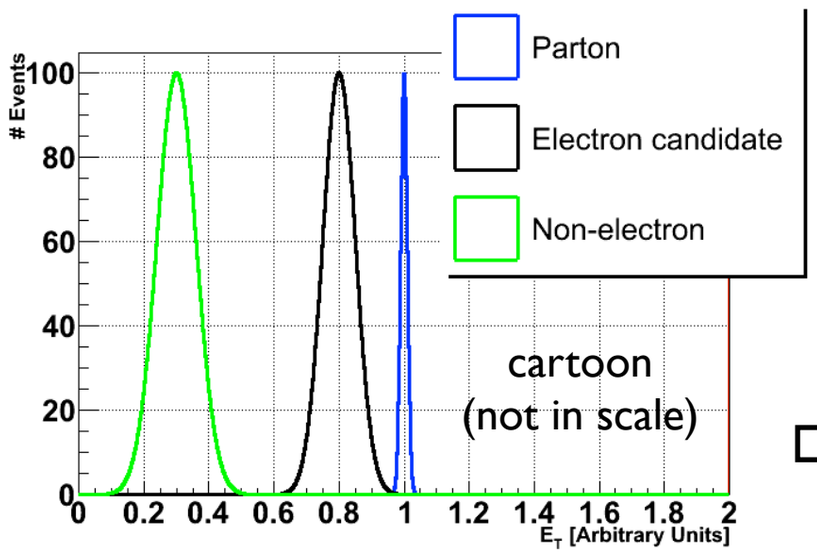
2. Extra activity outside the calorimetric cluster

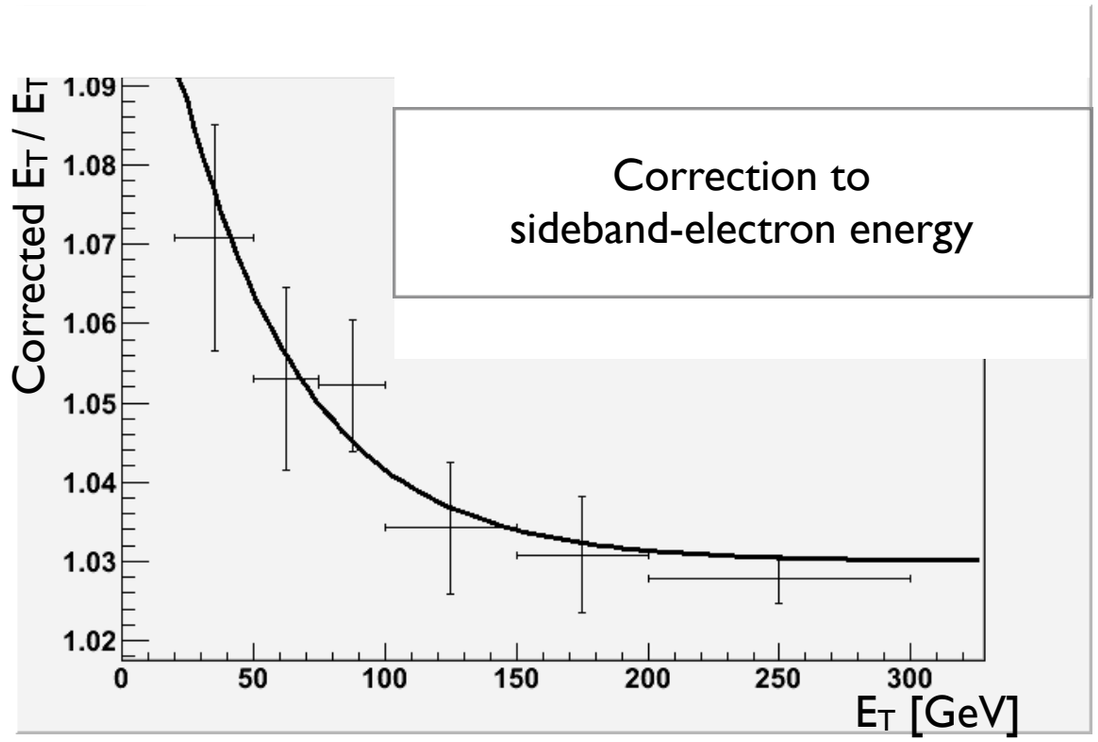


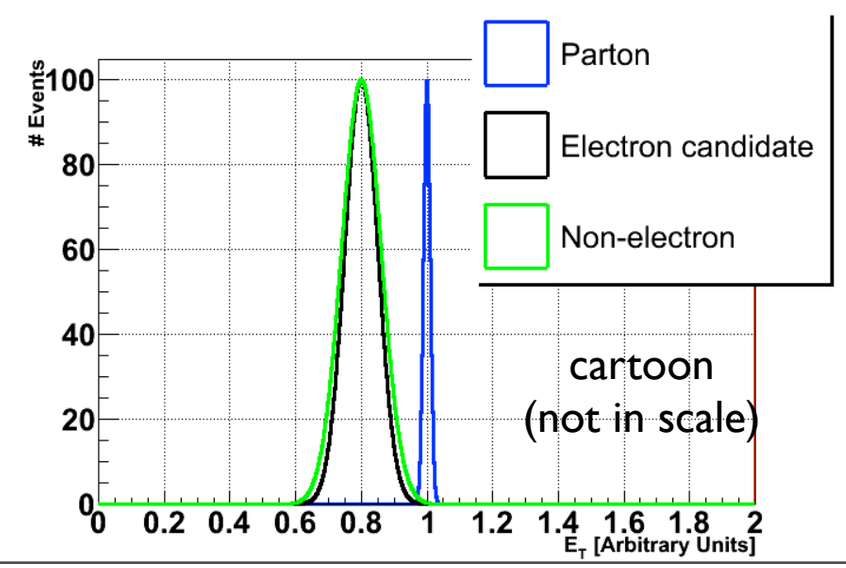


QCD multi-jets background

- *Question*: given the same parton energy, does non-electron energy properly model candidate-electron energy?
- *Answer*: from MC we see that non-electron energy needs to be corrected



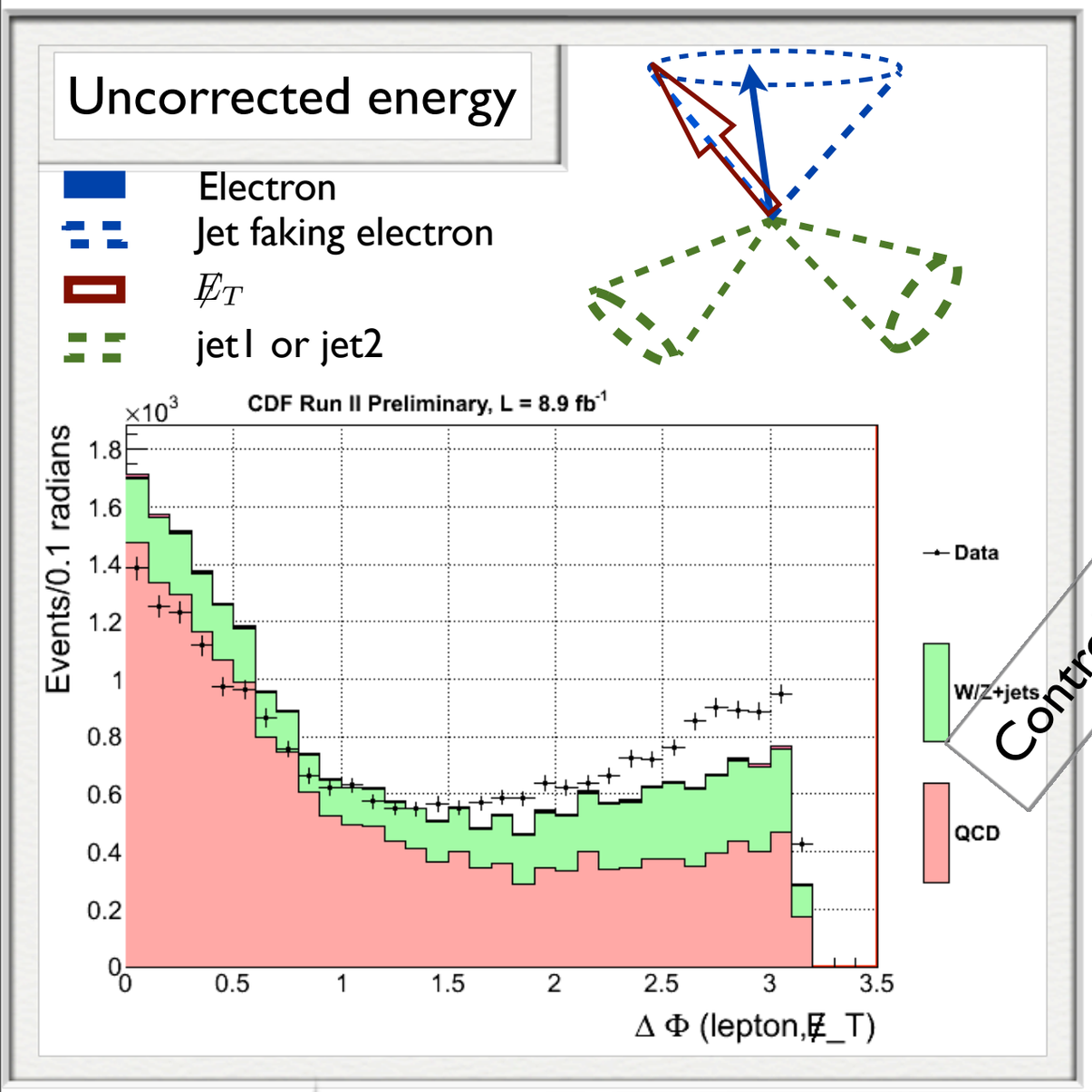




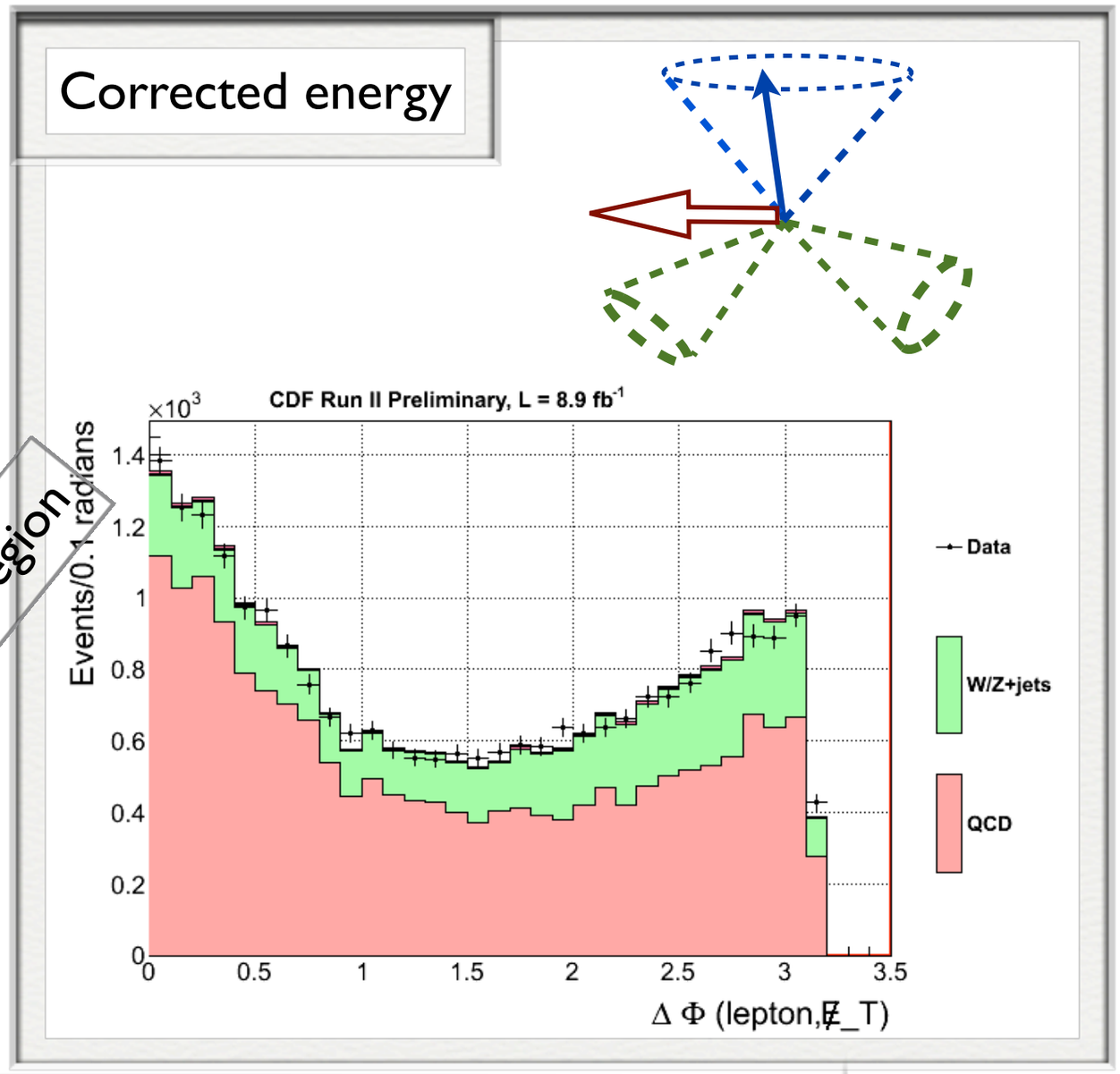
Corrections for non-electron energies: 3-8%

QCD multi-jets background

- Validation of the non-electron energy corrections
  - **Note**: Corrections affect the magnitude and the direction of the missing transverse energy



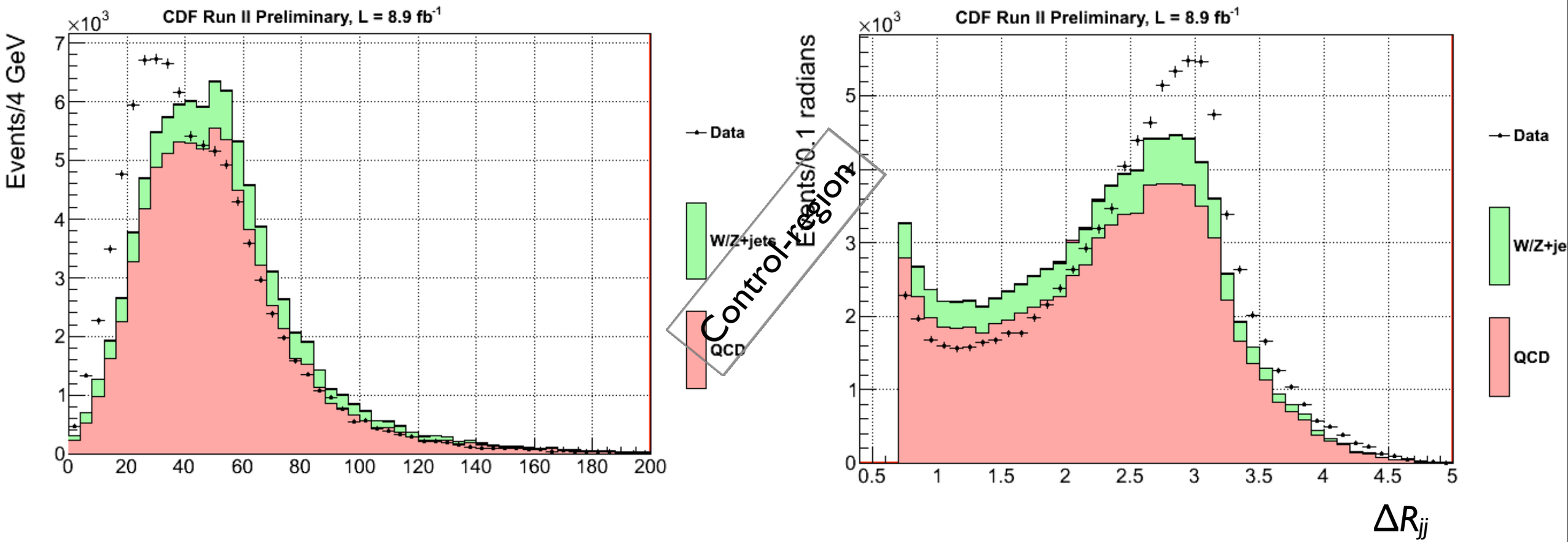
Control-region



Large improvements when corrections are applied

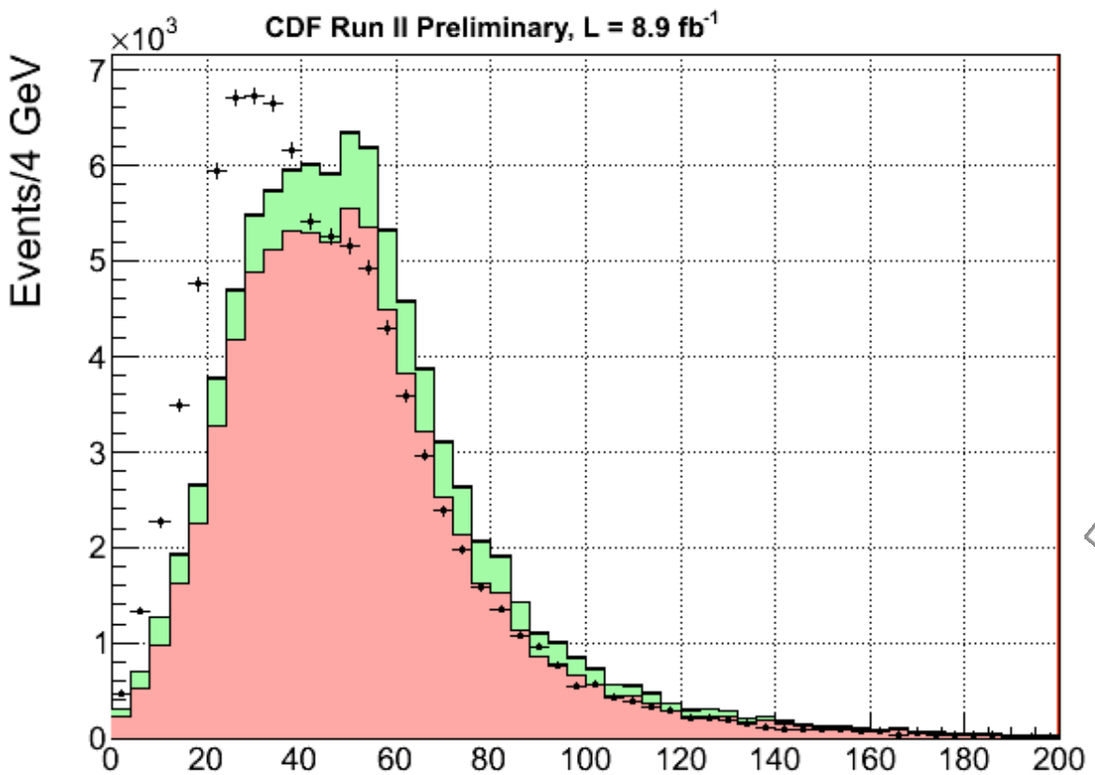
QCD multi-jets background

- Very nice, but disagreement in other distributions

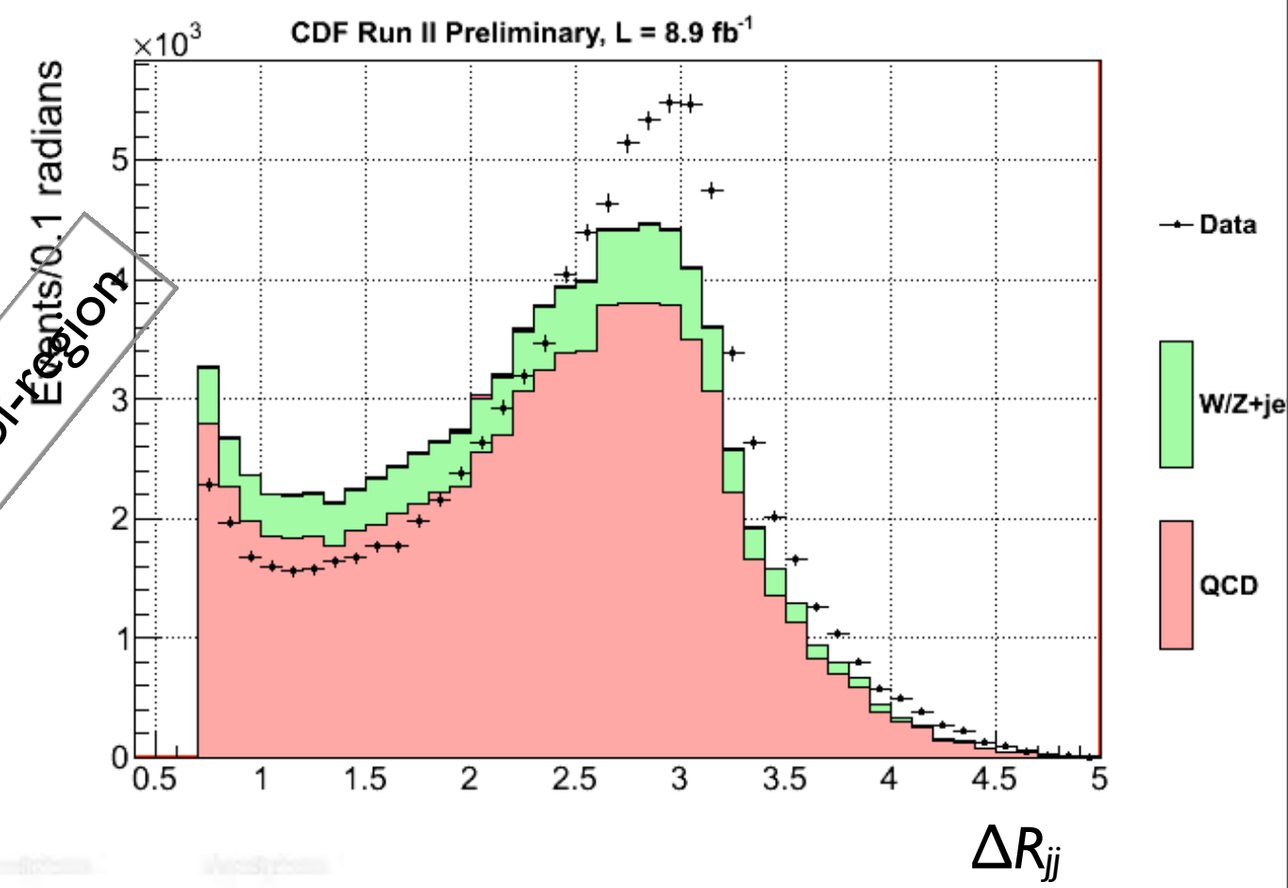


QCD multi-jets background

- Very nice, but disagreement in other distributions



Control-region



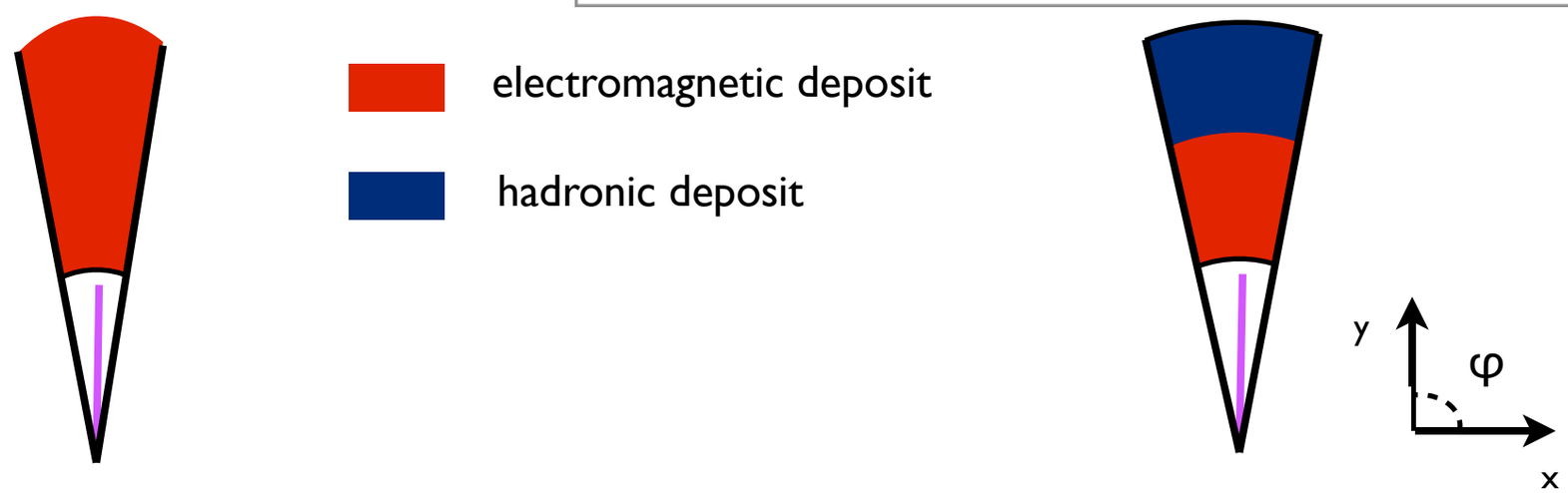
QCD multi-jets background

- *(Again) quick reminder:* differences between electron candidates and non-electrons

Electron candidate

Non-electron
(failing of two identification requirements)

1. Electromagnetic fraction



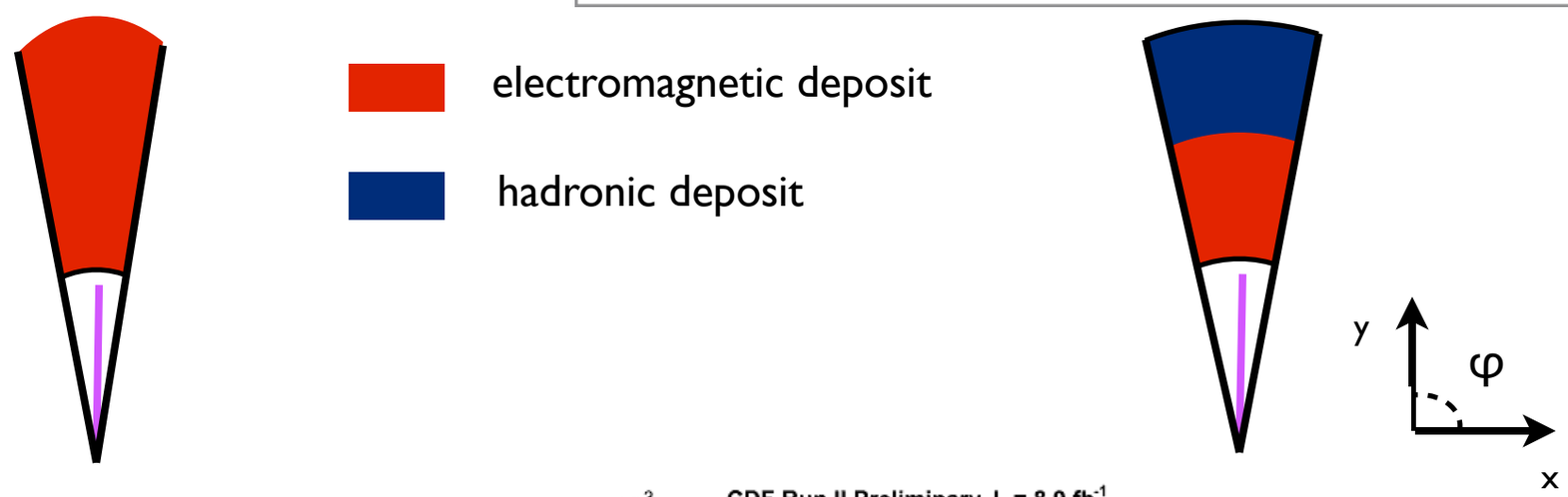
QCD multi-jets background

- *(Again) quick reminder:* differences between electron candidates and non-electrons

Electron candidate

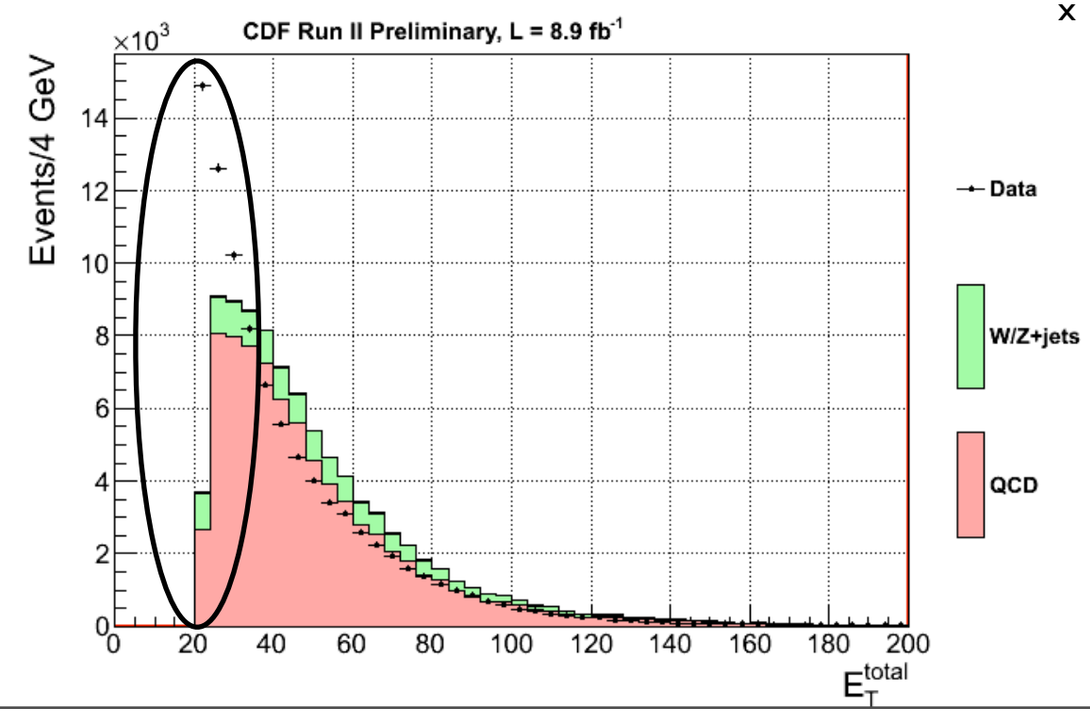
Non-electron
(failing two identification requirements)

1. Electromagnetic fraction



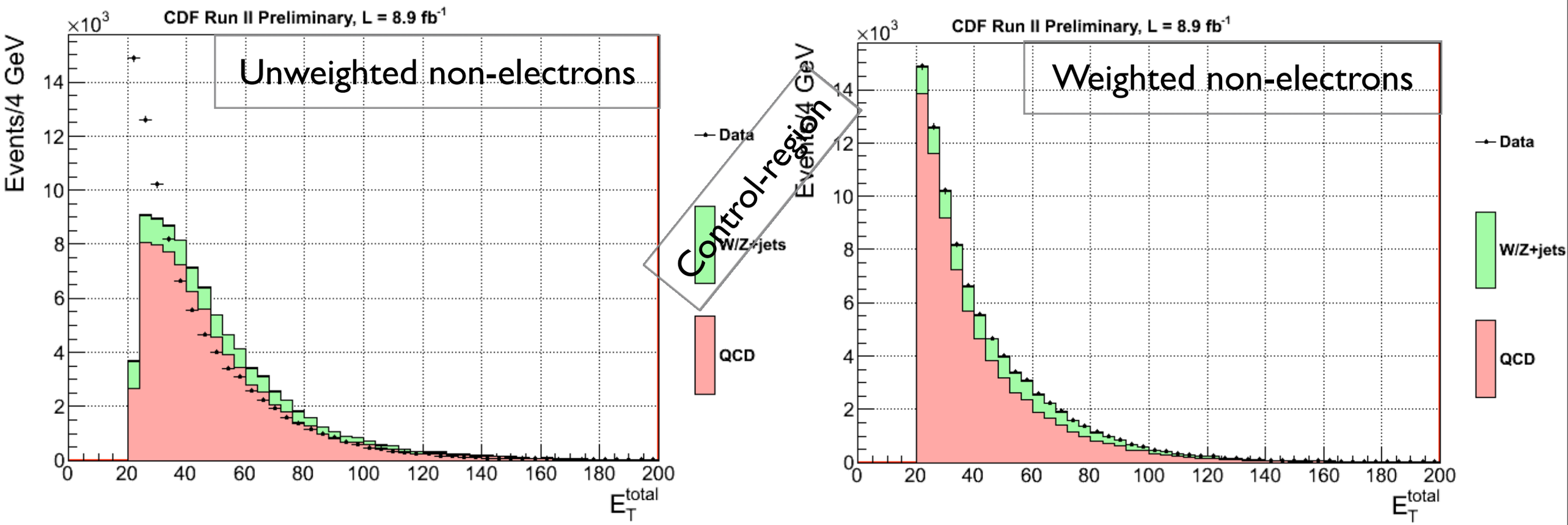
- Electron-trigger cuts on $E_T^{electromagnetic}$ rather than E_T^{total}

- lower trigger efficiency for QCD model wrt QCD when approaching the trigger threshold (~20 GeV)



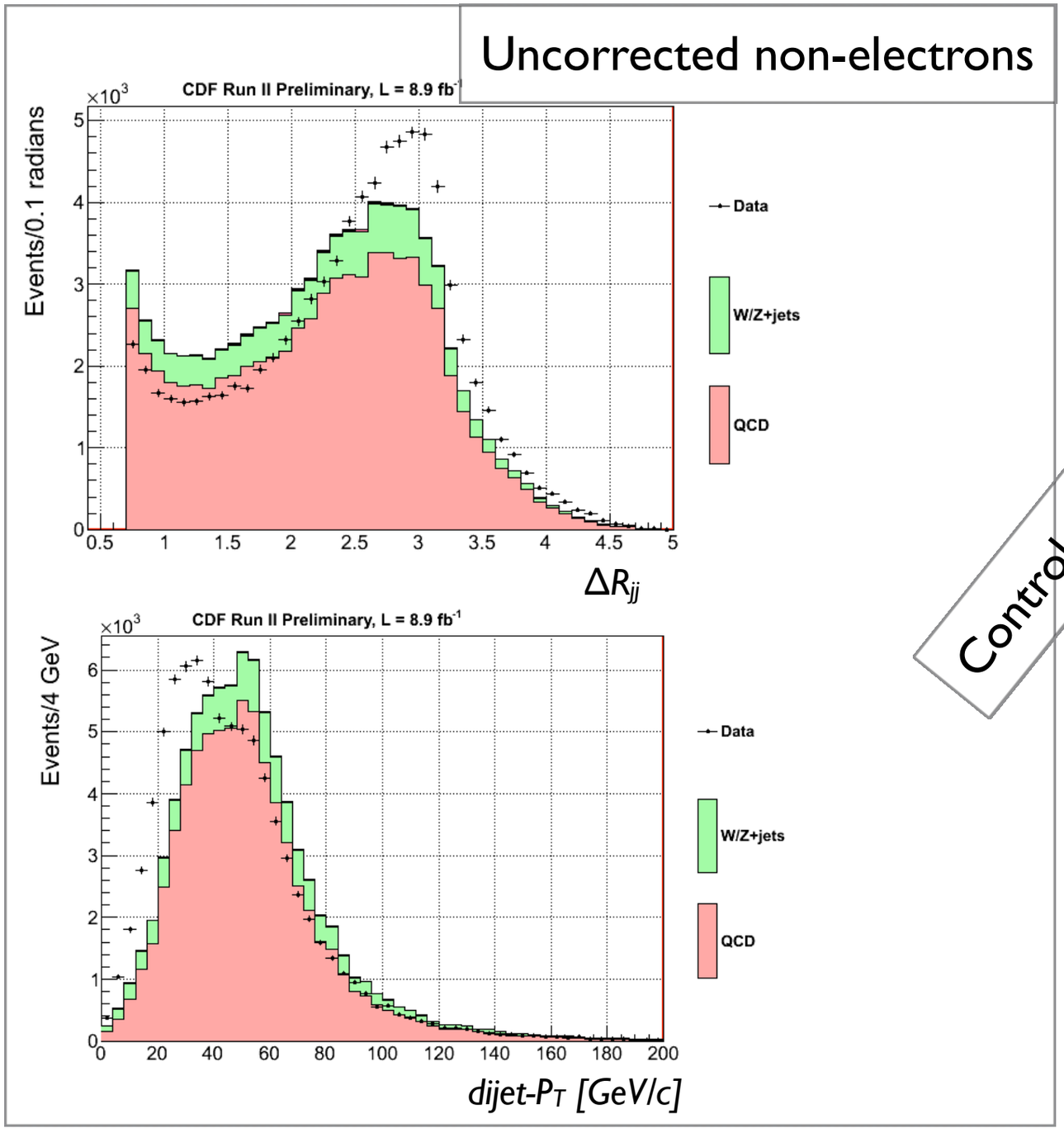
QCD multi-jets background

- To remove the trigger bias we determine a re-weighting in the control region and apply it in the signal region

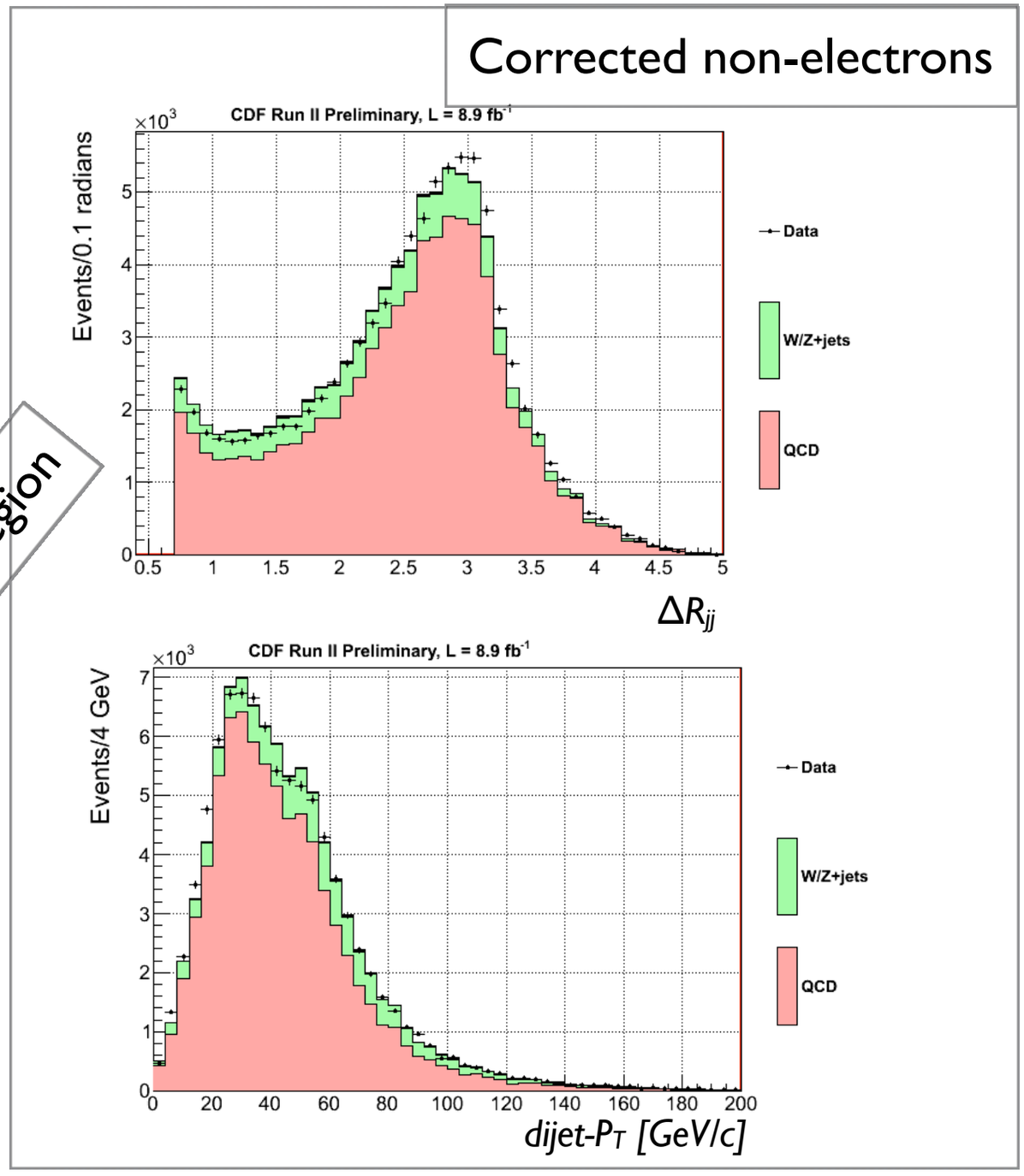


QCD multi-jets background

- We compare shapes before and after
  - non-electron energy corrections and
  - trigger-bias removal

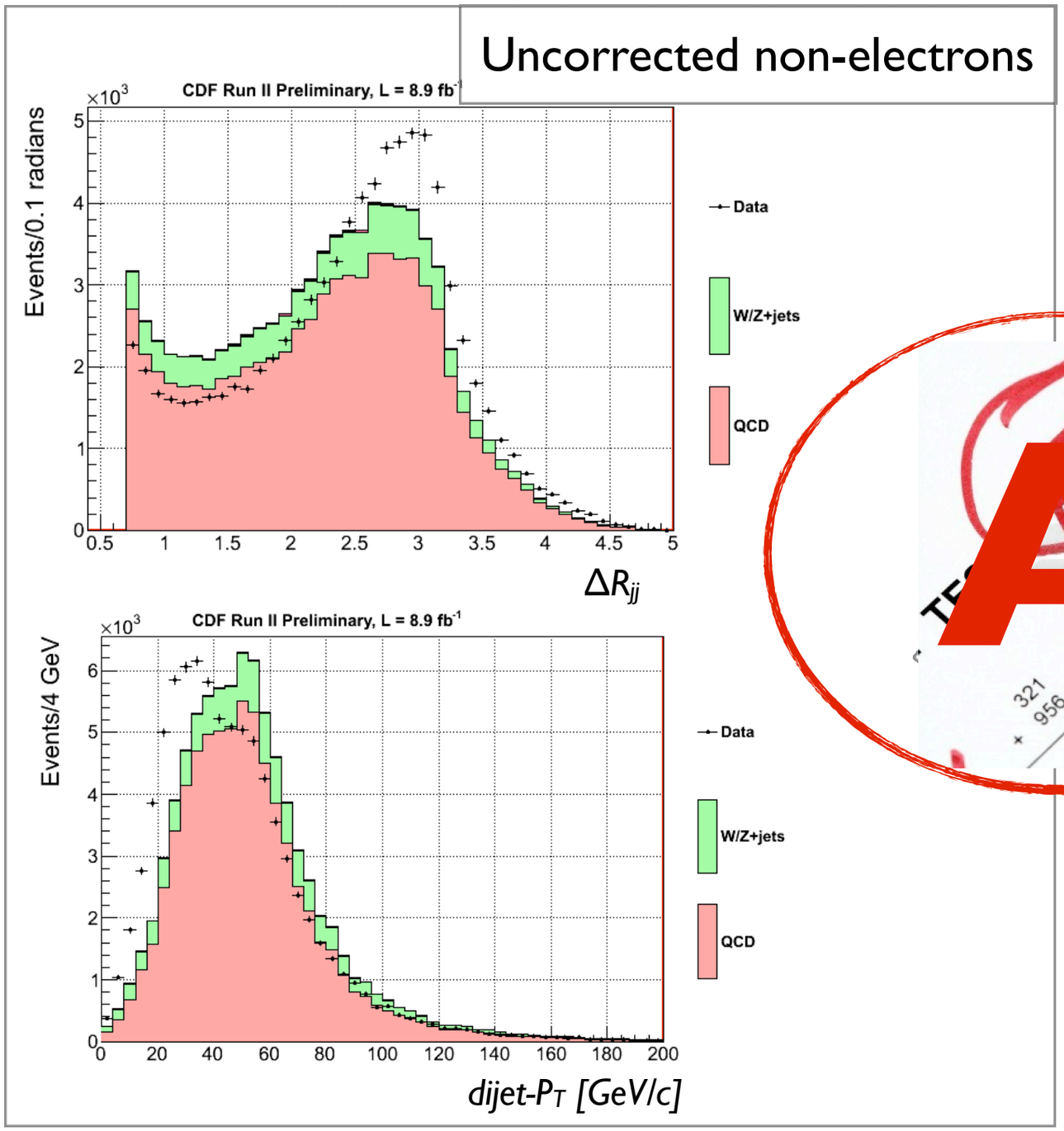


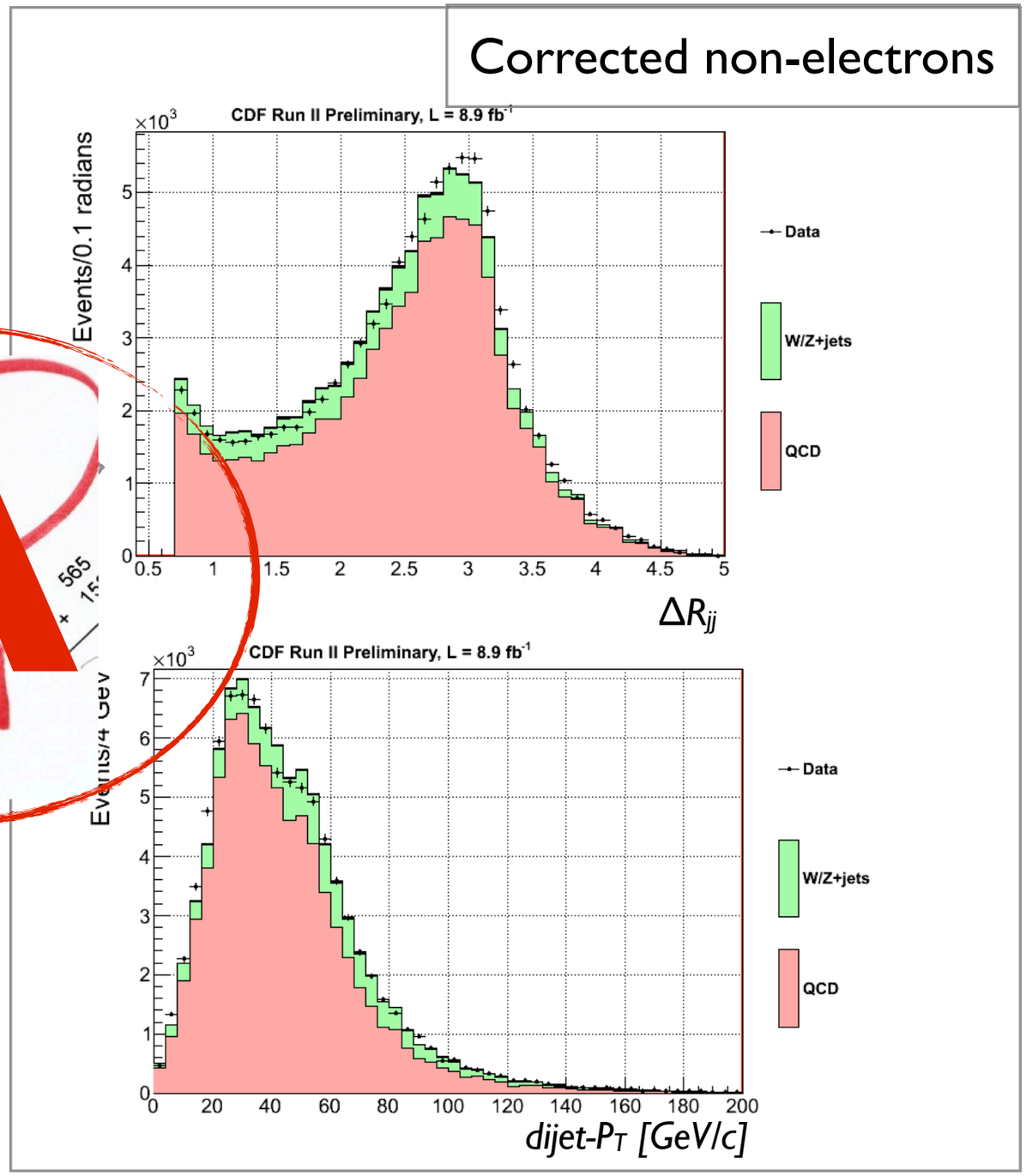
Control-region

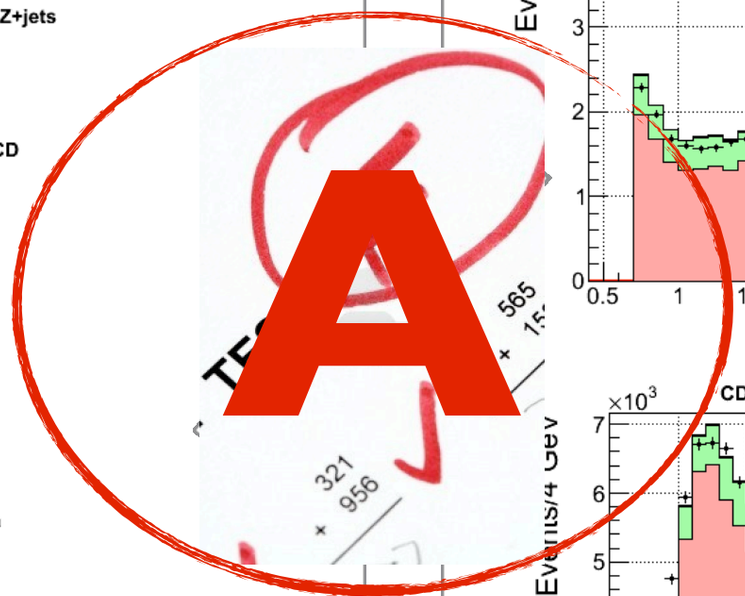


QCD multi-jets background

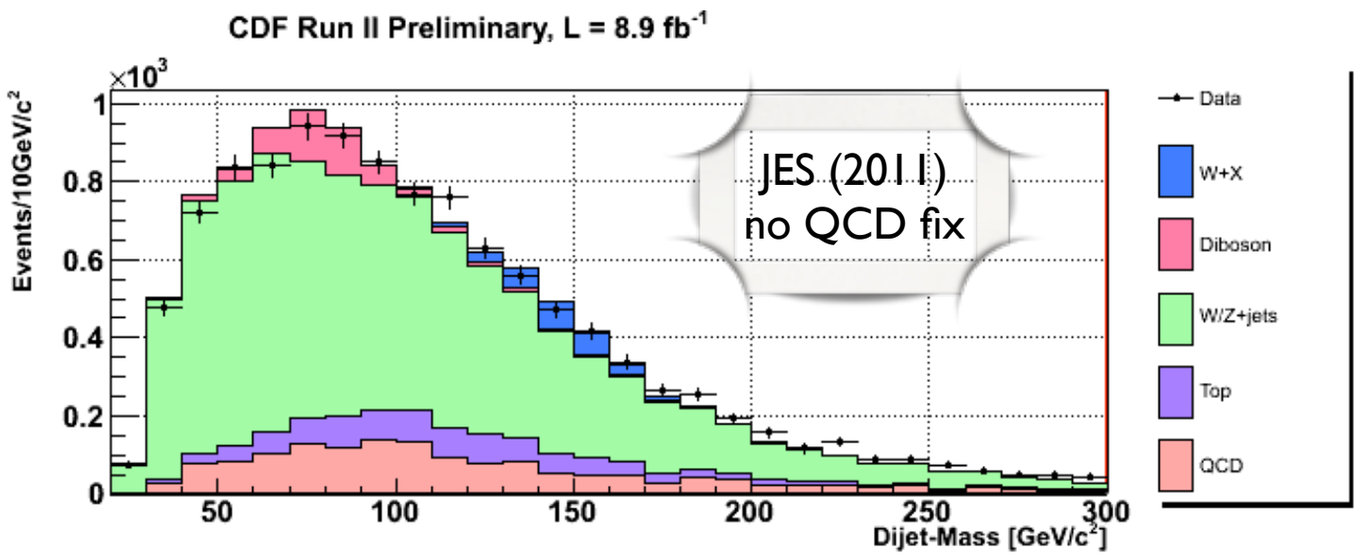
- We compare shapes before and after
  - non-electron energy corrections and
  - trigger-bias removal

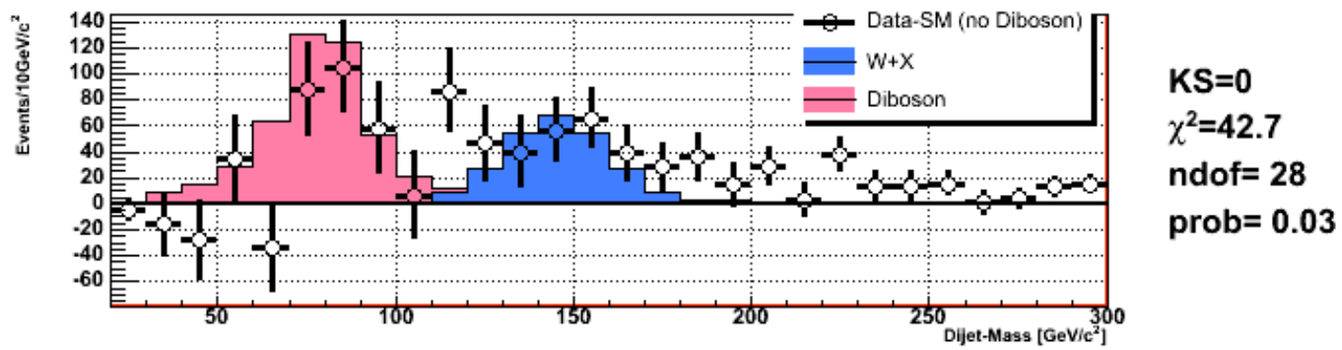


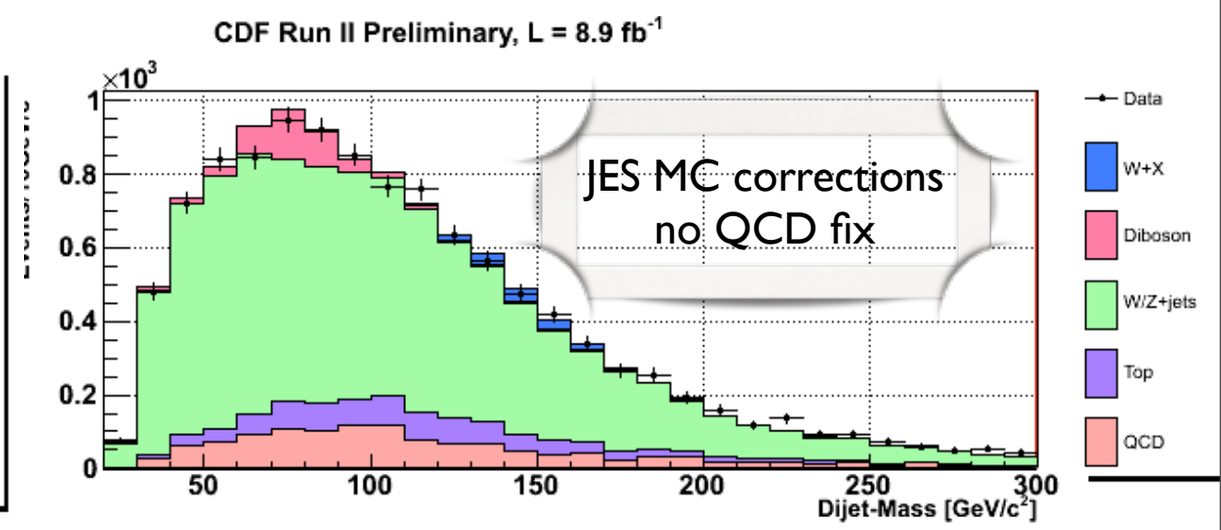


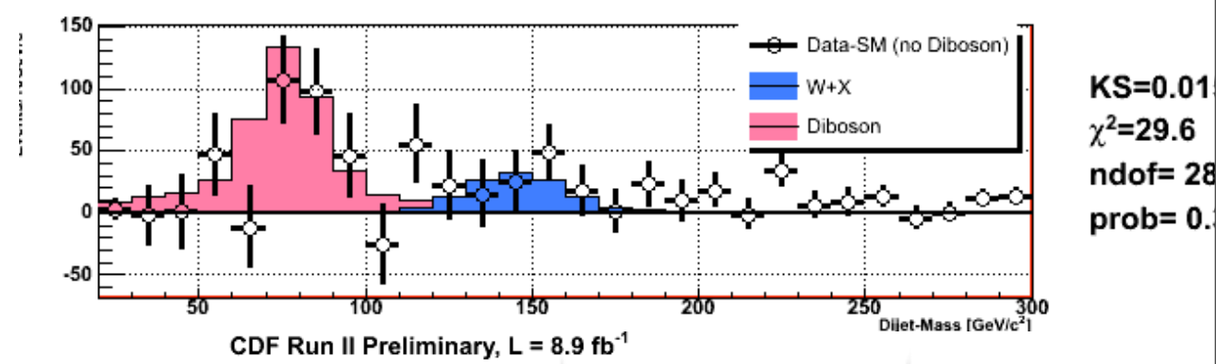


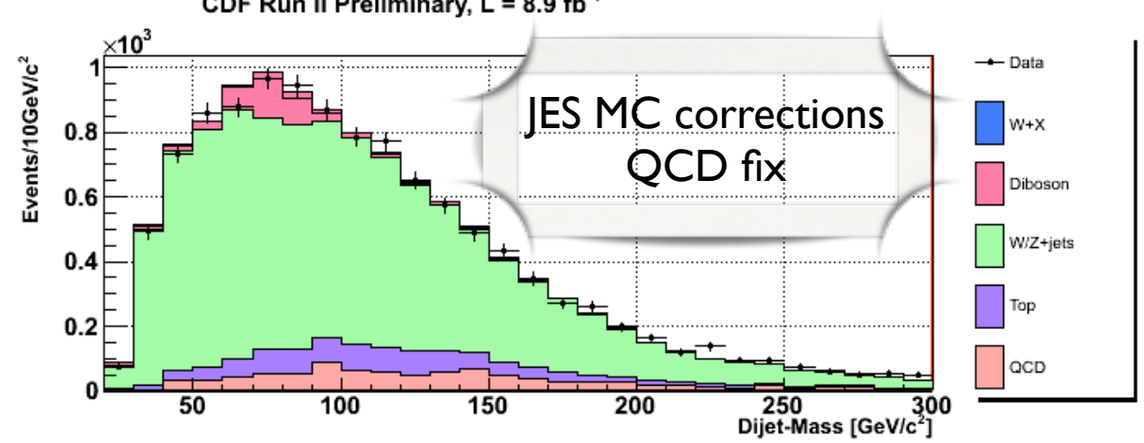
QCD multi-jets background

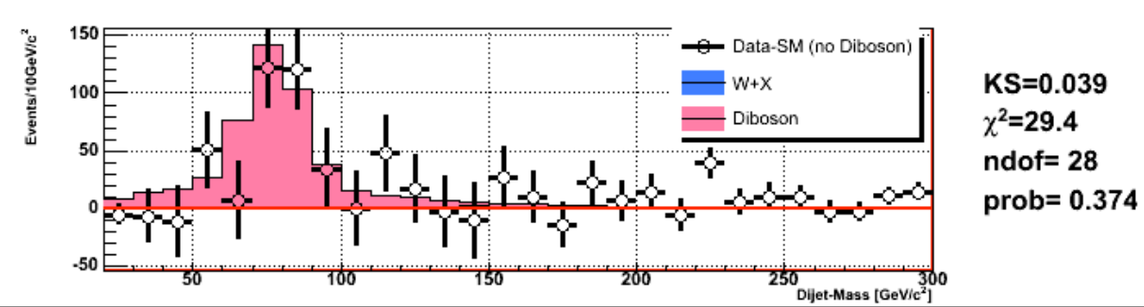












After:

- (JES) MC corrections
- QCD fix

consistency between the **electron and muon** samples

Fitting procedure

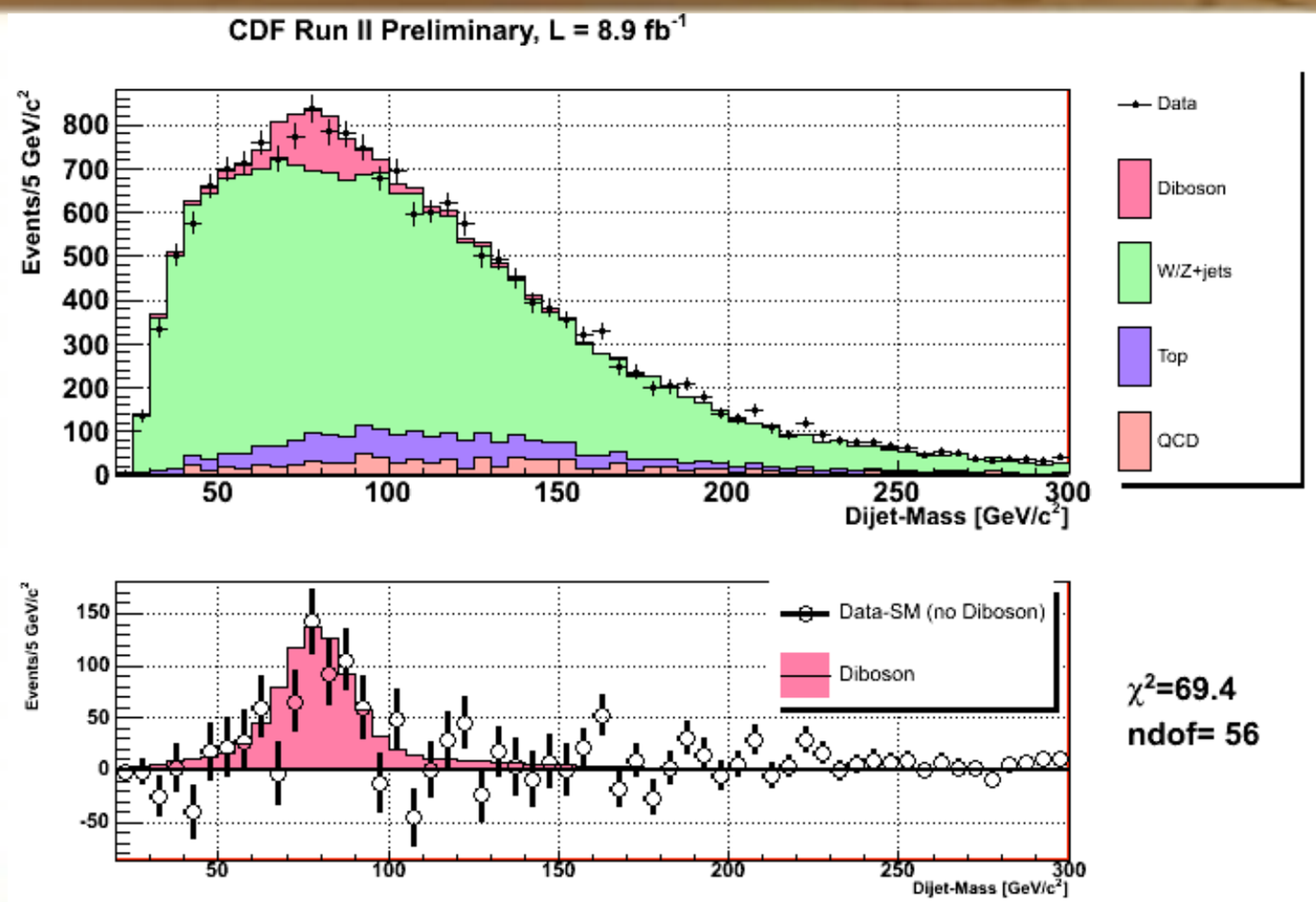
Now that all the pieces are in place
I'd like to define our fitting procedure:

- Fit dijet mass using 4 templates:
  - W/Z+jets
  - Top
  - Diboson
  - QCD multi-jets background

- Maximization of binned likelihood function

- Nuisance parameters in the fit
  - Jet energy scale
  - Background shapes and normalization

Results

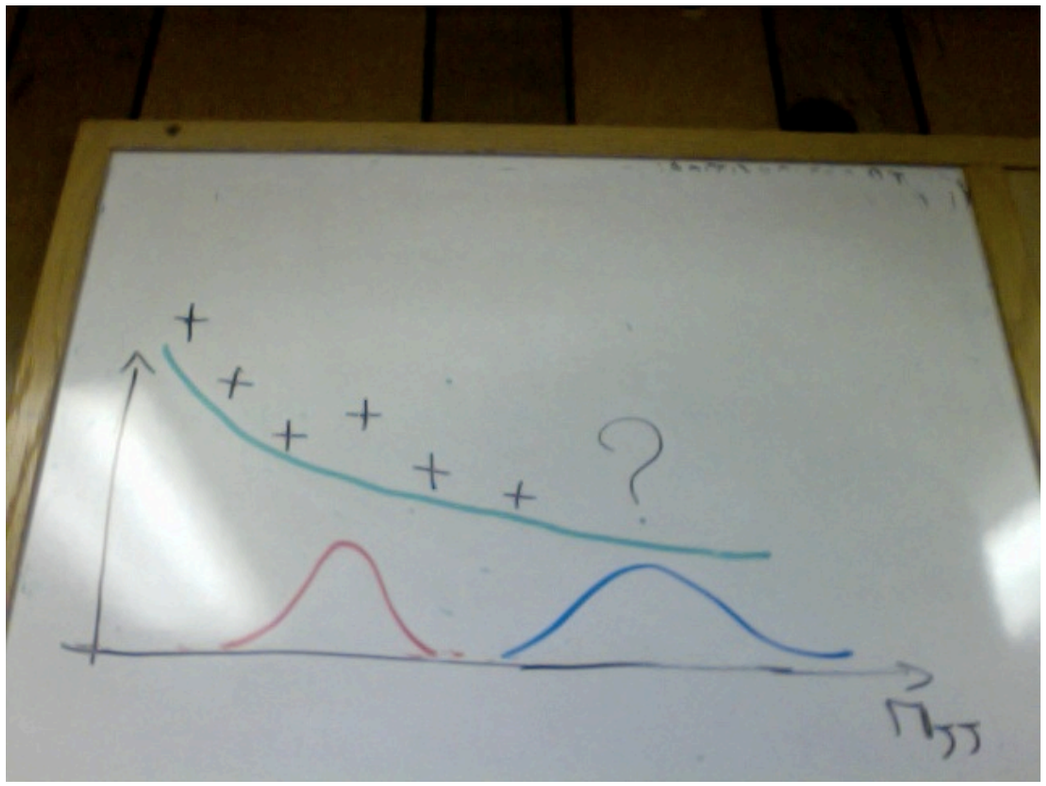


χ^2=69.4
ndof= 56

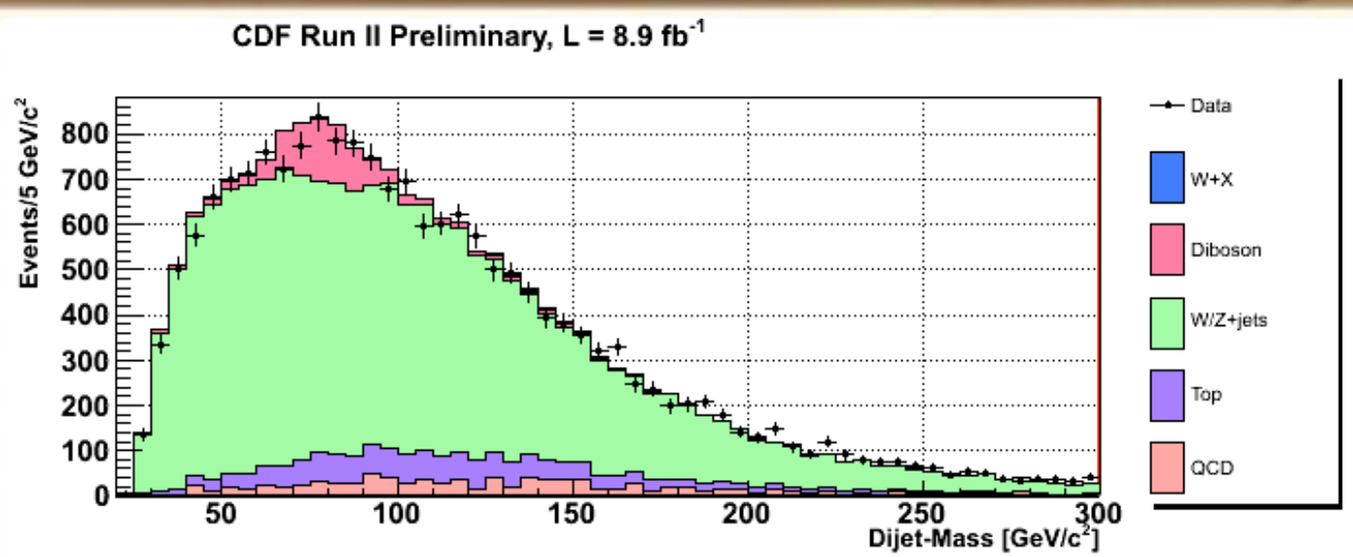
Good agreement between data and predictions

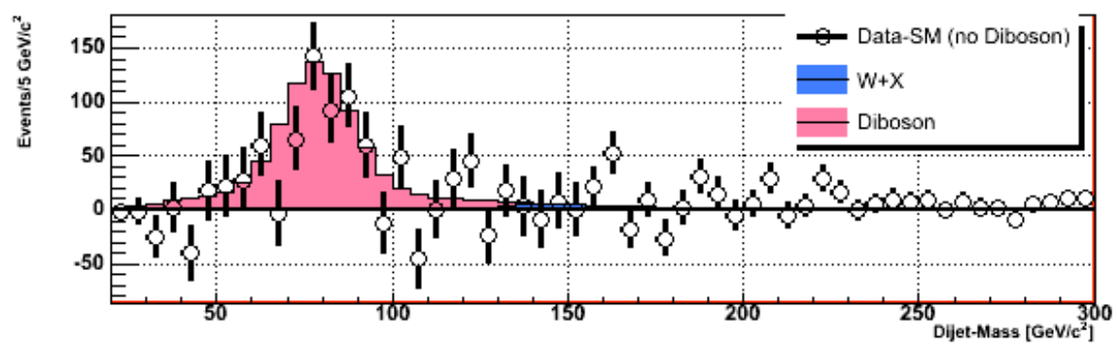
Fitting procedure

- Adding an additional template to the aforementioned 4 templates
  - Gaussian distribution with μ= 145 GeV, σ=14.3 GeV



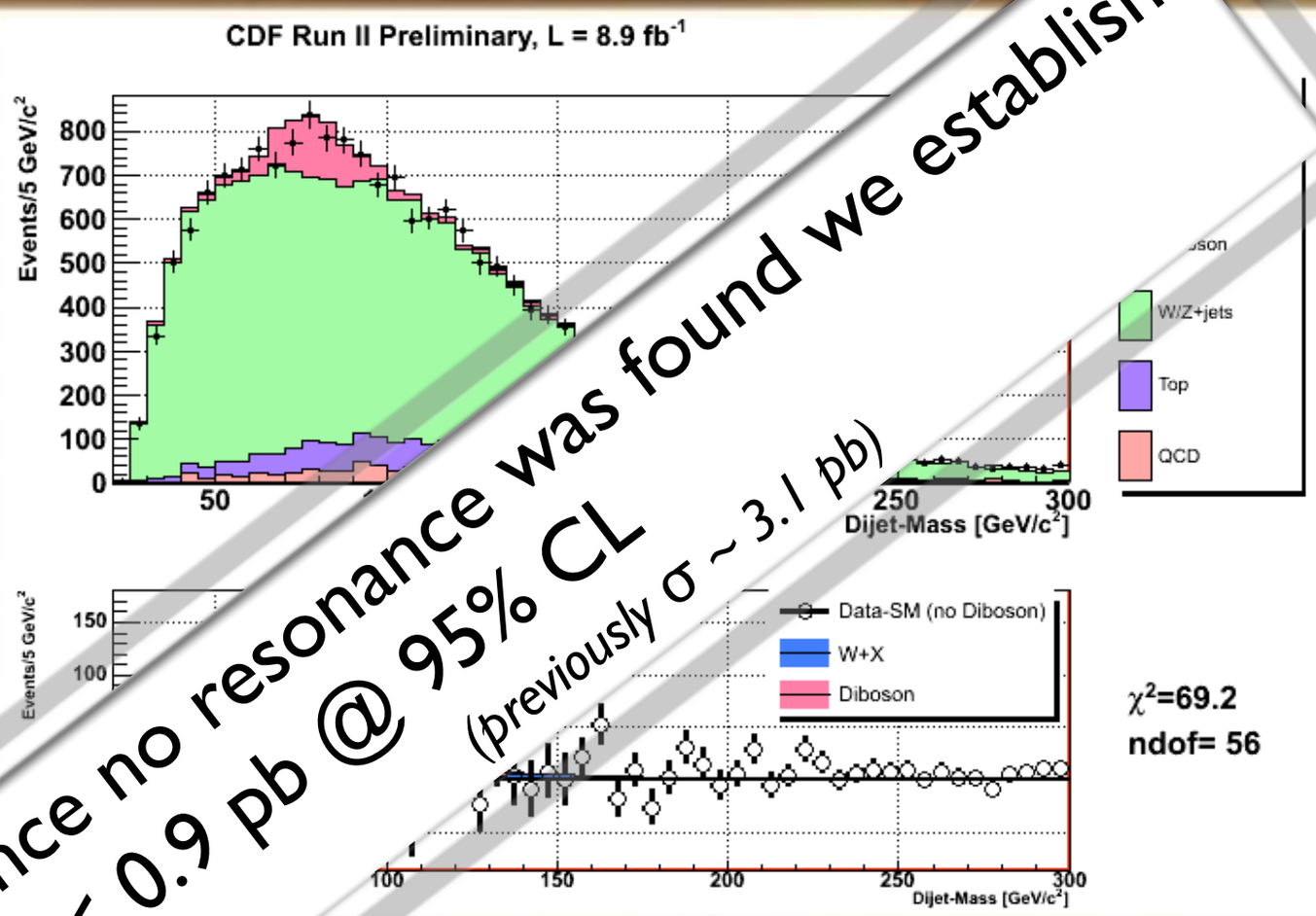
Results





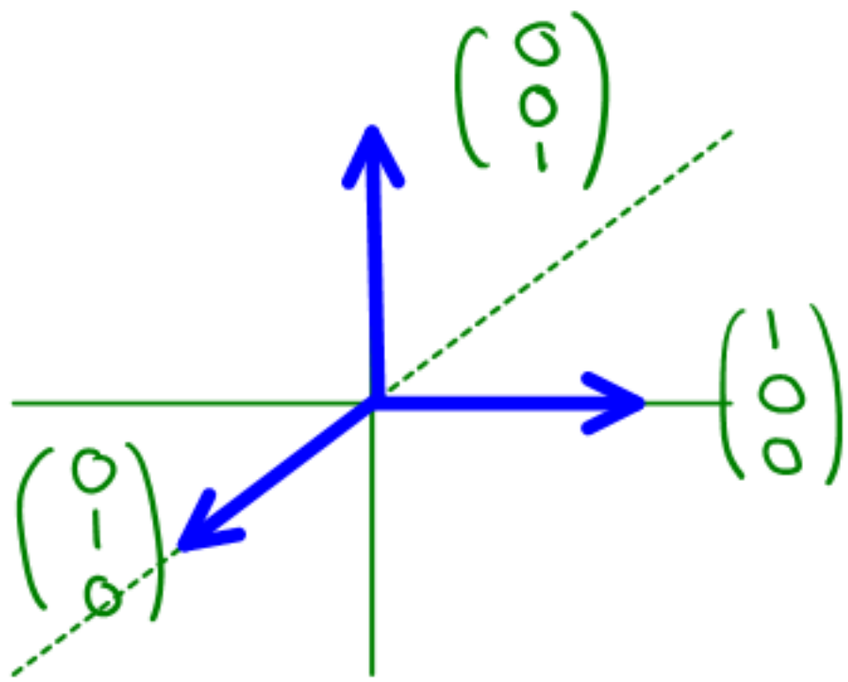
χ^2=69.2
ndof= 56

Results



Since no resonance was found we establish
$\sigma < 0.9$ pb @ 95% CL
(previously $\sigma \sim 3.1$ pb)

Studies in orthogonal samples



1. Two charged leptons + two jets

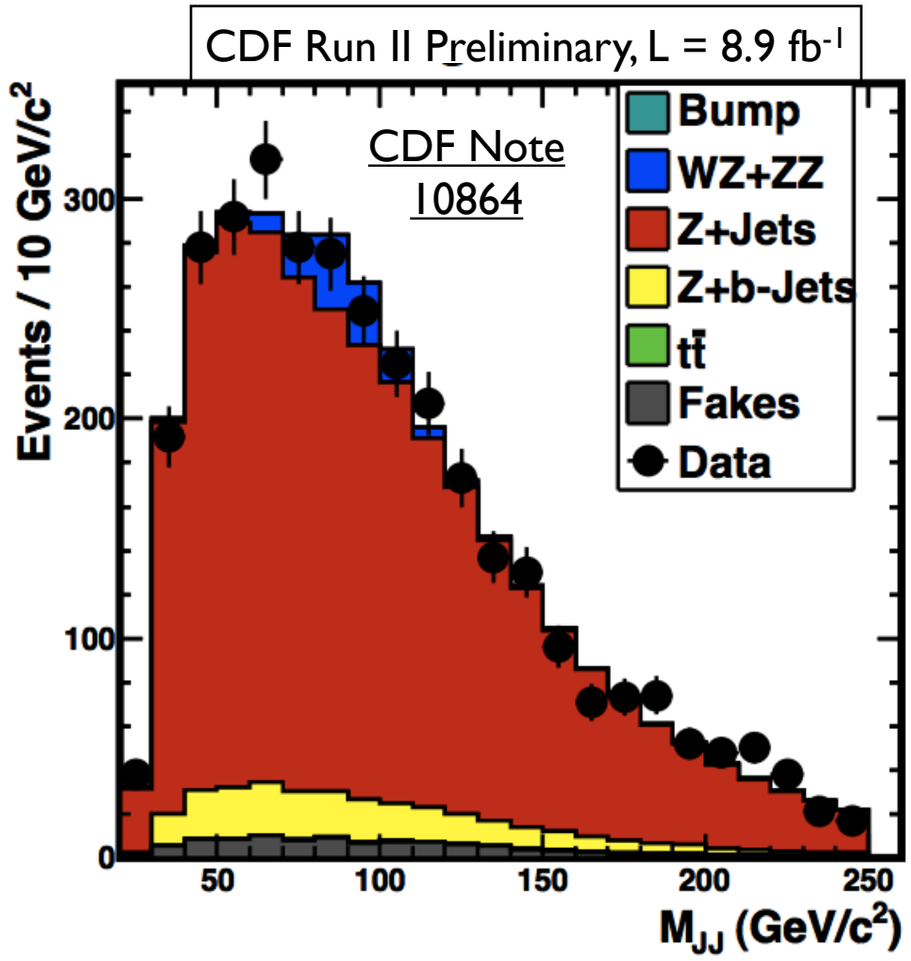
- Z+jets as dominating contributions

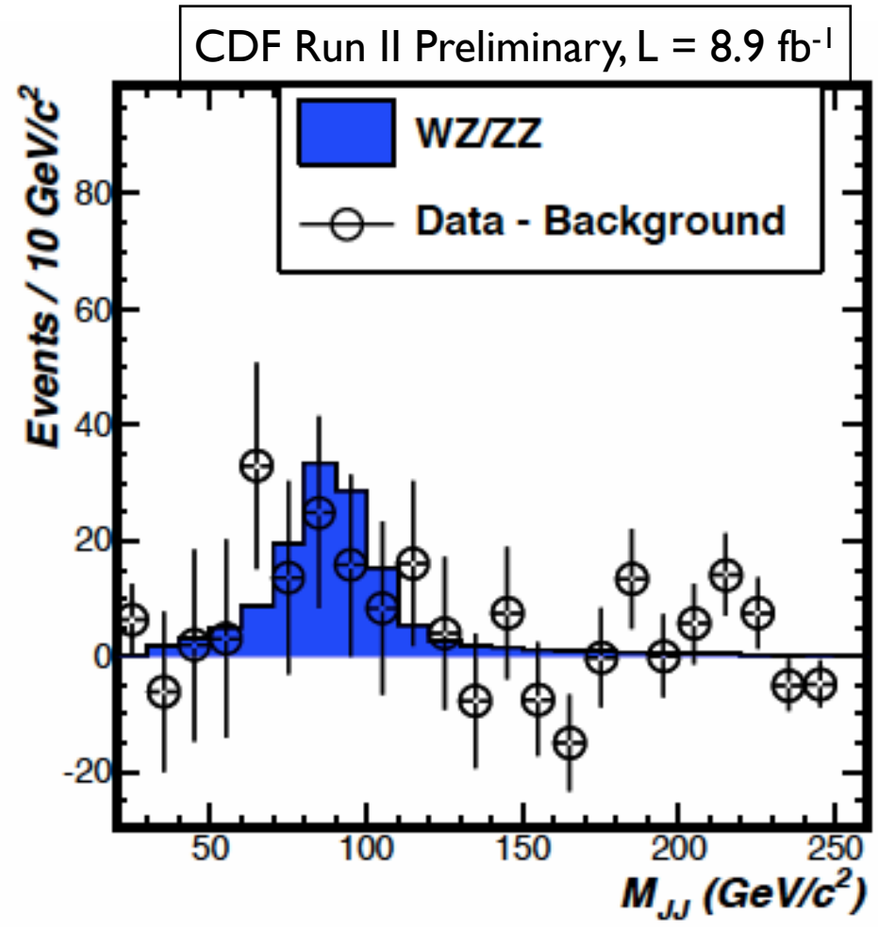
Selection

- =2 charged leptons with P_T >20 GeV/c
- 76<(M(lep1,lep2)/GeV/c2)<106
- $\not\!\!E_T$<20 GeV
- =2 jets with E_T >25 GeV
  - ΔR(jet1,jet2)>0.7

Summary of the aforementioned improvements: ***new-JES***

- new-JES is used as in W+2jets analysis
- different model for the small multi-jet background





1. Two charged leptons + two jets

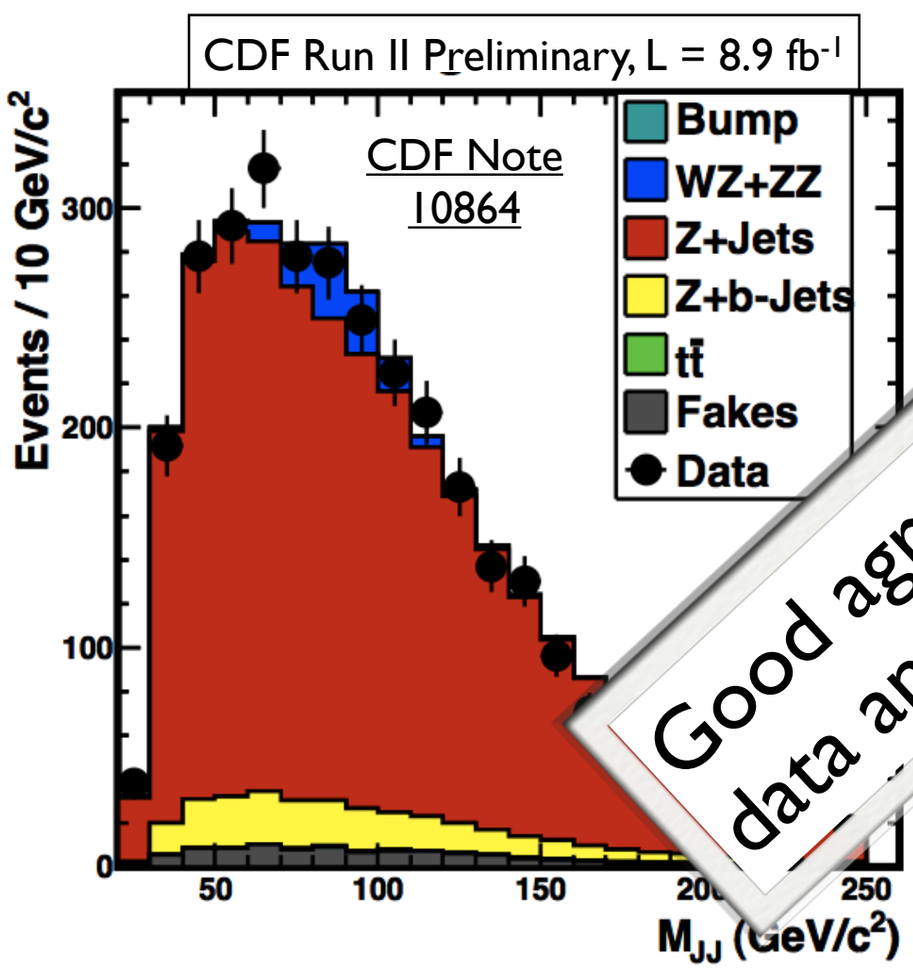
- Z+jets as dominating contributions

Selection

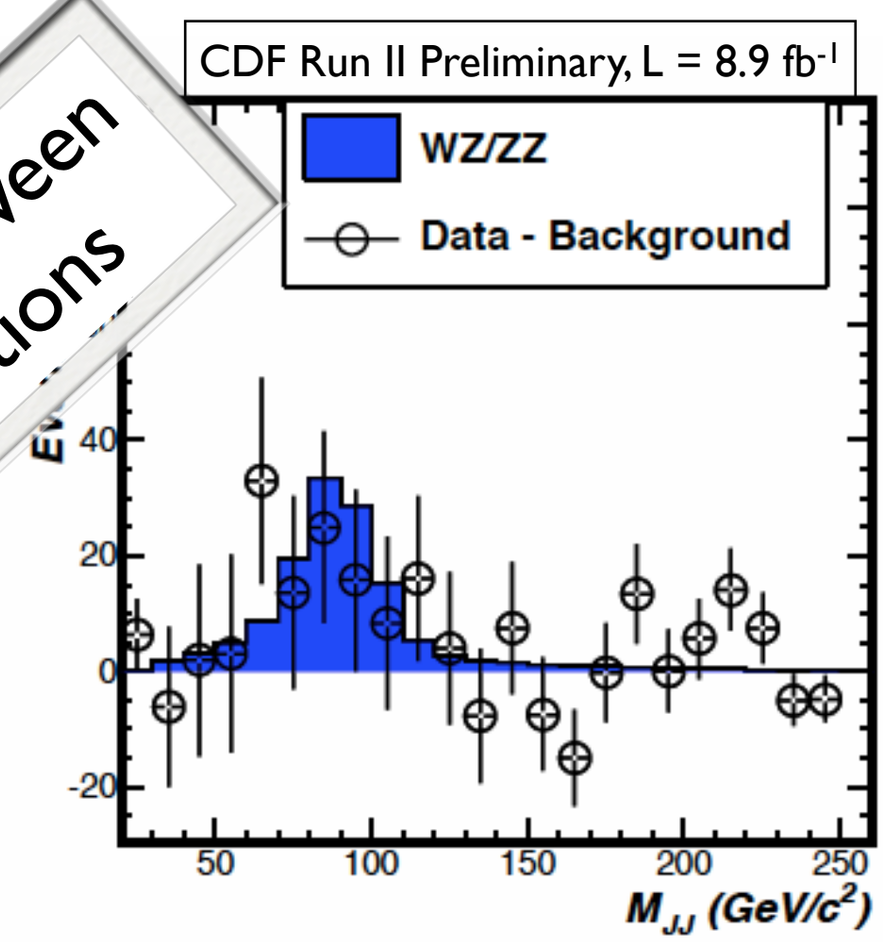
- =2 charged leptons with P_T >20 GeV/c
- 76<(M(lep1,lep2)/GeV/c2)<106
- $\not{E}_T$<20 GeV
- =2 jets with E_T >25 GeV
  - ΔR(jet1,jet2)>0.7

Summary of the aforementioned improvements: ***new-JES***

- new-JES is used as in W+2jets analysis
- different model for the small multi-jet background



Good agreement between data and SM predictions



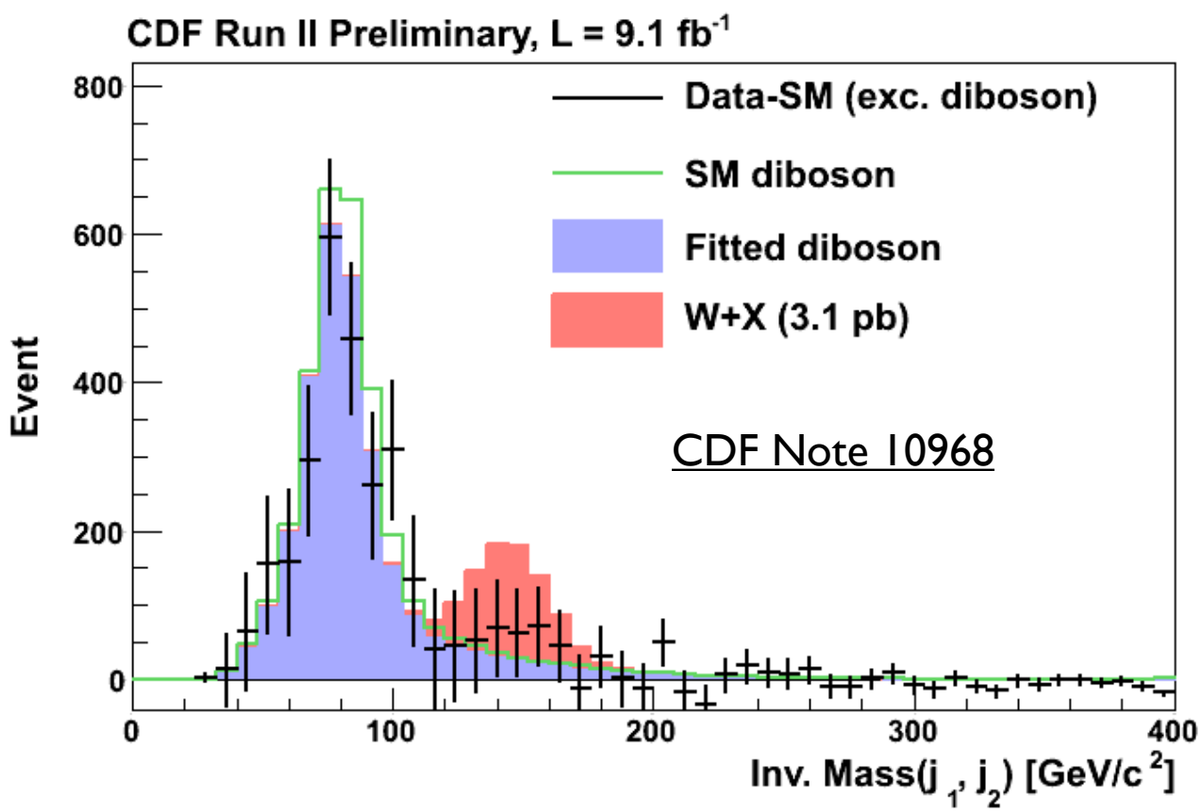
2. Large $\not{E}_T$ + two jets

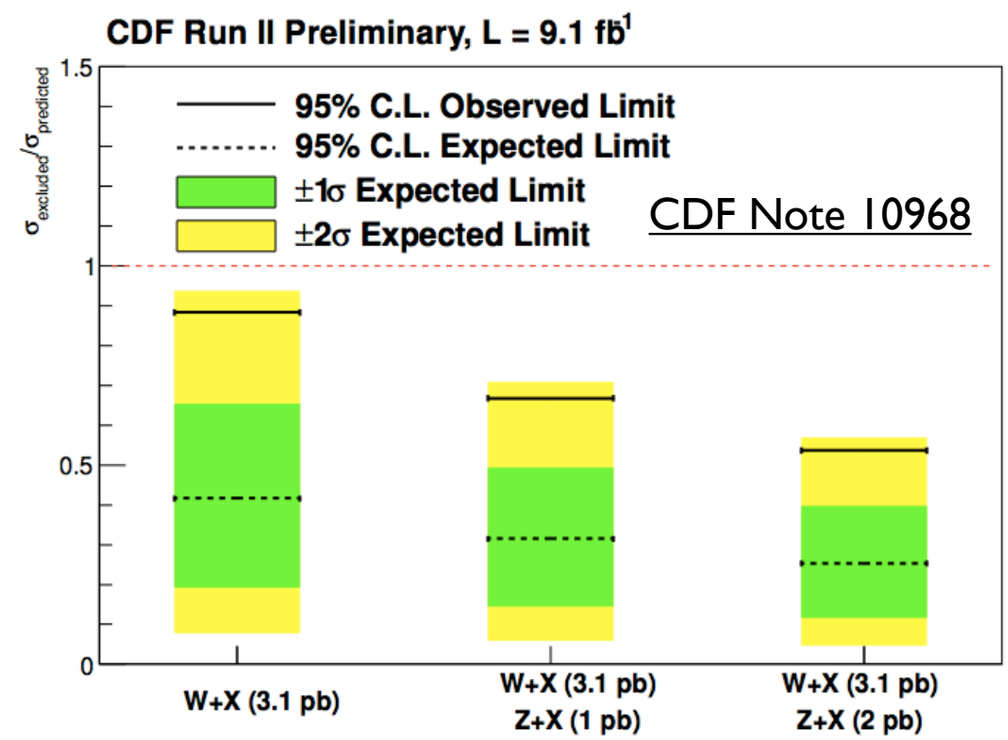
Selection

- no identified lepton
- MET>50 GeV
- =2, (3) jets with E_T >35, 25, (15) GeV
  - ΔR(jet1,jet2)>1, $|\eta_{j1 \text{ or } j2}|$<0.9

Summary of the aforementioned improvements: **none**

- No Q/G JES
  - lower gluon contamination due to the higher jet energies
- No QCD fix
  - different trigger exploited and no lepton to be faked





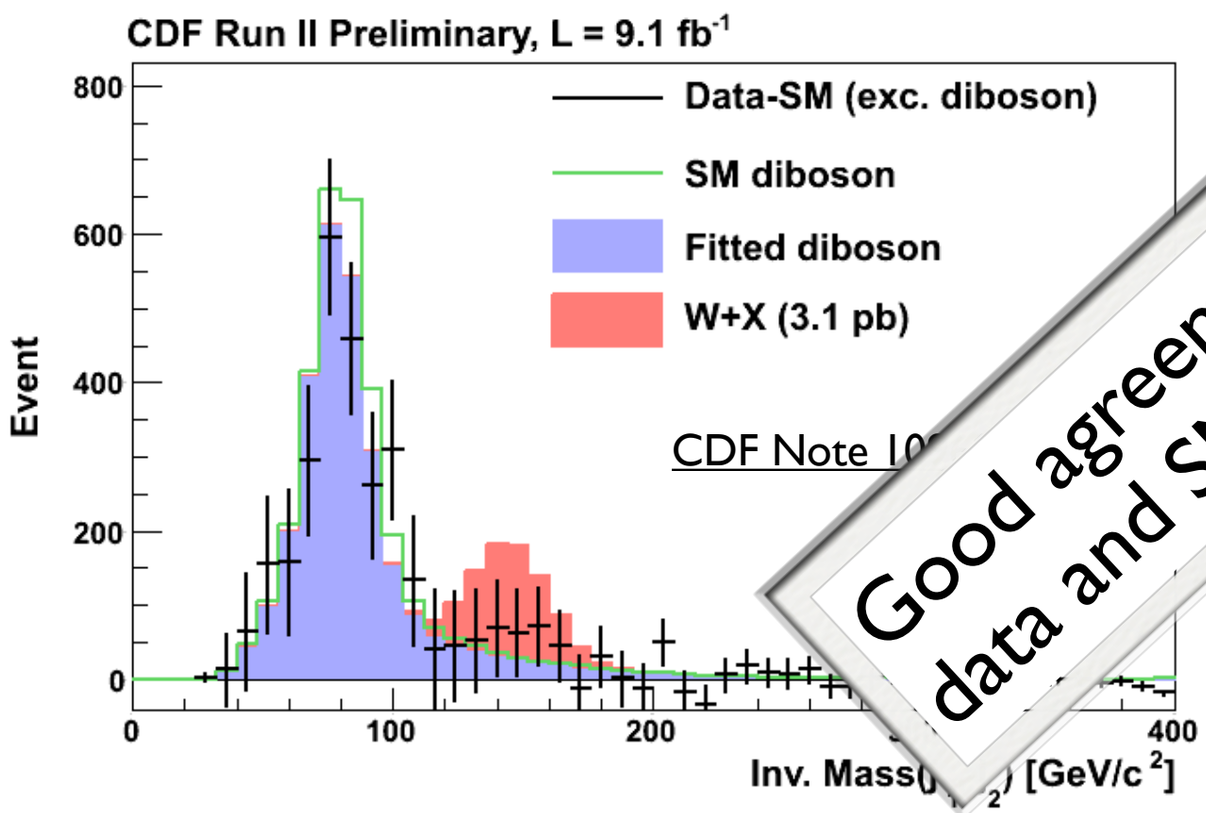
2. Large $\not{E}_T$ + two jets

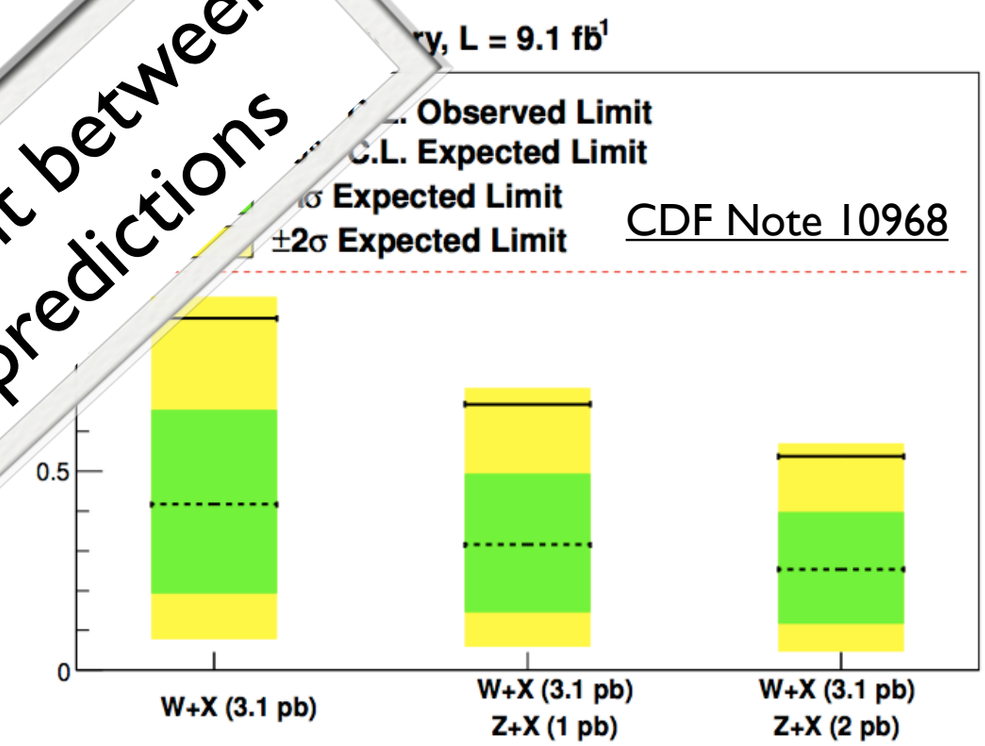
Selection

- no identified lepton
- MET>50 GeV
- =2, (3) jets with E_T >35, 25, (15) GeV
  - ΔR(jet1,jet2)>1, $|\eta_{j1 \text{ or } j2}|$<0.9

Summary of the aforementioned improvements: **none**

- No Q/G JES
  - lower gluon contamination due to the higher jet energies
- No QCD fix
  - different trigger exploited and no lepton to be faked





Good agreement between data and SM predictions

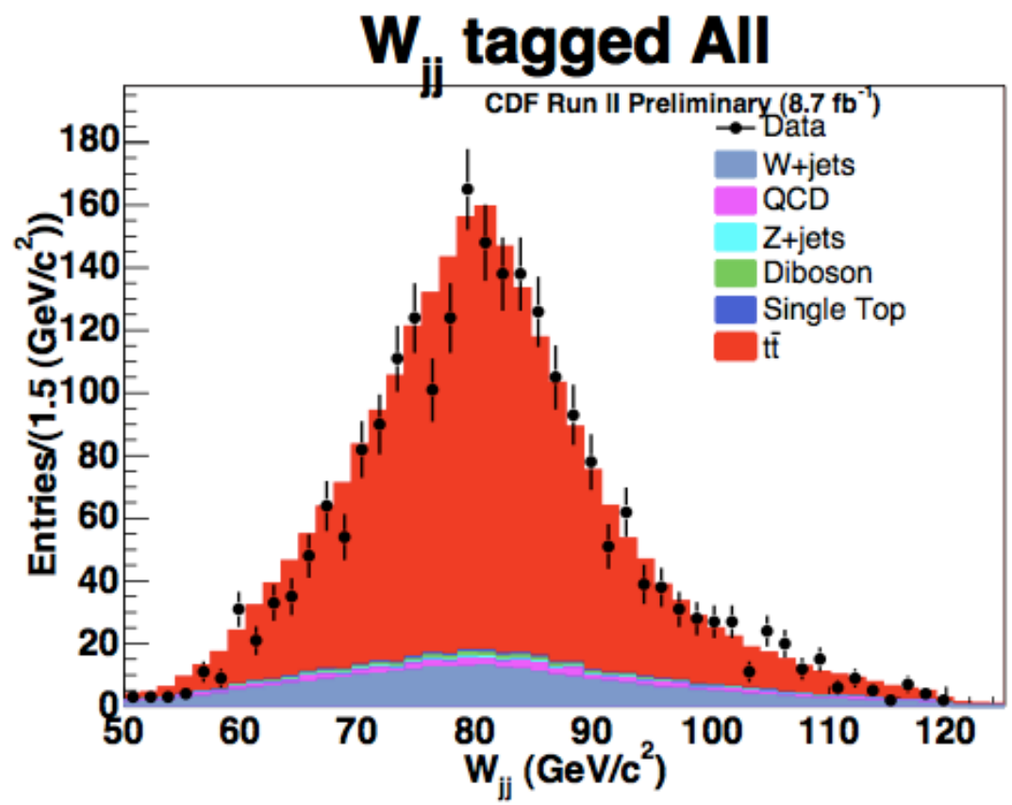
For a mixture of W+X (3.1 pb) and Z+X (1 pb) σ<2.1 pb @ 95% CL

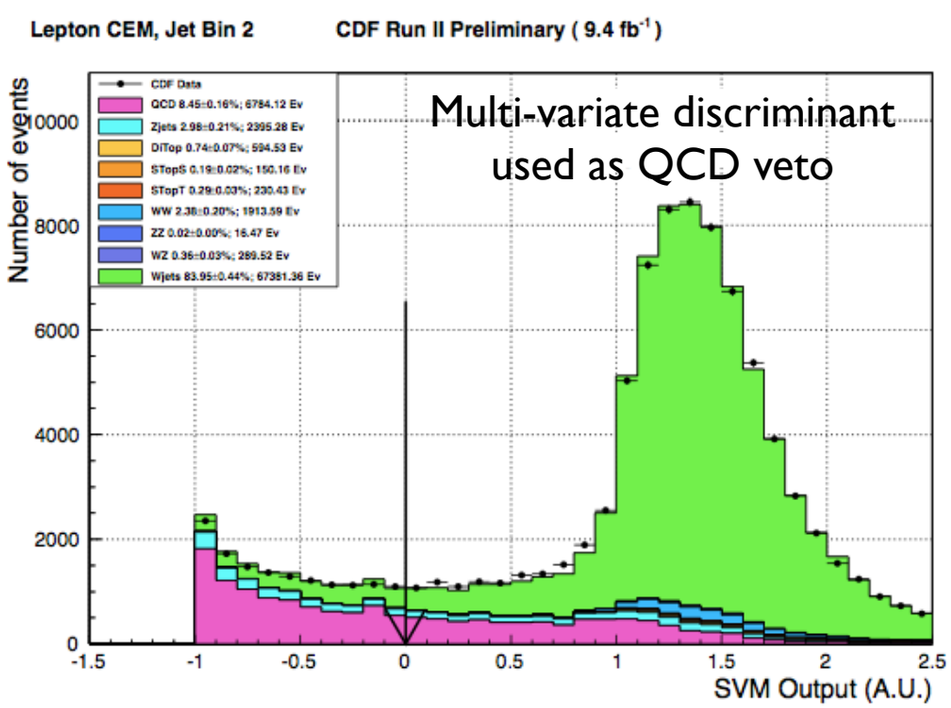
What is the impact of all these studies on other analyses @ CDF?

Effect of the changes on main analyses @CDF

1. Top-pairs:

- Effect of the quark/gluon-JES was checked - none
  - Top-pair signal extracted from ~gluon-free samples
- Effect QCD-model fix: negligible
  - QCD/Data < 5%





2. Higgs:

- Quark/gluon-JES already included
- Effect QCD-model fix: negligible
  - Applies only to one out of 3 main analyses
  - Was already partially included
  - QCD/Data ~ 2% thanks to a tight QCD veto

Effect of the changes on main analyses @CDF

3. QCD:

- cross-section analyses use R=0.7 cone-MIDPOINT or antiKT jets
  - those algorithms are less sensitive to soft radiation
  - Data-MC discrepancy even for JETCLU R=0.7 cone jets is within CDF JES uncertainty

Conclusions

- Big effort made by the CDF collaboration to reach this point
- Initially, data and expectations in the W+2jets sample were showing disagreement: deficiencies in models or new physics?
- Therefore we investigated a number of potential issues to understand the reason for the disagreement
- And we found that the disagreement was due to:
  - Important differences in the modeling of quark and gluon jet response
  - Insufficient modeling of one of the major backgrounds (QCD-multijets)
- It took us ~2 years to complete this task
  - Several subtle things were conspiring to prevent proper modeling of the W+2jets sample

Conclusions

- Big effort made by the CDF collaboration to reach this point
- Initially, data and expectations in the W+2jets sample were showing disagreement: deficiencies in models or new physics?

Thanks for the attention!

P.S: more details about these studies will be published soon in a PRD

- It took us ~2 years to complete this task
  - Several subtle things were conspiring to prevent proper modeling of the W+2jets sample