QCD Background and Systematic Uncertainties in Top Physics

Jin Wang
On behalf of ATLAS and CMS Collaboration

IHEP, China

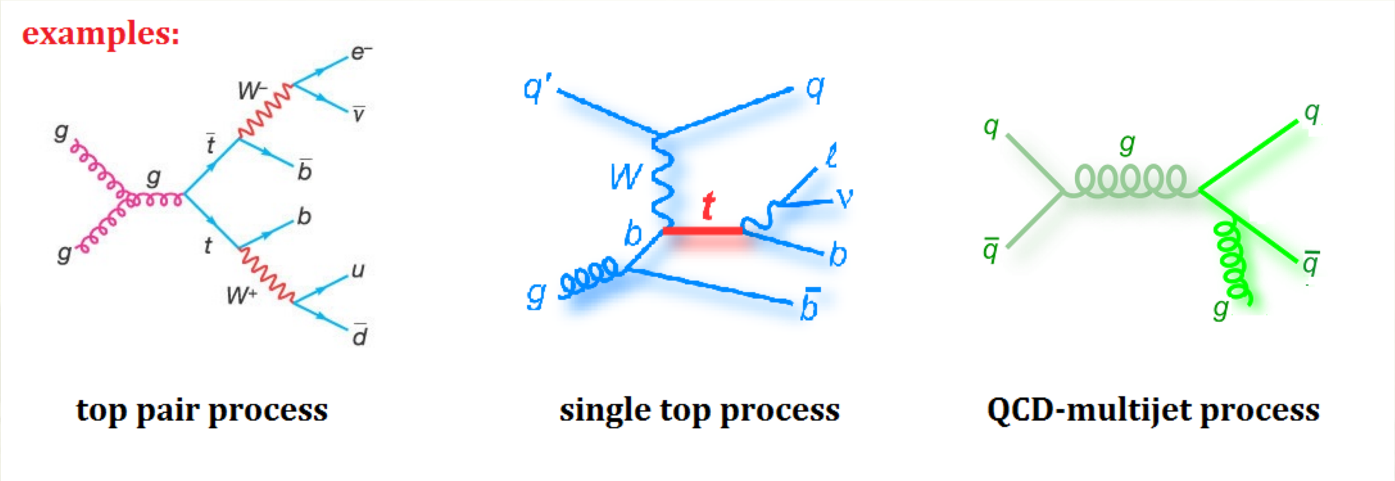
2013-09-14, Top2013

Outline

- Introduction of QCD Background in Top Physics
- Methods of QCD Background Estimation
- Latest Results from ATLAS, CMS, D0 and CDF
- Conclusion

Introduction

Top Events and QCD Backgrounds



- Top events signatures comparing to QCD backgrounds:
  - Isolated lepton with large transverse momentum from W decay
  - large missing transverse momentum (MET) if W decays leptonically
  - reconstructed W and top mass spectrum (m_W, m_T) from final state particles

How QCD-multijet events become Background?

- QCD-multijet events with misidentified and non-prompt leptons (collectively called fake leptons) could pass top event selection

sources of fake leptons

fake electons

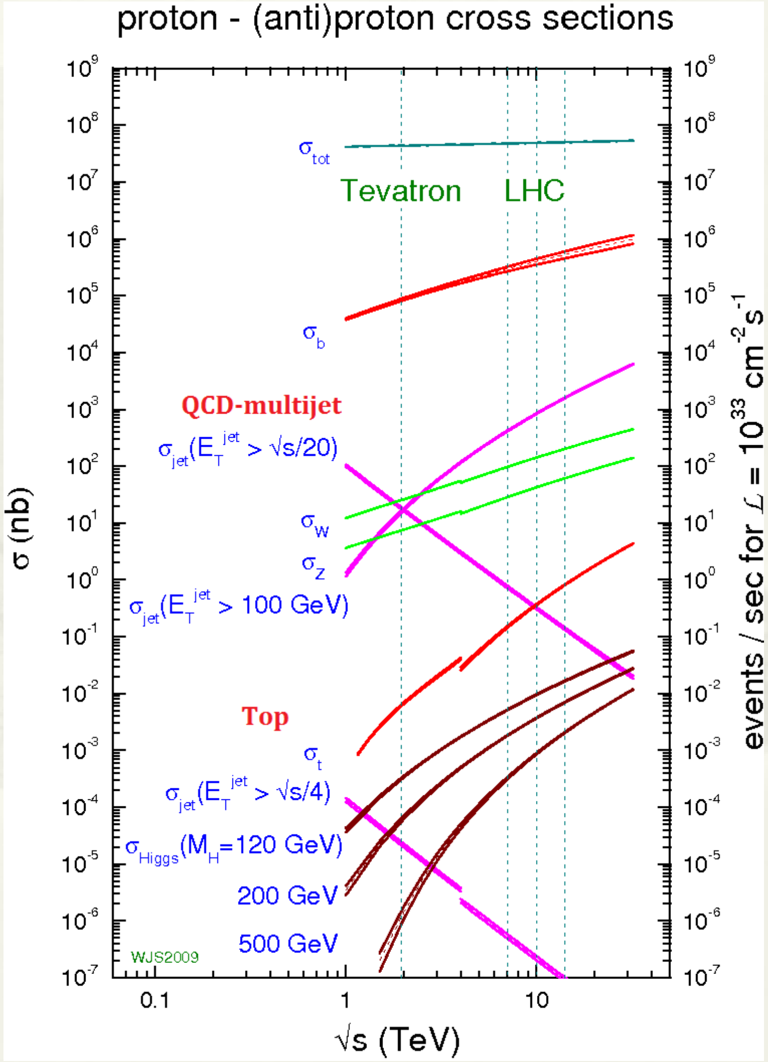
- b-quarks and c-quarks decay semileptonically
- jets misidentified as electrons
  - -- few charged tracks
  - -- little energy in hadronic compartments
- photon conversion

fake muons

- b-quarks and c-quarks decay semileptonically
- punch-through hadrons
- pions or kaons which decay in flight within tracking region

Top and QCD Cross Sections

- High QCD rejection factors from top event selections
- QCD multijet cross section (~ μb/mb) is many orders of magnitude higher than top cross section
  - usually contributes as one of the most important non-top backgrounds after Vector+jets
- Good estimation of QCD background is important for precision top analysis



Methods of QCD Background Estimation

Estimate QCD from Monte Carlo?

- To get good prediction of QCD background from MC, we need
  - high precision on fragmentation and hadronization models
  - high precision on description of interactions with matter and in shower model
  - good description of hadronic activity
  - precise estimate of the cross section

Such precision is not achievable

- Here comes data-driven approach for QCD background estimation
  - matrix method
  - template fit method where QCD templates from
    - jet-electron model
    - anti-electron/muon selection

Matrix Method

- real leptons: leptons from W (Z) decay
- tight leptons: standard selection in analysis
- loose leptons: looser identification and/or no isolation requirement

Solve Equation to Obtain N^{tight}_{fake}

$$N^{loose} = N^{loose}_{real} + N^{loose}_{fake}$$
$$N^{tight} = N^{tight}_{real} + N^{tight}_{fake} = \varepsilon_{real} N^{loose}_{real} + \varepsilon_{fake} N^{loose}_{fake}$$
$$N^{tight}_{fake} = \frac{\varepsilon_{fake}}{\varepsilon_{real} - \varepsilon_{fake}} (N^{loose} \varepsilon_{real} - N^{tight})$$

- N^{loose}: Number of events with one loose lepton
- N^{tight}: Number of events with one tight lepton
- Using data to obtain $\varepsilon_{real} = \frac{N^{tight}_{real}}{N^{loose}_{real}}$ and $\varepsilon_{fake} = \frac{N^{tight}_{fake}}{N^{loose}_{fake}}$.
- N^{tight}_{fake}: Estimated multijet events after tight selection

- this is an example of lepton +jets channel
- for dilepton channel: same idea with more categories

Matrix Method

Tag and Probe for ε_{real}

- Select control sample: $z \to ll$
  - Two Loose leptons with opposite signs
  - $80.2 < M(ll) < 100.2GeV,\ MET < 20GeV$
- Tag and probe method:
  - One lepton must be "tight"
  - Calculate the other lepton tight rate: ε_{real}

- ε_{fake} is measured from data control regions dominated by the contributions of fake leptons
  - example regions: low MET, low m_W or m_T, etc.
  - subtract *W+jets* and *Z+jets* backgrounds using Monte Carlo simulation

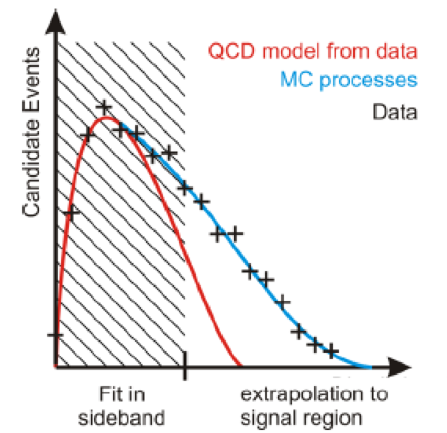
Estimate Systematics of Matrix Method

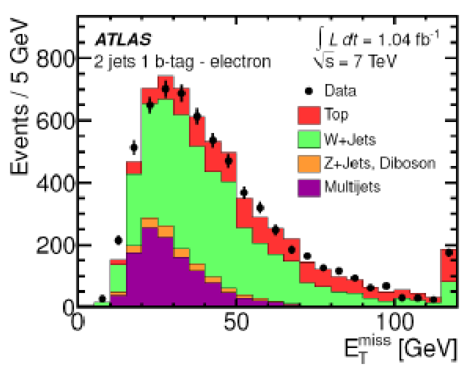
- Uncertainties of ε_{real} estimation are mostly statistical and negligible comparing to the ones of ε_{fake}

- Systematics Estimation of ε_{fake} :
  - varying the requirements of control regions selection
  - take into account used *W+jets* and *Z+jets* backgrounds uncertainties

Template Fit: Jet-electron Model

- Idea: use data samples with “electron-like” jet to simulate QCD with fake leptons





Multijet Model from Data

- 20 GeV jet trigger data
- Jets with kinematics of electrons
- With > 4 tracks
- Energy fraction requirement in electromagnetic calorimeter

QCD-multijet Estimation

- Binned likelihood fit on E_T^{miss}
  - Multijet shape from data
  - Other process shapes from MC

Template Fit: Anti-electron/muon Model

- Get QCD background template from a QCD enriched data sample where leptons failing some of the lepton cuts
  - reverse lepton identification(ID)/isolation
- Fit QCD templates and other process templates to data
  - fit with sensitive variable in control region orthogonal with signal region
  - or directly get QCD normalization from final measurement fit

Systematics of Template Fit Method

- For jet-electron model
  - cross check the fit with another variable, like the transverse *W* mass
  - could also evaluate the effect from pile-up
    - divide the jet-electron data sample into a high pile-up sample and a low pile-up sample
      - usually based on the number of primary vertices in events
    - compare estimated QCD of two samples

- For anti-lepton model
  - vary the QCD templates with different selections of control region

Latest Results from ATLAS, CMS, D0 and CDF

QCD Background Estimation in ATLAS
(arXiv:1207.5644)

- Measurements of top quark pair relative differential cross-sections with ATLAS
  - a paper published by The European Physical Journal C in January 2013
  - used 2.05 fb^{-1} data collected by ATLAS at $\sqrt{s}$=7TeV

- Used matrix method for QCD background estimation

Channel	μ + jets	e + jets
$t\bar{t}$	11100 ± 700	7400 ± 500
W+jets	1700 ± 700	1300 ± 500
Single top	490 ± 50	338 ± 32
Z+jets	192 ± 20	154 ± 26
Diboson	34 ± 4	21 ± 3
Fake-leptons	800 ± 800	250 ± 250
Signal+bkg	14400 ± 1700	9500 ± 1100
Observed	14416	9187

numbers of predicted and observed events after all selections

- Estimated QCD/signal+bkg is ~4.4%
- 100% uncertainty on QCD background estimation
  – Estimated by varying control region selection in fake rate calculation of matrix method

More Results in Latest ATLAS Publications

Channel	Goal	Luminosity (fb^{-1})	Reference	Method
Lepton+jets	search for ttbar resonances	4.7	Phys. Rev. D 88, 012004 (2013)	matrix method[1]
dilepton and lepton+jets	search for an excited bottom-quark b*	4.7	Phys. Lett. B 721 (2013) 171-189	matrix method[2]
Lepton+jets	Search for fourth-generation t' quarks	4.7	Phys. Lett. B 718 (2013)1284-1302	matrix method[3]
lepton	t-channel single top-quark production cross section	1.04	Phys. Lett. B 717 (2012) 330-350	jet-electron method[4]

QCD background estimation systematic uncertainties

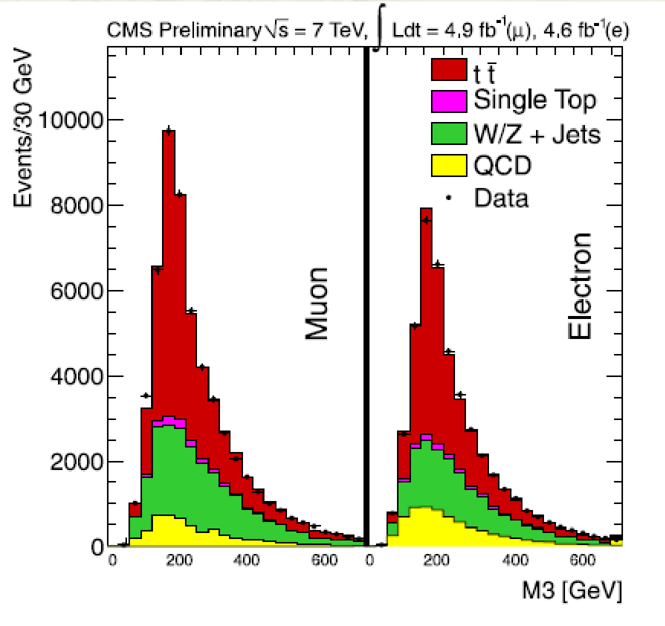
1. use jet-electron method to estimate the systematics
2. 50% uncertainties on QCD background estimation from fake rate calculation
3. 80% uncertainties: limited data sample (64%) + jet misidentification rate uncertainty (50%)
4. 50% uncertaintes: obtained from pile-up impact studies and alternative fit with $m_T(W)$

QCD Background Estimation in CMS

(arXiv:1212.6682)

- Measurement of the tt-bar production cross section in pp collisions at sqrt(s) = 7 TeV with lepton + jets final states
  - A paper published by Physics Letters B in March 2013
  - used 2.3 fb^{-1} data collected by CMS at $\sqrt{s}$=7TeV

- The mass of the three-jet combination distributions with estimated QCD backgrounds using anti-electron method



◈The QCD multijet normalization is obtained by fitting *missingEt* distribution in data
-- electron + jets channel: QCD distribution from MC
-- muon + jets channel: QCD distribution from anti-muon model

◈Use alternative QCD shapes with different selection for systematic estimation
◈different lepton ID/isolation/MET selection

More Results in CMS/D0/CDF

Experiment	Channel	Goal	Luminosity (fb^{-1})	Reference	Method
CMS	dilepton	σ(ttbar)	2.3	arXiv:1208.2671	matrix method
D0	lepton+jets	σ(ttbar)	0.9	PRL 100, 192004 (2008)	matrix method
D0	dilepton	σ(ttbar), m_t	1.0	PLB 679 , 177 (2009)	anti-electron
D0	single top	σ(t), t-ch	2.3	Fermilab-Pub-09/372-E	matrix method
CDF	lepton+jets	σ(ttbar)	2.7	Phys.Rev.D84:031101,2011	anti-electron
CDF	single top	σ(t)	3.2	Phys.Rev.D82:112005,2010	jet-electron anti-electron

Conclusion

Conclusion

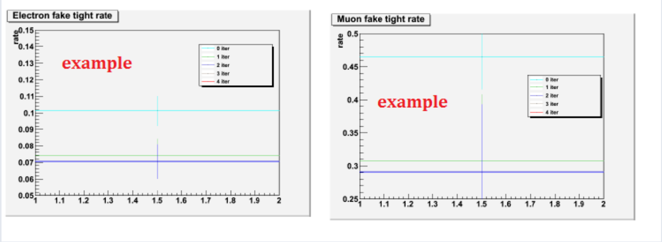
- QCD-multijet Background estimation is very important for precision top analysis
- In top analysis the data-driven methods are used to estimate QCD backgrounds
  - matrix method, templates fit method with jet-electron QCD model or anti-electron/muon QCD model
- Systematics estimation of QCD background usually comes from cross checks of alternative approaches in the method
- QCD background estimation method well fit in current top physics analysis
- Could get more ideas on QCD background estimation from W/Z studies and SUSY studies

Back Up

Cross Check Method in Matrix Method

Iteration Method for ε_{fake}

- Use multijet enriched control sample $E_T^{miss} < 10$ GeV



$$\varepsilon^0_{fake} = \left(\frac{N^{tight}}{N^{loose}}\right)_{E_T^{miss}<10}$$

$$N^{tight}_{fake0} = \frac{\varepsilon^0_{fake}}{\varepsilon_{real} - \varepsilon^0_{fake}}\left(N^{loose}\varepsilon_{real} - N^{tight}\right)$$

$$k^0_{W/Z} = \frac{N^{tight} - N^{0tight}_{fake} - N^{tight}_{t\bar{t},MC} - N^{tight}_{t,MC}}{N^{tight}_{W+jets,MC} + N^{tight}_{Z+jets,MC}}$$

$$\varepsilon^i_{fake} = \left(\frac{N^{tight} - k^{i-1}_{W/Z}(N^{tight}_{W+jets,MC} + N^{tight}_{Z+jets,MC}) - N^{tight}_{t\bar{t},MC}}{N^{loose} - k^{i-1}_{W/Z}(N^{loose}_{W+jets,MC} + N^{loose}_{Z+jets,MC}) - N^{loose}_{t\bar{t},MC}}\right)_{E_T^{miss}<10}$$