Measurement of the Jet Multiplicity in Top Quark Pair Events in Semi-Leptonic Decay Channel at CMS

6th Annual Workshop of the Helmholtz Alliance Physics at the Terascale

Institut für Experimentelle Kernphysik – Alexis Descroix

Outlook

- Introduction
- Jet multiplicity measurement
- Additional parton multiplicity measurement
- Results

Introduction

- $t\bar{t}$ total cross-section well established at LHC
- Differential cross-section measurement possible with 2011 data:
  - Jet multiplicity
  - Parton multiplicity: number of jet not originating from $t\bar{t}$ decay products

- Measurement important:
  - Constrain QCD scale parameter Q^2

    → large uncertainty in leading order MC (MadGraph + Pythia)
  - $t\bar{t}$ + (b-)jets is an important background for $t\bar{t}$ + Higgs, SUSY
- Recent theory calculations for $t\bar{t}$ + 1/2 jets exclusive @ NLO
- Powheg will allow $t\bar{t}$ +1 jet generation @ NLO

- Results and documentation recently public: CMS-PAS-TOP-12-018
  http://cdsweb.cern.ch/record/1494576

Jet Multiplicity

Analysis Overview

- Analysis in e + jets and μ + jets channels
- Whole 2011 dataset: 5.0 fb^{-1}
- Selection for lepton + jets events:
  - Exactly 1 isolated lepton (> 30 GeV)
  - Lepton veto (against other type leptons)
  - ≥ 3 jets (35 GeV)
  - 2 b-tagged jets (Combined tagger, ~1% mistagging rate)
- Data-driven background estimations:
  - QCD shape from data and normalization from fit to data in sideband
  - W + jets normalization from data (+ heavy flavor correction)
- Bin-by-bin unfolding back to generated jet multiplicity spectrum
- Normalized differential cross-section measurement: “cut & count”
- Combine results from e/μ channels with BLUE combination method

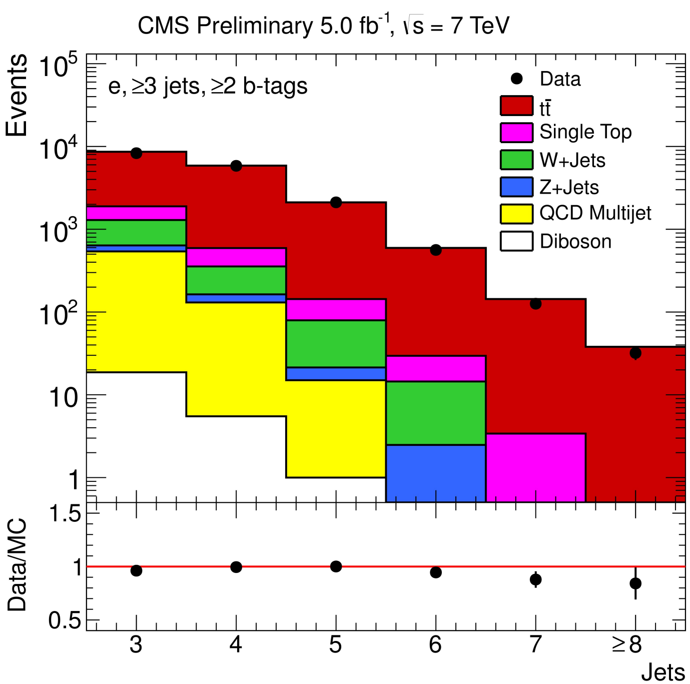
Data-Driven Background Estimation

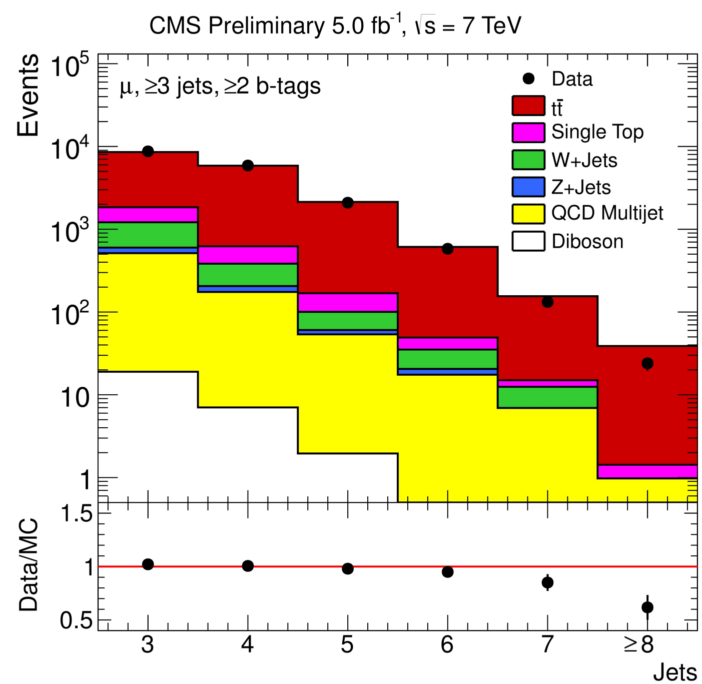
- The double b-tagging cut strongly reduces background contributions
- Remaining statistics in QCD MC too low → extraction from data:
  - QCD shape extracted from data in signal region (≥ 2 b-tagged jets) with an inverted relative isolation cut on lepton
  - Normalization extracted from fit to MET distribution in sideband region: (< 2 b-tagged jets)
  - QCD shape scaled with fit result

- W + jets most important background
  - Heavy flavor fractions corrected according to measurements in other top analyses
  - W + jets shape from MC. Normalization from charge asymmetry:
    - W + jets produces large charge asymmetry
    - Subtracting events (l^+ - l^-) provides good estimation of W+jets normalization

Jet Multiplicity Distribution in Signal Region

- MC simulated samples are corrected for:
  - Pileup
  - *b*-tagging scale factor
  - Lepton trigger, reconstruction, and isolation
- Good agreement between data and MC in signal region





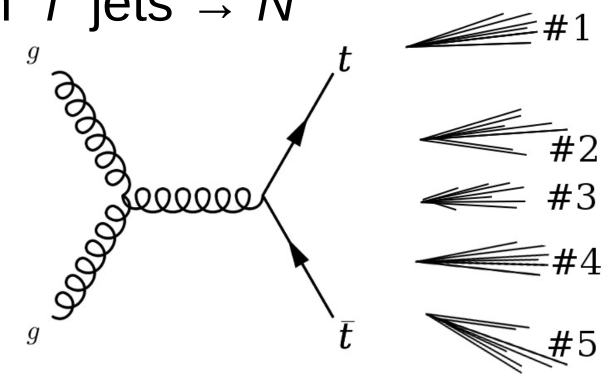
Uncertainties

- Total uncertainties are dominated by systematic uncertainties which increase with jet-multiplicity
- Dominant systematic uncertainties:
  - MadGraph generator parameters (2.5% to 35%):
    - Q^2 scale parameter (renormalization and factorization scale)
    - matrix-element to parton-showering matching threshold parameter
  - Jet energy scale and resolution (1.5% to 14%)
  - Parton distribution functions (2% to 10%)

- Smaller systematic uncertainties: (each of them is below 5%)
  - B-tagging scale factor uncertainty
  - Lepton trigger and selection efficiency
  - Pileup
  - Luminosity

Measurement Method

- Count jets in each event. Number of events with "*i*" jets → N^i
- Subtract background from data in each jet-bin:

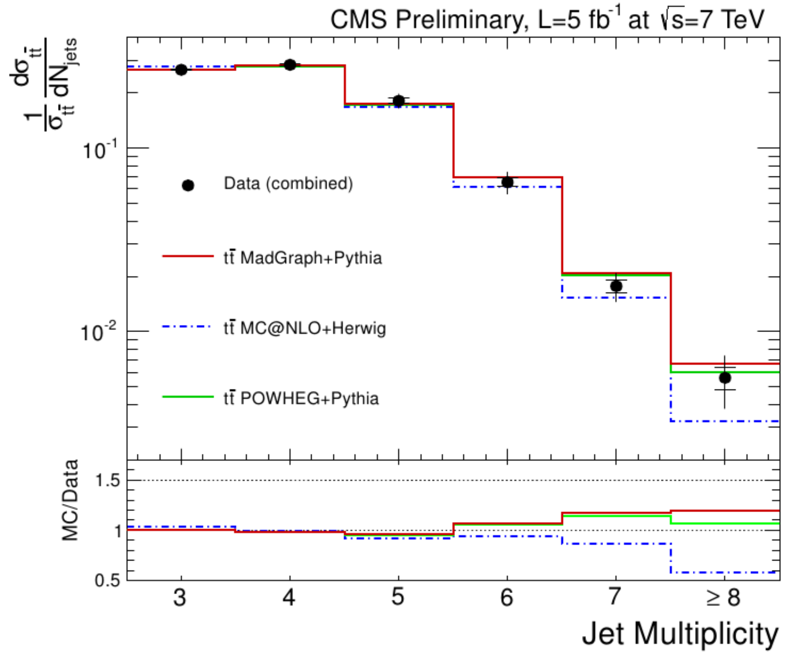


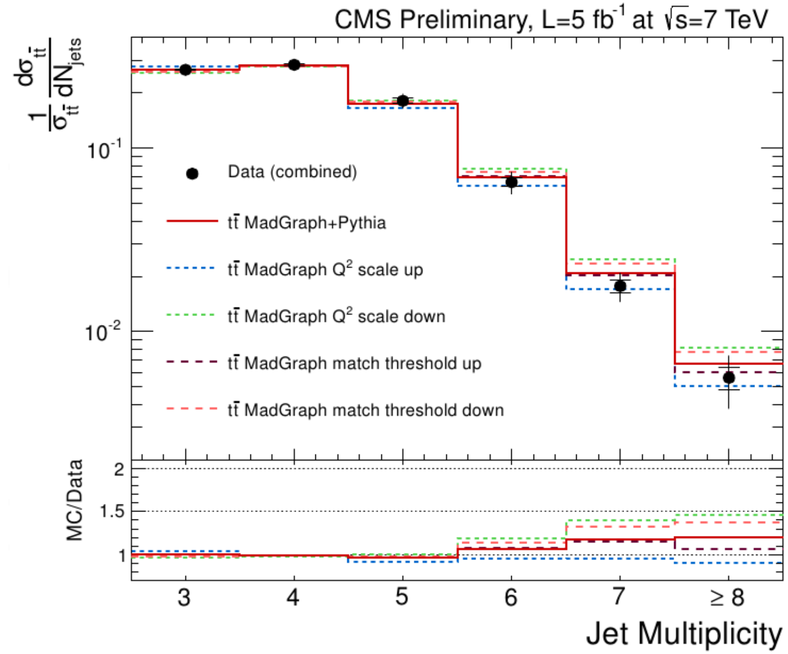
$$\frac{d\sigma_{t\bar{t}}^{measured}}{dN_{jets}} = \frac{N_{data}^{i} - N_{bkg}^{i}}{\epsilon^{i} \cdot Lumi} \quad \text{with:} \quad \epsilon^{i} = \frac{N_{rec}^{i}}{N_{gen}^{i}}$$

- ϵ^i enables to unfold from reconstructed back to generated jet multiplicity
  - Detector independent
  - Phase space : jet with p_T>35 GeV and $|\eta| < 2.4$
- Normalize differential cross-section to total measured cross-section
- Combination of e and μ + jets channels with BLUE method
  - Best Linear Unbiased Estimator
  - 0% correlation between lepton systematic uncertainties and statistical unc.
  - 100% correlation between other systematic uncertainties
  - Reduces mainly statistical uncertainties but systematics are dominating

Results

- Compare predictions from different MC generators:
  - MadGraph+Pythia, Powheg+Pythia, and MC@NLO+Herwig
- Compare to predictions from MadGraph with parameter variations:
  - Q2 scale up/down, ME/PS matching threshold up/down





- Agreement with MadGraph standard prediction
- Slight discrepancies with MC@NLO and MadGraph scale down

Additional Parton Multiplicity

Alternative Measurement

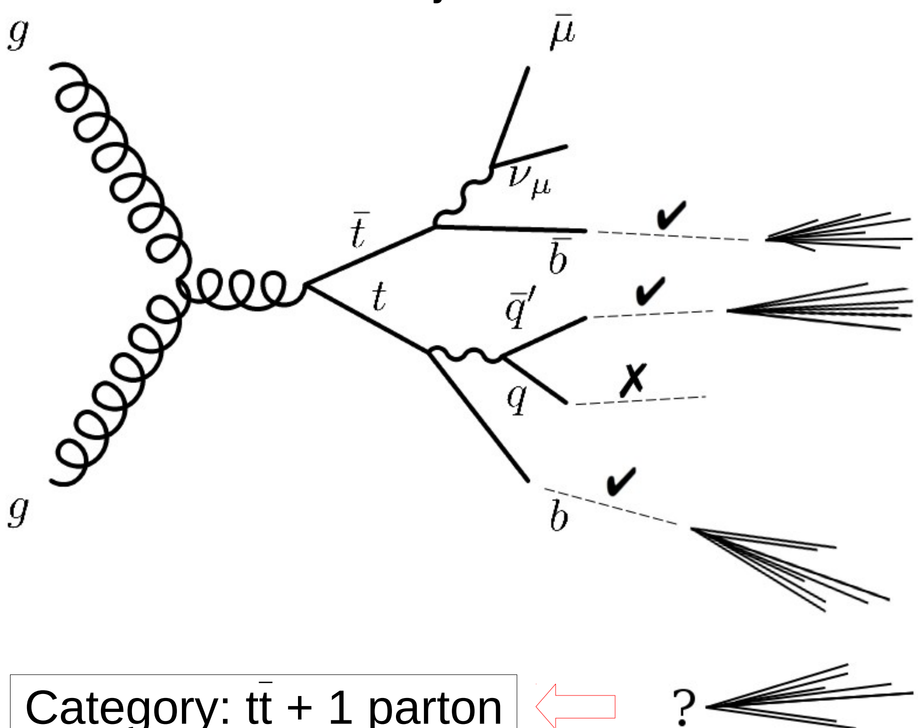
- Difference with cut & count measurement:
  - With MC information, remove jets from top decay products from counting
    $\rightarrow$ More robust way to define additional jets
- Count jets which are incompatible with top decay products
  - ΔR cut between top decay products and generator jets
- Incompatible jets are likely to originate from additional radiated partons

- Categorize all events:
  - $t\bar{t}$ + 0 partons
  - $t\bar{t}$ + 1 parton
  - $t\bar{t}$ + ≥ 2 partons
- Fit to data of 3 signal templates in χ, goodness of reconstruction

- Latest result of ATLAS employing a very similar method (ATLAS-CONF-2012-083)

Event Categorization with MC Information

- Categorize MC signal events depending on the number of jets not matching any top decay product jets:
  - ΔR cut between quarks from top decay and jets at 0.5 for match
  - Generator jets used



Category: $t\bar{t}$ + 1 parton

- $t\bar{t}$ event in μ + jets decay channel
- Using MC information:
  - One decay product quark does not correspond to any jet

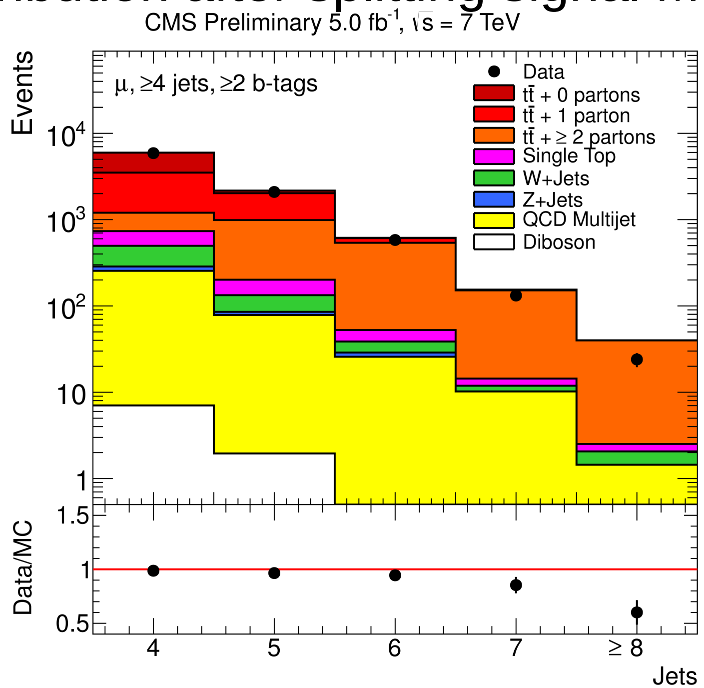
→ jet probably lost in detector or selection acceptance.

  - One jet does not match any decay product of the top

→ should originate from an additional parton

Analysis Setup

- Very similar configuration as the jet multiplicity analysis:
  - Utilizing same MC samples data-driven background estimation methods
  - Slight change of selection (jet-p_T>30 GeV instead of 35 GeV)
  - Same systematic uncertainty sources considered. Pseudo-experiment method employed for calculation
- Jet multiplicity distribution after splitting signal MC in three categories:



Template Fit Setup

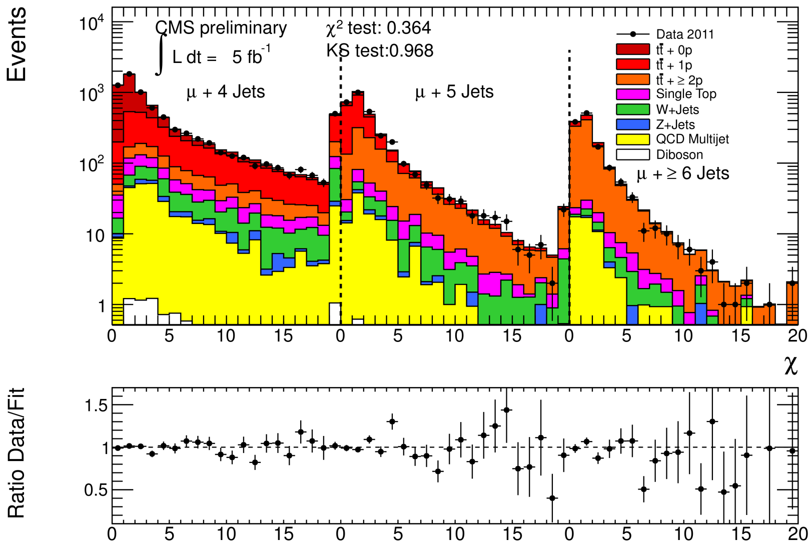
- Perform a full event reconstruction
- Consider only hypothesis with b-tagged jets assigned to b-quarks
- Calculate goodness of reconstruction χ:

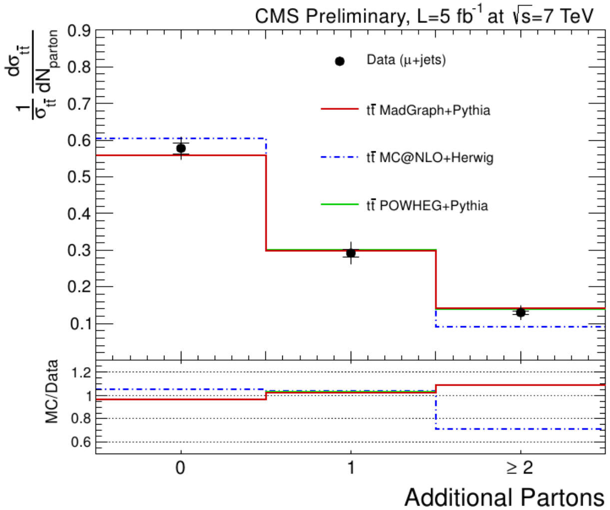
$$\chi = \sqrt{\left(\frac{m_{W^{had}}^{rec} - m_{W^{had}}^{true}}{\sigma_{W^{had}}}\right)^2 + \left(\frac{m_{t^{had}}^{rec} - m_{t^{had}}^{true}}{\sigma_{t^{had}}}\right)^2 + \left(\frac{m_{t^{lep}}^{rec} - m_{t^{lep}}^{true}}{\sigma_{t^{lep}}}\right)^2}$$

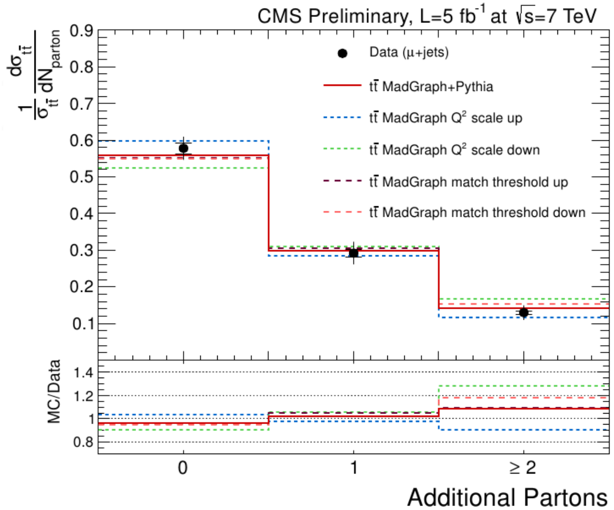
- With reconstructed mass of both tops and W boson decaying in quarks
- True masses and mass uncertainties taken from MC
- Distribution of χ gets split in three jet bins in addition (μ + 4, 5, ≥6 jets)
- χ enables to distinguish between events with or without additional partons
  - Example in μ + 4 jets bin:
    - No additional partons → all jets match top decay products → low χ
    - Additional partons → some jets from top decay products lost → high χ

Results

- Maximum likelihood template fit performs well (linearity checks performed)
- Apply fit result on MC predictions and normalize for final result
- Good agreement with predictions from MadGraph and Powheg
- Small discrepancies with MC@NLO and MadGraph scale down also observed with this measurement







Conclusion

- Two alternative methods measuring the normalized differential cross-section of $t\bar{t}$ events in additional jet multiplicity bins.
- Different measurement methods with similar result interpretations:
  - Good agreement between measurement and MadGraph and Powheg MC predictions
  - Slight discrepancies with MC@NLO and MadGraph scale down variation
  - Measurement uncertainties dominated by systematics