Constraining QCD radiation using ttbar events with the CMS experiment

María Aldaya

DESY

(for the CMS Collaboration)

TOPLHCWG Meeting, 29 November 2012

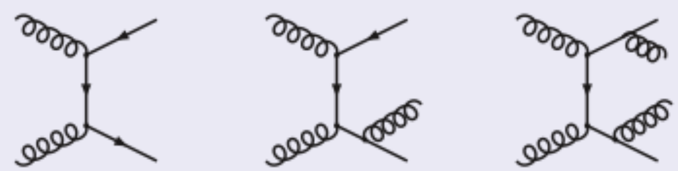
Quark and gluon radiation in ttbar events

- At LHC, the fraction of ttbar events produced with additional hard jets is high
- Precise understanding of these processes is important:
  - Test of perturbative QCD
  - Anomalous production of ttbar(+jets) could reveal new physics
  - Background for ttbarH and many BSM searches
- In general, sizeable uncertainty from QCD radiation for many top quark analyses
  - Theory predictions and models need to be tuned and tested with measurements
- Large samples of ttbar events provides a great opportunity to study the details of the ttbar production mechanisms
  - Potential of constraining QCD radiation at the scale of the top quark mass
    - Differential ttbar production cross section vs. p_T(top), p_T(ttbar)
    - Jet multiplicity in ttbar and associated jets
    - ttbar with veto on additional jets (a.k.a. "gap fraction")

Generator setups for ttbar at CMS

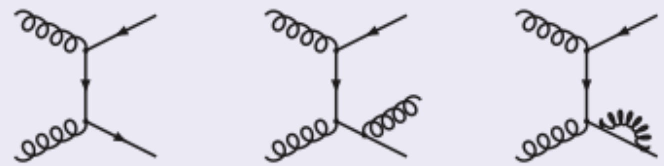
process	ME	PS	method	PDF	Tune
$t\bar{t}$ + jets	MadGraph v5.x	Pythia v6.42x	ME+PS	CTEQ6L1	Z2(*)
$t\bar{t}$	POWHEG-box 1.0	Pythia v6.42x	NLO	CTEQ6M	Z2(*)
$t\bar{t}$	MC@NLO v3.41	Herwig v6.520	NLO	CTEQ6M	

- Matrix Element + Parton Shower generators
  - Better description of high multiplicities
  - ISR/FSR modelling via ME from assumed Q^2 variation
  - Matching procedure to remove double counting between partons produced by ME and PS



tree level diagrams with up to 3 partons

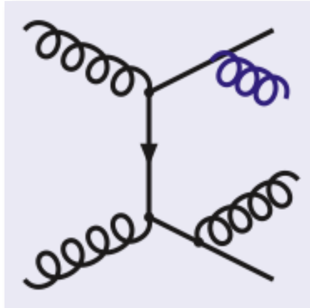
- Next to Leading Order generators
  - More accurate in normalization
  - Smaller uncertainty on Q^2



real + virtual corrections

- MadGraph(+Pythia) is the default for most of the analyses
  - Uncertainty on radiation covered by variations of Q^2 and ME-PS matching

Radiative corrections



- The 'Q^2 scale' variation addresses 2 aspects:
  - renormalisation and factorisation scale (ME)
  - amount of initial and final state radiation (ISR/FSR)

- For each event, Q^2 is defined as:

$$Q^2 = m_t^2 + \sum p_T^2 \text{ (MadGraph)}$$

$$Q^2 = m_t^2 \text{ (POWHEG/MC@NLO)}$$

  - Q^2 varied up (down) by a factor 4.0 (0.25)

- Parton showering:
  - p_T-ordered evolution scale of ISR/FSR
  - shares Q^2 factor α_S scale with ME
  - implicitly: starting scale changes with ΔQ^2

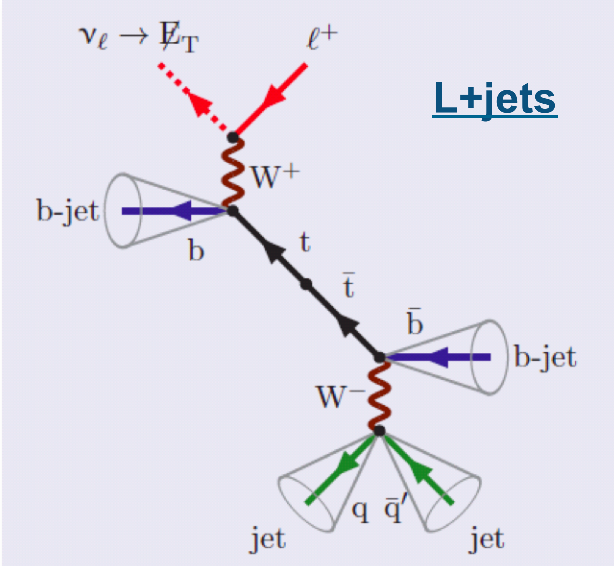
- MadGraph uses:
  - tree-level diagrams for hard radiation and interferences (up to 3 final-state partons for ttbar)
  - parton showering for soft and collinear region (with Pythia 6.42X)
  - matching via ktMLM, thresholds varied by factor 0.5 to 2.0 (nominal = 20 GeV)

Event selection

Lepton+jets:

- Exactly 1 isolated high-p_T lepton (μ or e)
  - μ, e: $p_T > 30$ GeV, $|\eta| < 2.1$ (also $|\eta| < 2.5$ for e)
- Veto additional leptons
- Analysis-dependent jet selection
  (≥ 3 jets, $p_T > 30$ GeV, $|\eta| < 2.4$)
- ≥ 2 b-tagged jets

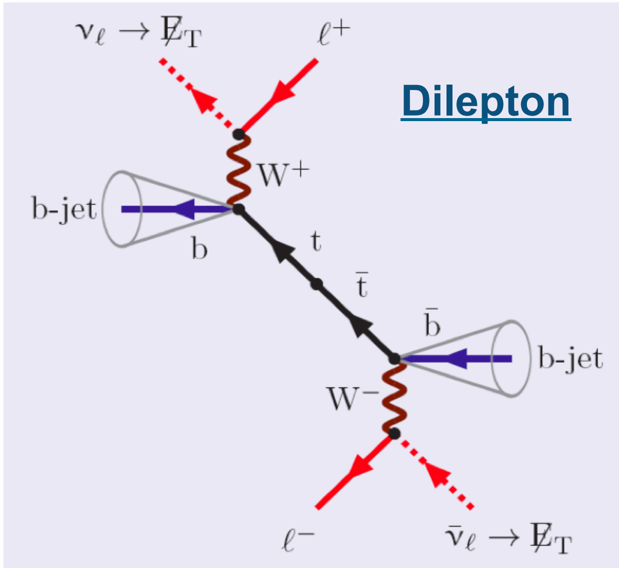
(Kinematic reconstruction of the ttbar system)



Dileptons:

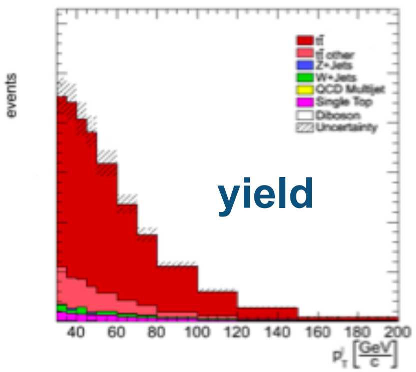
- 2 opp.-sign, high-p_T isolated leptons (ee, $\mu\mu$, μe)
  - μ, e: $p_T > 20$ GeV, $|\eta| < 2.4$
- QCD veto: $m_{ll} < 12$ GeV
- ≥ 2 jets, $p_T > 30$ GeV, $|\eta| < 2.4$
- ≥ 1 b-tagged jets
- ee, $\mu\mu$ channels: $E_T^{miss} > 30$ GeV, $|m_{ll} - m_Z| > 15$ GeV

Kinematic reconstruction of the ttbar system

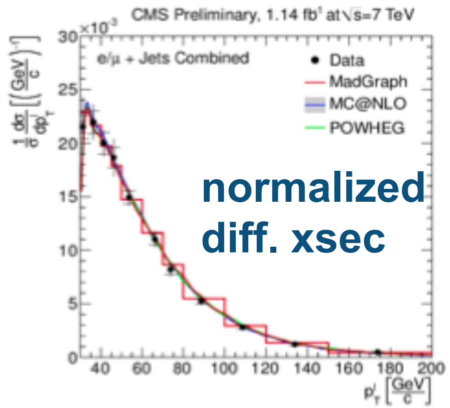


Normalized differential cross sections

- Measure ttbar(+jets) production differentially (X = N(jets), p_T(top), p_T(ttbar), …)



$$\frac{1}{\sigma}\frac{d\sigma^i}{dX} = \frac{1}{\sigma}\frac{N^i_{\mathrm{Data}} - N^i_{\mathrm{BG}}}{\Delta^i_{\mathrm{X}}\epsilon^i L}$$



- Background: data-driven and simulation
- Corrected back to particle or parton level, within visible phase space or extrapolated to full phase space
- Corrected for detector effects (finite experimental resolution) using MadGraph
- Normalized to inclusive cross section in corresponding phase space
- Compare to:
  - MadGraph+Pythia, MC@NLO+Herwig, POWHEG+Pythia
  - MadGraph with varied scales: Q^2, matching

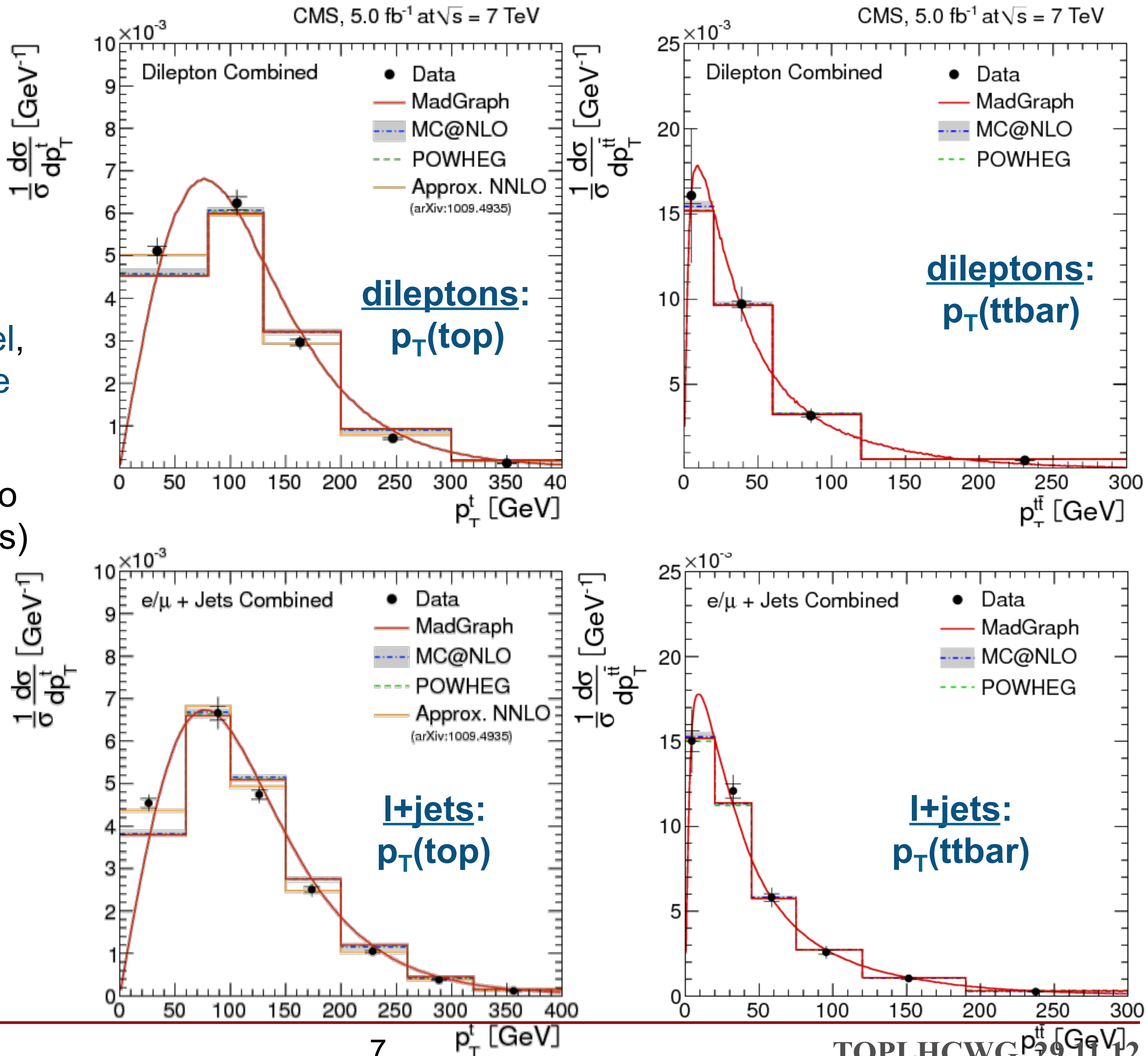
$1/\sigma\ d\sigma/dp_T(top)$, $1/\sigma\ d\sigma/dp_T(tt)$

arXiv:1211.2220

- l+jets: ≥ 4 jets, p_T > 30 GeV, |η| < 2.4, ≥ 2 b-tags
- Full kinematic reco. of the ttbar system for both channels
- Unfolded to parton level, extrapolated to full phase space
- Compare p_T(top) also to approx. NNLO (Kidonakis)

Softer top p_T spectrum in data, better described by approx. NNLO

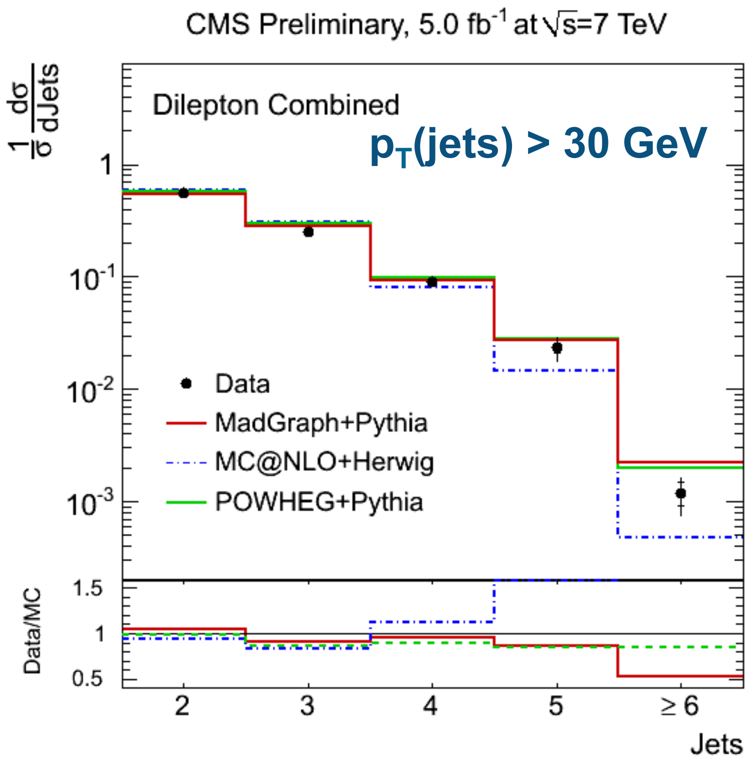
p_T(ttbar): good agreement btw:
- data and predictions
- different predictions



1/σ dσ/dN(jets) – dileptons

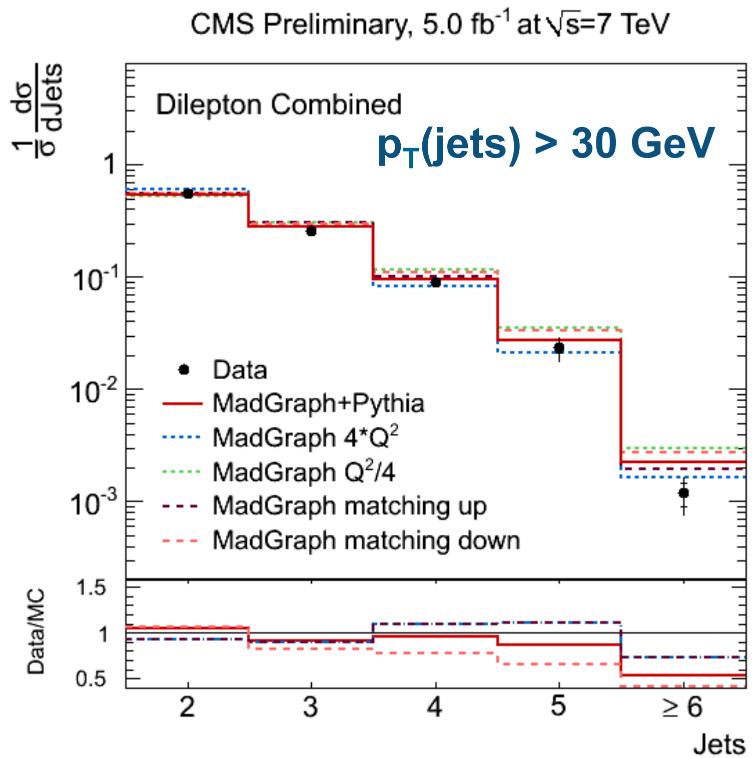
CMS-PAS TOP-12-023

- Additional jets: jets in the kinematic region NOT identified by the kinematic reconstruction as part of the ttbar system
- Unfolded to particle level, in visible phase space { jets: p_T > 30 GeV, $|\eta|$ < 2.4; leptons: p_T > 20 GeV, $|\eta|$ < 2.4 }



General good agreement between data and predictions

MC@NLO+Herwig underestimates large N(jets)

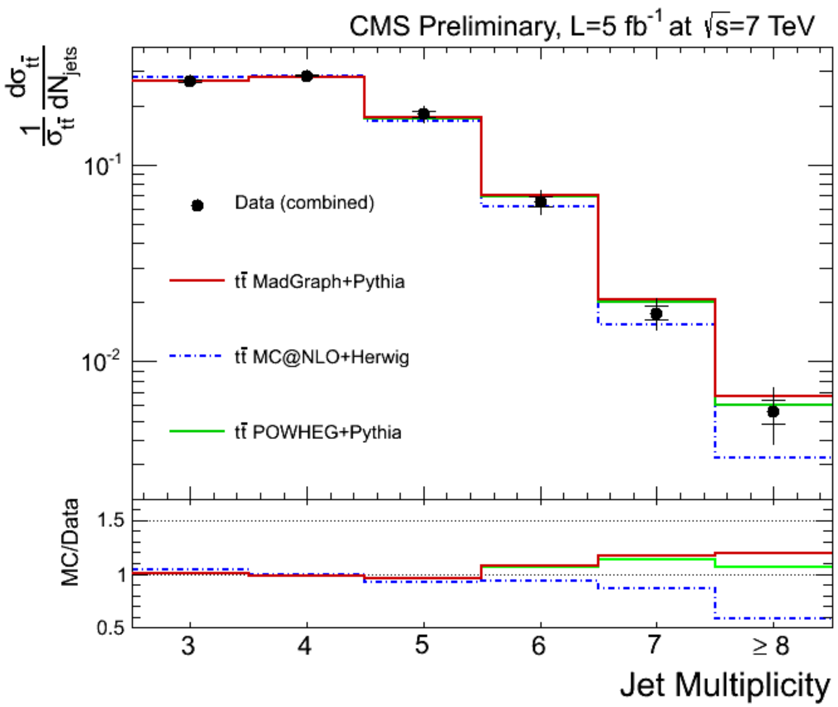


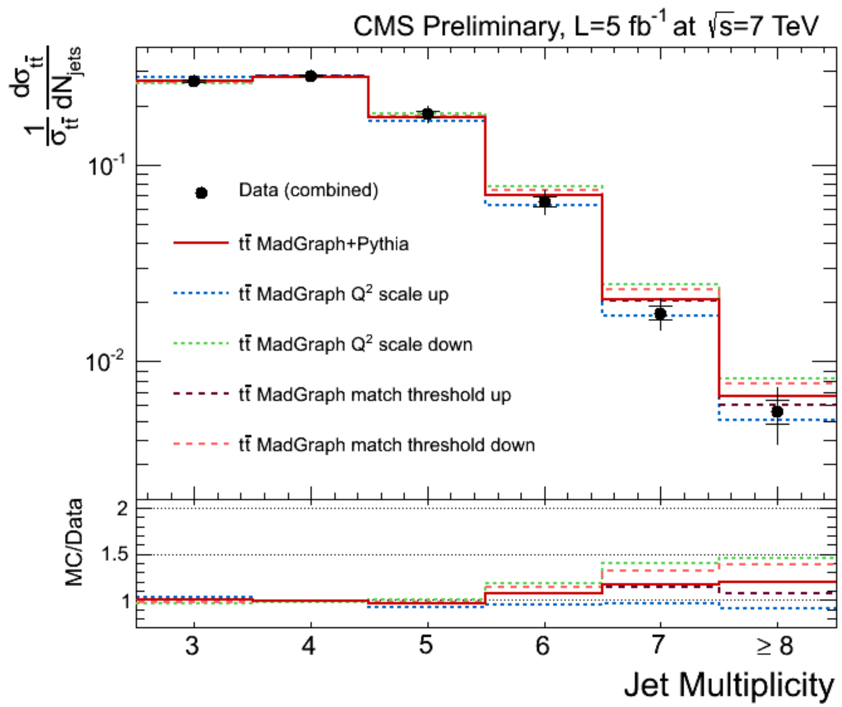
Sensitive to modelling of radiation in MadGraph: data seem to be better described by larger scales

$1/\sigma\ d\sigma/dN(jets)$ – l+jets

CMS-PAS TOP-12-018

- ≥ 3 jets, p_T > 35 GeV, $|\eta|$ < 2.4, ≥ 2 b-tags
- Corrected to particle level, in visible phase space → jets: p_T > 35 GeV, $|\eta|$ < 2.4





General good agreement between data and predictions

MC@NLO+Herwig underestimates large N(jets)

Sensitive to modelling of radiation in MadGraph: data seem to be better described by larger scales

$1/\sigma\ d\sigma/dN$(add. partons) – μ+jets

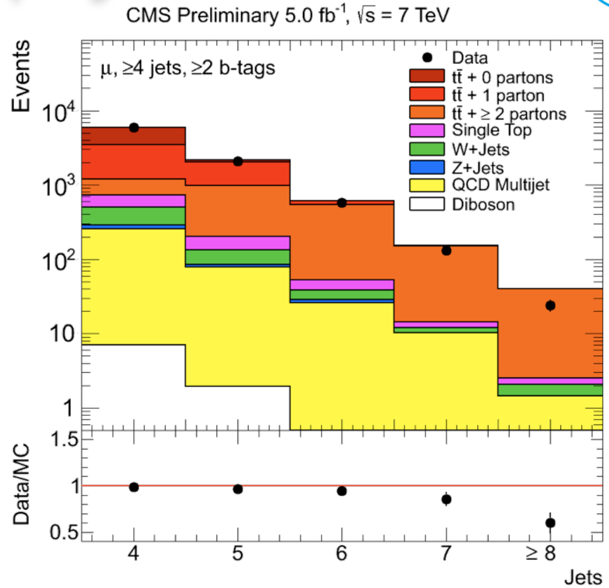
CMS-PAS TOP-12-018

- ≥ 4 jets, p_T > 30 GeV, $|\eta|$ < 2.4, ≥ 2 b-tags
- Categorize ttbar MC events with N(genJets) NOT matching any of the top decay products in ΔR > 0.5

→ originating from additional partons

  - 3 categories: ttbar + 0, 1, ≥ 2 add. partons

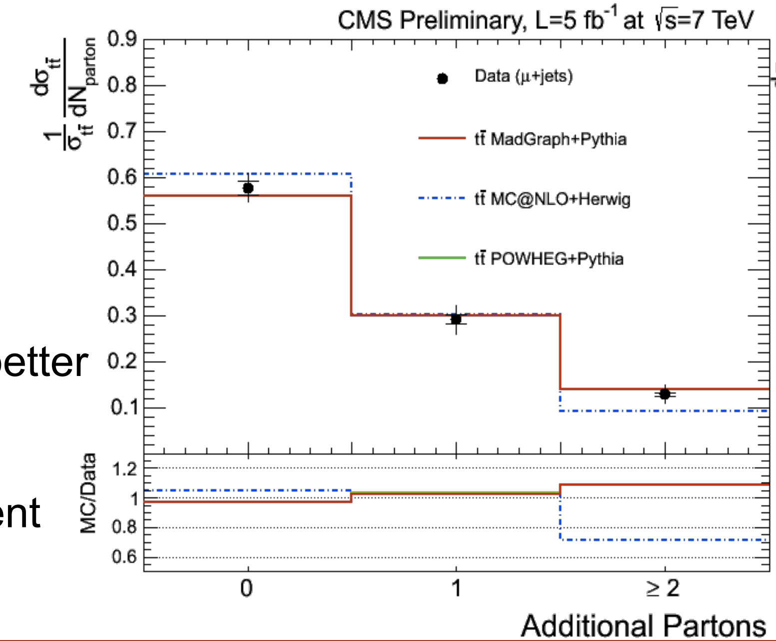
- Template fit for the 3 categories with background templates to data in 4, 5, ≥ 6 jet bins

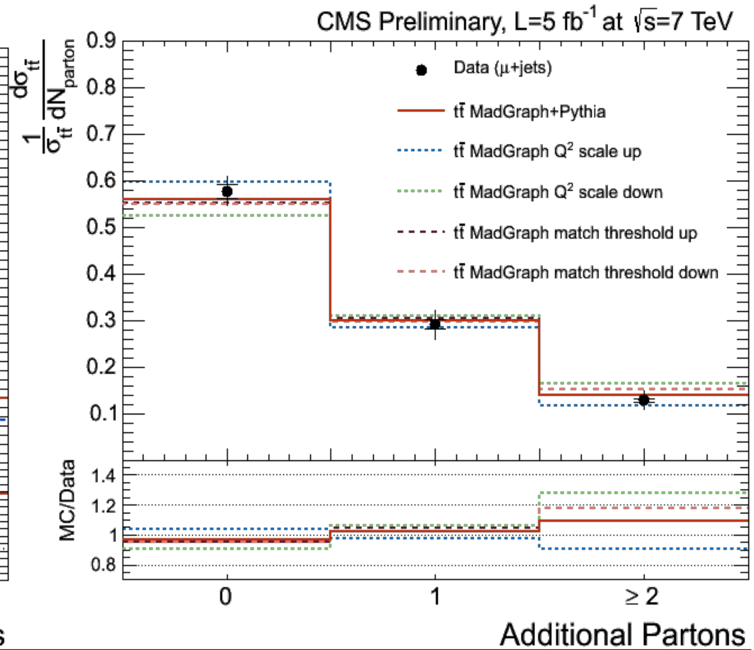


General good agreement between data and predictions

Higher Q^2 tends to describe data better

In agreement with N(jet) measurement





ttbar with veto on extra jets: "gap fraction"

- Quantify jet activity arising from quark and gluon radiation produced with the ttbar system with a jet veto:

$$f(p_T) = \frac{N(p_T)}{N_{total}}$$

$$f(H_T) = \frac{N(H_T)}{N_{total}}$$

N_{total}: total number of selected events

$N(p_T)$: events which do not contain 1 (2) additional jet p_T above a certain threshold

Sensitive to the leading- (2^{nd} leading-) p_T emission

$N(H_T)$: events in which the scalar p_T sum of all additional jets is below a certain threshold

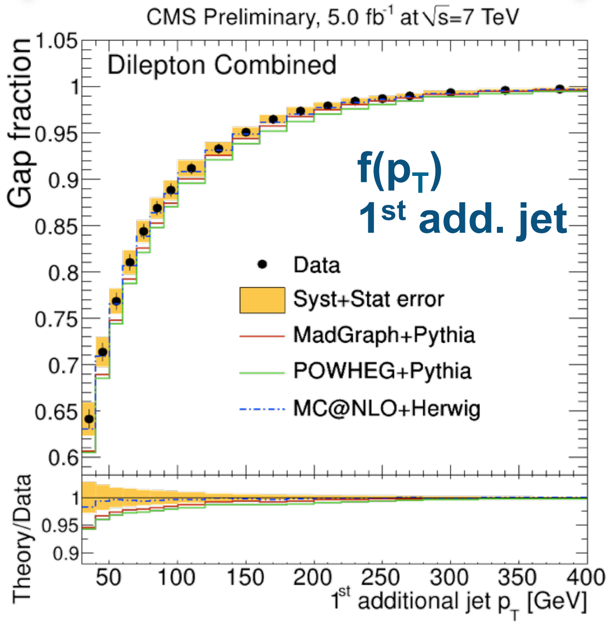
Sensitive to all hard emissions accompanying the ttbar system

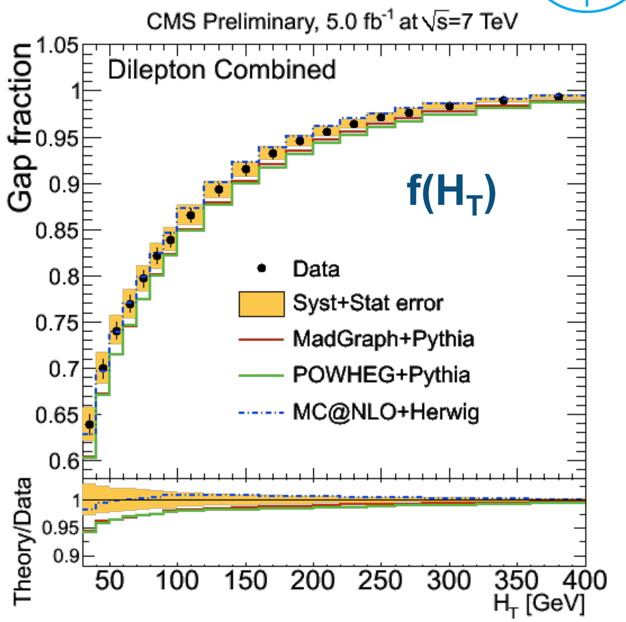
- Corrected back to particle level within visible phase space
- Corrected for detector effects using MadGraph
- Compare to:
  - MadGraph+Pythia, MC@NLO+Herwig, POWHEG+Pythia
  - MadGraph with varied scales: Q^2, matching

Gap fraction – dileptons

CMS-PAS TOP-12-023

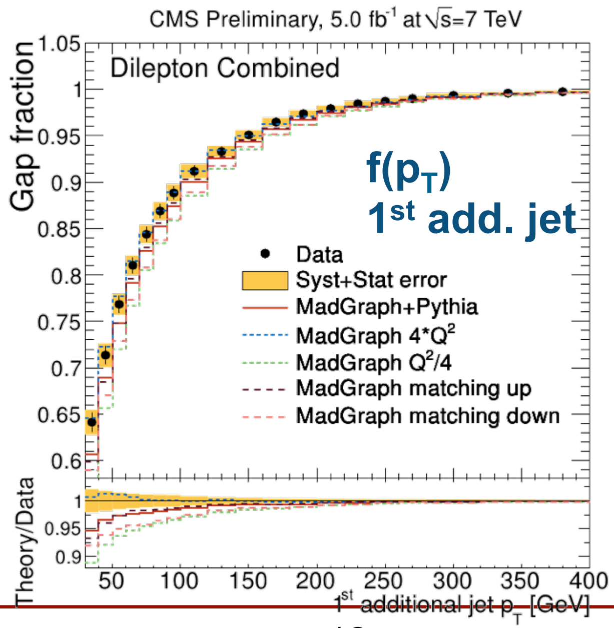
- Gap fraction vs. different generators:
  - General good agreement between data and predictions
  - MC@NLO+Herwig provides better description

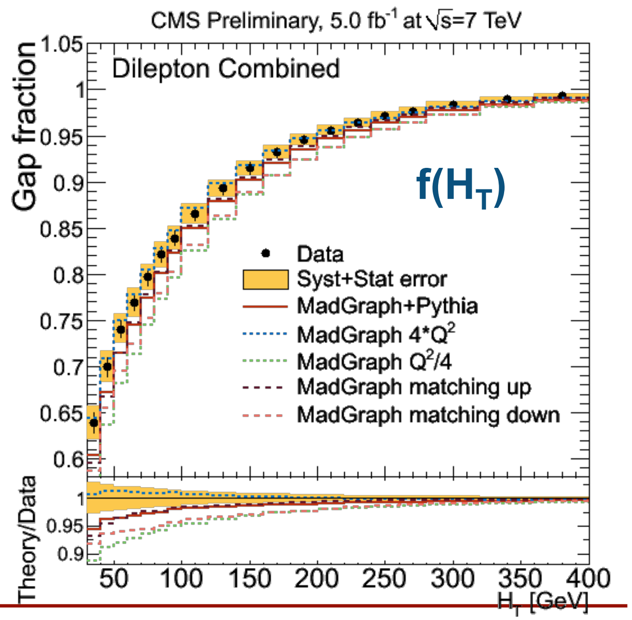




- Gap fraction vs. MadGraph varied scales:
  - Higher Q^2 seems to describe data better
  - Experimental precision smaller than spread due to parameter variation

    → variations could be significantly reduced



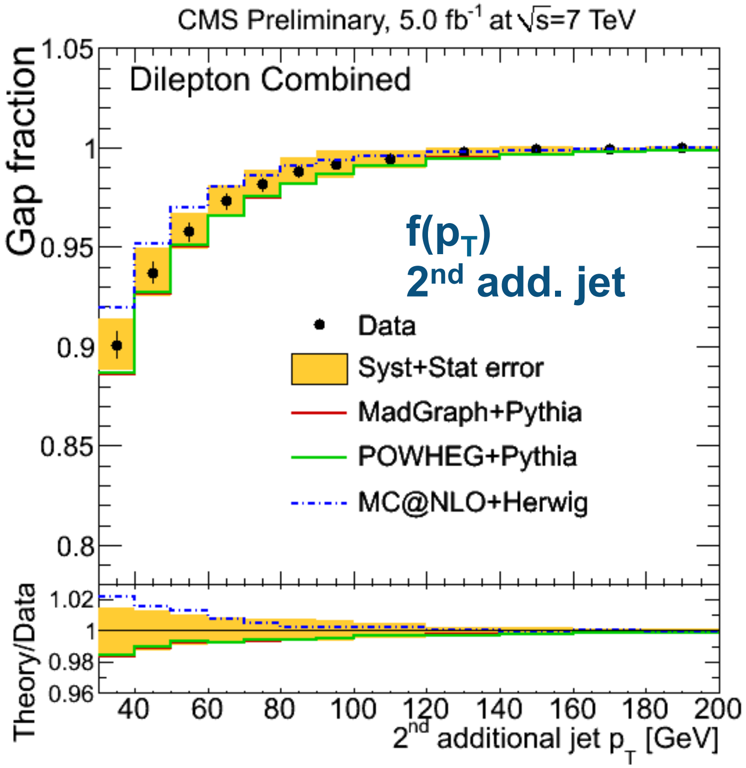


Gap fraction: 2nd add. jet – dileptons

CMS-PAS TOP-12-023

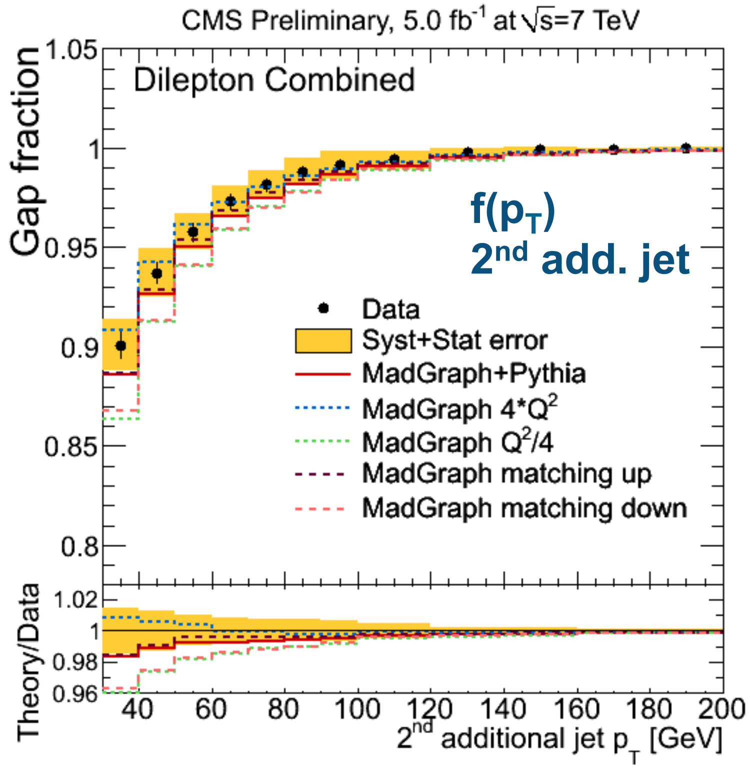
- First look into gap fraction for the 2nd additional jet

Gap fraction vs. different generators



- General good agreement between data and predictions
- MC@NLO+Herwig overestimates the gap fraction for the 2nd add. jet
→ Compare with POWHEG+Herwig

Gap fraction vs. MadGraph varied scales

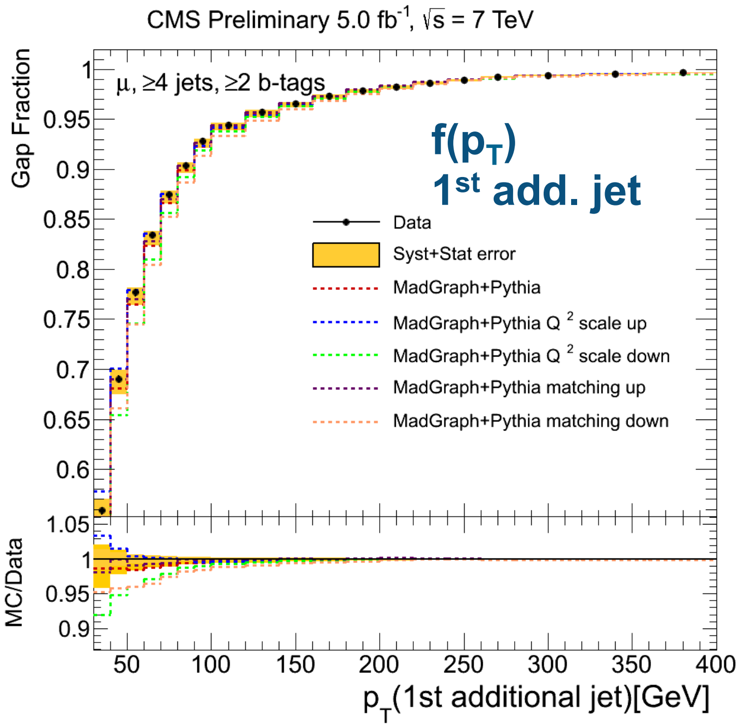


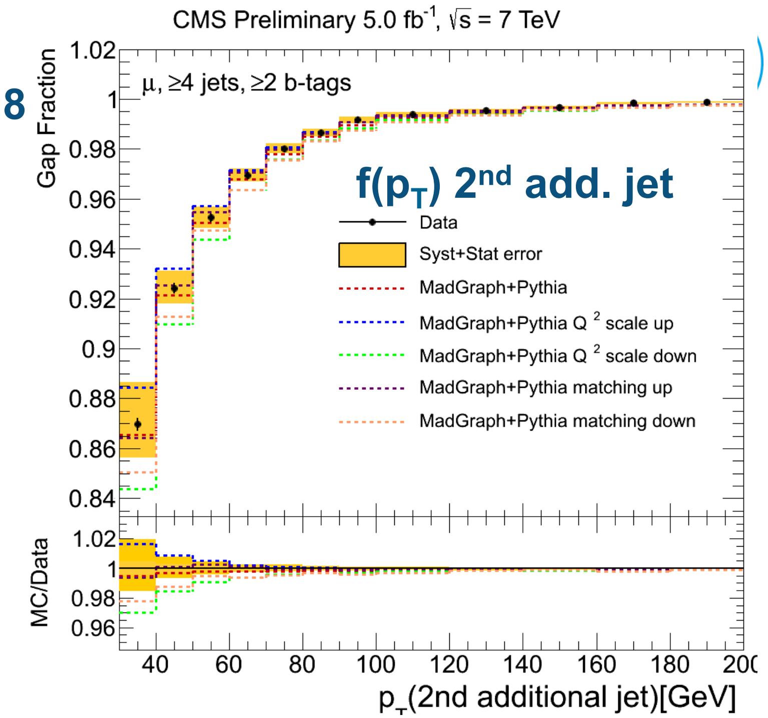
- Higher Q^2 seems to describe data better
- Experimental precision smaller than spread due to parameter variation

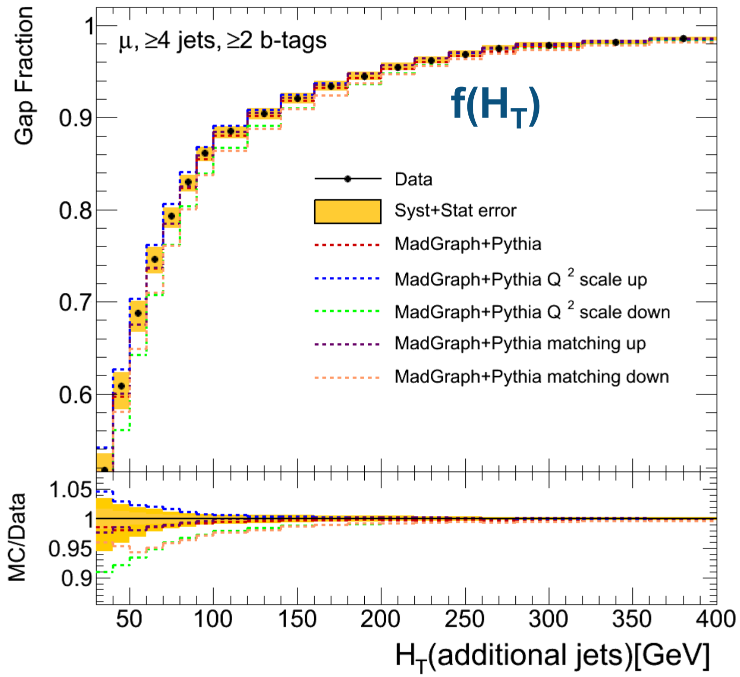
Gap fraction – μ+jets

CMS-PAS TOP-12-018

- Gap fraction vs. MadGraph varied scales





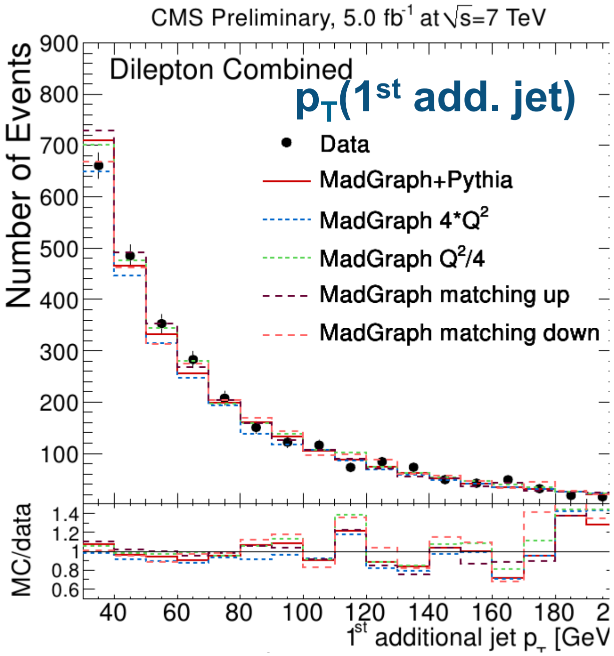


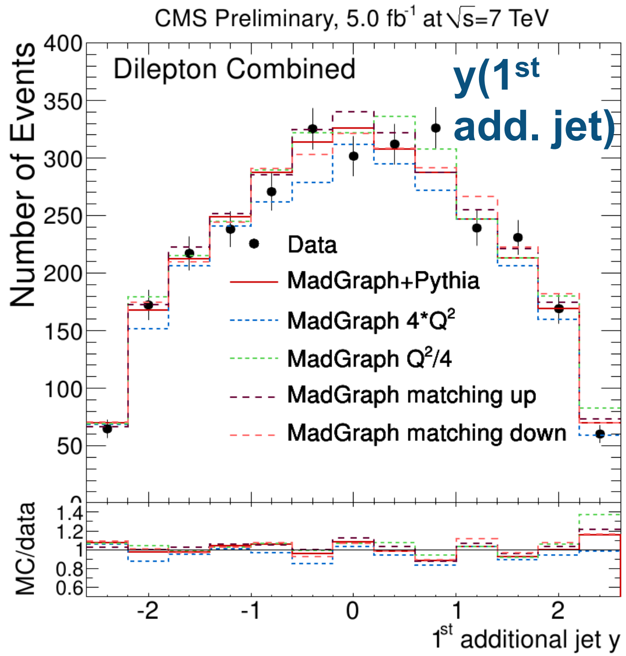
- 1st add. jet and H_T: higher scales seems to describe data better
- 2nd add. jet: seems better described by "nominal" MadGraph+Pythia
- All distributions: experimental precision smaller than spread due to parameter variation

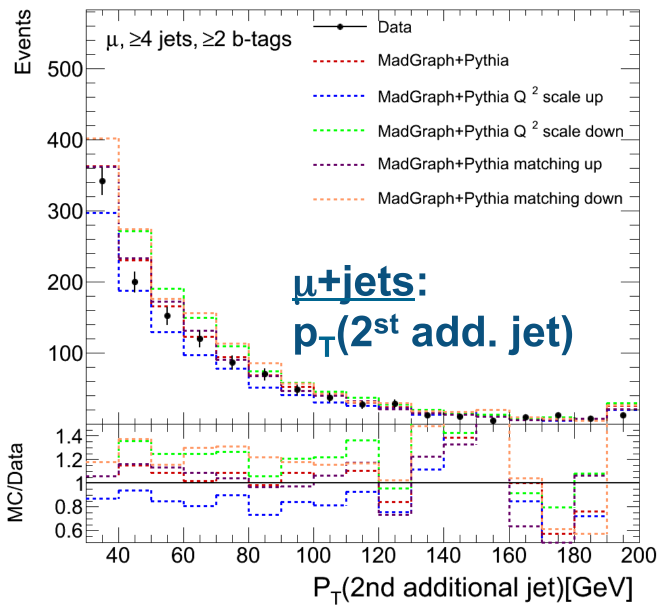
Kinematic variables add. jets

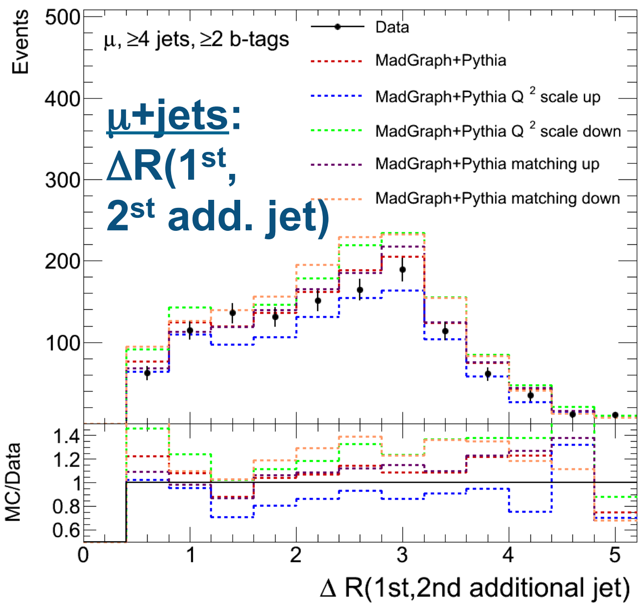
CMS-PAS TOP-12-018   CMS-PAS TOP-12-023

- Distributions at reconstructed level (no unfolding applied!)
- Background is subtracted from data
- Comparison to MadGraph varied scales
- All predictions normalized to approx. NNLO calculation









Work in progress !

Summary

- First studies towards constraining quark and gluon radiation in ttbar events
  - Differential ttbar production xsec vs p_T(top), p_T(ttbar)
  - Jet multiplicity in ttbar and associated jets
  - Gap fractions
- Compared to different MCs and parameter variations from Madgraph
  - General good agreement with different predictions
  - Often, experimental precision smaller than spread due to parameter variation

  → Variations could be significantly reduced
- Studies ongoing to find the best way to constrain the MC radiation parameters (Q^2, matching) with data
  - Comparison with POWHEG+Herwig would be useful
- On a larger timescale, interesting to look into new NLO matching tools
  - aMC@NLO, Sherpa, …

Additional information

$1/\sigma\ d\sigma/dX$ – Phase space definitions

reconstructed quantities:
top quarks, $t\bar{t}$ system

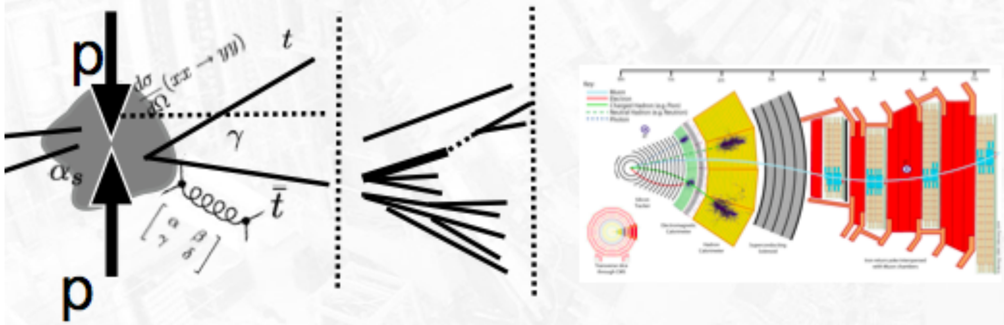
→ extrapolated parton level PS

correct for

detector effects
hadronization effects
extrapolate to full PS

→ as close to theory as possible

parton **particle** **reconstructed**



directly measurable quantities:
lepton(s), b-jets

→ visible particle level PS

correct for

detector effects
no hadronization correction
visible PS, no extrapolation:

$p_T^{jets} > 30$ GeV, $\eta^{jets} < 2.4$,

$p_T^{lep} > 20\ (30)$ GeV, $\eta^{lep} < 2.4\ (2.1)$

for dilepton (lepton+Jets) PS

→ **object definition on generator level:**

- particles after radiation & hadronization
- jets: same jet algorithm
- b-Jets: identified by B-hadron
- leptons: from W, $\Delta R(\text{lep,genJet}) > 0.4$

→ as model independent as possible

1/σ dσ/dX – Systematic uncertainties

- Determined individually for each bin of the measurement
- Normalized cross sections: only shape uncertainties contribute, correlated uncertainties cancel

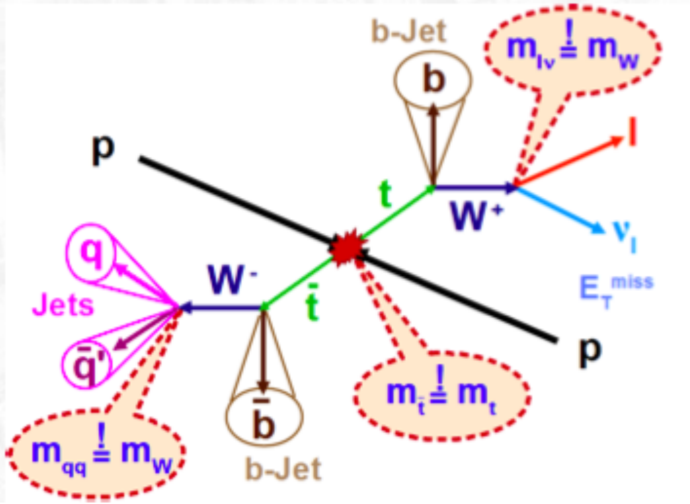
Typical values per bin

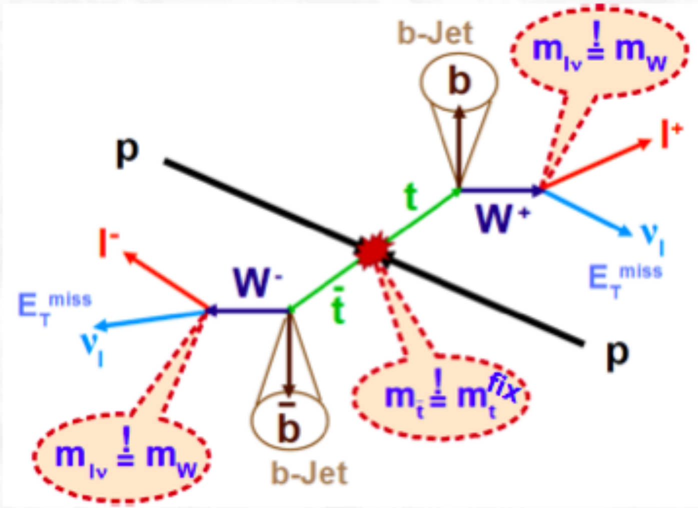
	Source	Method	Systematic uncertainty (%)	
			ℓ+jets	dileptons
Experimental	Background	vary with 30%-50%	3.5	0.5
	Trigger eff.	p_T-η dependent	0.5	1.5
	Lepton sel.	p_T-η dependent	0.5	2.0
	Jet energy scale	p_T-η dependent	1.0	0.5
	Jet energy resolution	p_T-η dependent	0.5	0.5
	Pileup	vary $\sigma_{\text{inel.}}(\text{pp}) \pm 8\%$	0.5	0.5
	b tagging	p_T-η dependent	1.0	0.5
	Kinematic reco	p_T-η dependent	–	0.5
Model	Q^2	vary factor 0.25–4	2.0	1.0
	ME/PS threshold	vary factor 0.5–2	2.0	1.0
	Hadronisation	PYTHIA vs. HERWIG	2.0	2.0
	Top-quark mass	172.5 ± 0.9	0.5	0.5
	PDF choice	PDF4LHC	1.5	1.0

1/σ dσ/dX – Kinematic reco of the ttbar pair

Lepton + jets: Kinematic fit

- vary measured 4-momenta for lepton, jets and neutrino
- to fullfil constrains:
  - $m_W \equiv$ 80.4 GeV
  - $m_t \equiv m_{\bar{t}}$
- neutrino: E_t^{miss}, p_z unmeasured as initial value
- consider 5 leading jets
- use b-tag information for b-jet association
- choose permutation with lowest variation wrt. object resolution (minimum χ^2)





Dilepton: Kinematic reco (~MWT)

- underconstrained (2 neutrinos)
- constraints:
  - $m_W \equiv$ 80.4 GeV
  - $m_t \equiv m_{\bar{t}}$ = fixed
  - $p_{\nu 1}(x,y) + p_{\nu 2}(x,y) = E_t^{miss}(x,y)$
- vary m_t (1 GeV steps): 100 – 300 GeV
- prefer solutions with b-tagged jets
- choose solution with best reconstructed neutrino energy wrt. MC spectrum

$1/\sigma\ d\sigma/dN$(add. partons) – μ+jets

Template Fit Setup

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

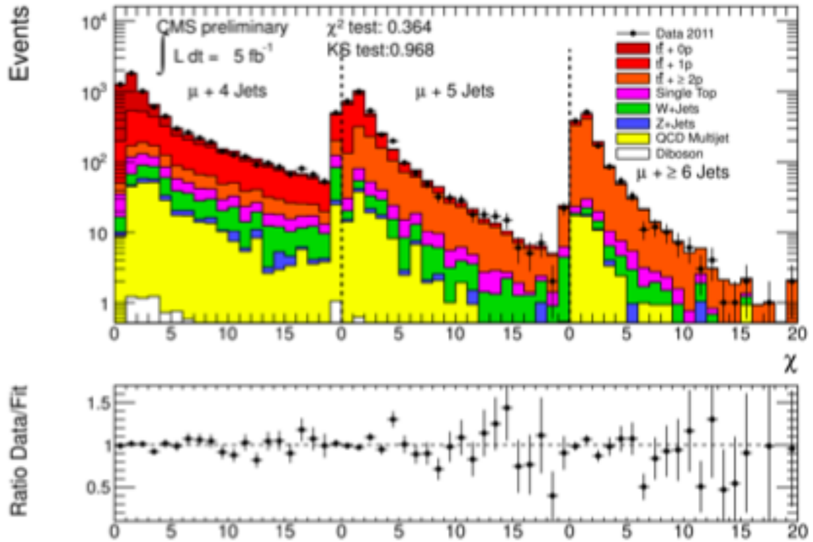
$$\chi=\sqrt{\left(\frac{m^{rec}_{W^{had}}-m^{true}_{W^{had}}}{\sigma_{W^{had}}}\right)^2+\left(\frac{m^{rec}_{t^{had}}-m^{true}_{t^{had}}}{\sigma_{t^{had}}}\right)^2+\left(\frac{m^{rec}_{t^{lep}}-m^{true}_{t^{lep}}}{\sigma_{t^{lep}}}\right)^2}$$

- 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

1/σ dσ/dN(add. partons) – μ+jets

Results

- 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 up also observed with this measurement