Hard QCD @ CMS

Piergiulio Lenzi – CERN

QCD@LHC 2011
Overview of the QCD physics program at the LHC

- CMS detector
  - Detection techniques for jets

CMS has produced a large amount of QCD measurements on the 2010 data sample

- Jet inclusive spectra
- Di-jet mass, angular correlations
- Event shapes
- Forward jets
- Inclusive photon production differential spectra
- W/Z + jets, Z+ heavy flavor
Hard QCD at LHC

- Hard QCD processes are important for two broad classes of reasons
- They represent a ubiquitous source of background for virtually any signal (both SM and searches) at a hadron collider
- They provide a tool to test the predictions of perturbative QCD
  ▪ The current understanding of our detectors allows both ATLAS and CMS collaborations to do **precision** QCD measurements
Available predictions

- Accurate predictions for dijet production, W/Z/gamma + jets production at the LHC are available
  - Monte Carlo event generators
    - NLO + parton shower (MC@NLO, POWHEG)
    - LO (many legs) + parton shower (Alpgen, MadGraph, Sherpa)
  - Parton level codes for distributions at NLO
- Modern parton distribution functions
- 4 T solenoid
- Pixel + SiStrip tracker
- Scintillating crystals (PbWO$_4$) electromagnetic calorimeter
- Brass/plastic hadron calorimeter (non-compensating)
- Muon spectrometer in the magnet iron return yoke
Jet reconstruction

- Jets are reconstructed with the anti-kt algorithm, with radius of 0.5 or 0.7
- 3 available algorithms for jet reconstruction
  - Calo-Jets: use only the calorimeter towers
  - Jet-Plus-Track Jets: improve the calorimeter jets using the tracks in the jet cone
  - Particle-Flow jets: uses particle flow candidates as input to the clustering algorithm
    - Particle flow reconstruction:
      - global event reconstruction
      - Identifies muons, electrons, taus, photons, charged hadron, neutral hadrons
      - Combines the information from all detectors
Jet reconstruction

- Jets are reconstructed with the anti-kt algorithm, with radius of 0.5 or 0.7
- 3 available algorithms for jet reconstruction
  - Calo-Jets: use only the calorimeter towers
  - Jet-Plus-Track Jets: improve the calorimeter jets using the tracks in the jet cone
  - Particle-Flow jets: uses particle flow candidates as input to the clustering algorithm

  ▪ Particle flow reconstruction:
    ▪ global event reconstruction
    ▪ Identifies muons, electrons, taus, photons, charged hadron, neutral hadrons
    ▪ Combines the information from all detectors
Jet energy scale

- We use a multi-step procedure to correct the energy of our jets

\[ p_{\mu}^{\text{cor}} = C \cdot p_{\mu}^{\text{raw}}. \]

\[ C = C_{\text{offset}}(p_{T}^{\text{raw}}) \cdot C_{\text{MC}}(p_{T}, \eta) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p_{T}) \]

- \( C_{\text{offset}} \) accounts for detector noise and pile-up

- The method uses correction factors extracted from the full simulation of CMS, \( C_{\text{MC}} \)

- Residual differences with respect to data are accounted for as further scaling factors
  - \( C_{\text{rel}} \) accounts for non-uniformity in eta. It is obtained applying on data and MC the di-jet balance method
  - \( C_{\text{abs}} \) accounts for residual absolute scale differences between data and MC. It is obtained applying on data and MC the \( \gamma \)+jet and \( Z \)+jet pT balancing

- In this MC + residual method effects like the presence of additional radiation spoiling dijet or \( \gamma \)+jet and \( Z \)+jet balancing enter only at second order
Jet energy scale

- Total systematic uncertainty on the energy scale for particle-flow jets
- The main sources of uncertainty are:
  - The photon energy scale, known at 1%
  - The relative response across detector regions
  - Pile-up effects
  - Extrapolations down to 0 for the additional activity in the balance methods
  - Dependency on jet flavor in the MC used
Jet energy resolution

- Determined with di-jet and γ+jet pT balance
- Plots show two example regions in η
- Resolution is of the order of 10% around 50 GeV
Inclusive jets

- Jet $p_T$ spectra are measured in the 18-1100 GeV range
- In 6 rapidity intervals, up to 3
- Resolution effect are unfolded
- Main systematic: jet energy scale
- Data are compared with the predictions at NLO, including non-perturbative (NP) corrections obtained with a shower MC

CERN-PH-EP-2011-053
Inclusive jets

- Data/theory ratios for the 6 rapidity bins
  - Experimental uncertainty represented by shaded area
  - Theoretical uncertainty as solid lines
    - The envelope of predictions from CT10, MSTW08 and NNPDF2.0 is used
    - The central values for the three PDF sets are also shown

- Data and theory agree within systematic uncertainty
- Predictions are systematically above data
- Shapes of data and of theory central predictions are similar
3-jets over 2-jets ratio

Measurement of the ratio of events with 3 or more jets over events with 2 or more jets, as a function of $H_T$ (scalar sum of jets' pT)

- Jets: $p_T > 50$ GeV, $|y|<2.5$
- Provides a stringent test of hard gluon radiation and higher order effects
- Several systematic effects cancel (largely or completely)
  - Luminosity
  - Jet energy scale

3-jets over 2-jets ratio

- Data fully corrected for detector effects with bin-by-bin corrections

- Main systematics:
  - Jet energy scale, unfolding uncertainties

- Comparison to several MC models:
  - Madgraph is the closest to data
    - Matched sample with up to 4 partons
  - Alpgen doesn't do quite as good
    - Why? Could the difference between Madgraph and alpgen be regarded as an estimate of the theory uncertainty?

- Pure shower models overestimate the ratio for $H_T < 0.5$ TeV
Azimuthal decorrelation

- \( \Delta \phi \) between the two leading jets in the event

- It is very sensitive to additional radiation effects (hence to higher order corrections) but also to MPI and hadronization

- Anti-kt (0.5) jets are required to have \( p_T > 30 \) GeV and \(|y| < 1.1\)

- Five bins of leading jet \( p_T \)

- Data corrected to hadron level

- Main sources of systematics
  - Jet energy scale
  - Transverse momentum resolution
  - Unfolding
Azimuthal decorrelation

Comparison to several MC models

- Pythia6 and Herwig++ provide the best description of data
- Madgraph (Pythia8) predict less (more) decorrelation
- Surprisingly, the matched calculation implemented in Madgraph doesn't provide a good description of data
  - Might be due to interplay between higher order corrections and tuning aspects
    - Might learn something about tuning
    - It would be useful to compare to other ME + PS models
Event shapes

- Distributions of central transverse thrust and thrust minor, using central ($|\eta|<1.3$) jets as input, in the transverse plane

\[ \tau_{\perp,C} \equiv 1 - \max_{\hat{n}_T} \sum_i |\vec{p}_{\perp,i} \cdot \hat{n}_T| \]

- Is a measurement of radiation along the thrust axis
- A dijet event has small values of central transverse thrust, while an isotropic multi jet has large values

\[ T_{m,C} \equiv \frac{\sum_i |\vec{p}_{\perp,i} \times \hat{n}_{T,C}|}{\sum_i p_{\perp,i}} \]

- Is a measurement of the radiation out of the plane defined by the thrust axis and the beams
- A dijet event has small values of central thrust minor, while an isotropic multi jet has large values

- Jets are reconstructed with the anti-kt algorithm
- $p_T > 30$ GeV
- 3 bins of leading jet $p_T
Event shapes

- 90 GeV < $p_T$(leading) < 125 GeV
- $p_T$(leading) > 200 GeV
Event shapes

- Pythia6 and Herwig++ do a good job in all bins
- Pythia8 tends to underestimate high values, i.e. very busy multi-jet events
- Both Alpgen and Madgraph are worse than the pure shower models
  - Why?
    - A pattern seems to emerge: it looks like ME+Shower are in general good at describing rates, but not as good at describing angles
- Does tuning play a role here?
- Checks with other tools are needed
Di-jet mass


- A measurement of the di-jet mass in 5 bins of leading jet rapidity, ranging from 0.2 to 2.5 TeV
- Anti-kt 0.7 jets, |y|<2.5
- Experimental resolution unfolded to hadron level with MC correction factors
- Comparison with pure NLO + non perturbative corrections
  - Theory prediction with CT10, MSTW2008, NNPDF2.0, folded according to PDF4LHC prescription
- Main systematic is the Jet energy scale
  - Experimental error comparable to theory uncertainty
- With improved energy scale systematic it will be possible to constrain PDFs

\[ d^2\sigma/dM_{JJ}dy_{\text{max}} = \frac{(pb/TeV)}{10^{17}} \]

\[ \sqrt{s} = 7 \text{ TeV} \]

\[ L_{\text{int}} = 36 \text{ pb}^{-1} \]

\[ l_y \text{ max} < 0.5 \]
\[ 0.5 < l_y \text{ max} < 1.0 (\times 10^4) \]
\[ 1.0 < l_y \text{ max} < 1.5 (\times 10^5) \]
\[ 1.5 < l_y \text{ max} < 2.0 (\times 10^6) \]
\[ 2.0 < l_y \text{ max} < 2.5 (\times 10^6) \]

pQCD at NLO \( \otimes \) Non Pert. Corr.

PDF4LHC

\[ \mu_F = \mu_R = p_T^{\text{ave}} \]
- Data show good agreement with predictions in all rapidity bins.
- The experimental uncertainty is comparable with the theoretical uncertainty.
- Data can be used to constrain PDFs.
**Inclusive forward jets**

- Inclusive measurement of the rate of jets in the forward region $3.2 < |\eta| < 4.7$
- Sensitive to PDFs
- Also sensitive to tuning aspects

- With more statistics and improved JES we will become more and more sensitive to PDFs
Forward-central jets

- An even more complicated topology:
  - One central jet ($|\eta| < 2.8$) and one forward jet ($3.2 < |\eta| < 4.7$)
  - $\text{PT} > 35 \text{ GeV}$
- It is sensitive to the details of the UE model and on the details of the shower
- Several MC generators were compared to the data
  - A particularly tough topology to get right
Forward-central jets

- All models overestimate the total rate
- Herwig seems to be best at describing both spectra
- Pythia8 and Pythia6 tune Z2 describe data better than D6T
- Powheg + Herwig is ok in shape but doesn't get the normalization right
- HEJ (pure parton level) describes data reasonably well
Inclusive photon production

- Prompt photon production is a stringent test of pQCD
- Measurement of differential production rate as a function of pT in bins of η
- The prompt photon signal is defined at particle level through an isolation cut of 5 GeV on the scalar sum of charged and neutral particles in a cone of 0.4 around the photon

- Analysis strategy:
  - Fit of the isolation distributions (non converted component)
  - Fit of the ratio Et in calorimeters to pT of the electrons from conversions (converted component)

- Main systematics:
  - Signal and background modeling in fits
  - Photon identification efficiency

The measurement has been performed in 4 photon rapidity bins, for transverse energies between 25 and 400 GeV.

Good agreement with NLO predictions from JETPHOX:
- Predictions are corrected for non-perturbative effects.
- MC predictions show a slight tendency to overshoot the data at low pT.
- Important as background for searches and as testing ground for higher order corrections in pQCD
- Detector's jet energy scale is the main systematic effect.
- CMS measured rates of events with jets accompanying the vector boson
  - Results are given within the kinematic acceptance for leptons, unfolding detector effects
  - Jets are reconstructed with the anti-kT algorithm, with a radius of 0.5, $p_T > 30$ GeV in CMS
- Pure parton shower (Pythia) is not able to describe multi jet rates
- Several Matrix Element + shower predictions compared to data
  - General agreement with these predictions is found
W/Z+jets

- CMS measured the associated production of Z + b-jets
- Z selection plus high purity b-tagging
- Main systematics: JES, b-tagging efficiency and mistag rate
- The ratio between the Z+ b jets and Z + any jet has been measured for both electron and muon decay channels

| Sample      | $\mathcal{R}(Z \rightarrow ee)$ (%) , $p_T^e > 25$ GeV, $|\eta^e| < 2.5$ | $\mathcal{R}(Z \rightarrow \mu\mu)$ (%) , $p_T^\mu > 20$ GeV, $|\eta^\mu| < 2.1$ |
|-------------|-------------------------------------------------|-------------------------------------------------|
| Data HE     | $4.3 \pm 0.6$ (stat) $\pm 1.1$ (syst)          | $5.1 \pm 0.6$ (stat) $\pm 1.3$ (syst)          |
| Data HP     | $5.4 \pm 1.0$ (stat) $\pm 1.2$ (syst)          | $4.6 \pm 0.8$ (stat) $\pm 1.1$ (syst)          |
| MADGRAPH    | $5.1 \pm 0.2$ (stat) $\pm 0.2$ (syst) $\pm 0.6$ (theory) | $5.3 \pm 0.1$ (stat) $\pm 0.2$ (syst) $\pm 0.6$ (theory) |
| MCFM        | $4.3 \pm 0.5$ (theory)                          | $4.7 \pm 0.5$ (theory)                          |
Conclusion

- The CMS QCD program is progressing very well!
- CMS produced an large number of results with 2010 data
  - Cross sections
  - Differential distributions
  - Associated production of vector boson with jets (and b-jets)
  - Forward jet measurements
- Plenty of data to test different codes and different models
- And more results are coming from the 2011 data!
- W polarization for large transverse momentum
- Effect unique to pp collisions!
- CMS measured the effect for $p_T > 50$ GeV and found that $W$s are predominantly left-handed in pp collisions, as predicted by the SM
- Since the kinematic is not closed, the lepton-projection (LP) variable was used and fitted to data

$$L_P = \frac{\vec{p}_T(\ell) \cdot \vec{p}_T(W)}{|\vec{p}_T(W)|^2}$$

![Graphs showing LP distribution for different scenarios]
Inclusive photon production

- Data to theory ratios in the four rapidity bins
- Shaded area is the data uncertainty
- PDF and scale uncertainties on the predictions are also shown