Recent CMS results on the Soft QCD & Forward Physics

Ankita Mehta
(on behalf of the CMS Collaboration)

Eötvös Loránd University, Budapest

27th Workshop on Deep-Inelastic Scattering & Related Subjects
Turin, Italy
Bird’s Eye View

Understanding particle production at the LHC → Important to realize it’s physics goals

- hard interaction & UE
- Diffractive processes dominate in forward regions
- Measurements of UE activity in central regions
  - Beam-beam remnants
  - Multiple-parton interactions (double-parton scattering)
  - Initial & final state radiations
  - Sensitive to interplay between perturbative & non-perturbative regions
  - Non-perturbative phenomenological models → free parameters to tune
  - Forward energy drives development of cosmic ray induced air showers

Expt. measurements → MB, UE, total and diffractive cross section & particle correlations

Soft interactions: Why to study them?
- Responsible for a very large fraction of the total cross section
- Their modeling impacts all high-$p_T$ measurements
- Indispensable ingredients to improve background estimates for SM & BSM processes
Soft QCD & forward physics at CMS → Facilitated by the forward instrumentation

- CASTOR calorimeter in the very forward region of the CMS ($-6.6 < \eta < -5.2$)
- Cherenkov sampling calorimeter, consisting of quartz and tungsten plates, with an overall depth of 10 interaction lengths

This talk covers results on energy measurements using CASTOR & DPS WW production
Average Very Forward Energy @13 TeV (CMS PAS FSQ-18-001)

- Energy carried by particles produced in the very forward region powerful probe to study UE activity
- Increase of energy with multiplicity is driven by MPI → Model validation & tuning
- Relation between electromagnetic & hadronic energy can constrain muon production in air showers
- First correlation study of hadron activity at very forward & central rapidities performed at 13 TeV
- Results with 0.22 nb$^{-1}$ of low pileup pp data selected using Zerobias triggers at Zero Tesla

Rich variety of MC samples compared with data

- PYTHIA8 (CUETP8M1, 4C+MBR, CP5)
- QGSJETII.04
- EPOS LHC
- SIBYLL (2.1,2.3c)
- HERWIG7.1
Analysis Ingredients

- Event selection:
  - Activity in at-least one tower of HF calorimeter
  - At-least one track reconstructed in CMS tracker with $|\eta| < 2.0$
  - Cut on reconstructed vertex multiplicity → reduce pileup contributions

- Pixel-based track reconstruction → straight line tracking & vertexing
- Tracking efficiency $\sim 76\%$ & misreconstruction probability $\sim 5\%$ for charged particles with $p_T > 200$ MeV
- Event classification based on number of reconstructed tracks ($N_{\text{tracks}}$)
- CASTOR energy scale → Dominating source of uncertainty

<table>
<thead>
<tr>
<th>Source</th>
<th>Total energy</th>
<th>Electromagnetic energy</th>
<th>Hadronic energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASTOR energy scale</td>
<td>17%</td>
<td>17%</td>
<td>17%</td>
</tr>
<tr>
<td>CASTOR intercalibration</td>
<td>2–3%</td>
<td>-8%</td>
<td>+15%</td>
</tr>
<tr>
<td>HF energy scale</td>
<td>&lt;0.5%</td>
<td>&lt;0.5%</td>
<td>&lt;0.5%</td>
</tr>
<tr>
<td>Tracking efficiency</td>
<td>1–5%</td>
<td>1–5%</td>
<td>1–5%</td>
</tr>
<tr>
<td>Pileup rejection</td>
<td>1–8%</td>
<td>1–8%</td>
<td>1–10%</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>0.05–1.6%</td>
<td>0.06–1.9%</td>
<td>0.06–1.8%</td>
</tr>
<tr>
<td>Total</td>
<td>18–19%</td>
<td>18–20%</td>
<td>20–26%</td>
</tr>
</tbody>
</table>

- Novel forward folding technique:
  - Model/theory → Detector level
  - Particle multiplicity and CASTOR energy are smeared
Total energy deposited in CASTOR;

\[ E_{\text{reco}}^{\text{tot}} = \sum_{i=\text{towers}} E_i; E_i > \text{Noise threshold} \]

- \( \langle E_{\text{reco}}^{\text{tot}} \rangle \) increases with \( N_{\text{tracks}} \)
- Only Sibyll 2.X & Herwig 7.1 describe the relative increase well
- Mismatch strongest for EPOS LHC & PYTHIA8 CP5
Electromagnetic & Hadronic Energy Components

Relevant for simulation of cosmic ray induced extensive air showers
Point towards the modeling accuracy for neutral vs charged pions

\[ \langle E_{\text{em reco}} \rangle \] described well by all models except SIBYLL2.3c

PYTHIA8 4C+MBR slightly underestimates \[ \langle E_{\text{em reco}} \rangle \] at low values of \[ N_{\text{tracks}} \]

\[ \langle E_{\text{had reco}} \rangle \rightarrow \] overestimated by all but SIBYLL2.3c & PYTHIA8 4C+MBR models
- Sensitive to differences in underlying final state hadron production mechanisms

- Ratio is almost constant over the whole track multiplicity range → No dramatic change of the particle production mechanism in forward regions

- All model predictions are lower than the data

- Energy ratio best described by QGSJETII.04, SIBYLL2.1, & HERWIG7.1
Double-parton scattering (DPS) ⇒ Two separate hard parton-parton interactions in a single pp collision → Grows more rapidly as compared to SPS with $\sqrt{s}$

\[
\sigma_{AB}^{\text{DPS}} = \frac{m}{2} \sum_{i,j,k,l} \int \Gamma_{ij}(x_1, x_2, y; Q_A, Q_B) \Gamma_{kl}(x'_1, x'_2, y; Q_A, Q_B) \hat{\sigma}_{ik}^{A}(x_1, x'_1) \hat{\sigma}_{jl}^{B}(x_2, x'_2) \, dx_1 \, dx'_1 \, dx_2 \, dx'_2 \, d^2y
\]

$m = 2$ if $A \neq B$, else 1

doouble-parton distribution functions (dPDFs)

Pocket formula: $\sigma_{AB}^{\text{DPS}} = \frac{m}{2} \frac{\sigma_A \sigma_B}{\sigma_{\text{eff}}}$ ⇒ Used in all phenomenological calculations

$\sigma_{\text{eff}} \rightarrow$ transverse profile of partons → Assumed to be process & energy independent

Experimental measurements: $\sigma_{\text{eff}} \rightarrow 15–25$ mb with uncertainties $\approx 30$

**Importance of DPS**

- Possible to explore at colliders → even using high scale process at the LHC
- Provides information about hadron structure in transverse plane
- Understanding of background contributions to interesting SM & BSM processes
DPS With Same-Sign WW @13TeV

- WW production → Golden channel for DPS production
- Quark initiated → Sensitive to longitudinal quark polarizations
- Non-factorization models predict spin, color, momentum ... interference effects (Phase-2 Upgrade of CMS Muon Detectors)
- SPS $W^\pm W^\pm$ production suppressed at matrix element level
- Insensitive to pileup effects & clean final state with fully leptonic W decays

PYTHIA8 predicts a cross section value of 1.9 pb for inclusive WW production via DPS @13TeV → calculated with $\sigma_{\text{eff}} = 28$ mb which is also generator tune dependent!!

$\sigma_W(\text{NNLO}) \oplus \sigma_{\text{eff.}} = 20.7 \pm 6.6$ mb (CMS)→ $\sigma_{\text{factorized}}^{\text{DPSWW}} = 0.87$ pb

Comparison of measured cross section with predictions → Important input for development and testing of existing models of dPDFs → Improved MC models
77 fb$^{-1}$ of data from combined 2016 & 2017 at $\sqrt{s} = 13$ TeV

**Signal** ⇒ two same-sign leptons (dimuon or electron-muon pairs) $\oplus p_T^{\text{miss}}$

**pythia8 & herwig++** signal samples

Broad spectrum of background processes & few variables to play with!!

Dominant backgrounds: WZ & non-prompt leptons

Signal & background discrimination based on BDT classifiers; trained separately against dominant backgrounds

Two BDT classifiers → 1D classifier with bins ordered in S/B for statistical analysis
BDT Classifier Training

- Explore topological differences b/w DPS & background processes
- No correlations expected in leptons’ kinematic phase space for signal
- Leptons from background processes share the boost → correlations in $\eta \phi$
- Two different BDTs trained, one against WZ & another against fakes

$\rightarrow$ $p_T^{l_1, 2}$
$\rightarrow$ $p_T^{\text{miss}}$
$\rightarrow$ $M_{T2}^{\text{ll}}$
$\rightarrow$ $|\eta_1 \times \eta_2|$
$\rightarrow$ $|\eta_1 + \eta_2|$

BDT classifier outputs

$\rightarrow$ $m_T(l_1, p_T^{\text{miss}})$
$\rightarrow$ $m_T(l_1, l_2)$
$\rightarrow$ $\Delta\phi(l_1, l_2)$
$\rightarrow$ $\Delta\phi(l_2, p_T^{\text{miss}})$
$\rightarrow$ $\Delta\phi(l_1 l_2, l_2)$

$\rightarrow$ two BDT classifiers → one final discriminant variable

$\rightarrow$ Bins are defined in a way to have few bins dominated by signal & few by WZ & fakes
Maximum likelihood fit to the final classifier
Fitting is performed in 4 different lepton charge & flavor categories → Benefits from asymmetry in W production → better signal sensitivity (by ~10%)
Expected to be more sensitive to ++ configuration than --

Preliminary
CMS (13 TeV) 1−77 fb

Results
**First evidence of DPS WW**

Results from same-sign WW are extrapolated to the inclusive WW phase space.

Observed cross section is used to extract $\sigma_{\text{eff}}$.

---

**Table:**

<table>
<thead>
<tr>
<th>obtained value</th>
<th>significance (standard deviations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{\text{DPSWW}}$, exp</td>
<td>1.92 pb</td>
</tr>
<tr>
<td>$\sigma_{\text{DPSWW}}$, factorized</td>
<td>0.87 pb</td>
</tr>
<tr>
<td>$\sigma_{\text{DPSWW}}$, obs</td>
<td>1.41 ± 0.28 (stat) ± 0.28 (syst) pb</td>
</tr>
<tr>
<td>$\sigma_{\text{eff}}$</td>
<td>12.7 ± 5.0 mb</td>
</tr>
</tbody>
</table>

**Graph:**

- **Observed**
  - $\mu^-\mu^-e^-\mu^-$
  - $\mu^+\mu^+e^+\mu^-$
  - $\mu^-\mu^+e^-\mu^-$

- **Predictions:**
  - PYTHIA 8 (CP5)
  - Factorization approach

- **Total, stat, syst**
  - $1.96 ± 0.74 (±0.54, ±0.51) \text{ pb}$
  - $1.36 ± 0.46 (±0.33, ±0.32) \text{ pb}$
  - $1.41 ± 0.40 (±0.28, ±0.28) \text{ pb}$

---

**CMS Preliminary**

77 fb$^{-1}$ (13 TeV)

---

**$\sigma_{\text{eff}}$ extractions (vector boson final states):**

- CDF $\gamma+3$jets (1.8 TeV)
- D0 $\gamma+3$jets (1.96 TeV)
- D0 $\gamma+b/c+2$jets (1.96 TeV)
- ATLAS $W+2$jets (7 TeV)
- CMS $W+2$jets (7 TeV)
- ATLAS $Z+J/\gamma$ (8 TeV)
- CMS $W^+W^-$ DPS (8 TeV)
- CMS $W^+W^-$ (13 TeV)
- ATLAS $4l$ (13 TeV)
- CMS $W^+W^-$ (13 TeV)
Data from the LHC provide a new energy scale for studying soft QCD & forward physics

Soft QCD processes → Test predictions from phenomenological models ⊕ input for their improvement

Still quite a few unresolved problems, but we possess a wealth of data

Model parameters tuned to UE data at central rapidities are consistent with the very forward data within experimental uncertainties

Energy measurements in the very forward $\eta$ regions indicate some interesting potential to further improve the underlying event model predictions

DPS measurements → Important to understand partonic structure of hadrons & for new physics searches @ LHC; very sensitive to non-factorization models

First evidence for DPS WW production using 2016+2017 CMS data

Could do some interesting DPS physics with full Run2 data (differential cross sections, correlation studies ....) other than a DPS WW observation

*source: Internet