CMS: Jet and missing ET reconstruction

Milos Dordevic
Vinca Institute of Nuclear Sciences, University of Belgrade
On behalf of the CMS Collaboration

25-30 May 2020
Jets at CMS clustered using the **anti-k\_T algorithm** (mostly using the R = 0.4, 0.8)

- **Particle-level jets**: stable and visible particles in gen.evt; **Calo jets**: from energy deposits in calorimeter towers; **Particle Flow jets**: by clustering PF candidates; **PF + CHS** (Charged Hadron Subtraction) and Pile Up Per Particle ID (PUPPI) jets
Jet and M(issing)ET reconstruction at HLT

- The adapted jet and MET reco runs at HLT level and is speed/performance optimized
- Particle Flow@HLT x100 faster than offline
- At the HLT tracking reduced to 3 iterations
- L1L2L3 JEC derived for HLT (only from MC)

https://twiki.cern.ch/twiki/bin/view/CMSPublic/HLTplotsSummary2016

Efficient HLT reconstruction of jets & MET

- The jet HLT efficiency as a function of the offline jet $p_T$, measured using the single-muon sample
- The $E_T^{\text{miss}}$ trigger efficiency as a function of offline $E_T^{\text{miss}}$, measured using the single-electron sample
Pileup mitigation techniques at CMS

- **Pile-up** became an ever growing challenge in LHC physics
- The LHC Run 2: ~29 interactions/evt (Run 3 exp up to 50)
- Charged hadrons subtraction (CHS) algo uses the tracking info to remove particles associated to the pileup vertices

![Diagram of pileup mitigation](image)

- **Pile Up Per Particle ID (PUPPI)** use distribution of neighboring particles to estimate probability of neutral particles to originate from pileup

- After CHS some PU jets remain -> **PU Jet ID: MVA** to reject jets from pileup particles

---

M. Dordevic (Vinca Institute, UB) 26 May 2020, LHCP2020
Pileup mitigation techniques: performance

- Jet energy resolution as function of ptcl-level jet pT for PF, PF+CHS & PUPPI jets in QCD MC
- PUPPI has better performance than PF (+CHS) (neutral PU ptcl. contribute more to AK8 jets)

- Ratio of the total number of jets with $|\eta| < 2.5$, $p_T > 20$ GeV over the corresponding number of hard scatter jets before and after applying the PU ID WP corresponding to the 95% efficiency
- Ratios for PU jets before and after the PU jet ID

--- NEW for LHCP 2020
Pileup mitigation: efficiency and purity

- Measured in MC: CHS (+PU JetID) vs PUPPI

- **Efficiency** defined as a fraction of particle-level jets with $p_T > 30$ GeV matched ($\Delta R < 0.4$) with reconstruction jets of $p_T > 20$ GeV

- **Purity**: fraction of reco-level jets ($p_T > 30$ GeV) matched ($\Delta R < 0.4$) with generator jets ($p_T > 20$ GeV) from main interaction

- CHS efficiency over 95%, but purity drops to 70% in barrel

- PUPPI has an improved efficiency and purity overall perform. in cent. region
Jet energy corrections (JEC) at CMS

- JEC procedure: a factorized approach to correct the jets to particle jet level

- Pileup correction in order to account for offset energy coming from pileup
- Correction to the particle level jet vs $p_T$ and $\eta$ obtained from MC simulation
- Small residual corrections to data for pileup, relative vs $\eta$, absolute vs $p_T$ -> full physics analysis to derive residuals

NEW for LHCP 2020 ---v

CMS Preliminary

Markers: Data, Histograms: MC
- Photons
- EM Deposits
- Neutral Hadrons
- Hadronic Deposits
- Unassoc. Charged Hadrons
- Assoc. Charged Hadrons

CMS DP-2020/019

2017 41.5 fb$^{-1}$ (13 TeV)
Jet energy corrections: performance

- Jet response, \( \langle p_T^{\text{RECO}} \rangle / \langle p_T^{\text{ptcl}} \rangle \), corrections in bins of \( p_T^{\text{jet}}, |\eta_{\text{jet}}| \)
- Stable in the barrel (BB) region
- N. had. resp. 0.6, 15% of \( p_T^{\text{ptcl}} \)
- Stronger depend. in EC and HF
- EC2-> calorimeter degradation

\[ \chi^2 / \text{NDF} = 63.6 / 32 \]

- Data-to-simulation comparison for the jet response dependence on the jet \( p_T \)
- Combination of \( \gamma + \text{jet}, Z + \text{jet} & \text{Multijet} \) (2016)
- Yellow band indicates absolute scale uncertainty
Jet energy (scale) uncertainties and resolution

- The Jet Energy Scale (JES) uncertainty sources and total as function of jet $p_T$
- Run I result without the flavour and time sources is shown for comparison

Jet Energy Resolution (JER) measured in dijet and $Z/\gamma +$ jet simulated events vs $p_T^{ptcl}$, $\eta$ and $\mu$ and data to MC scale factors from di-jet applied in addition

- SFs of 1.1-1.2, larger in the EC-HF transition region of $|\eta| \in [2.5, 3]$
MET reconstruction, cleaning and performance

- **PF/PUPPI MET definition:**
  \[ \vec{p}_T^{\text{miss}} = \vec{p}_T^{\text{miss, raw}} - \sum_{i \in \text{jets}} (\vec{p}_{T,jet}^{\text{corr}} - \vec{p}_T^{\text{jet}}) \]
  \[ p_T^{\text{miss}} = p_T^{\text{miss, raw}} - \sum_{\text{jets}} (p_{T,jet}^{\text{corr}} - p_T^{\text{jet}}) \]
  \[ p_T^{\text{miss}} (\text{Type-I MET}) \rightarrow \]

- **Jet energy corrections propagated** \( p_T^{\text{miss}} \) (Type-I MET) \( \rightarrow \)

- **Anomalous MET events** \( \rightarrow \) mostly due to detector noise

Response \( \sim 1 \) for \( q_T > 100 \text{ GeV} \)

PUPPI MET has **20% better resolution** for avg. Run 2 PU

Stable performance vers PU
Standard and ML heavy object tagging

- Standard heavy object taggers -> groomed mass & N-subjettiness

Machine-learning based taggers -> large performance improvements vs non–ML

ML : N$_3$-BDT, BEST, ImageTop & DeepAK8

NEW for LHCP: DeepAK8-DDT, ParticleNet

Talk on boosted objects by Pantelis Kontaxakis
Summary and Outlook

- Demonstrated the ability to deal with the pileup conditions expected in Run 3 with mitigation techniques exercised in Run 2
- Significant gain in MET performance using the new PU mitigation techniques (PUPPI)
- Further evolving boosted object taggers

Performance of missing transverse momentum in pp collision at 13 TeV (arXiv:1903.06078v2)
Performance of the pile up jet identification in CMS for Run 2 (CMS DP-2020/020)
Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018 (CMS DP-2020/019)
Mitigation of anomalous missing transverse momentum measurements in data collected by CMS at √s=13 TeV during the LHC Run 2 (CMS DP-2020/018)
Identification of highly Lorentz-boosted heavy particles using graph neural networks and new mass decorrelation techniques (CMS DP-2020/002)

https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsJME

--- New for LHCP 2020
Backup Slides
The CMS detector at CERN

CMS DETECTOR
- Total weight: 14,000 tonnes
- Overall diameter: 15.0 m
- Overall length: 28.7 m
- Magnetic field: 3.8 T

STEEL RETURN YOKE
- 12,500 tonnes

SILICON TRACKERS
- Pixel (100x150 μm) ~16m² ~66M channels
- Microstrips (80x180 μm) ~200m² ~9.6M channels

SUPERCONDUCTING SOLENOID
- Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
- Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
- Endcaps: 540 Cathode Strip, 576 Resistive Plate Chambers

PRESHOWER
- Silicon strips ~16m² ~137,000 channels

FORWARD CALORIMETER
- Steel + Quartz fibres ~2,000 Channels

CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
- ~76,000 scintillating PbWO₄ crystals

HADRON CALORIMETER (HCAL)
- Brass + Plastic scintillator ~7,000 channels
Particles in the CMS detector

Key:
- Muon
- Electron
- Charged Hadron (e.g., Pion)
- Neutral Hadron (e.g., Neutron)
- Photon

Transverse slice through CMS

Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid

Iron return yoke interspersed with Muon chambers

M. Dordevic (Vinca Institute, UB)  26 May 2020, LHCP2020
Pileup mitigation techniques at CMS

\[ \alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < 0.4} \left( \frac{p_{Tj}}{\Delta R_{ij}} \right)^2 \begin{cases} & \text{for } |\eta_i| < 2.5, \ j \text{ are charged PF candidates from PV} \\ & \text{for } |\eta_i| > 2.5, \ j \text{ are all kinds of reconstructed PF candidates} \end{cases} \]

\[ \chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{\text{PU}})^2}{\text{RMS}^2_{\text{PU}}} \]  - to determine the probability that the PF candidate is from pile-up

\[ w_i = F_{\chi^2_{\text{NDF}=1}}(\chi_i^2) \]  - the weight being zero (one) if the PF candidate is from pileup (PV)
Pileup mitigation techniques at CMS (NEW)
### Pileup Jet ID: input variables

<table>
<thead>
<tr>
<th>Input variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>β</strong></td>
<td>Fraction of $p_T$ of charged particles associated with the LV, defined as $\sum_{i=LV} p_{T,i} / \sum_{i} p_{T,i}$ where $i$ iterates over all charged PF particles in the jet</td>
</tr>
<tr>
<td>$N_{\text{vertices}}$</td>
<td>Number of vertices in the event</td>
</tr>
<tr>
<td>$\langle \Delta R^2 \rangle$</td>
<td>Square distance from the jet axis scaled by $p_T^2$ average of jet constituents: $\sum_i \Delta R^2 p_{T,i}^2 / \sum_i p_{T,i}^2$</td>
</tr>
<tr>
<td>$f_{\text{ringX}}$, $X = 1, 2, 3, \text{and } 4$</td>
<td>Fraction of $p_T$ of the constituents ($\sum p_{T,i}^{\text{jet}} / p_T^{\text{jet}}$) in the region $R_i &lt; \Delta R &lt; R_{i+1}$ around the jet axis, where $R_i = 0, 0.1, 0.2$, and 0.3 for $X = 1, 2, 3, \text{and } 4$</td>
</tr>
<tr>
<td>$p_T^{\text{lead}} / p_T^{\text{jet}}$</td>
<td>$p_T$ fraction carried by the leading PF candidate</td>
</tr>
<tr>
<td>$p_T^{\text{lead}, \text{ch.}} / p_T^{\text{jet}}$</td>
<td>$p_T$ fraction carried by the leading charged PF candidate</td>
</tr>
<tr>
<td>$</td>
<td>\vec{m}</td>
</tr>
<tr>
<td>$N_{\text{total}}$</td>
<td>Number of PF candidates</td>
</tr>
<tr>
<td>$N_{\text{charged}}$</td>
<td>Number of charged PF candidates</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>Major axis of the jet ellipsoid in the $\eta$-$\phi$ space</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>Minor axis of the jet ellipsoid in the $\eta$-$\phi$ space</td>
</tr>
<tr>
<td>$p_T^{D}$</td>
<td>Jet fragmentation distribution, defined as $\sqrt{\sum_i p_{T,i}^2 / \sum_i p_{T,i}}$</td>
</tr>
</tbody>
</table>
Jet energy scale uncertainties

CMS Preliminary

2016 35.9 fb\(^{-1}\) (13 TeV)

R=0.4 PF+CHS
\(p_T = 30\) GeV

CMS Preliminary

2017 41.5 fb\(^{-1}\) (13 TeV)

R=0.4 PF+CHS
\(p_T = 30\) GeV

CMS Preliminary

2018 59.7 fb\(^{-1}\) (13 TeV)

R=0.4 PF+CHS
\(p_T = 30\) GeV

Total uncertainty
Excl. flavor, time
Run I
Absolute scale
Relative scale
Pileup (\(\langle \eta \rangle = 25\))
Method & sample
Jet flavor (QCD)
Time stability
Jet PF composition studied from dijet events using fully corrected jets.

Cross-check comparison between data and simulation for monitoring the stability of JES.

All categories considered: Photons, Leptons, Neutral and Charged Hadrons.

Fraction of energy removed by CHS before jet clustering is overlaid.
An event rejected by the HCAL noise filter

Jet
- $p_T = 9789$ GeV
- $\eta = 0.57$
- $\phi = 0.66$

HCAL energy deposit 1
- $E = 9728$ GeV
- $\eta = 0.53$
- $\phi = 0.66$

HCAL energy deposit 2
- $E = 1730$ GeV
- $\eta = 0.82$
- $\phi = 0.66$

$p_T^{miss}$
- $p_T = 10054$ GeV
- $\phi = -2.49$