CMS Physics Analysis Summary

Underlying Event Measurements with Leading Particles and Jets in pp collisions at \( \sqrt{s} = 13 \) TeV

The CMS Collaboration

Abstract

A measurement of the underlying event (UE) activity is performed in proton-proton collisions at the centre-of-mass energy of 13 TeV with the CMS experiment at the LHC. The measurement is performed using leading charged-particles as well as leading charged-particle jets as reference objects. A leading charged-particle or charged-particle jet is required to be produced in the central pseudorapidity region (\( |\eta| < 2 \)) and with transverse momentum \( p_T \geq 0.5 \) (\( p_{T,jet} \geq 1 \) GeV for leading charged-particle (charge particle jet). The UE activity is measured in terms of the average multiplicity and scalar sum of \( p_T \) of the charged-particles, with \( |\eta| < 2 \) and \( p_T \geq 0.5 \) GeV, in the azimuthal region transverse to the direction of the leading reference object. The measurements are compared to predictions of several Monte Carlo event generators providing constraints on the modelling of the UE dynamics.
1 Introduction

The underlying event (UE) in a high-energy hadron-hadron collision is defined as the activity that is not part of the final-state activity originating from the most energetic parton scattering of the collision. Unlike electron-positron collisions, hadron production in a proton-proton (pp) collision is understood to originate from multiple scatterings of the partonic constituents in a collision region of extended spatial dimensions. Theoretically, the noncolliding partons emitted in the very forward direction also contribute to the final-state radiation. The produced partons reduce their virtuality through gluon radiation and quark-antiquark splittings, and finally fragment into hadrons at scales approaching $0.2 \text{ GeV (}$Λ_{\text{QCD}}$)). Ideally, the produced hadrons in an event should be identified as those coming directly from the fragmentation of partons resulting from the scattering with the largest momentum transfer (hard scattering), and the rest which forms the UE. In reality, the final-state hadrons cannot be uniquely identified as coming from any one of these classifications. The UE activity is then comprised of hadrons coming from initial- and final-state radiation (ISR, FSR) from the hard scattering, softer partonic scatters in the same pp collision (multiple parton interactions, or MPI) possibly with their own initial-and final-state radiation, and proton remnants concentrated along the beam direction. However, experimentally “clean” observables that are relatively uncontaminated by the radiation from the hardest partonic scattering can be defined for the UE activity.

An accurate understanding of the UE is required for precise measurements of standard model processes at high energies and searches for new physics. The UE affects measurements of isolated high transverse momentum $p_T$ leptons or photons, and because it dominates most of the hadronic activity from the multiple pp collisions taking place in a given bunch crossing (pileup/PU) at the high luminosities achieved by the CERN LHC, it is important to understand the impact of its activity. Since the semi-hard and low-momentum partonic processes, which dominate the UE cannot be adequately calculated with perturbative Quantum Chromodynamics (pQCD) methods alone, the UE activity requires a phenomenological description containing parameters that must be tuned to data.

The experimental observables sensitive to the UE activity can be defined by exploiting the topological structure of pp collisions with a hard scattering. One example is the study of particle properties in regions away from the direction of the products of the hard scattering. At a given centre-of-mass energy, the UE activity is expected to increase with the momentum transfer between the interacting partons (hard scale). On average, increasingly hard parton interactions result from pp collisions with decreasing impact parameters between the two protons, which in turn enhance the overall hadronic activity originating from MPI until a saturation is reached for central collisions with maximum overlap [1, 2]. At the same time, the activity related to the ISR and FSR components also increases with the hard scale. For events with the same hard scale, the MPI activity rises with $\sqrt{s}$ as more partons are expected in the protons at increasingly smaller parton fractional momenta $x \sim 2p_T/\sqrt{s}$ [1, 2]. Hence, studying the UE as a function of the hard scale at several centre-of-mass energies provides an insight into the UE dynamics and its evolution with the collision energy, further constraining the model parameters.

The hard scale of the interaction and the UE activity with respect to the probe object can be probed by several methods. At the Tevatron, the CDF experiment measured UE observables using inclusive jet and Drell-Yan (DY) events as probes in pp collisions at the centre-of-mass energies of $\sqrt{s} = 0.63, 1.8,$ and 1.96 TeV [3–5]. In pp collisions at the LHC, the ALICE, ATLAS, and CMS experiments have carried out UE measurements at $\sqrt{s} = 0.9, 2.76,$ and 7 TeV using events containing a leading (highest $p_T$) charged-particle jet [6–9], a leading charged particle [10–12], a DY lepton pair [13], or a $t\bar{t}$ system [14]. In this analysis, the UE activity is studied
in pp collisions at \( \sqrt{s} = 13 \) TeV by measuring the average multiplicity and scalar transverse momentum sum \( (\Sigma p_T) / \text{average energy densities of charged-particles in the region orthogonal to the azimuthal direction of the leading charged-particle or jet, referred to as the transverse region.} \)

This note is organized as follows. The main features of the Monte Carlo models used are described in section 2 followed by section 3 which describes the CMS detector. Section 4 describes the data sample and the event selection and track selections. This is followed by section 5 which discusses the UE observables and the analysis methodology. Section 6 then covers the systematic uncertainties of the measurements, followed by the results and conclusion in sections 7 and 8.

## 2 Monte Carlo models

In this analysis PYTHIA\(^8\) (version 8.153) [15], HERWIG++ (version 2.7.0) [16], and EPOS (version 1.99) [17] generators are used with various tunes. Events are simulated with no pileup as well as with average pileup 1.3 required for the data correction and the calculation of the systematic uncertainties. Generated events are passed through detector simulation using GEANT [18]. These various event generators differ in the treatment of initial- and final-state radiation, hadronisation, colour reconnections, and the regularisation of the infrared cross-section divergence (including that for MPI). PYTHIA\(^8\) uses the Lund string hadronisation model [19], EPOS uses a similar string fragmentation hadronisation model [20], while HERWIG uses a cluster hadronisation model [21]. The MC generators were tuned by the optimisation of predictions to describe existing UE data, especially those measured with LHC pp collisions. However, minimum bias observables and data from the Tevatron, collected at lower centre-of-mass energies, were also used to obtain some of these tunes.

In PYTHIA\(^8\), the \( 2 \rightarrow 2 \) non-diffractive processes (including those in MPI), are described by lowest-order perturbative QCD with the divergence of the differential cross-section \( (d\sigma / dp_T^2) \propto \alpha_s^2 (p_T^2 / p_T^4) \) as \( p_T \rightarrow 0 \) regulated by the introduction of a sharp infrared \( p_T \) cutoff at \( p_{T0} \). With this cutoff, the differential cross-section becomes \( \alpha_s^2 (p_{T0}^2 + p_T^2) / (p_{T0}^2 + p_T^2)^2 \) which is no longer divergent in the infrared limit. The cutoff \( p_{T0} \) then becomes a free parameter which scales with energy in the same way as the total cross-section, \( p_{T0} (\sqrt{s}) = p_{T0}^\text{REF} \times (\sqrt{s} / E_0)^\epsilon \). Measurement of the UE at various energies is needed to obtain the centre-of-mass energy dependence exponent, \( \epsilon \). In HERWIG++, the energy dependence of MPI is modelled as in the same way [16].

The ordering of the parton showers differ among the various Monte Carlo generators. In PYTHIA\(^8\), \( p_T \)-ordered parton showers are used and the MPI and parton showers are interleaved in one common sequence of decreasing \( p_T \) values. For PYTHIA\(^8\), interleaving is done between the initial-state, final-state showers, and MPI. Since the \( p_T \)-ordered showers and interleaving with MPI [22] are considered to be a model improvement [23], this is the only configuration available in PYTHIA\(^8\). HERWIG++ on the other hand, uses an angular-ordered parton shower [16]. It has an MPI model similar to the one used by the PYTHIA generators, with tunable parameters for regularising the behaviour at very low momentum transfer, but does not include the interleaving with the parton showers. Inclusive hadron-hadron collisions are simulated by applying the MPI model to events with no hard scattering. It is therefore possible to generate an event with zero \( 2 \rightarrow 2 \) partonic scatters, in which only beam remnants are produced, with nothing in between them. While HERWIG++ has no explicit model for diffractive processes, these zero-scatter events will look similar to double-diffractive dissociation. EPOS MC describes soft-parton dynamics by Gribov-Regge theory [20], with the exchange of virtual quasi-states with
multi-pomeron exchanges. The hard-parton processes at high energies are described similarly but generalised to include hard-pomeron scattering, which are equivalent to a leading-order pQCD Dokshitzer-Gribov-Lipatov -Altarelli-Parisi (DGLAP) evolution approach as used by both PYTHIA and HERWIG++.

The introduction of MPI models (extending the perturbative description in a soft regime) in Monte Carlo tools, improved the description of past observations related to the final-state multiplicity of particle production. Only a phenomenological approach can be used, due to the fact that its dynamics are not well understood since they are governed by soft and semi-hard interactions. In almost all cases, MPI are implemented by assuming a Poissonian distribution of elementary partonic interactions, with an average number depending on the impact parameter of the hadronic collision [1, 2]. The cross section is characterised by a stronger dependence on the incoming parton flux (as compared with the single pair case). Therefore MPI becomes increasingly important at higher centre-of-mass energies [1, 2] of the colliding hadrons, where partons with smaller and smaller fractional momenta play an active role. The dependence of the UE activity on the energy scale is well described by Monte Carlo [6, 7, 9–12], illustrating the universality of MPIs in different topologies and production processes. This universality is also indicated by the similarity between the UE activity in DY [13] and hadronic events [6, 7, 11, 12], although these events have different types of radiation.

In this analysis, several event generators tunes are used for comparison with the data. These are the PYTHIA8 CUETP8M1 and Monash [24] tunes, HERWIG++ CUETHS1 tune, as well as the EPOS version 1.99. On top of the model differences between HERWIG++, PYTHIA8, and EPOS, the main differences in the tunes lie in the different data sets used for tuning, giving different values for the centre-of-mass energy dependence.

### 3 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume there are several complementary detectors: a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. The silicon tracker measures charged particles within the pseudorapidity range $|\eta| < 2.5$. For non-isolated particles of $1 < p_T < 10$ GeV and $|\eta| < 1.4$, the track resolutions are typically 1.5% in $p_T$, and 25–90 (45–150) $\mu$m in the transverse (longitudinal) impact parameter [25]. The CMS subdetectors acting as LHC beam monitor, Beam Pick-up Timing for the eXperiments (BPTX) devices, are used to trigger the detector readout. The two BPTX devices, located inside the beam pipe at distances of 175 m from the IP, are designed to provide precise information on the bunch structure and timing of the incoming beams, with a time resolution better than 0.2 ns. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [26].

### 4 Event selection and tracking validation

This section describes the event and track selection criteria. They are similar to the selections in the previous analyses [6, 7, 9, 10].

The data used in this analyses forms part of the data sample collected by the CMS detector in July 2015 with low PU. The collected data sample has an average PU of 1.3 and correspond to an
integrated luminosity of 281 nb$^{-1}$. Events are selected from the ZeroBias datasets, which are required to pass the ZeroBias triggers (triggering on events with a coincidence of both BPTX devices, signalling a beam crossing in the CMS detector). Events with exactly 1 primary vertex within 10 cm of the beamspot in the $z$-direction, 2 cm in the $xy$-plane ($\rho \leq 2$ cm), and at least 5 degrees of freedom ($dof > 4$) are chosen for the analysis. The vertex selection results in selecting 36% of events from the triggered data sample, 67% from the CUETP8M1 sample, and 60% from the CUETHS1 sample.

Tracks and charged-particles used in this analysis are within the phase space region of $|\eta| < 2$ and $p_T \geq 0.5$ GeV. This is to ensure that only tracks with good reconstruction efficiency are used. All MC generated charged-particles with $c\tau < 1$, where $c$ is the speed of light and $\tau$ is the lifetime of the particle, are included for the analysis. They are then passed through the GEANT package to simulate the detector response, after which tracks can be reconstructed. Fake tracks from mis-reconstruction are removed by requiring the tracks to pass the highPurity [27] selection, which makes a selection based on the minimum fraction of consecutive hits depending on the pseudorapidity region. Secondary decays are removed by requiring the impact parameter significance $d_0/c\sigma_d$ and the significance in the $z$-direction $d_z/c\sigma_{dz}$ to each be less than 3. To keep only tracks with good momentum measurement, only tracks with a relative uncertainty of $c\sigma_{p_T}/p_T < 0.05$ are selected.

The track/charged-particle jets in this analysis are constructed by clustering reconstructed tracks/charged-particles using the Seedless Infrared-Safe Cone (SISCone) jet algorithm [28]. Although the Anti-$k_T$ algorithm [29] is currently the more preferred jet clustering algorithm in CMS, this analysis is done with SISCone jets to be compatible with previous UE results at 0.9, 2.76, and 7 TeV [6, 7, 9]. Tracks/charged-particles within the interval of $|\eta| < 2.5$ and with $p_T > 0.5$ GeV are used to construct the track jets with a distance parameter of 0.5. These tracks/charged particles have to pass the selection criteria outlined above. Only jets within $|\eta| < 2$ and $p_T^{jet} \geq 1$ are considered for the analysis.

5 Phase-space and observables definition

For each event, the direction of the leading track/charged-particle (or track jet/charged-particle jet) is used to define two transverse regions in $\eta$-$\phi$ space. They are defined by $60^\circ < |\Delta\phi| < 120^\circ$, where $\Delta\phi = \phi - \phi_1$ is the azimuthal angle difference, $\phi$ is the direction of the leading track/charged-particle (track/charged-particle jet), and $\phi_1$ is the direction of any other track/charged-particle. Depending on the sign of $\Delta\phi$, the phase space is divided into two separate regions, each with a solid angle of $4\pi/3$ in $\eta$-$\phi$ space. The transverse regions are sensitive to the UE activity, being relatively uncontaminated by radiation from the partonic interaction with the largest momentum transfer.

In each of these regions we construct two variables that are useful for describing underlying event activity:

- The average charged-particle multiplicity density, which is the number of tracks in the region divided by its area in $\eta$-$\phi$ space.
- The average scalar transverse momentum ($\Sigma p_T$) / average energy density, which is the scalar sum of track transverse momenta in the region divided by its area in $\eta$-$\phi$ space.

The two transverse regions are used to construct a total of four observables for each of these variables:
• **TransMAX** is the density of the transverse region with a higher activity.
• **TransMIN** is the density of the transverse region with a lower activity.
• **TransDIF** is the difference of transMAX and transMIN densities.
• **TransAVE** is the average density of both transverse regions.

The average energy and particle densities are measured as function of the $p_T$ of the leading track/charged-particle (track/charged-particle jet).

The measurement of the UE activity in terms of the transMAX, transMIN, and transDIF densities improve the differential power in the quantification of the activity coming from MPI which allows for the better tuning of the model parameters describing the MPI. On an event-by-event basis, the fluctuation of the event topology results in an asymmetry in the activity of the two transverse regions. The dominant di-jet cross-section should not have a significant contribution to this asymmetry due to momentum conservation. Three- and higher jet contributions with increasingly smaller cross-sections on the other hand, can still contribute to the difference. The most significant contributions come from the three-jet topology where radiation from the third jet can be captured in one of the transverse sides. This results in the transMAX region being sensitive to both activity coming from the MPI as well as radiation from the jet. The transMIN region will largely contain radiation coming from MPI. The difference between these two activities should give results compatible with the activity coming from the third jet.

6 Correction and Systematics Uncertainties

Measured distributions are corrected for detector effects and the selection efficiency to reflect the UE activity of charged-particles as defined in section 4, in order to compare the data to various theory predictions. In the present analysis, detector level distributions are corrected to final state particles using the Bayesian unfolding method [30] which properly accounts for bin migration effects. At hadron level, the leading charged-particle (charged-particle jet) is required to be within $|\eta| < 2$ and have $p_T \geq 0.5$ (or $p_{jet}^T \geq 1$) GeV. The average particle and energy densities are constructed using all final state charged-particles with $|\eta| < 2$ and $p_T \geq 0.5$ GeV. The 4-dimensional response matrix required for unfolding is constructed using detector and hadron level observables with simulated events using **PYTHIA** 8 CUETP8M1 tune. Corrected distributions are compared with various theory predictions.

Systematic uncertainties are quantified by a systematic variation in the different conditions that could give rise to a bias in the final results. In general, the systematic uncertainties vary between the densities in the different regions (transAVE/transDIF/transMAX/transMIN), and as a function of $p_T$ ($p_{jet}^T$). The uncertainties considered are described below:

• **Model dependency**: The effect due to the model dependency of the correction method is calculated by taking the ratio of bin-by-bin correction factors. Due to the HERWIG++ sample not describing data well up to 5 GeV as a result of a lack of a proper simulation of diffractive events, and EPOS not describing the data above 5 GeV likely because of the hard-pomeron scattering approach, a hybrid systematic uncertainty is used where the systematic is taken between EPOS and CUETHS1 against CUETP8M1 above that. The systematic uncertainty varies between 1–4% (1–4%) above 5 GeV as a function of the leading charged-particle (charged-particle jet) $p_T$, where the range reflects the dependence of the densities in various regions. In the lowest few $p_T$ ($p_{jet}^T$) bins, devia-
Results

The corrected distributions of the average particle and energy densities as a function of the $p_T$ of the leading jet/track are compared with the predictions by the several simulations. Figure 1 shows the average energy density, in the transverse region, as a function of the $p_T$ of the leading charged particle. The agreement between the measurements and the simulations is quantified by the ratio plots shown in the bottom panels. The measurements are better described by the Monash tune of PYTHIA8. The PYTHIA8 CUETP8M1 describe the measurements within 10–20%. The predictions by the CUETHS1 tune of HERWIG fails in the low $p_T$ region. EPOS describes the low $p_T$ rising region well but fails to describe the plateau region between 20–40%. Figure 2 reports similar distributions but for the average particle density as a function of the leading charged particle $p_T$. The amount of agreement of the measured average particle density with
the simulations is the same as that for the average energy density. The best description is again by the Monash tune of PYTHIA8.

Figure 3 and Figure 4 report the average energy and particle densities as a function of $p_{T}^{jet}$. The measured distributions for average energy and particle densities are best described by the Monash tune of the PYTHIA8. Predictions by other simulations describe the measurements with a maximum deviation of up to 30%.

In all plots, the densities increase sharply with $p_{T}$ ($p_{T}^{jet}$) up to 5 (12–15) GeV and slow rise afterward. It is obvious from the evolution that the transMIN densities are flatter compared to the transMAX and transDIF densities, whereas the transMAX and transDIF densities show a similar trend in the plateau region. Simulations describe the qualitative behavior of the measurements, i.e. the sharp rise and the flattening of the UE activity, and a larger rise in transMAX and transDIF in the plateau region. The level of agreement between simulations and the measurements fall within 10–20% in the plateau region but differ in the low $p_{T}$ region. The sharp rise with $p_{T}$ is interpreted in the MC models as due to an increase in the MPI contribution which reaches a plateau at high $p_{T}$. A slow increase in large $p_{T}$ region is mainly due to the increase in the initial- and final-state radiation contribution. As MPI activity is expected to be uniform in the whole phase-space, the transMIN densities capture mainly the activity coming from MPI whereas transDIF densities gives the evolution of the radiation with $p_{T}$ of the reference object.

Comparisons between various MC simulated samples and data across centre-of-mass energies of 0.9, 2.76, 7, and 13 TeV are made for transAVE, and at 2.76 and 13 TeV for transDIF, transMAX, and transMIN densities as a function of $p_{T}^{jet}$ as shown in figures 5 and 6. There is a strong rise in the UE activity as a function of the centre-of-mass energy as predicted by the MC tunes. This is attributed to an increase in the of number partons with smaller fractional momenta $x \sim 2p_{T}/\sqrt{s}$. The transMIN densities show a stronger $\sqrt{s}$ dependence than the transDIF density, indicating that the activity coming from MPI grows more with $\sqrt{s}$ than that from ISR and FSR.

8 Summary

A measurement of the underlying event activity is made for proton-proton collisions at a centre-of-mass energy of 13 TeV. Average particle and energy densities are calculated in the transverse region and measured as a function of the $p_{T}$ of the leading charged-particle as well as of the leading charged-particles jet. All charged particles with $p_{T} \geq 0.5$ GeV and $|\eta| < 2.0$ are used to calculate the UE activity. Measured distributions are corrected for the detector effects and the selection efficiencies. The corrected distributions are compared with several theory predictions and Monash tune of PYTHIA8 is found to be in best agreement with the measurements. A comparison of the UE densities at 0.9, 2.76, 7, and 13 TeV shows a strong growth in the activities with the centre-of-mass energy, with transMIN densities having a stronger $\sqrt{s}$ dependence than transDIF.
Figure 1: Comparisons of corrected (a) transAVE, (b) transDIF, (c) transMAX, and (d) transMIN average energy densities with the various simulations as a function of $p_T$. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels show the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.
Figure 2: Comparisons of corrected (a) transAVE, (b) transDIF, (c) transMAX, and (d) transMIN average particle densities with the various simulations as a function of $p_T$. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.
Figure 3: Comparisons of corrected (a) transAVE, (b) transDIF, (c) transMAX, and (d) transMIN average energy densities with the various simulations as a function of $p_T^{\text{jet}}$. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.
Figure 4: Comparisons of corrected (a) transAVE, (b) transDIF, (c) transMAX, and (d) transMIN average particle densities with the various simulations as a function of $p_T^{jet}$. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.
Figure 5: Comparisons of corrected (a) transAVE, (b) transDIF, (c) transMAX, and (d) transMIN particle densities with various simulations at $\sqrt{s} = 0.9, 2.76, 7, \text{and} 13 \text{ TeV}$ for (a), and $\sqrt{s} = 2.76 \text{and} 13 \text{ TeV}$ for (b,c,d) as a function of $p_T^{\text{jet}}$. 

- **transAVE**
- **transDIF**
- **transMAX**
- **transMIN**
Figure 6: Comparisons of corrected (a) transAVE, (b) transDIF, (c) transMAX, and (d) transMIN energy densities with various simulations at $\sqrt{s} = 0.9, 2.76, 7,$ and $13$ TeV for (a), and $\sqrt{s} = 2.76$ and $13$ TeV for (b,c,d) as a function of $p_T^{\text{jet}}$. 
References


References


