Study of multiple parton interactions in Z+jets events using an event mixing technique

Matheus Da Costa Geraldes
CERN Summer Student 2025
Supervised by:
Patrick L.S. Connor
Ankita Mehta
Kyle Cormier

September 12, 2025

Abstract

The study of multiple parton interactions (MPI) is essential for precision measurements at the Large Hadron Collider [1]. This work presents an analysis of double parton scattering (DPS), a key MPI process, using the experimentally clean Z+jets channel [1]. A event-mixing technique is employed to construct a pure sample of synthetic DPS events from collider data, allowing for the isolation of its kinematic properties. Kinematic observables sensitive to the DPS dynamics are defined and their behavior is studied, comparing the data with the mixed sample [2]. We demonstrate the efficacy of the technique in isolating the characteristic signatures of DPS. This report establishes a framework for a future quantitative measurement of the DPS contribution, aimed at providing data to constrain MPI models [1, 3].

Contents

1 Introduction

High-energy proton-proton collisions at the Large Hadron Collider (LHC) are inherently complex processes. The common model to describe these events involves a single hard interaction between constituent partons (quarks and gluons), a scenario known as single parton scattering (SPS) [1]. However, the high density of partons within protons makes it probable that multiple parton pairs interact in the same collision, a phenomenon known as multiple parton interactions (MPI) [2]. A major challenge in particle physics is the difficulty of unambiguously separating the primary hard scattering from the additional activity generated by MPI [3]. The study of MPI is therefore crucial for two main reasons: it provides fundamental information about the correlations and spatial distribution of partons within the proton [2], and it constitutes a critical background for many standard model measurements and searches for new physics [1, 3].

The study of MPI, and its specific process, double parton scattering (DPS), has been an active area of research at various hadron colliders [2]. Previous experimental measurements have successfully identified DPS signatures in a variety of final states, such as W+jets, 4-jet production, and WW boson pair production [1].For example, in W+2-jets events at 7 TeV, the DPS fraction was measured to be around 5.5% [1]. Building on these results, the Z+jets channel has been identified as an excellent laboratory for precision studies [3]. Furthermore, strategies to enhance the sensitivity to MPI, such as applying an upper limit on the Z boson's transverse momentum, have been validated in prior analyses and are a cornerstone of modern approaches [3].

This work focuses on isolating the DPS contribution within the Z+jets final state. The analysis of jets in these events reveals a characteristic energy scale for MPI, with the average transverse momentum of the jets reaching a plateau around 10 GeV in events where radiation from the primary interaction is suppressed [3]. The methodology central to this analysis is an event-mixing technique, which constructs a mixed sample of events designed to emulate the properties of pure DPS. The mixed sample serves as a pure empirical model, or "template," for the DPS signal [2]. This allows for a direct comparison with the data to disentangle the two components and enables a robust test of theoretical models used in event generators. The objective is to develop and validate an analysis workflow to measure kinematic observables sensitive to DPS, using methodologies and observables implemented in the studies presented in this Ref. [2, 3]. By isolating the DPS contribution, this demonstrates the viability a future measurement that can constrain the phenomenological parameters used in MPI models [1, 3], improving the precision of future analyses at the LHC.

This report is structured as follows. Section 2 introduces the theoretical concepts of the Parton Model, differentiating between Single and Multiple Parton Interactions. Section 3 justifies the choice of the Z+jets final state as the analysis channel, highlighting its experimental advantages. Section 4 details the experimental strategy, explaining the kinematic signatures used to distinguish DPS from SPS and introducing the event-mixing technique. Section 5 describes the data samples and the full workflow for event preparation. Section 6 focuses on the final measurement, defining the kinematic regions and detailing the custom software developed. Section 7 presents the preliminary results, validating the methodology by comparing the obtained distributions with reference studies. Finally, section 8 concludes the report, summarizing the key findings and discussing future steps towards a quantitative measurement.

2 The parton model and multiple interactions

To understand the underlying physics of this study, it is essential to first revisit the model that describes the internal structure of protons and how they interact in high-energy collisions.

2.1 Single Parton Scattering

The modern view describes protons not as elementary particles, but as bound states of fundamental constituents called partons, which are quarks and gluons. The model to describe a proton-proton collision is single parton scattering. In this scenario, a single parton from one of the incident protons interacts with a single parton from the other proton [1]. This single "hard" interaction (with high momentum transfer) is responsible for producing the final state of interest, such as a Z boson and jets [1]. Figure 1 illustrates this idealized process.

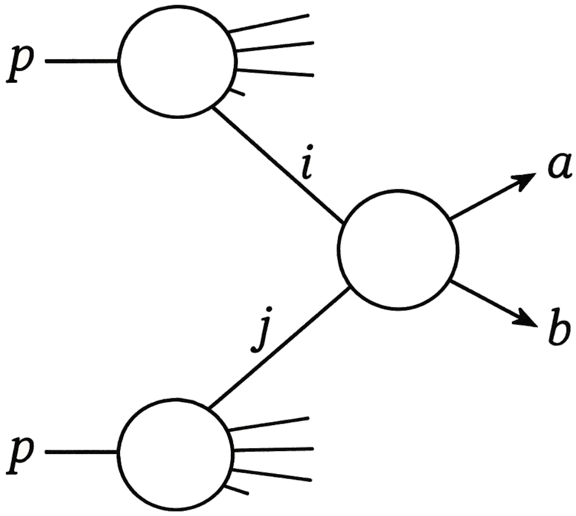


Figure 1: Theoretical view of a proton-proton (p-p) collision dominated by a Single Parton Scattering (SPS). A parton i from the first proton interacts with a parton j from the second to produce a final state $a + b$ [4].

2.2 MPI and the underlying event

A realistic description. of a high-energy proton-proton collision, as illustrated in Figure 2, reveals a far more complex environment. The event is initiated by a single high-momentum transfer interaction, known as the hard scatter, which produces the primary particles of interest. However, this is just one component. The high density of partons inside the colliding protons makes it likely that one or more additional, lower-energy parton pairs will interact simultaneously. This phenomenon is known as multiple parton interactions (MPI) [1].

The collective activity from MPI, along with radiation from the initial and final state partons (ISR and FSR), the remnants of the protons, and the subsequent cascade of parton splittings (parton showering) and formation of hadrons (hadronization), constitutes the underlying event—a sea of additional particles that accompanies the primary interaction [5].

Within the broad phenomenon of MPI, this analysis focuses on a particular case of interest: double parton scattering (DPS). DPS occurs when two distinct, semi-hard parton interactions happen in the same proton-proton collision [1, 4]. For example, in a single p-p (proton - proton) collision, one parton pair can produce a Z boson, while a second pair independently produces a pair of jets [1]. The general diagram for DPS is shown in Figure 3.

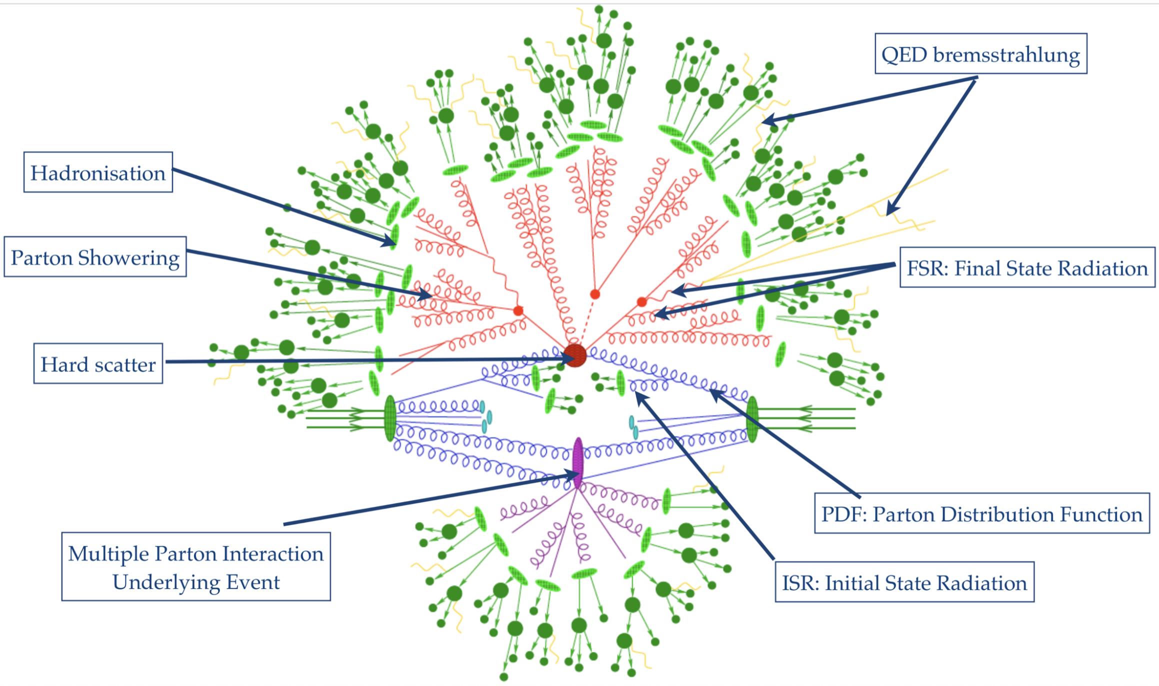


Figure 2: Illustration of the various components in a realistic proton-proton collision. The event includes a primary hard scatter, additional Multiple Parton Interactions (MPI), initial and final state radiation (ISR/FSR), and the subsequent parton showering and hadronization that form the final state particles. Source: Smart [5].

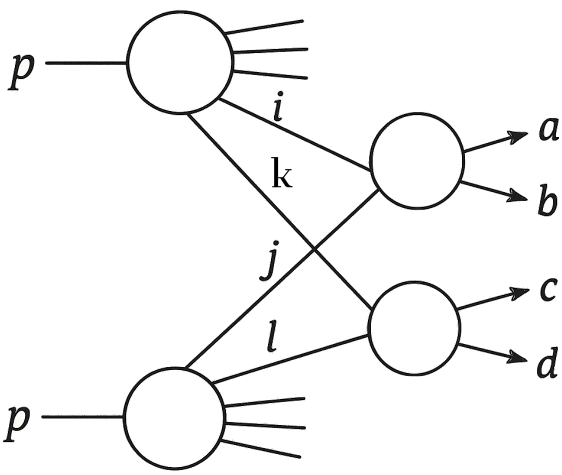


Figure 3: Representation of a double parton scattering (DPS), where two independent parton interactions occur in a single p-p collision, producing two distinct final states ($a + b$ and $c + d$) [4].

2.3 The effective sigma (σ_{eff})

The probability of a DPS event occurring is governed by the spatial distribution of partons within the proton [4]. This dependence is parameterized by a phenomenological term called the effective sigma, σ_{eff} [4]. The cross section for a DPS process producing final states A and B can be expressed in a simplified form as [4]:

$$\sigma_{\text{DPS}}^{A+B} = \frac{\sigma_A \cdot \sigma_B}{\sigma_{\text{eff}}}$$

where σ_A and σ_B are the cross sections for the individual production of states A and B via SPS.

The parameter σ_{eff} has dimensions of area and effectively measures the size of the transverse region within the proton where the partons are confined [4]. A small value of σ_{eff} implies that the partons are more concentrated, thus increasing the probability of a DPS interaction [4]. Precise measurement of σ_{eff} is crucial to the physics of MPI, as this parameter is assumed to be universal (independent of the specific process) and it provides valuable information about parton-parton correlations [4].

3 The analysis channel: Z boson production with jets

The choice of an appropriate analysis channel is a critical step for any measurement in particle physics. The process of producing a Z boson in association with jets (Z+jets) constitutes an excellent laboratory for the study of Multiple Parton Interactions (MPI) due to a combination of an exceptionally clean experimental signature, robust theoretical understanding, and kinematics that are particularly sensitive to MPI effects [1].

The main advantage of this channel lies in its clear golden signature. This analysis focuses on the decay of the Z boson into a pair of oppositely charged muons ($Z \to \mu^+\mu^-$). Muons traverse most of the detector, allowing them to be reconstructed with very high efficiency and excellent momentum precision [1]. Furthermore, requiring the invariant mass of the muon pair to be within a narrow window around the known Z boson mass (≈ 91.2 GeV) provides a powerful tool for background rejection. This selection results in a final sample with very high purity, where the dominant background from top quark pair production contributes only 2 to 5% [1].

This experimental cleanliness is precisely what allows for a sensitive probe of the underlying QCD dynamics, particularly MPI. In Single Parton Scattering (SPS) events, jets are primarily produced from initial-state radiation (ISR), which imparts a transverse momentum (p_T) to the Z boson. This creates a strong kinematic correlation: a high p_T^Z is typically associated with high-p_T jets originating from the same interaction [2]. This correlation is the key to isolating MPI. As proposed in recent studies, applying a strict upper cut on the Z boson's transverse momentum (e.g., $p_T^Z < 10$ GeV) drastically suppresses the contribution of jets from the primary interaction. In the sample of events that survives this cut, the observed jets cannot be explained by radiation from the Z system and, therefore, originate almost exclusively from secondary, independent parton interactions—that is, from MPI [3]. This strategy effectively transforms the MPI activity from a background into a high-purity signal. The two distinct production mechanisms (SPS and DPS) for the Z+2 jets topology are illustrated in Figure 4.

Finally, this well-motivated strategy is experimentally feasible due to the large production rate of Z bosons at the LHC. The high cross section guarantees a vast sample of events, even after applying strict selection cuts.

4 Experimental signatures and analysis strategy

Although theoretically distinct, experimentally separating SPS events from DPS events is a considerable challenge. In a hadron collider environment, detectors record a final state that is a mixture of both processes. The key to isolating the DPS contribution lies in exploiting its fundamental characteristic: the occurrence of two largely independent and uncorrelated parton interactions within the same proton-proton collision [1, 4]. In contrast, the products of an SPS event are, by nature, kinematically correlated, as they originate from a single high-energy interaction (the *hard scatter*) [4]. The experimental goal is, therefore, to define kinematic observables

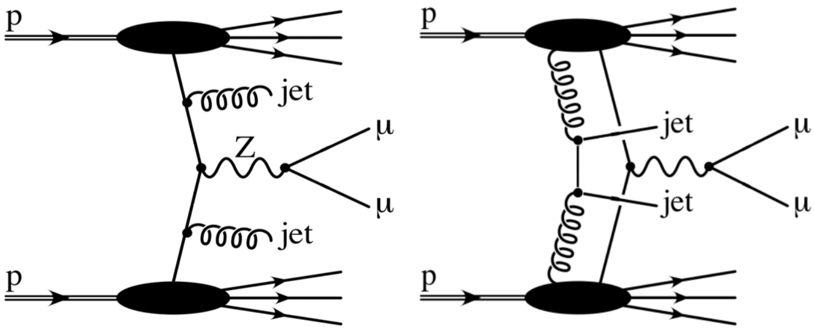


Figure 4: Representative Feynman diagrams for the production of a Z boson and two jets. On the left, the Single Parton Scattering (SPS) process, where all particles originate from a single interaction. On the right, the Double Parton Scattering (DPS) process, where the Z boson and the dijet system are produced by two independent parton interactions in the same collision [1].

that are sensitive to this lack of correlation, such that the distributions for SPS and DPS events exhibit markedly different shapes [1].

This inherent mixing of processes is the central experimental challenge. To precisely extract the properties of pure DPS, a method is required to isolate its signature or construct a reference sample free from SPS correlations. It is precisely this challenge that motivates the event-mixing technique adopted in this work. The strategy is to artificially construct a sample of pseudo-events that, by construction, have no correlation between the Z boson and the jet systems. This sample serves as a pure empirical model for DPS, allowing for a direct comparison with the data to disentangle the two contributions.

4.1 Kinematic signatures of DPS

To probe the differences between SPS and DPS, specific kinematic signatures related to angular correlations and momentum balance are investigated.

One of the most direct ways to test the correlation between particle systems is through their angles in the transverse plane. The azimuthal difference $(\Delta\phi)$** between the Z boson and the jet system is a key example. For SPS events, momentum conservation dictates that the Z boson and the recoiling jet system tend to be emitted in opposite directions, resulting in a $\Delta\phi$ distribution with a sharp peak around π radians. For DPS events, the lack of correlation results in an expected flat distribution [1, 4]. Figure 5 illustrates this behavior, where the data clearly shows a mixture of a peak and a flat component, while simulations without MPI fail to describe the flat region.

Another powerful probe is the relative transverse momentum balance $(\Delta_{rel}p_T)$**. This variable quantifies how well the p_T vectors of a particle system cancel each other out. A particularly interesting observable for identifying DPS is the balance between the two leading jets, $\Delta_{rel}p_T(j_1, j_2)$. Since the two jets in a DPS event originate from their own $2 \to 2$ interaction, they are expected to be produced back-to-back, resulting in a distribution of $\Delta_{rel}p_T(j_1, j_2)$ with a pronounced peak near zero, which is a classic signature of DPS contributions [1].

4.2 Definition of physical observables

The key observables measured in this analysis are defined below, categorized by the physics goals inspired by the reference literature [2, 3].

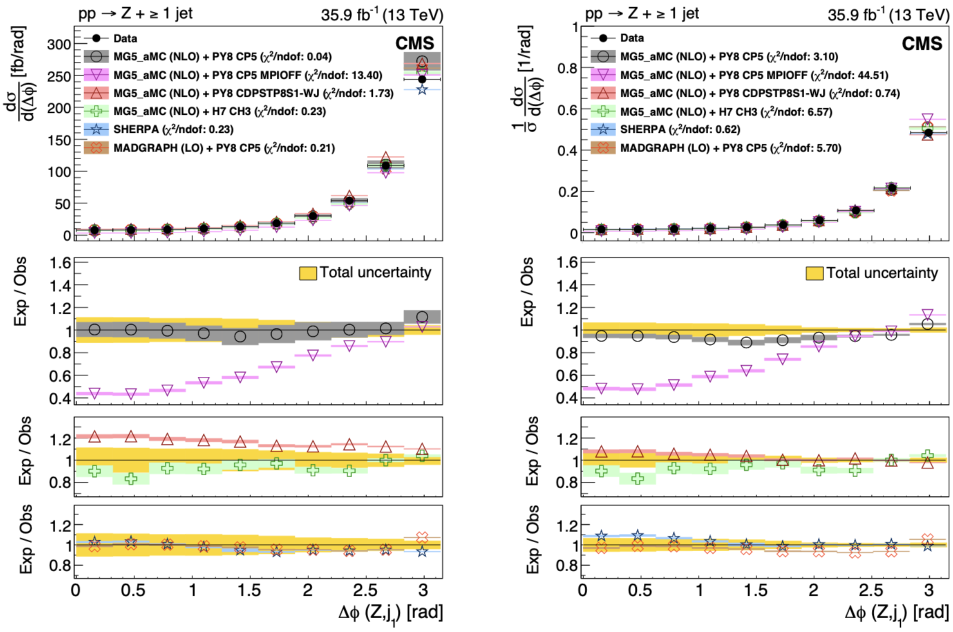


Figure 5: Distributions of the azimuthal angle difference, $\Delta\phi(Z, j_1)$, for Z+≥1 jet events. The pink line (triangles) represents a simulation with MPI turned off, which fails to describe the data in the low $\Delta\phi$ region [1].

4.2.1 Classic DPS observables

These observables are designed to statistically separate SPS and DPS contributions in inclusive samples:

- **Azimuthal separation ($\Delta S(Z, \mathbf{dijet})$):** The angle between the p_T vectors of the Z system and the dijet system, expected to be flat for DPS. It is defined as:

$$\Delta S = \arccos\left(\frac{\vec{p}_T(Z)\cdot\vec{p}_T(\text{dijet})}{|\vec{p}_T(Z)||\vec{p}_T(\text{dijet})|}\right)$$

- **Relative transverse momentum balance ($\Delta p_T^{\mathbf{rel}}(j_1, j_2)$):** The momentum balance between the two leading jets.

$$\Delta p_T^{\text{rel}}(j_1, j_2) = \frac{|\vec{p}_T(j_1) + \vec{p}_T(j_2)|}{|\vec{p}_T(j_1)| + |\vec{p}_T(j_2)|}$$

4.2.2 High-Purity MPI observables

These observables are measured in a kinematic region where MPI is dominant, achieved by a strict cut on p_{tZ}:

- **Average leading jet p_T vs. Z p_T:** A profile histogram that measures $\langle p_T^{j_1}\rangle$ as a function of p_{tZ}. A plateau at low p_{tZ} is a classic signature of MPI.
- **Cumulative jet spectrum ($\langle n(p_{tj,min})\rangle$):** The average number of jets with p_T above a given threshold. This becomes a direct probe of MPI activity when primary radiation is suppressed.
- **Azimuthal difference ($\Delta\phi(j_1, j_2)$):** The angle in the transverse plane between the two leading jets, providing direct information on the kinematics of the secondary parton scattering.

5 Data samples and event preparation

The present analysis uses data from proton-proton collisions collected by the CMS (Compact Muon Solenoid) detector at the Large Hadron Collider (LHC) at CERN. The data correspond to a special operating period during **Run 2**, specifically **Run2017H**, which was conducted under low collision pileup (Low PU) conditions.

The DasAnalysisSystem (DAS) software framework was used for this analysis. It is a modular C++ system that allows the sequential execution of atomic analysis steps through shell commands. This approach facilitates development, debugging, and analysis reproducibility. In this work, DAS was used to apply all physical corrections, prepare data samples, and implement the event mixing technique.[6]

5.1 Datasets

- **Muon Signal Sample:** To select events containing the signal of interest (a Z boson decaying into two muons), the following dataset was used:

  - `/SingleMuon/Run2017H-09Aug2019_UL2017_LowPU-v1/MINIAOD`

  This dataset is composed of events that were selected and recorded by the CMS trigger system due to the presence of at least **one muon** with high transverse momentum. This sample is, therefore, the primary source for the reconstruction of Z bosons and serves as the basis for the selection of signal events in the event mixing technique.

- **Jet Sample for Mixing:** For the *event mixing* technique, an independent sample of jet events is required. To cover a wide range of transverse momentum, two datasets collected with jet triggers with different energy thresholds were used:

  - `/HighEGJet/Run2017H-09Aug2019_UL2017_LowPU-v1/MINIAOD`
  - `/LowEGJet/Run2017H-09Aug2019_UL2017_LowPU-v1/MINIAOD`

  The `HighEGJet` dataset contains events selected by jet triggers with a **high energy** threshold, while the `LowEGJet` was selected with a **low energy** threshold. The combination of these two datasets provides a statistically robust sample of jets, which are subsequently mixed with the events from the `SingleMuon` sample to simulate the DPS signal. The fact that this data was collected during the same period (`Run2017H`, `LowPU`) ensures the consistency of experimental conditions between the signal sample and the jet sample.

5.2 Event preparation and mixing workflow

To implement the event-mixing technique, a multi-stage workflow is applied to the datasets to produce two pure, mutually exclusive samples: a Dimuon sample to isolate the $Z \to \mu^+\mu^-$ process, and a Dijet sample to isolate the secondary hard scatter.

The workflow begins with the application of physical corrections and quality criteria:

- **Muon momentum calibration:** Rochester corrections are applied to calibrate muon momentum for a precise Z boson mass resolution.

- **Jet energy calibration:** Jet Energy Scale (JES) corrections are used to adjust the measured jet energy.

- **Event normalization:** Events are normalized to account for integrated luminosity and trigger efficiencies.

- **Muon quality and trigger:** A strict selection is applied to ensure that only well-reconstructed, tightly identified, and isolated muons are used, and that events were triggered by a single high-p_T muon.

Following these corrections, the crucial selection of the two independent samples is performed:

- **Dimuon sample selection:** Events are selected from the `SingleMuon` data, requiring a pair of oppositely charged muons with $p_T^\mu > 25$ GeV, $|\eta_\mu| < 2.4$, and an invariant mass $60 < m_{\mu\mu} < 120$ GeV.

- **Dijet sample selection:** Events are selected from the `EGJet` data, requiring at least two jets and a strict veto on any muon with $p_T^\mu > 20$ GeV to ensure statistical independence.

Finally, the central tool of the analysis, `applyEventMixing`, combines one event from the Dimuon sample (which is required to have no jets with $p_T^{jet} > 20$ GeV), to obtain purer muon events, with one from the Dijet sample to create a new dataset of mixed events where the dimuon and dijet systems are, by construction, independent and uncorrelated, thus simulating the properties of pure DPS.

6 Measurement and software implementation

6.1 Definition of measurement regions

The physical observables defined previously are measured in specific kinematic regimes, or signal regions, designed to enhance sensitivity to different aspects of MPI.

- **Classic DPS selections:** Based on Kumar et al. [2], an inclusive region is defined with at least two jets ($p_T > 20$ GeV, $|\eta| < 5.0$). A second, MPI-enhanced region imposes an additional cut on the Z boson's transverse momentum, $p_{tZ} < 10$ GeV, to suppress the SPS background.

- **High-Purity MPI selections:** Based on Andersen et al. [3], more stringent cuts are applied, requiring two muons with $p_T > 27$ GeV and $|\eta| < 2.5$. To create a sample where MPI activity is almost completely dominant, a very restrictive cut of $p_{tZ} < 2.0$ GeV is used.

6.2 Software development and framework contribution

An integral part of this project was the contribution to the analysis framework. The final stage of the analysis required the creation of a new tool, the `getDPSobservable` executable, which was developed from scratch. This program reads the mixed-event n-tuples, reconstructs the physical objects, applies the selection criteria for the measurement regions, and calculates the final kinematic observables, saving them as ROOT histograms.

Furthermore, existing tools were extended and adapted for this specific analysis:

1. **Kinematic correction for muons:** The functionality of `applyPVshift` was expanded to apply vertex corrections to muons, increasing the kinematic precision of the dimuon system.

2. **Adaptation of event mixing:** The `applyEventMixing` executable was adapted to handle the specific topology of the analysis, explicitly mixing the dimuon systems from one sample with the dijet systems from another.

This development work demonstrates the ability to extend the framework for new physics analyses, implementing the necessary logic to translate theoretical concepts into quantitative results.

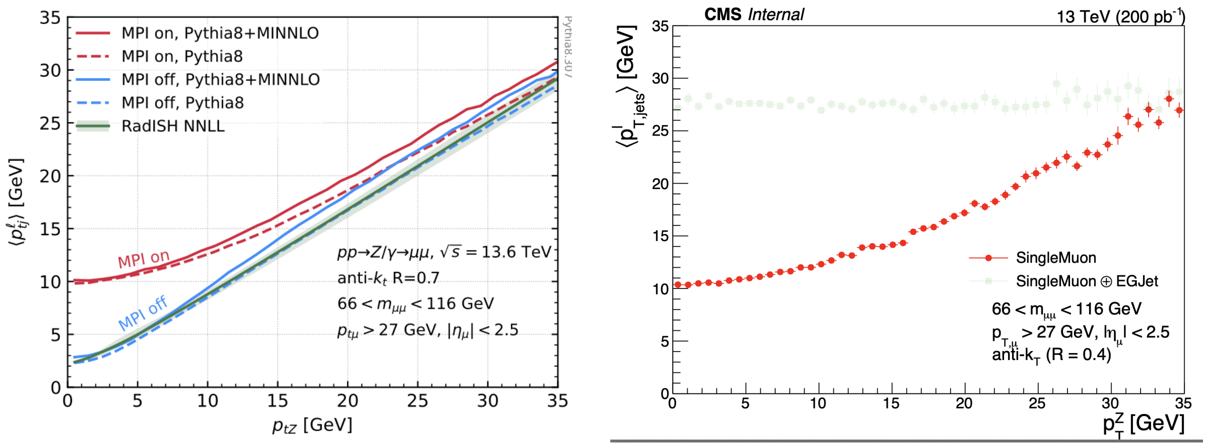
7 Preliminary results

The preliminary results of this analysis are compared with key findings from the reference literature to validate the methodology and the implementation. The following subsections detail the expected behavior for each observable based on the reference studies and the results obtained in this work.

7.1 Correlation between jet and Z boson momenta

Based on the reference study by Andersen et al. [3], it was expected that the average leading jet p_T would show two distinct behaviors as a function of the Z boson's p_T. For high p_T^Z, where SPS dominates, a linear correlation is expected as the jet system recoils against the Z boson. Conversely, for low p_T^Z, where MPI dominates, this correlation should vanish, causing $\langle p_T^{j_1} \rangle$ to reach a plateau at a characteristic energy scale of MPI, predicted to be around 10 GeV.

The analysis performed in this work successfully reproduces this behavior, as shown in Figure 6. The plot obtained (right) shows a clear qualitative agreement with the reference plot (left). In the `SingleMuon` data (red points), for $p_T^Z > 15$ GeV, $\langle p_T^{j_1} \rangle$ increases linearly, as expected for SPS. For low $p_T^Z < 10$ GeV, the distribution reaches the expected plateau, confirming the presence of a strong MPI component decoupled from the Z production. The synthetic DPS sample (light green) correctly models the addition of an energetic dijet system, showing a consistently higher $\langle p_T^{j_1} \rangle$.



(a) Reference from Andersen et al. [3]

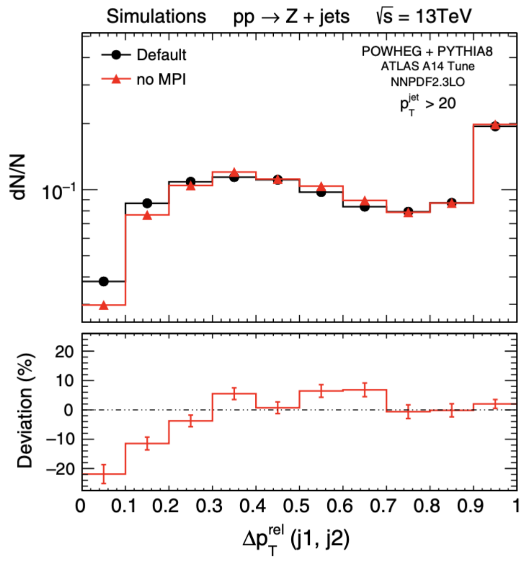
(b) Result obtained in this work

Figure 6: Comparison of the average leading jet momentum versus the Z boson's momentum. Both plots show the characteristic plateau at low p_T^Z when MPI is present.

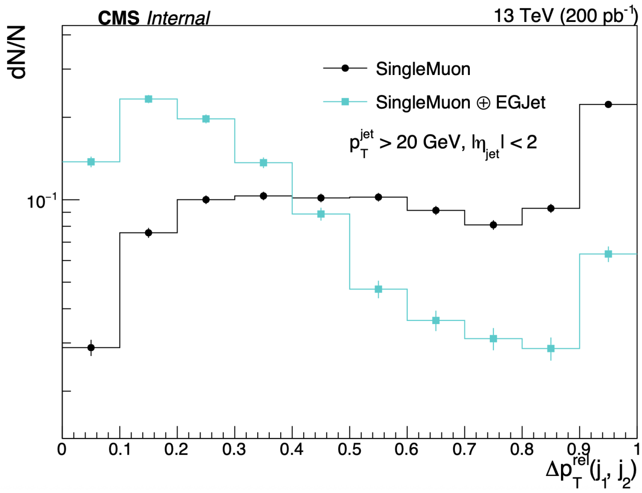
7.2 Dijet momentum balance in inclusive events

For the relative transverse momentum balance between the two leading jets, $\Delta p_T^{rel}(j_1, j_2)$, it is expected that DPS events, where jets are produced back-to-back, would create a peak at low values. In contrast, the inclusive data, dominated by SPS, should show a much flatter distribution, as the two leading jets are not necessarily correlated with each other [1].

This is precisely what was observed. Figure 7 compares the result from this analysis (right) with the measurement from the CMS Collaboration (left). The synthetic DPS sample (cyan) in our result exhibits the expected sharp peak at low values ($0.1 < \Delta p_T^{rel} < 0.3$), confirming it as a good template for the DPS signal. The `SingleMuon` data (black) displays the characteristic flat shape of an SPS-dominated sample, in clear qualitative agreement with the reference measurement from CMS.



(a) Reference from CMS Collaboration [1]



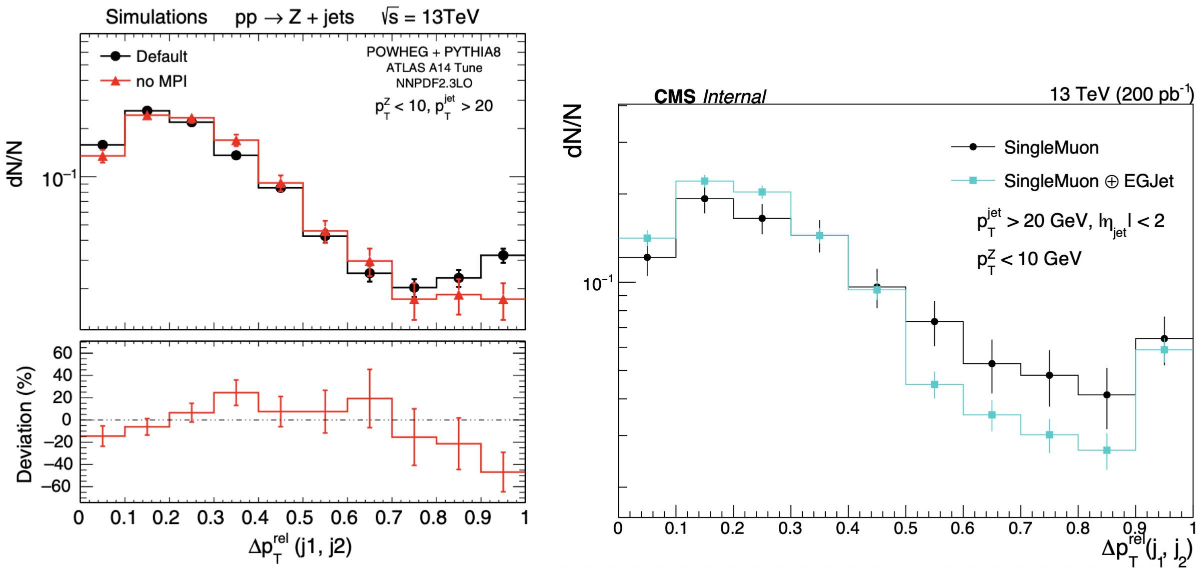
(b) Result obtained in this work

Figure 7: Comparison of the inclusive $\Delta p_T^{rel}(j_1, j_2)$ distribution. The reference data (left, black) is flat, similar to the data obtained in this work (right, black). The synthetic DPS sample (right, cyan) shows the expected peak at low values.

7.3 Dijet momentum balance in the MPI-Enhanced region

As proposed by Kumar et al. [2], applying a cut of $p_T^Z < 10$ GeV is expected to drastically suppress the SPS background. This should enhance the visibility of the DPS component in the data, causing the Δp_T^{rel} distribution to shift from a flat shape towards a more peaked structure at low values.

The results obtained in this work, presented in Figure 8, powerfully confirm this expectation. The reference simulation (left) shows a large difference between the model with MPI (black) and without MPI (red), indicating high sensitivity in this region. The result from this analysis (right) demonstrates the effect on data: the `SingleMuon` data (black), with the SPS contribution now suppressed, shows a clear enhancement in the low Δp_T^{rel} region, moving towards the shape of the pure DPS template (cyan). This validates that suppressing jet radiation from the primary interaction is an effective strategy for isolating MPI signatures.



(a) Reference from Kumar et al. [2]

(b) Result obtained in this work

Figure 8: Comparison of the $\Delta p_T^{rel}(j_1, j_2)$ distribution after applying a cut of $p_T^Z < 10$ GeV. Both the reference simulation (left) and the result on data (right) show a significant MPI effect in this phase space.

These preliminary findings confirm that the event-mixing technique is a powerful tool for isolating DPS signatures and provide a solid qualitative foundation for a future quantitative measurement of MPI in the Z+jets channel.

8 Conclusion and future outlook

This work successfully established and validated a comprehensive analysis workflow for the study of Double Parton Scattering (DPS) in the Z+jets channel. By implementing an event-mixing technique within the `DasAnalysisSystem` framework, a synthetic DPS sample was constructed from low-pileup CMS data. The preliminary physics results robustly confirm the entire analysis strategy. Key MPI signatures predicted in the reference literature, such as the decorrelation between the jet and Z boson momenta at low p_T^Z and the back-to-back nature of the dijet system in DPS, were successfully reproduced [2, 3]. Crucially, the application of a low-p_T^Z cut was shown to effectively suppress the SPS background and enhance the visibility of the DPS signature in the data, proving the efficacy of the methodology. The development of new analysis tools and contributions to the existing software infrastructure were vital to achieving these results.

This work lays a solid foundation for a future quantitative measurement. The immediate next steps involve a detailed statistical analysis, including a template-fitting procedure to extract the DPS fraction from the data and measure the effective sigma, σ_{eff}. A thorough evaluation of systematic uncertainties associated with the experimental corrections and the event-mixing technique itself will be essential for the precision of the final result. Ultimately, this analysis paves the way for a novel measurement at the LHC, with the potential to significantly advance our understanding of the internal structure of the proton and improve the modeling of high-energy hadron collisions.

References

[1] The CMS Collaboration. "Study of Z boson plus jets events using variables sensitive to double-parton scattering in pp collisions at 13 TeV". In: *JHEP* 10 (2021), p. 176. DOI: 10.1007/JHEP10(2021)176. arXiv: 2105.14511 [hep-ex].

[2] R. Kumar et al. "New observables for multiple-parton interactions measurements using $Z+$ jets processes at the LHC". In: *Phys. Rev. D* 93 (5 Mar. 2016), p. 054019. DOI: 10.1103/PhysRevD.93.054019. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.054019.

[3] Jeppe R. Andersen et al. "Exploring High-Purity Multiparton Scattering at Hadron Colliders". In: *Phys. Rev. Lett.* 132 (4 Jan. 2024), p. 041901. DOI: 10.1103/PhysRevLett.132.041901. URL: https://link.aps.org/doi/10.1103/PhysRevLett.132.041901.

[4] Chris Jackson. *Double Parton Scattering at the LHC*. Presentation in collaboration with E. Berger & G. Shaughnessy. Based on Phys. Rev D81, 014014. 2010.

[5] Ben Smart. *Searching for associated Higgs boson production at the LHC*. Presentation at IOP HEPP & APP Conference. Accessed via https://indico.cern.ch/event/214998/contributions/437825/attachments/341915/477057/BenSmart_IOP_2013.pdf. 2013.

[6] P. L. S. Connor et al. *DasAnalysisSystem*. Software repository. Accessed: 2025-09-10. 2024. URL: https://gitlab.cern.ch/cms-analysis/general/DasAnalysisSystem.

A Analysis Workflow Commands

This appendix provides the sequence of shell commands used within the `DasAnalysisSystem` framework to execute the analysis workflow. The workflow is divided into two main parts: the preparation of the Dimuon sample and the preparation of the Dijet sample, followed by the final event mixing.

A.1 Dimuon Sample Preparation (`SingleMuon` dataset)

Step 1: Merge N-tuples

The individual ROOT files from the dataset are merged into larger, more manageable files.

```
# Create output directory and run the merge in parallel
mkdir -p D1_muon
parallel -j$(nproc) mergeNtuples /path/to/SingleMuon/ D1_muon/ -f
```

Step 2: Apply Rochester Corrections

Momentum corrections are applied to all muons in the sample.

```
mkdir -p D2_muon_rochester
parallel -j$(nproc) applyRochesterCorr D1_muon/ D2_muon_rochester/ \
  $DARWIN_TABLES/Rochester/RoccoR2017UL.txt 0 -f
```

Step 3: Apply Normalization Factor

Events are weighted by the inverse of the integrated luminosity (L = 219.04 /pb).

```
mkdir -p D3_muon_Norm
parallel -j$(nproc) applyNormFactor "D2_muon_rochester/*.root" \
  D3_muon_Norm/ 0.0045653056 -f
```

Steps 4-6: Apply Muon Selections and Skim

A sequence of selections is applied for muon reconstruction quality, identification, and isolation. A skim is then applied to select events with two muons forming a Z candidate and a veto on jets.

```
# Note: Steps for reconstruction, ID, and isolation selections
# ... (commands for applyMuonSelection) ...

# The final skim for the Z-pure sample
mkdir -p D6_DPSskim_ISO
parallel -j$(nproc) applyDPSskim D5_MuonSelection_ID D6_DPSskim_ISO \
  CleanJets -f
```

Steps 7-9: Trigger and Kinematic Selections

The trigger requirement is applied, followed by kinematic cuts on the muons and the reconstructed Z boson.

```
mkdir -p D7_muon_trigger
parallel -j$(nproc) applyTFormula D6_DPSskim_ISO D7_muon_trigger \
  "'@muonTrigger.Bit[2] == 1'" -f

mkdir -p D8_DPSskim_Muon
parallel -j$(nproc) applyDPSskim D7_muon_trigger D8_DPSskim_Muon Muoncut -f

mkdir -p D9_DPSskim_Zcut
parallel -j$(nproc) applyDPSskim D8_DPSskim_Muon D9_DPSskim_Zcut Zcut -f
```

Step 10: Apply Primary Vertex Correction

The pseudorapidity of muons and jets is recalculated relative to z=0 to remove dependence on the primary vertex position. This creates the final, analysis-ready Dimuon sample.

```
mkdir -p D11_muon_pvshift
parallel -j$(nproc) applyPVshift D9_DPSskim_Zcut D11_muon_pvshift/ -f
```

A.2 Dijet Sample Preparation (EGJet datasets)

Steps 1-2: Merge and Apply Jet Corrections

The HighEGJet and LowEGJet N-tuples are merged separately and then undergo Jet Energy Scale (JES) corrections.

```
# Merge HighEGJet and LowEGJet samples (example for High)
mkdir -p D1_Ntuples_high
parallel -j$(nproc) mergeNtuples /path/to/HighEGJet/ D1_Ntuples_high -f

# Apply JES corrections (example for High)
mkdir -p D2_Jcorrections_high
parallel -j16 applyJEScorrections D1_Ntuples_high/ D2_Jcorrections_high/ \
  /path/to/Summer20UL17_RunH_V5_DATA_AK4PFchs.json Summer20UL17_RunH -f
```

Steps 3-4: Trigger Efficiency and Normalization

Trigger turn-on curves are generated, and efficiencies are calculated. Events are then normalized based on trigger efficiency and integrated luminosity.

```
# (Commands for getTriggerCurves, hadd, getTriggerTurnons)
# ...

# Apply data normalization (example for High)
mkdir -p D3_Dnormalised_high
parallel -j16 applyDataNormalisation D2_Jcorrections_high/ D3_Dnormalised_high/ \
  $TRIGGER_INFO $LUMI_UNC_INFO $POINTS_FILE $EFF_FILE pt lumi -f
```

Step 5: Apply Primary Vertex Correction

The PV shift is applied to the jet samples.

```
mkdir -p D4_PVshift_high
parallel -j16 applyPVshift D3_Dnormalised_high/ D4_PVshift_high/ -f
```

Steps 6-8: Merge and Final Selection

The Low and High pT jet samples are merged. A final selection is applied to veto events containing high-pT muons, creating the final, analysis-ready Dijet sample.

```
# Merge low and high pT jet samples
hadd Merged_Jets/EGJets_merged.root low/D5_.../*.root high/D4_.../*.root

# Apply muon veto
mkdir -p D6_Muon_Veto
parallel -j16 applyTFormula EGJets_merged.root D6_Muon_Veto/ \
  "'@recMuons.size() == 0 || @recMuons.at(0).CorrPt() < 20'" -f
```

A.3 Final Mixing and Observable Extraction

Step 11: Event Mixing

The final Dimuon sample and the final Dijet sample are combined to create the synthetic DPS sample.

```
mkdir -p mixed_SingMuon_EGJet_events
parallel -j$(nproc) applyEventMixing D6_Muon_Veto/ D11_muon_pvshift/ \
  mixed_SingMuon_EGJet_events/ none true 0 -f
```

Step 12: Extraction of Physical Observables

The new executable `getDPSobservable` is run on the mixed sample to calculate the final physics observables and create the plots.

```
# Create directory for final plots
mkdir -p DPSobs

# Run observable extraction
getDPSobservable mixed_SingMuon_EGJet_events/ DPSobs/DPS_observables.root
    DPSplot
```