DISSERTATION

submitted to the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Falk Bartels

born in Hildesheim

Oral examination: 21.06.2024

Search for Low-Mass Dijet Resonances Using Trigger-Level Jets at the ATLAS Detector

Referees: apl. Prof. Dr. Monica Dunford
Prof. Dr. Stephanie Hansmann-Menzemer

Suche nach Dijet-Resonanzen niedriger Masse mithilfe von Jets auf Trigger-Level am ATLAS Detektor

Konventionelle Suchen nach Dijet-Resonanzen am Large Hadron Collider sind für Resonanzmassen unterhalb von etwa 1 TeV durch die Bandbreite der Datenerfassungssysteme statistisch limitiert. Die hier vorgestellte Suche erschließt diesen Massenbereich mit noch nie erreichter statistischer Präzision durch die Verwendung von Jets, die vom ATLAS High-Level Trigger aus Kalorimeterinformationen rekonstruiert und aufgezeichnet wurden. Es werden Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 13\,\text{TeV}$ analysiert, die einer integrierten Luminosität von bis zu $132\,\text{fb}^{-1}$ entsprechen. Eine speziell für diese Analyse konzipierte Kalibrierung für die Jets auf Trigger-Level wird angewandt und zwei komplementäre, datengetriebene Untergrundabschätzungen mit sehr hoher statistischer Genauigkeit werden vorgestellt. Keine signifikante Abweichung vom erwarteten Untergrund wird beobachtet. Dementsprechend werden Ausschlussgrenzen für Wirkungsquerschnitte von Dijet-Resonanzen abgeleitet. Diese stellen die strengsten Grenzen für schmale, Gauß-förmige Dijet-Resonanzen im Massenbereich von 375 – 1200 GeV im Allgemeinen und im Bereich von 375 – 1500 GeV für einen Axialvektor-Mediator, der an Quarks und Dunkle Materie koppelt, im Besonderen dar.

Search for Low-Mass Dijet Resonances Using Trigger-Level Jets at the ATLAS Detector

Conventional dijet resonance searches at the Large Hadron Collider are statistically limited for sub-TeV resonance masses by the bandwidth of the data acquisition systems. The presented search explores this mass range at unprecedented statistical precision by utilizing a partial event readout consisting of jets reconstructed by the ATLAS High-Level Trigger using only calorimeter information. Proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ corresponding to an integrated luminosity of up to $132\,\text{fb}^{-1}$ are analyzed. A dedicated calibration for the trigger-level jets is applied and two complementary, data-driven background estimates with very high statistical precision are presented. No significant excess over the background expectation is observed and exclusion limits on dijet resonance cross sections are derived. These represent the most stringent constraints on narrow, Gaussian-shaped dijet resonances in the mass range of 375–1200 GeV in general and in the range of 375–1500 GeV for an axial-vector Dark Matter mediator coupling to quarks in particular.

Contents

1 Introduction

The field of particle physics strives to establish a complete theory of the fundamental interactions in nature. The best current understanding is encoded in the Standard Model of particle physics (SM), which describes nature as a set of quantum fields that interact with each other. Predictions of the SM have been confirmed at unparalleled precision in many experiments at various energy scales.

Despite its considerable success, the SM is known to be incomplete, leaving several fundamental questions unanswered. One of these is the nature of Dark Matter. Several sources of complementary astrophysical evidence indicate the existence of a form of matter that does not, or at most very weakly, interact electromagnetically. This Dark Matter (DM) is expected to make up 84 % of the matter content of the Universe. It is likely to be a fundamental particle in nature, however, the SM does not provide a suitable candidate to explain the observations. The existence of DM and other phenomena unexplained by the SM, such as gravity or the observed matter-antimatter asymmetry in the Universe, suggest the existence of new fundamental particles that have yet to be discovered.

Searches at the Large Hadron Collider (LHC) play a crucial role in the pursuit of new discoveries. The LHC produces proton-proton (pp) collisions at an unprecedented energy and luminosity, enabling the exploration of physics beyond the Standard Model (BSM) at both the energy and the precision frontiers.

If DM can be produced at the LHC, it is likely due to the existence of a new mediator particle that couples both to DM and to quarks or gluons, either directly or via loop interactions. Such a mediator, if accessible at LHC energies, could be discovered via its decay back into quarks or gluons, producing two collimated jets of hadrons in the final state – a so-called dijet event. This process would result in a resonance in the dijet invariant mass spectrum at the mediator mass.

Dedicated searches for dijet resonances are sensitive to a wide range of potential DM and other BSM models, as the observable signature is relatively independent of many model parameters other than the mediator mass and its decay width. However, the strong interaction of the SM produces a very large background to the dijet signature that rises approximately exponentially towards low energies. In fact, jet production constitutes the majority of the hard interaction cross section at the LHC.

The high rate at which dijet events occur poses a limitation to conventional dijet resonance searches. At low energies, the dijet event rate exceeds the bandwidth of the data acquisition systems of the LHC experiments. Therefore, the readout rate of events with only low-energy jets is limited, which results in reduced statistical precision of the recorded dijet invariant mass spectrum in the sub-TeV range.

In the ATLAS experiment at the LHC, this limitation is overcome with partial event readouts. An approximate jet reconstruction is performed during data-taking to trigger a full readout of the event if high-energy jets are identified. Additionally, partial information

of the reconstructed trigger-level jets can be recorded regardless of their energy, since this information constitutes a sufficiently small amount of data as to not exceed the readout bandwidth at the full dijet event rate.

The dijet resonance search using trigger-level jets that is presented in this thesis faces two key challenges. Firstly, a custom calibration for the trigger-level jets must be derived for this analysis, since only full-readout jets are centrally calibrated by the ATLAS Collaboration. It is crucial for this calibration to be a smooth function of the jet energy to not induce any resonance-like structures in the observed dijet invariant mass spectrum. Secondly, the very high statistical precision achieved with the dijet signature must be matched with a background estimate of at least equal precision to avoid a systematic limitation of the search sensitivity. This precision cannot be achieved with typical Monte Carlo background estimates. Instead, two complementary fit-based methods are employed.

Overcoming these challenges allows the constraint of a leptophobic Z' DM mediator model in specific and, more generally, any BSM models that predict a new, narrow dijet resonance. To the author's knowledge, the presented dijet resonance search is the first analysis to constrain the intermediate resonance mass range of 375–1800 GeV using the (almost) full LHC Run-2 data set of pp collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, which corresponds to an integrated luminosity of up to $132\,\text{fb}^{-1}$. This improves upon previous publications of trigger-level dijet resonant searches by the ATLAS [1] and CMS Collaborations [2], which only considered approximately one quarter of the Run-2 data set. At high resonance masses above 1.5 TeV, conventional dijet (and di-b-jet) resonance searches already provide the full Run-2 sensitivity [3, 4]. At very low resonance masses in the range of 100–450 GeV, alternative signatures with dijets in addition to a photon from initial state radiation [5–8] provide the currently strongest sensitivity. This is achieved by utilizing the low energy thresholds required to trigger photons, although this comes at the cost of reduced signal cross sections due to the enforced initial state radiation.

The presentation of the trigger-level dijet resonance search in this thesis is structured as follows: Chapter 2 provides a theoretical introduction to the SM and BSM phenomena, with an emphasis on DM and the leptophobic Z' model that is constrained in this thesis. The application of the theory to simulate background and signal processes in pp collisions is presented in Chapter 3. Chapter 4 summarizes the experimental setup of the LHC and the ATLAS detector, linking to the special considerations for an analysis at trigger level and the jet reconstruction in Chapter 5. Chapter 6 details the steps of the custom jet calibration for this analysis as well as the constraints on its smoothness. In Chapter 7, the event selection criteria for the signal region of this analysis are outlined. Chapter 8 discusses the signal predictions and systematic uncertainties. The two employed methods for the background estimate and their validation are presented in Chapter 9. In Chapter 10, the compatibility of the observation with the background hypothesis is investigated, and constraints on BSM physics are derived. Finally, the presented work is summarized in Chapter 11.

1.1 Author's Contributions

The work presented in this thesis was performed as part of an analysis team within the ATLAS Collaboration. The operation of the LHC and the ATLAS detector requires many dedicated people to produce and record the proton collisions that are analyzed in this thesis. Additionally, CERN and the ATLAS Collaboration provide central reconstruction and analysis software tools that were used in the presented search. Without these large, dedicated teams, experimental high-energy physics with a scope as presented in this thesis would likely not be possible.

As part of the ATLAS Collaboration, the author contributed to the Level-1 Calorimeter Trigger, specifically to the planned Phase-II Upgrade of the jet Feature EXtractor (jFEX) that performs the jet reconstruction [9]. Significant improvements were made to the jFEX simulation, and the prospects of more refined jet calibration methods were analyzed for potential resolution improvements.

The author was also the (initially deputy) lead administrator of the Heidelberg ATLAS computing cluster during the course of this thesis. Computations performed on this cluster are a key contribution to the search presented in this thesis and to numerous other analyses performed in the Heidelberg ATLAS group. During this period, a series of enhancements were implemented, including significant performance upgrades and the incorporation of new functionalities to meet current demands.

The author provided the daily supervision for, among others, two students working on a Bachelor's and an integrated Master's thesis [10, 11]. These students assisted in the optimization of the event selection criteria applied in the presented search. Their contributions are cited in the relevant discussions.

For the presented dijet resonance search, the author developed the analysis framework for performing the fits and validations for the background estimate as well as the statistical analysis of the results. This framework has been adopted by other ATLAS analyses since, with active support provided by the author. At the time of writing this thesis, two of these analyses are published [5, 12], while multiple others are still in progress.

The author provided major contributions to almost all parts of the presented search. The analysis of trigger-level jets requires a custom jet calibration specifically derived for this analysis. The author validated the calibration smoothness (presented in Section 6.3) and contributed to the development of strategies to ensure the smoothness of individual calibration steps. The kinematic requirements of the event selection presented in Section 7.2 were primarily derived by the author and partially by the supervised students mentioned above. The work presented in Chapters 8 to 10 was performed solely by the author. This includes the derivation of signal predictions, the implementation of two independent background estimates (including the theoretical cross section calculations presented in Section 3.3 for one of them), the application and validation of these background estimates with the analyzed data set, and the statistical interpretation of the results to constrain BSM physics.

Finally, the author also serves as the analysis contact of the presented search within the ATLAS Collaboration. This involves the general coordination of the group efforts as well as regular exchange with the ATLAS physics groups. The analysis is expected to be ready for the internal ATLAS review process soon.

2 Theoretical Background

Measurements and searches at the LHC rely on solid theoretical foundations to compare observations to expectations. This foundation is provided by the SM, which contains the best current knowledge of fundamental interactions.

Section 2.1 provides a summary of the SM with emphasis on the strong interaction, which is crucial for the phenomenology of dijet signatures studied in this thesis. The description is mainly based on Refs. [13–16].

However, the SM is known to be incomplete as it fails to explain multiple, mainly astrophysical observations. One of them is DM, which is expected to make up 84 % of the matter content of the Universe [17]. Despite solid evidence for its existence via gravitational interactions, its exact nature could not yet be identified. Section 2.2 presents a selection of the evidence for DM and current experimental efforts to determine its nature.

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics is a relativistic quantum field theory that combines the *electroweak theory* developed by Glashow, Salam and Weinberg [18–20] in the 1960s with *quantum chromodynamics* (QCD) describing the strong interaction [21, 22] from the 1970s.

For each particle type, the SM contains one underlying quantum field of which individual particles can be regarded as excitations. Figure 2.1 summarizes these fundamental particles, their electrical charge, spin and the best current knowledge of their mass [17]. The particle content is categorized into two sectors based on the spin: the fermions with half-integer spin and the bosons with integer spin.

The bosonic sector contains the interaction fields. The spin-1 gauge bosons mediate the three fundamental forces: The electromagnetic force is mediated by the massless photon (γ), the weak force by the massive W^+, W^-, and Z bosons, and the strong force by the eight massless gluons (g). All massive particles also interact with the spin-0 Higgs boson (H). The gravitational force could not yet be included in the SM, which remains a process of active theoretical research. However, it is sufficiently weak compared to the other fundamental forces to be negligible in most particle physics experiments.

The fermionic sector contains the matter fields. It consists of twelve fundamental particles of spin $1/2$ and their respective antiparticles with the same mass and spin but otherwise opposite quantum numbers. They are grouped into three generations which behave identically besides an increase in masses and different flavour states. The fermions are divided into quarks, which are affected by the strong force, and leptons, which are not.

Each of the three lepton generations consists of one charged lepton – the electron (e),

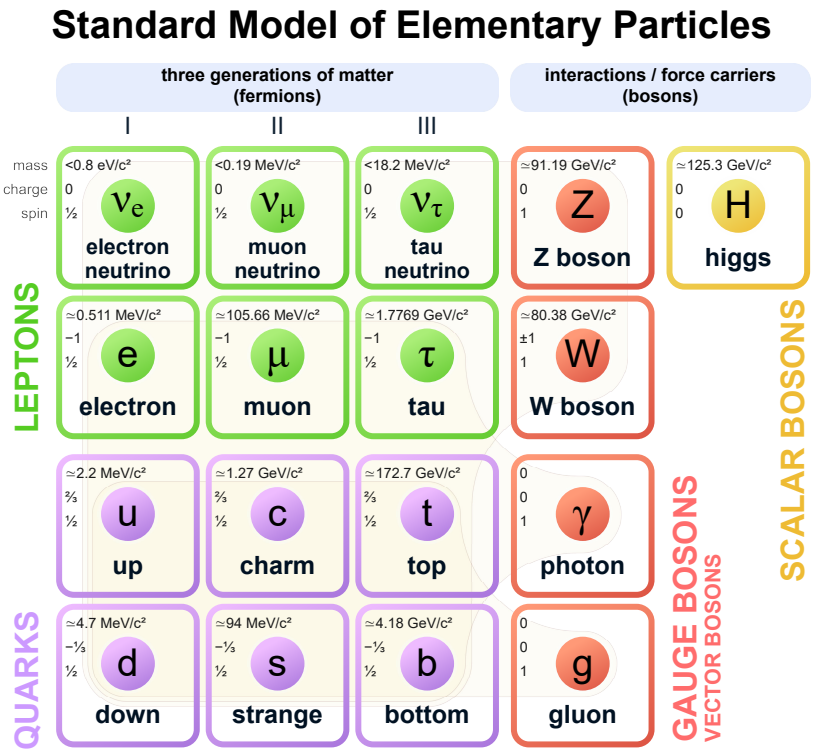


Figure 2.1: Particle content of the SM. The electrical charge, spin and mass are indicated for each particle. Each of the fermions also has an antiparticle with opposite charge that is not shown. The lightly shaded regions indicate which gauge bosons each fermion interacts with. Adapted from Ref. [23] to reflect the best current knowledge of the particle masses [17].

muon (μ), or tau (τ) – and an associated electrically neutral neutrino ($\nu_{\{e,\mu,\tau\}}$).[1] All leptons are subject to the weak interaction, while only the charged leptons interact electromagnetically.

The six quarks are the up (u), down (d), charm (c), strange (s), top (t) and bottom (b) quark. Each of the three generations contains an up-type quark with an electrical charge of $+2/3$ and a down-type quark with an electrical charge of $-1/3$. They also carry a colour charge and are therefore subject to the strong force in addition to the electromagnetic and weak force. Quarks are never observed freely but only in colour-neutral bound states due to the confinement of QCD. These bound states are called *hadrons*, which are divided into *mesons* consisting of a quark-antiquark pair and *baryons* consisting of three quarks or antiquarks.[2]

The description of the SM is based on a Lagrangian density $\mathcal{L}$, from which equations of motions are derived using the principle of minimal action.

[1] While neutrinos were assumed massless in the original formulation of the SM, the observation of neutrino oscillations [24–26] indicates non-vanishing masses. Accordingly, the SM was extended from 19 to 26 free parameters to include their masses, mixing angles and a CP-violating phase. However, the exact nature of the neutrino mass term is still unknown [27].

[2] Evidence for bound states of four or five quarks has been observed [28–31]. Tetraquarks are considered mesons, as they consist of an equal number of valence quarks and antiquarks, while pentaquarks are considered baryons, as they have an odd number of valence quarks.

This Lagrangian density is required to obey global Poincaré symmetry to make the SM a relativistic theory. Additionally, invariance under local $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge transformations is required. This symmetry generates the three fundamental forces described by the SM with gauge bosons mediating the interactions. Following Noether's theorem, energy, momentum, and angular momentum are conserved due to the Poincaré invariance, while the gauge symmetries also necessitate conserved charges.

2.1.1 The Electroweak Sector

The electromagnetic and the weak interaction are unified in the Glashow-Salam-Weinberg theory that imposes the local $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge symmetry.

The $\mathrm{U}(1)_Y$ group generates one gauge field, labelled B_μ, which couples to the weak hypercharge Y. The $\mathrm{SU}(2)_L$ group generates three fields W^a_μ, $a \in \{1, 2, 3\}$, which couple to the weak isospin T. Setting $T = 0$ for right-handed fermions encodes the behaviour observed in nature that only left-handed fermions interact with these fields [32, 33]. Hence, the subscript L for the $\mathrm{SU}(2)_L$ group. The non-abelian nature of $\mathrm{SU}(2)_L$ leads to self-interactions of the weak gauge bosons.

The Higgs doublet ϕ, consisting of two complex scalar fields, is introduced to allow for massive gauge bosons without breaking $\mathrm{SU}(2)_L$ gauge invariance. It is assigned a potential of fourth order that has infinite global minima for $\phi \neq 0$. The specific choice of one of those minima in nature leads to spontaneous symmetry breaking in which the Higgs field acquires a vacuum expectation value. In this process, the massive Higgs boson and three massless Goldstone bosons are generated.

The physically observable $W^\pm$ bosons can then be expressed as linear combinations of the charged $W^{(1)}_\mu$ and $W^{(2)}_\mu$ fields, while the neutral B_μ and $W^{(3)}_\mu$ fields mix to form the photon and the Z boson. The $W^\pm$ and Z bosons absorb the massless Goldstone bosons as longitudinal polarizations and thereby acquire their masses without violating the gauge invariance.

Fermion masses are introduced in a $\mathrm{SU}(2)_L$ gauge-invariant way by adding a Yukawa coupling of the fermion to the Higgs field that is proportional to each fermion's observed mass.

2.1.2 The Strong Interaction

The strong interaction is described by the theory of QCD, which obeys the $\mathrm{SU}(3)_C$ gauge symmetry. This symmetry gives rise to eight gluon fields G^a_μ, $a \in \{1, ..., 8\}$, which couple to the associated conserved *colour* charge carried by quarks and gluons.

Quarks are hence represented as colour triplets, whose three orthogonal states are labelled red, green, and blue. Antiquarks are assigned the corresponding anti-colours. All other fermions are treated as colour singlets as they do not participate in the strong interaction.

To express the Lagrangian density of QCD, the gauge covariant derivative of a quark triplet q is defined as

$$\mathrm{D}_\mu\, q = \left(\partial_\mu + i\frac{g_S}{2}\lambda^a G^a_\mu\right) q\,, \tag{2.1}$$

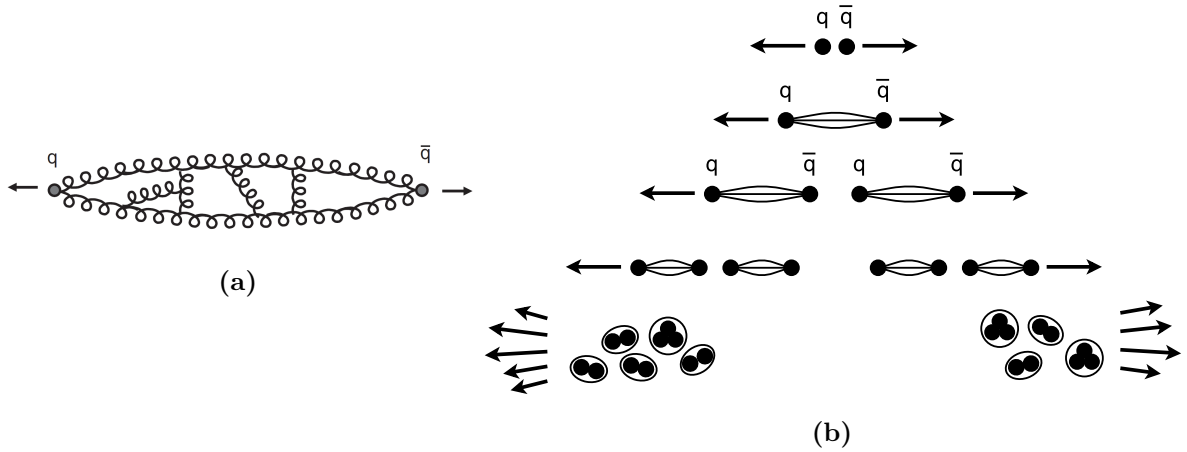


Figure 2.2: (**a**) Self-interaction of gluons mediating the strong interaction between two quarks, leading to a flux tube. (**b**) The process of hadronization: The energy stored in the flux tube increases as quarks separate from each other. At some point, the creation of a $q\overline{q}$-pair from the vacuum becomes energetically favourable over further increasing the tube's length. A repetition of this process dissipates the initial momentum difference into a shower of newly formed colour-neutral hadrons. Adapted from Ref. [13].

where g_S is the coupling strength[3] and λ^a are the Gell-Mann matrices, to which the $\mathrm{SU}(3)_C$ generators are proportional. The gluon field strength tensor is given by

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_S f^{abc} G^b_\mu G^c_\nu \,, \tag{2.2}$$

where f^{abc} are the structure constants arising due to $\mathrm{SU}(3)_C$ being non-abelian. They are defined in terms of the commutators of the Gell-Mann matrices: $\left[\lambda^a, \lambda^b\right] = 2if^{abc}\lambda^c$.

With these definitions, the QCD Lagrangian is:

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a + \sum_{\substack{q \in \\ \{u,d,s,c,b,t\}}} i\overline{q}\gamma^\mu \,\mathrm{D}_\mu \, q \,. \tag{2.3}$$

Quark mass terms are excluded here, although they would not break the $\mathrm{SU}(3)_C$ symmetry. In the full SM Lagrangian, they enter via the Yukawa coupling to the Higgs field to preserve $\mathrm{SU}(2)_L$ gauge invariance.

In Equation (2.3), $\overline{q}$ and q are colour triplets and, accordingly, D_μ is a 3×3 matrix. The colour states are summed over to form the scalar Lagrangian density, which reduces the appearance of the Gell-Mann matrices to scalar *colour factors* in QCD matrix elements.

The non-abelian nature of $\mathrm{SU}(3)_C$ gives rise to a defining feature of QCD: Gluons carry a colour charge and self-interact in three- and four-gluon vertices. In contrast to the analogous self-interaction of the weak vector bosons, gluons are massless, such that the strong interaction range is not limited and large numbers of gluon self-interactions

[3] In the literature, both g_S and $\alpha_S = g_S^2/4\pi$ are commonly referred to as the strong coupling constant. In this thesis, that term is used for the latter.

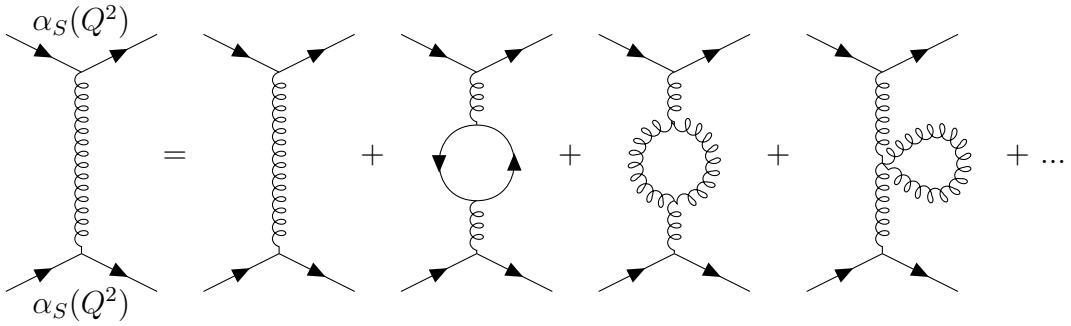


Figure 2.3: Renormalization of QCD. Next-to-leading order corrections to the tree-level gluon propagator are shown, which are absorbed into the running coupling $\alpha_S(Q^2)$. Adapted from Ref. [13].

are not mass-suppressed at low energies. Phenomenologically, this leads to the concept of *colour confinement*.

Gluons mediate an attractive force between quarks. As the gluons are also attracted to one another, they form a *flux tube* [34], which is visualized in Figure 2.2a. Using the analogy of electromagnetic field lines, this leads to the field lines of the strong interaction being compressed into a tube. Consequently, the attractive force between the quarks approaches a (very large) constant for large distances. Thus, the potential energy stored in the gluon field between them rises linearly with their distance: $V(r) \propto r$. At some threshold, the formation of a $q\overline{q}$-pair via vacuum polarization becomes energetically favourable over further increasing the tube's length. The creation of $q\overline{q}$-pairs repeats until the relative momenta between (anti-)quarks are sufficiently low to form colour-neutral hadrons. This hadronization process is illustrated in Figure 2.2b. Since only colour-neutral hadrons do not experience the strong attractive force at large distances, solely they are observed freely in nature. Consequently, quarks or gluons produced in a hard interaction will be observed as collimated sprays of hadrons – so-called *jets*.

Another defining aspect of QCD is the running of the strong coupling α_S. When perturbatively calculating cross sections of QCD processes, an infinite number of higher-order corrections affect the quark and gluon propagators as well as the quark-gluon and gluon-gluon vertices. An example of the next-to-leading order corrections to the gluon propagator is shown in Figure 2.3. However, the contributions of the quark and gluon loops can diverge towards infinitely small or large momenta of the virtual particles.

To restore finite observable cross sections, the process of *renormalization* isolates the ultraviolet divergences and cancels them with an infinite counterterm derived at an arbitrarily chosen renormalization scale μ_R. In this process, α_S acquires a dependence on the momentum transfer Q^2:

$$Q^2 \frac{\mathrm{d}\alpha_S}{\mathrm{d}Q^2} = \beta(\alpha_S) = -\alpha_S^2 \cdot \left(\beta_0 + \beta_1 \alpha_S + \beta_2 \alpha_S^2 + ...\right), \tag{2.4}$$

where $\beta(\alpha_S)$ is the beta-function that can be expressed perturbatively and includes the corrections of higher orders in α_S. While the renormalization does not provide an absolute value for α_S, it allows it to be expressed relative to its value at the renormalization scale.

At leading order, this is a logarithmic dependence:

$$\alpha_S\left(Q^2\right) \approx \frac{\alpha_S\left(\mu_R^2\right)}{1 + \beta_0\, \alpha_S\left(\mu_R^2\right) \ln\left(\frac{Q^2}{\mu_R^2}\right)} \,. \tag{2.5}$$

The constant β_0 depends on the number of colours N_c and contributing quark flavours N_f:

$$\beta_0 = \frac{11N_c - 2N_f}{12\pi} \,. \tag{2.6}$$

In the SM, β_0 is positive, such that $\alpha_S\left(Q^2\right)$ decreases with increasing momentum transfer.

This behaviour gives rise to a phenomenon referred to as *asymptotic freedom*: At the high energy scales probed at colliders ($|Q| \gtrsim 10\,\text{GeV}$), α_S is of $\mathcal{O}(0.1)$. This means that the strong interaction is sufficiently weak to be treated perturbatively.[4] Conversely, at $|Q| \sim 1\,\text{GeV}$ α_S is of $\mathcal{O}(1)$. Therefore, QCD becomes non-perturbative at the low energy scales relevant in the hadronization process. This regime is still experimentally driven, although, more recently, numeric lattice QCD simulations have improved their predictive power [17].

Details on the computation of QCD cross sections are presented in Section 3.1.

2.2 Physics Beyond the Standard Model

Although the SM is arguably the most precisely tested theory of modern physics[5], it must be embedded in a yet-to-be-found, more general theory. A variety of observed phenomena indicate BSM physics – a selection of which is presented in this section.

Gravity. The SM does not provide a description of gravity. While general relativity as a theory of gravity makes highly accurate predictions in accordance with observations, it cannot be quantized in a renormalizable way without extensions [36]. Reconciling quantum field theory and general relativity is an active field of theoretical research – for example in the form of loop quantum gravity [37], string theory [38], or asymptotically safe gravity [39]. However, these theories are still limited in their predictive power. In experimental particle physics, collider searches for e.g. signatures of extra spatial dimensions, quantum black holes, or gravitons are performed to guide this development [40].

Matter-antimatter asymmetry. The Universe appears to be primarily composed of matter rather than antimatter. Our local neighbourhood is known to be dominated by matter. If other regions of the Universe were dominated by antimatter, gamma-ray signatures would be expected to originate at the matter-antimatter boundaries. The absence of such observations places strict constraints on the antimatter content of the Universe [41, 42].

Generating the observed asymmetry from symmetric initial conditions necessitates, among others, a violation of both baryon number conservation and CP symmetry [43].

[4] However, α_S is still so large that higher-order corrections are significant contributions, making precise predictions challenging.

[5] The most precise prediction of the SM – the electron's magnetic moment – agrees with experiments at a relative precision of 10^{-12} [35].

The SM allows for baryon number violation in non-perturbative processes [44] and includes CP violation in the electroweak Lagrangian density. The Cabibbo-Kobayashi-Maskawa matrix [45, 46], which relates the quark flavour and mass eigenstates, contains a CP-violating phase. The corresponding leptonic mixing matrix (Pontecorvo-Maki-Nakagawa-Sakata matrix [47, 48]) is still associated with large experimental uncertainties, but CP conservation is excluded at 95 % confidence level from neutrino oscillations [49].[6] While the mechanism of baryogenesis is still speculative, it is widely believed that the sources of CP violation in the SM are insufficient to result in the observed matter-antimatter asymmetry [44]. Experimentally, searches for signs of baryon or lepton number violation as well as other CP-violating processes are conducted [52].

2.2.1 Dark Matter

The existence of DM is well-supported by a vast amount of independent astrophysical evidence. It is based on the comparison of the gravitational potential as, for example, deduced from stellar velocities or gravitational lensing [53, 54] to the density of non-dark matter observable via electromagnetic radiation.

One of the most prominent pieces of evidence are the rotational velocities of stars and gas in disc galaxies. In these galaxies, the visible mass is concentrated in the central bulge of the galaxy. Outside of the bulge rotational symmetry can be assumed, such that the velocity v of a star at radius r should be approximately given by equating its centripetal acceleration and the gravitational force of the galaxy's mass $M(r)$ contained within this radius:

$$v^2 \approx \frac{G}{r} M(r)\,. \tag{2.7}$$

Since the majority of the galaxy's mass in the form of stars and gas is concentrated at its centre, $M(r)$ only rises slowly at large radii. Accordingly, $v(r)$ should decrease with an approximate $1/\sqrt{r}$ dependence.

Observations of the stellar velocities in a large number of galaxies do not match this expectation [56, 57]. Figure 2.4 shows the galaxy M33 as an example. Here, $v(r)$ is increasing instead of decreasing towards large radii, with the expectation of the mass from the stellar disc and interstellar gas shown as the short-dashed and long-dashed lines. Additional non-luminous matter that is not concentrated at the centre of the galaxy is required to describe the observation according to Equation (2.7) – a DM halo surrounding the galaxy. This halo must amount to a majority of the galaxy's mass. Its contribution is shown as the dot-dashed line in Figure 2.4 and a good description of the observation is achieved in a combined fit (solid line) with the luminous matter.

Further evidence is present on cosmological scales. The distribution of matter in the early Universe at the time of the decoupling of matter and radiation is embedded in the cosmic microwave background. This shows an almost isotropic distribution with local fluctuations in the order of 10^{-5} [58]. The angular correlation of these anisotropies is

[6] QCD also theoretically allows for a CP-violating phase. However, this is experimentally found to be $|\theta_{\mathrm{QCD}}| < 4 \cdot 10^{-11}$ [50, 51]. Why this is the case is another unanswered question of particle physics known as the strong CP problem.

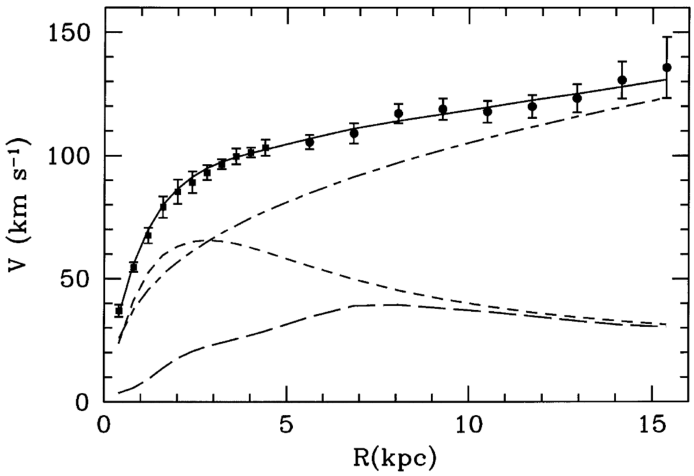


Figure 2.4: Rotational velocity of stars in the galaxy M33 as a function of their distance from the centre. The points show the observation and the solid line shows a fit containing the different components contributing to the galaxy's mass: the stellar disc (short-dashed line), interstellar gas (long-dashed line), and the DM halo (dot-dashed line). Taken from Ref. [55].

very sensitive to cosmological parameters that govern the large-scale structure formation in the Universe. The ΛCDM model of cosmology provides an excellent description of the observed power spectrum of the cosmic microwave background [59]. It assumes a cosmological constant Λ, which is now associated with Dark Energy, and non-relativistic, i.e. *cold* Dark Matter, which interacts at most weakly with itself and SM particles [54].

Alternative theories that alter the effect of gravity, like Modified Newtonian Dynamics [60], have been proposed in an attempt to explain the observations without DM. However, they can only provide explanations at limited scales but not to all phenomena simultaneously [17].

Despite the vast amount of evidence for the gravitational influence of DM in the Universe, its nature remains yet unknown. For a long time, non-luminous macroscopic objects (MACHOS) with masses in the order of several solar masses were considered viable candidates. However, they have been excluded as the sole constituent of DM for masses $m \gtrsim 10^{-7} M_\odot$ recently by microlensing observations and stability constraints on dwarf galaxies [61, 62].

This suggests one or more fundamental particles as the most likely candidates for DM. These must be stable over time scales of the age of the Universe and must not (or only very weakly) interact electromagnetically. While neutrinos fulfil these requirements, they are expected to contribute only between 0.5 % and 1.6 % to the DM mass in the Universe [17]. This is deduced from them behaving relativistically in the early Universe due to their low mass, which would inhibit the gravitational clustering of local density fluctuations and thus delay structure formation [63, 64]. Additionally, the Pauli exclusion principle places an upper limit on the fermionic density in a galaxy, known as the Tremaine-Gunn limit [65]. This requires fermionic DM to have a mass $m > 70\,\mathrm{eV}$, which also excludes neutrinos as the only source of DM [66].

Many DM models include a coupling between the DM and SM particles besides gravity to achieve a thermal equilibrium between them in the early Universe. One class of

these models are weakly interacting massive particles (WIMPs), which could explain the observed DM relic density with new particles at the GeV–TeV scale. WIMP DM is the focus of this analysis.

2.2.2 The Simplified Z' Model

Fully self-consistent DM models, like the Minimal Supersymmetric Standard Model [67], often contain a rich sector of new particles and interactions. The large number of viable complete models and the number of free parameters in them impede the interpretation of search results for all of them.

Simplified DM models are an approach to facilitate the interpretation and potential combination of different experimental results. They can be understood as integrating out the majority of the particle content of more complete theories, after which a phenomenological description via an s- or t-channel mediator exchange and a stable DM particle remains [68].[7] They can be constructed to be renormalizable, obey Lorentz invariance and the SM gauge symmetries and not violate the accidental global symmetries of the SM, e.g. the baryon and lepton number conservation.

In this thesis, a leptophobic axial-vector Z' model is considered [68, 69].[8] It introduces a new U(1) symmetry, under which the DM Dirac fermions χ and quarks are charged. A charge for leptons is possible, but it would barely affect the dijet decay channel studied in this analysis and is already strictly constrained by measurements of the Drell-Yan process [69].

A spontaneous breaking of the U(1) symmetry gives rise to a massive Z' boson, which serves as the mediator of the interaction between the quarks and DM fermions. This interaction is described by the Lagrangian density

$$\mathcal{L}_{\text{int}} \supset g_q \sum_{\substack{q \in \\ \{u,d,s,c,b,t\}}} Z'_\mu \overline{q} \gamma^\mu \gamma^5 q + g_\chi Z'_\mu \overline{\chi} \gamma^\mu \gamma^5 \chi \,, \tag{2.8}$$

where g_q is assumed to be a universal coupling to all quarks and g_χ is the coupling to the fermionic DM. It allows the Z' mediator to be produced in quark-antiquark annihilations and to decay into $\chi\overline{\chi}$ or $q\overline{q}$ as shown in the tree-level Feynman diagrams in Figure 2.5.

The potential existence of a portal between DM and the SM as shown in Figure 2.5a is used in a multitude of DM searches using different detection strategies.

WIMP Detection Strategies

The process $\chi + \overline{\chi} \to X_{\text{SM}} + X_{\text{SM}}$, where X_{SM} denotes any SM particle, is studied in *indirect detection* experiments. They search for astrophysical evidence of DM annihilation

[7] This is a less reductionist approach than effective field theories, which integrate out all mediators to only leave contact interactions. These only provide a good approximation if the potential mediators are beyond the energy scale probed in collider experiments, since otherwise the kinematics of the resonance are disregarded.

[8] In collider searches, the phenomenology of a vector and an axial-vector mediator are very similar [69]. This analysis only considers the axial-vector model, but the constraints on vector mediators are expected to be nearly identical.

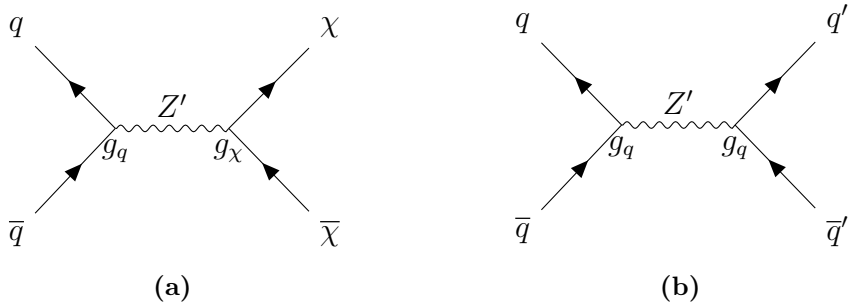


Figure 2.5: Feynman diagram for a hypothetical axial-vector Z' boson decaying into (**a**) a DM $\chi\overline{\chi}$-pair or (**b**) a $q\overline{q}$-pair, producing a dijet signature.

products or their secondary decays, for example in the form of gamma rays at the Fermi Large Area Telescope [70], neutrinos at Super-Kamiokande [71] and IceCube [72], or antiprotons and positrons in the AMS experiment [73].[9] The different detection channels provide sensitivity over many orders of magnitude of potential DM masses. However, the small annihilation cross sections expected from cosmological models to explain the observed relic density could not yet be excluded in a wide range of the parameter space. One limiting factor in many searches is the large uncertainty on the propagation of cosmic rays and potential background sources [74]. A detailed overview of current experimental limits can be found in Ref. [75].

Alternatively, DM in our galaxy could be *directly detected* by scattering on nuclei of ordinary matter in the process $\chi + q \to \chi + q$ (or on the electrons if the respective coupling is non-zero). This process is searched for mostly in underground observatories with large amounts of active material to detect nuclear or electron recoil. Many experiments, like LUX-Zeplin [76] or XENONnT [77], use xenon as the active material for its large nucleon number. The interaction cross section of vector mediators (in the context of direct detection typically referred to as spin-independent interaction) scales with the nucleon number, which results in an enhanced sensitivity. This is not the case for axial-vector mediators (spin-dependent interaction) [78], where the cross section only scales with the total nucleus spin. Therefore, direct detection experiments provide a sensitivity to vector mediators 5-7 orders of magnitude higher than to axial-vector mediators. For vector-mediators, direct detection experiments provide the currently strongest exclusion limits for DM masses $m_\chi \gtrsim 5\,\text{GeV}$. A detailed comparison to collider searches is given in Ref. [79].

In collider experiments, searches for the production of DM are performed. Since the produced DM fermions are not expected to interact with the detectors, at least one additional visible object X is required in the final state to observe the interaction, typically originating as initial state radiation. In these $q\overline{q} \to \chi\overline{\chi} + X$ processes, the DM fermions escape the detector unseen, which leaves X unbalanced in the plane transverse to the beam. Due to momentum conservation, the momenta of all final state particles must add up to zero in this plane, where any deviation from zero is labelled missing

[9] These processes can be realized via small couplings of X_{SM} to the mediator itself or via loop corrections involving, for example, a quark loop coupling to the mediator.

transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$.[10] This type of collider search is hence typically referred to as mono-X or $E_{\mathrm{T}}^{\mathrm{miss}} + X$ search. $E_{\mathrm{T}}^{\mathrm{miss}}$+jet searches are often the most sensitive channel due to the large probability of QCD initial state radiation, but $E_{\mathrm{T}}^{\mathrm{miss}} + V$ ($V = W, Z, \gamma$) searches can also provide complementary sensitivity [40].

Dijet searches are an alternative approach in collider experiments to observe a new mediator. Any mediator that can be produced in pp collisions can also decay into partons, resulting in a pair of jets in the final state – a dijet event. For the Z' model, this happens at tree-level in the $q\overline{q} \to Z' \to q'\overline{q}'$ process shown in Figure 2.5b. Due to this being an on-shell s-channel process, it is detectable as a resonance at $m_{Z'}$ in the invariant mass of dijet events.

The decay width of the leptophobic Z' resonance depends on the masses of and the couplings to the fermions it can decay into:

$$\Gamma_{Z'} = \frac{g_\chi^2\, m_{Z'}}{12\pi}\beta_\chi^3\,\Theta\big(m_{Z'} - 2m_\chi\big) \;+ \sum_{\substack{q \in \\ \{u,d,s,c,b,t\}}} \frac{3g_q^2\, m_{Z'}}{12\pi}\beta_q^3\,\Theta\big(m_{Z'} - 2m_q\big)\,, \tag{2.9}$$

$$\beta_f = \sqrt{1 - \frac{4m_f^2}{m_{Z'}^2}}\,, \tag{2.10}$$

where β_f is the velocity of a fermion with mass m_f in the mediator rest frame and the Heaviside function Θ ensures an on-shell decay [69]. The additional factor 3 for the quark decays is their colour factor.

This width is an important metric for resonance searches as a narrow resonance is more easily distinguished from the QCD background than a wide resonance. If $2m_\chi > m_{Z'}$, the decay into $\chi\overline{\chi}$ is kinematically suppressed and the width only depends on g_q. From experimental constraints (shown below), $g_q \lesssim \mathcal{O}(0.1)$ is to be expected. This makes $\Gamma_{Z'}/m_{Z'}$ small in comparison to the relative detector resolution, which is 4–7 % as shown in Section 6.5. If $2m_\chi \leq m_{Z'}$, the Z' width is increased by the possible decay into DM. This can reduce the sensitivity of resonance searches, since for $g_\chi \gtrsim 1$ the intrinsic Z' width dominates the detector resolution. Additionally, the branching ratio $\mathcal{B}(Z' \to q\overline{q})$ will be decreased, further reducing the sensitivity.

The interpretation in this analysis assumes $m_\chi \gg m_{Z'}$, such that decays into DM are negligible. A reinterpretation for smaller m_χ and arbitrary g_χ is possible as long as the width does not significantly surpass the detector resolution.

For the case of a narrow on-shell resonance, the total cross section of the $q\overline{q} \to Z' \to q'\overline{q}'$ process scales approximately like [69]:

$$\sigma \propto \frac{g_q^4}{\Gamma} \propto \begin{cases} \frac{g_q^4}{g_q^2+g_\chi^2} & \text{if } 2m_\chi \leq m_{Z'}\,, \\ g_q^2 & \text{if } 2m_\chi > m_{Z'}\,. \end{cases} \tag{2.11}$$

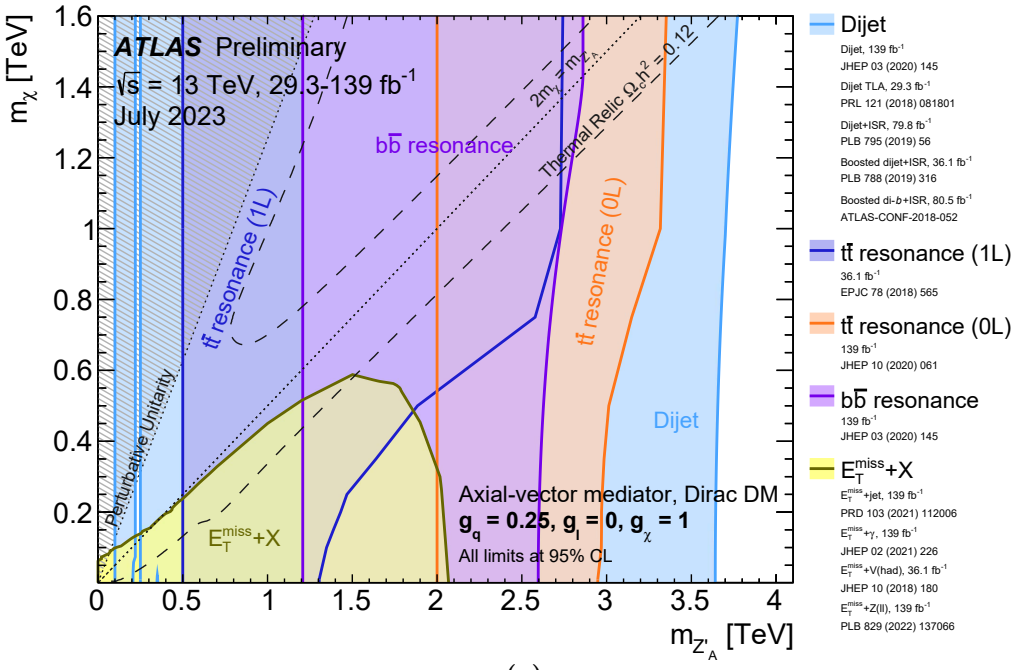
This relation is used to translate constraints on the total cross section into constraints

[10] At collider experiments, the ultrarelativistic limit $|\boldsymbol{p}| \approx E$ provides a good approximation for most final state particles. Since momenta are only measured directly for charged particles, the determination of $E_{\mathrm{T}}^{\mathrm{miss}}$ depends strongly on the more inclusive energy measurement that captures most final state particles.

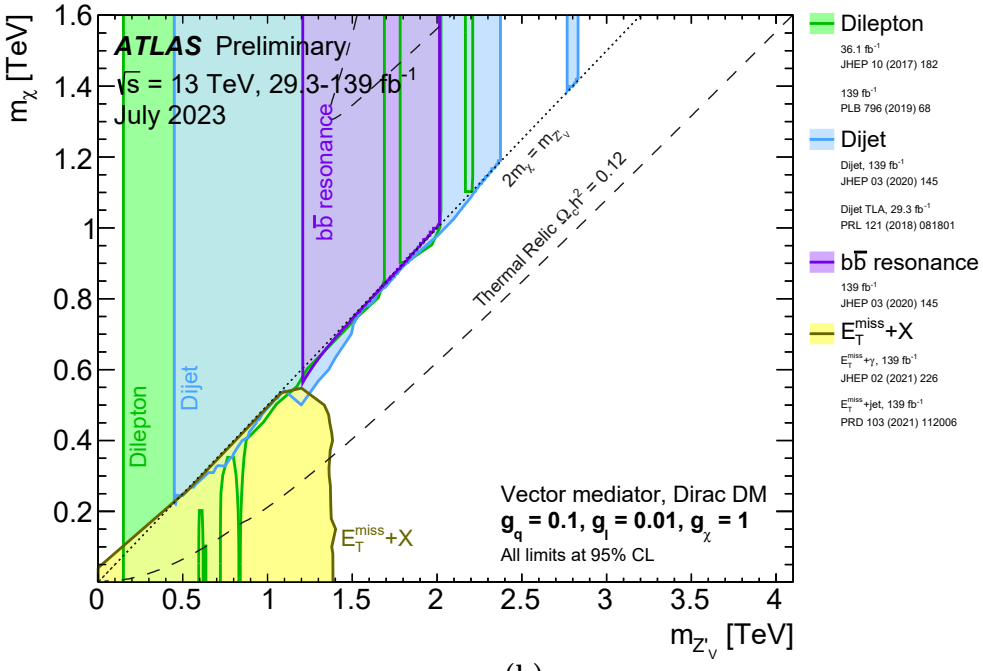
on g_q in Chapter 10.

Figure 2.6 shows current experimental constraints on the parameter space of the simplified Z' model for two different choices of the coupling parameters. The $E_{\mathrm{T}}^{\mathrm{miss}} + X$ searches shown in yellow are sensitive to the parameter space where $2m_\chi \leq m_{Z'}$, allowing for on-shell DM production. In Figure 2.6a, a relatively large quark coupling of $g_q = 0.25$ is assumed. In this case, resonance searches like the inclusive dijet resonance search or the exclusive $b\bar{b}$ and $t\bar{t}$ resonance searches provide sensitivity approximately independent of the DM mass. Figure 2.6b shows the constraints for a smaller quark coupling of $g_q = 0.1$. In this case, the decrease of $\mathcal{B}(Z' \to q\bar{q})$ reduces the sensitivity to the parameter space where $2m_\chi \leq m_{Z'}$. A small coupling $g_\ell = 0.01$ to leptons is assumed in this figure as well. This results in a high sensitivity of dilepton searches, which have a much lower SM background compared to searches with hadronic final states. These constraints illustrate the strong complementarity of $E_{\mathrm{T}}^{\mathrm{miss}} + X$ searches and resonance searches as they provide sensitivity to different regions of the possible parameter space.

One noteworthy benefit of dijet resonance searches, in particular, is the large variety of BSM models they are sensitive to. While DM mediators are the primary motivation for this search, a dijet resonance search can also constrain numerous other BSM models that predict a new resonance coupling to quarks or gluons [3, 4]. For such an interpretation, this analysis provides constraints on generic Gaussian-shaped resonances in the invariant dijet mass spectrum. These provide a sufficient approximation for a variety of resonant processes [80]. Relative Gaussian resonance widths between 5 % and 15 % are tested to cover a wide range of BSM models.



(a)



(b)

Figure 2.6: Constraints on the parameters of the Z' model obtained by various searches performed by the ATLAS experiment. The coloured regions indicate areas in the $(m_\chi, m_{Z'})$ plane that are excluded at 95 % confidence level. The couplings to SM and DM fermions are assumed to be (**a**) $g_q = 0.25$, $g_\ell = 0$, $g_\chi = 1$ and (**b**) $g_q = 0.1$, $g_\ell = 0.01$, $g_\chi = 1$. In each figure, the parameters which would explain the observed DM relic density are shown as the long-dashed line. Taken from Ref. [79], where other coupling choices are shown as well.

3 Simulation of Proton-Proton Collisions

Simulations of *pp* collision events, the emergence of jets, and their interaction with the detector are essential for multiple aspects of this analysis. They are used to calibrate the observed jets, optimize the event selection criteria, simulate potential BSM signals, and are employed in a novel background estimation method.

As Monte Carlo techniques are employed in calculating the high-dimensional phase space integrals occurring in the scattering cross sections and in the simulation of particle shower and hadronization processes, these simulations are referred to as Monte Carlo (MC) simulations.

This chapter provides an overview of the full simulation chain from the scattering process to the detector interaction for the MC samples used in this search, based on Refs. [81–83]. Furthermore, the computation of parton-level dijet cross sections specifically for the background estimate of this analysis is described. The emphasis is on QCD calculations, as this analysis focuses solely on jet production processes.

3.1 Production of Dijet Monte Carlo Samples

Due to the non-perturbative nature of the strong interaction at energy scales below 1 GeV, calculations involving *pp* collisions rely on factorization theorems. These separate cross section calculations into a short-distance interaction, described by perturbative parton-level matrix elements, and long-distance processes such as the parton composition of the proton or final state fragmentation, which need to be described with heuristic models tuned to data.

3.1.1 Hard Interaction

Factorized, the differential cross section of a *pp* interaction resulting in a final state X can be expressed as a convolution of parton-level cross sections $\widehat{\sigma}_{ab\to X}$ with parton distribution functions (PDFs) $f_a(x_a, \mu_F)$:

$$\mathrm{d}\sigma_{pp\to X} \approx \sum_{a,b} \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \, f_a(x_a, \mu_F) f_b(x_b, \mu_F) \, \mathrm{d}\widehat{\sigma}_{ab\to X}(\Phi_n, \mu_F, \mu_R) \,. \tag{3.1}$$

The PDFs universally describe the probability of a parton a with momentum fraction x_a to interact in a proton hard scatter process. They depend on the factorization scale μ_F at which the factorization into short-distance and long-distance interaction is performed. They are obtained from global fits to a large number of measurements [84, 85] at various momentum fractions and factorization scales. The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations can be used to analytically transfer them between

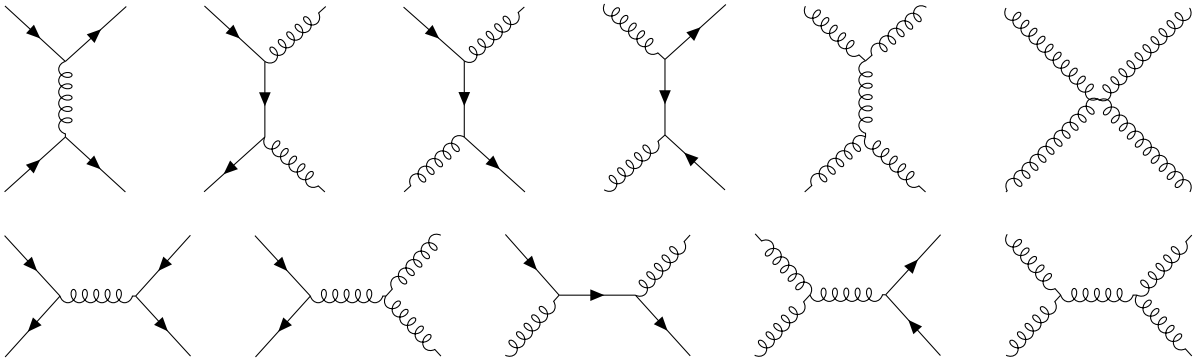


Figure 3.1: Selection of leading-order partonic subprocesses contributing to the dijet production in pp collisions.

scales.

The parton-level cross section is obtained from the spin- and colour-averaged matrix element $\mathcal{M}_{ab\to X}$:

$$\mathrm{d}\widehat{\sigma}_{ab\to X}(\Phi_n, \mu_F, \mu_R) = \frac{1}{2\widehat{s}}\left|\mathcal{M}_{ab\to X}(\Phi_n, \mu_F, \mu_R)\right|^2 \mathrm{d}\Phi_n \,, \tag{3.2}$$

where $\widehat{s}$ denotes the centre-of-mass energy between partons a and b, Φ_n the phase space momenta of the n outgoing particles, and μ_R the renormalization scale.

The matrix element is calculated perturbatively at fixed order in α_S. In this thesis, leading order (LO) and next-to-leading order (NLO) matrix elements of the $2 \to 2$-jet production process are considered.

Leading-Order Calculation

At LO, the dijet production corresponds to the sum of all $2 \to 2$-parton processes. This already corresponds to a large number of diagrams, since all combinations of both initial and final state partons being (anti-)quarks or gluons must be considered. Figure 3.1 shows a selection of LO Feynman diagrams to be computed.

In order of decreasing total contribution, these are $gg \to gg$, $qg \to qg$, $qq \to qq$, $q\overline{q} \to q\overline{q}$, $q\overline{q} \to gg$, and $q\overline{q} \to q\overline{q}$ processes. This order is mainly driven by the dominance of gluons and valence quarks in the proton PDF. All of these processes contain at least one t-channel diagram that enhances the cross section at small deflection angles [15]. This angular behaviour is utilized in Section 7.2 to enhance the separation of the QCD background from the s-channel resonances targeted by this search.

Next-to-Leading Order Calculation

The dijet production at NLO involves two types of corrections: The virtual correction accounts for one-loop corrections to vertices and self-energy and includes box diagrams in the $2 \to 2$-parton process. Some examples of this are shown in Figure 3.2a. The real correction introduces a third parton in the final state in the form of gluon radiation or splitting. Figure 3.2b shows exemplary Feynman diagrams of this.

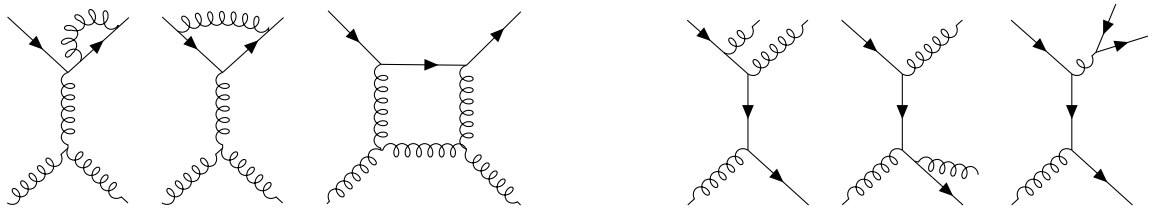


Figure 3.2: Selection of next-to-leading-order Feynman diagrams contributing to the dijet production in proton-proton collisions. (**a**) Virtual corrections to the $2 \to 2$-parton process and (**b**) real corrections involving a third final state parton are shown.

When integrating these correction terms, divergences are encountered. The real correction exhibits a divergence for the soft emission of a parton and for the collinear splitting of a parton. These divergences are addressed by constructing infrared- and collinear-safe observables, i.e. observables that are invariant under soft emission or collinear splitting. Typically, this is achieved by defining observables not in terms of individual final state particles, but by introducing an algorithm that clusters the final state partons into jets and only considering observables as functions of jet properties. The anti-k_t jet clustering algorithm used in this analysis is described in detail in Section 5.2.

The virtual correction exhibits an infrared divergence when the loop propagator becomes on-shell, which cancels exactly with the infrared divergence of the real correction according to the Kinoshita-Lee-Nauenberg theorem [86, 87]. In numerical, differential cross section calculations, this cancellation is achieved with a dipole subtraction method [88].

To calculate cross sections differentially or in a limited phase space matching the experimental event selection, $\mathrm{d}\widehat{\sigma}_{ab\to X}$ needs to be integrated numerically and convolved with the PDFs. As these integrals are high-dimensional, Monte Carlo integration techniques are employed which scale favourably to higher dimensions in comparison to alternative methods. Events are sampled throughout the available phase space and assigned weights proportional to the squared matrix element for the specific configuration of particle momenta.

Since real and virtual corrections can be large while cancelling each other out to a large degree, numerical NLO calculations need to generate more events than LO calculations to achieve the same statistical precision.

Cross sections computed with the factorized approach exhibit a residual dependence on the artificially introduced scales μ_F and μ_R. This dependence is a remnant of truncating the perturbation theory at a fixed order and decreases as higher orders are included. There is no fundamentally best choice for these scales. They are typically set to a representative momentum scale for a given interaction and the impact of a variation of these scales is used as an estimate for the uncertainty of the calculation due to missing higher orders [89].

3.1.2 Parton Shower and Hadronization

The momenta of jets reconstructed from the two or three final state partons generated in the hard interaction at (next-to-)leading order already provide a good approximation of the leading jets that would be observed in a detector. This description, however, fails to model events with higher jet multiplicities in the final state, as well as the complex structure of a real jet composed of numerous hadrons.

Currently, it is not feasible to describe these phenomena in fixed-order perturbation theory. Instead, probabilistic parton shower models are used to simulate the radiation of gluons or the splitting of gluons into $q\overline{q}$-pairs in the final state that are missed when truncating at a fixed order.[1] This is done in an evolution from the energy scale of the partons in the hard interaction down to a scale of the order of 1 GeV, where perturbation theory breaks down. A matching and merging procedure is implemented to prevent overlaps between splitting or radiation accounted for by NLO corrections and in the parton shower simulation.

With two partons of the two incoming protons being part of the hard process, the remnants of the protons are no longer colour-neutral and hence participate in additional strong interactions. These multi-parton interactions are complex and are approximated by empirical models as the so-called underlying event.

The partons generated in the parton shower are not free as they still carry colour charge. The subsequent hadronization is an involved process that is described via phenomenological models in which colour connections are identified and a clustering into colour-neutral hadrons is performed. Heavy flavour decays are simulated phenomenologically until only stable final state particles remain.

The models for the parton shower, the underlying event, and the hadronization process involve sets of free parameters that are tuned to match measurements.

3.1.3 Detector Simulation

A detailed simulation of the used detector is necessary to compare MC samples directly to real data. It enables the use of MC samples to predict event rates or for calibration purposes.

For this analysis, a simulation of the complete ATLAS detector based on GEANT4 [90, 91] is employed for which the stable final state particles[2] generated as described above are used as input. The simulation accounts for interactions of particles with both active and inactive detector material and emulates readout electronics. The resulting event format is identical to that of a real recorded event and can be analyzed using the same reconstruction algorithms employed for data (see Section 5.2).

In addition to this *reconstruction-level* information, *truth-level*[3] information about the

[1] Gluons shower systematically more than quarks due to their larger average colour factors. These are determined from averaging over the colour states of the initial state partons and summing over those of the final state partons. For parton splitting, they are $C_{g\to gg} = 3$, $C_{g\to qq} = 4/3$, and $C_{q\to qg} = 1/2$, which is why gluons are more likely to split or emit radiation than quarks [15].

[2] In ATLAS simulations, 'stable' refers to particles with a lifetime $c\tau > 10\,\text{mm}$, as they could interact with material.

[3] Often also referred to as *particle-level* to distinguish it from the parton-level information before hadronization.

underlying generated particles is kept in every event.

3.2 Samples Used in this Analysis

This analysis uses two sets of MC samples: dijet QCD samples for the jet calibration and signal samples simulating the leptophobic Z' simplified DM model introduced in Section 2.2.

The inclusive dijet QCD samples are generated with Pythia 8.235 at LO [92]. Pythia computes both the hard scatter event and performs the consequent shower and hadronization simulation. It is tuned with the A14 parameter set [93] and uses the Nnpdf2.3lo PDF [84]. Heavy flavour decays are simulated with EvtGen 1.6.0 [94].

Alternative dijet QCD samples are produced with Sherpa 2.2.5, which simulates the hard scatter event, the shower, the hadronization, and hadron decays [95]. To evaluate the impact of the hadronization model, two different samples are generated using the cluster fragmentation [96] and Lund string model [97]. The CT14nnlo PDF [85] and the default CT10 tune are used.

For the Z' model, the hard interaction is generated with MadGraph5_aMC@NLO 2.2.3, with the coupling to SM quarks set to $g_q = 0.1$. The mass of the fermionic dark matter χ is set to 10 TeV to have the mediator decay exclusively into quarks.[4] With these parameters, the width of the Z' resonance including the parton shower and the detector resolution is approximately 9 % at a mass of 350 GeV, decreasing to 5 % at 1800 GeV. The generator is interfaced to Pythia 8.210 for the parton shower and EvtGen 1.2.0 for heavy flavour decays. The A14 tune and the Nnpdf2.3lo PDF set are also used for these samples.

3.3 Calculation of Parton-Level Cross Sections

For the NLOFit background estimate detailed in Section 9.2, the parton-level dijet cross section is calculated with NLOJet++ 4.1.3 [98]. The aim is to generate a nominal dijet mass spectrum prediction and its systematic theoretical variations (from uncertainties of the PDF set, α_S and the scale choices) to use as templates in a fit to the observed data. To achieve a statistical precision better than that in data with available computational resources, only the calculation of the hard parton interaction at NLO is performed. Parton showering, hadronization, and detector simulation are omitted due to their computational expense.

The up to three generated real partons are clustered into jets with the anti-k_t algorithm with radius parameter $R = 0.4$ using FastJet 3.2.2 [99] to avoid infrared divergences. The two leading jets are required to pass the kinematic selection criteria on rapidity and transverse momentum discussed in Chapter 7.

For a given number of generated events, the LO calculation has a higher statistical precision than the NLO calculation due to cancelling corrections. Both are stored

[4] See the discussion of Equations (2.9) and (2.11) regarding the reinterpretation for the parameter space where a decay of the Z' into DM is possible.

separately to make use of the LO calculation's high statistical precision and the NLO calculation's more accurate description of the shape of the dijet mass spectrum.

The computationally intensive cross section calculation with NLOJet++ only provides the nominal prediction for one chosen PDF set. To avoid repeating this process for each systematic variation, it is interfaced to Applgrid 1.6.17 [100]. This software enables the export of the weights calculated in NLOJet++ on a grid in momentum fractions x_1, x_2, momentum transfer Q^2, partonic subprocess (i.e. which types of partons a and b interact), and order of α_S in a specified binning of the chosen observable (the invariant mass of the dijet system m_{jj}). For this thesis, a grid consisting of $40 \times 40 \times 25 \times 7 \times 2 \times 300$ bins is selected. The binning in m_{jj} is chosen specifically to align with the bins used throughout this analysis, whereas the grid range in x and Q^2 are automatically determined by Applgrid in a first run with $\approx$1 % of the final statistics.

The created grid is then convolved a posteriori with arbitrary PDF sets using Applgrid. The PDFs are interpolated between their published discrete points in x using Lhapdf 6.2.3 and DGLAP-evolved to arbitrary Q^2 with Hoppet 1.2.0.

For this thesis, the nominal PDF set CT14nnlo, along with its 56 eigenvector variations, which parametrize the PDF uncertainties via the Hessian method [101], are chosen. Additional varied templates are generated by changing α_S within its uncertainty and by rerunning the convolution with the factorization and renormalization scale shifted up and down by a factor of two both individually and simultaneously.

4 The ATLAS Experiment

For the dijet resonance search presented in this thesis, pp collisions produced by the LHC at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ are analyzed. The data were recorded by the ATLAS detector between 2016 and 2018 during the LHC Run-2.

This chapter provides a description of the LHC and its pre-accelerator complex at CERN[1] in Section 4.1 and of the ATLAS detector and its subsystems in Section 4.2, mainly based on Refs. [102, 103]. It covers the status of the experiment during Run-2 when the data for this analysis were collected. Several upgrades for the currently ongoing Run-3 have been performed since, but these are not relevant to this thesis.

4.1 The Large Hadron Collider

The LHC is the largest and most powerful particle accelerator in the world, having recently broken its own record with proton collisions at a centre-of-mass energy of $\sqrt{s} = 13.6\,\text{TeV}$ in Run-3 [104]. It is located in a 26.7 km long circular tunnel underground the Geneva area in Switzerland and France at a depth between 45 m and 170 m. This tunnel previously accommodated the Large Electron–Positron Collider (LEP) [105], the most powerful lepton collider ever built, that reached a centre-of-mass energy of up to $\sqrt{s} = 209\,\text{GeV}$.

Since the energy loss via synchrotron radiation is strongly mass-suppressed for protons in comparison to electrons,[2] the LHC is capable of accelerating protons (as well as heavy ions [107]) to much larger energies than LEP was previously able to achieve with electrons.

The protons are sourced from hydrogen gas by stripping its electrons and are accelerated in a first linear accelerator (Linac2) to 50 MeV.[3] Subsequently, the protons are successively accelerated in the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS), where they reach an energy of 450 GeV before being injected into the LHC. Figure 4.1 shows a schematic overview of the pre-accelerator complex.

The protons are injected as *bunches* of $\mathcal{O}(10^{11})$ protons into the LHC with a spacing of 25 ns (7.5 m). These bunches are later brought to collision in so-called *bunch crossings*. The filling scheme of the pre-accelerator chain, in combination with finite switching times of the injection and dumping magnets, results in regular patterns of filled and empty bunches. While the circumference of the LHC allows for 3564 bunches, at most 2556 were filled during Run-2 [110].

[1] Conseil Européen pour la Recherche Nucléaire (European Organization for Nuclear Research)

[2] The power P radiated off a charged particle bent on a circular trajectory scales like $P \propto {}^{E^4}/_{m^4}$ [106]. At the same beam energy, the synchrotron radiation of protons is thus $\mathcal{O}(10^{13})$ lower than that of electrons.

[3] In 2020, the Linac2 has been replaced by the Linac4 in preparation of future LHC runs with higher luminosity [108]. It accelerates negative hydrogen ions whose electrons are now stripped after the linear accelerator.

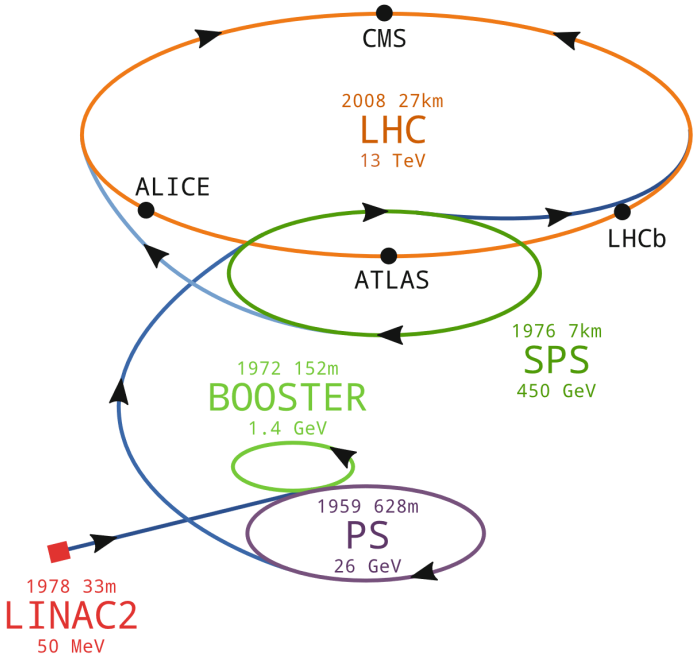


Figure 4.1: Schematic overview of the CERN accelerator complex. The completion date, circumference, and the maximum reached proton energy during Run-2 for each of the accelerators and the four main interaction points of the LHC are shown. Adapted from Ref. [109].

Due to the limited space in the tunnel, the two counter-rotating beams of protons are contained in two rings within the same twin-bore magnet. It consists of superconducting dipole magnets, capable of reaching 8 T at design energy, and quadrupole magnets to focus the beam and thus increase the luminosity at the interaction points. The magnets have not been ramped to their full potential yet to reduce the risk of quenching, limiting the centre-of-mass energy to 13 TeV in Run-2 instead of the design value of 14 TeV [111].

The counter-rotating beams are crossed at four main interaction points to produce the proton-proton or heavy-ion collisions. At these points, the four large LHC experiments are located: ATLAS [103] and CMS [112] are both general-purpose detectors built with the goal of detecting the Higgs boson – which was achieved in 2012 [113, 114] – as well as searching for (hints of) BSM physics. The ALICE experiment [115] mainly studies the deconfinement properties of quarks and gluons in a hot plasma produced in heavy-ion collisions. LHCb [116] is a forward detector designed to study hadron decays involving bottom or charm quarks with a focus on measurements of CP violation.

At the interaction points, the *instantaneous luminosity* $\mathcal{L}$ is a measure of the density of the colliding beams. It determines the rate of collision events of a process with cross section σ:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \sigma\mathcal{L}\,. \tag{4.1}$$

The expected total number of events in a given data-taking period is accordingly proportional to the *integrated luminosity* $L = \int \mathcal{L}\,\mathrm{d}t$ accumulated over this period.

For Gaussian beam profiles with standard deviations σ_x and σ_y in the transverse plane,

the instantaneous luminosity is given by

$$\mathcal{L} = \frac{N_p^2 \, N_b \, f}{4\pi\sigma_x \, \sigma_y} \, S \, , \tag{4.2}$$

where N_p is the number of protons per bunch, f is the LHC revolution frequency of 11.2 kHz, N_b is the number of filled bunches, and S is a geometric correction factor in the order of 70 % to account for the crossing angle between the colliding beams [117–119]. With the well-performing focussing optics, and hence small σ_x and σ_y, the LHC surpassed its design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ by a factor of 2 [118].

A high instantaneous luminosity comes with the challenge of *pile-up*, i.e. multiple pp interactions per bunch crossing. During Run-2, the average number of simultaneous interactions per bunch crossing (μ) varied between approximately 10 and 60, depending on the run conditions, with an overall average of 34 [120]. Pile-up collisions pose challenges to the trigger and event reconstruction to distinguish their effects from the interaction of interest. Pile-up effects on the jets studied in this thesis are considered in Chapter 6.

4.2 The ATLAS Detector

The ATLAS detector is a general-purpose particle detector with high granularity. With a diameter of 25 m, a length of 46 m, and a mass of 7000 t, it is the largest LHC experiment as well as the largest general-purpose particle detector ever constructed. The ATLAS detector covers the same physics goals as the CMS detector but relies on different technical implementations to cross-verify discoveries. The design goal of discovering the Higgs boson required an excellent photon, electron, and muon reconstruction for the most sensitive Higgs decay channels. A good jet resolution and vertex reconstruction are necessary for the efficient identification of potential BSM signatures involving (b-)jets or $E_{\mathrm{T}}^{\mathrm{miss}}$.

The ATLAS detector offers an almost 4π solid angle coverage. It employs a layered structure composed of the Inner Detector, electromagnetic and hadronic calorimeters and the Muon Spectrometer as depicted in Figure 4.2. These subsystems are described in the following sections with particular emphasis on the calorimeters as mainly their information is used in the dijet resonance search presented in this thesis.

4.2.1 Coordinate System

A right-handed coordinate system originating from the interaction point is used for the description of the detector. The z-axis aligns with the beam direction, the x-axis points towards the centre of the LHC ring, and the y-axis points upwards. Transverse quantities, such as the transverse momentum, are projections onto the (x, y) plane. The azimuthal angle ϕ represents the angle in this plane relative to the positive x-axis, and the polar angle θ describes the angle towards the positive z-axis. The polar angle is typically expressed in terms of the pseudorapidity $\eta = -\ln\left[\tan(\theta/2)\right]$, which approaches the rapidity $y = 1/2 \ln\left[(E + p_z)/(E - p_z)\right]$ in the ultrarelativistic limit. This choice is beneficial because rapidity differences are invariant under Lorentz boosts, which typically

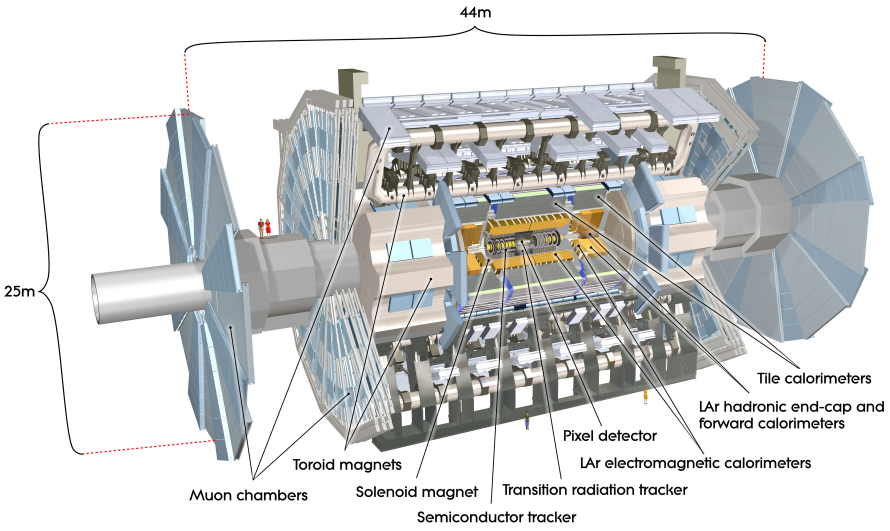


Figure 4.2: Schematic overview of the ATLAS detector. Taken from Ref. [121].

occur in hadron-hadron collisions due to the different momentum fractions x of the interacting partons. Distances between two objects are measured in the (η, ϕ) plane as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

4.2.2 Inner Detector

The Inner Detector (ID) is the central component of the ATLAS detector and is designed to track the trajectory of charged particles to determine their charge and momentum. It is immersed in a 2 T magnetic field generated by the central ATLAS solenoid magnet (see Figure 4.2) to bend the trajectories of charged particles due to the Lorentz force. The curvature radius is proportional to the particle momentum and its direction distinguishes positive from negative charges [13]. The detected particle tracks allow for the reconstruction of primary collision vertices, which is important to distinguish pile-up collisions from the collision of interest, and of secondary decay vertices of longer-lived particles, which is crucial for the identification of e.g. B mesons or τ leptons.

The ID covers the central detector region of $|\eta| < 2.5$. It consists of several layered sub-detectors that are shown in Figure 4.3: the *Silicon Pixel Detector*, the *Semiconductor Tracker*, and the *Transition Radiation Tracker*. Each sub-detector is composed of central Barrel layers that surround the beam pipe and perpendicular, disc-shaped End-Caps to cover the more forward $|\eta|$ regions.

The Pixel Detector contains sensitive silicon pixels with sensor sizes down to 50 µm × 250 µm.[4] Each layer of the Semiconductor Tracker consists of pairs of silicon microstrips oriented at a small angle to each other to obtain two-dimensional information regarding the hit position along the microstrips. The outermost part of the ID, the Transition Radiation Tracker, is built of layers of gaseous drift tubes. It typically provides the largest

[4] In 2014, the Pixel Detector was upgraded with a fourth Barrel layer closest to the beam pipe. This *Insertable B-Layer* [123] offers the highest spatial resolution of the ATLAS ID.

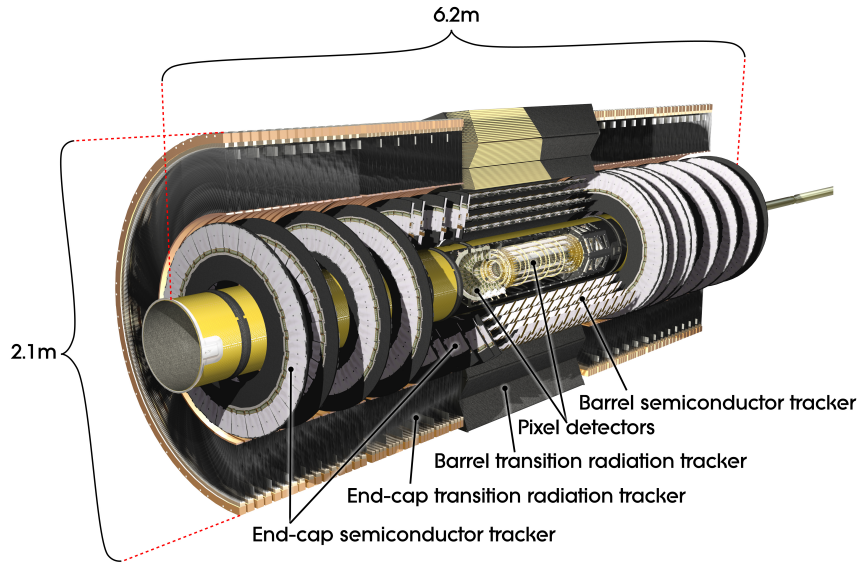


Figure 4.3: Schematic overview of the ATLAS Inner Detector. Taken from Ref. [122].

amount of hits of charged particles for the track reconstruction. The space between the tubes is filled with polymers to induce transition radiation of traversing electrons to assist the particle identification [124].

The material budget of the ID is kept low to minimize the energy loss of light particles – mainly electrons and pions – traversing the ID and to reduce the probability of photons converting into electron-positron pairs.

The ID achieves a primary vertex resolution of approximately 50–300 µm along the beam-axis, depending on the pile-up conditions, and 10 µm in the transverse direction [125]. The resolution of the momentum determination from the bending radius degrades towards large momenta due to the particle trajectory approaching a straight line. The ID is designed to achieve a relative transverse momentum resolution of:

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.05\,\% \cdot p_\mathrm{T}\,[\mathrm{GeV}] \oplus 1\,\%\,, \tag{4.3}$$

which was confirmed for the high-energy limit in cosmic ray measurements [124]. In this description, $\oplus$ refers to the summation of uncertainties in quadrature.

4.2.3 Calorimeters

The ATLAS calorimeters surround the ID and the solenoid magnet. They are sampling calorimeters, consisting of alternating layers of dense absorber material to stop the majority of particles (with the exception of muons and neutrinos) and active material to measure the energy deposited by the particle showers. A fine granularity allows for the measurement of shower shapes, mainly to distinguish electromagnetic (narrow) from hadronic (wide) showers, which is important for particle identification.

The ATLAS calorimeters consist of two main layers: The inner *electromagnetic* (EM) layer, designed to detect electrons and photons, and the outer *hadronic* layer, designed

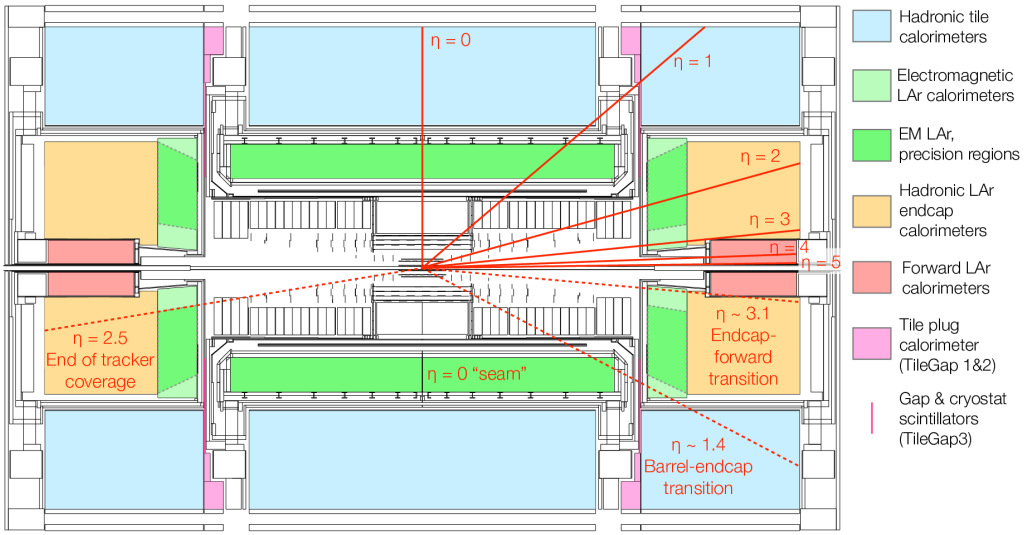


Figure 4.4: Schematic overview of the different subsystems making up the ATLAS calorimeters. The EM layer is shown in green, and the multiple systems of the hadronic layer are shown in blue, orange, and red. The $|\eta|$ values of relevant transition regions are indicated. Taken from Ref. [126].

to detect hadrons. Both layers consist of a central Barrel surrounding the beam pipe and perpendicular End-Caps covering the more forward regions. Figure 4.4 shows a cross section of the calorimeters, indicating the individual subsystems and their relevant $|\eta|$ regions.

Electromagnetic Calorimeter

The EM layer specializes in the detection of electrons and photons, which deposit their energy in a relatively dense shower. At energies above ≈10 MeV, electrons mainly interact with material via bremsstrahlung, and photons mainly interact via pair production of electron-positron pairs. Both of these processes increase the number of electrons and photons amongst which the incidental energy is distributed, resulting in a particle shower. This shower dissipates the energy down to the MeV scale, at which the energy loss via ionization of the detector material becomes dominant [127].

In the particle shower, the energy decreases exponentially with the depth, characterized by the radiation length X_0 as decay constant. Since the cross section of bremsstrahlung and pair production scales quadratically with the atomic number Z, the radiation length scales as [13]

$$X_0 \propto \frac{m_e^2}{Z^2}\,. \tag{4.4}$$

To achieve a high stopping power, lead is chosen as absorber material in the ATLAS EM calorimeter. It is arranged in accordion-shaped strips longitudinally from the interaction point. The gaps between the lead strips are filled with liquid argon (hence the name *LAr* calorimeter) as an active material with a high voltage applied to detect the ionization

in the argon due to traversing charged particles. Voltages of 1–2.5 kV result in ion drift times of approximately 450 ns. To optimize the signal-to-noise ratio, the signal is electronically modulated into a bipolar shape of a positive peak of approximately 100 ns duration and a following negative valley of approximately 400 ns duration.[5]

The EM calorimeter consists of the EM Barrel covering $|\eta| < 1.475$ (shown as the horizontal green areas in Figure 4.4) and the EM End-Caps covering $1.375 < |\eta| < 3.2$ (shown as the vertical green areas in Figure 4.4). The latter are further divided into the Outer Wheels in the region $1.375 < |\eta| < 2.5$ and the Inner Wheels in the region $2.5 < |\eta| < 3.2$ (inner, lightly green shaded areas in Figure 4.4). The EM Barrel and the End-Cap Outer Wheels offer a granularity of three layers longitudinally and up to 0.025×0.025 in the (η, ϕ) plane, depending on the layer and $|\eta|$. The End-Cap Inner Wheels have a reduced granularity of two layers and up to 0.05×0.025 in (η, ϕ).

In total, the EM calorimeters amount to a material budget of 20 to 38 X_0, depending on $|\eta|$. This ensures complete confinement of most electromagnetic showers, such that the total energy of the original particle can be measured. Muons only deposit a small amount of energy in the EM calorimeters due to the inverse quadratic scaling of the energy loss via bremsstrahlung.

Most hadrons start showering in the EM calorimeters but are not fully contained due to the mass suppression of bremsstrahlung. They enter the hadronic layer that is specialized for their detection.

Hadronic Calorimeter

High-energy hadrons primarily interact with material via inelastic nuclear scattering. This develops a cascade that is much more complex than an electromagnetic shower. The process is schematically illustrated in Figure 4.5a and described briefly in the following paragraphs. Additional details are provided in Ref. [127].

Spallation. Predominantly, high-energy hadrons interact with a single nucleon in a nucleus via the strong force. The produced hadrons can escape the nucleus or further interact with other nucleons in an intranuclear cascade on timescales of 10^{-22} s. Ultimately, nucleons or light fragments at energy scales in the order of 100 MeV to few GeV escape the nucleus (see Figure 4.5a). Charged spallation products mainly deposit their energy via ionization of the surrounding material, while neutral spallation products can interact with other nuclei.

Internuclear cascade. If the spallation products are of sufficiently high energy, they can again inelastically interact with other nuclei. This leads to a cascade of new spallation processes, increasing the number of particles participating in the hadronic shower.

Nucleus deactivation. The nucleus typically remains in a highly excited state following spallation. After timescales in the order of 10^{-18} s, the energy is released by evaporation of photons, nucleons, or nuclear fragments at energy scales in the order of a few MeV. Sometimes, the energy is released via nuclear fission instead. Both processes are illustrated in Figure 4.5b. The energy expended in breaking up the nuclear binding energy is lost for the detection process and referred to as *invisible energy*. Additionally,

[5] This pulse duration is long in comparison to the time of 25 ns between collisions. Consequently, the effect of previous collisions distorts the energy measurements. The effect of this *out-of-time pile-up* must be corrected for in the object reconstruction.

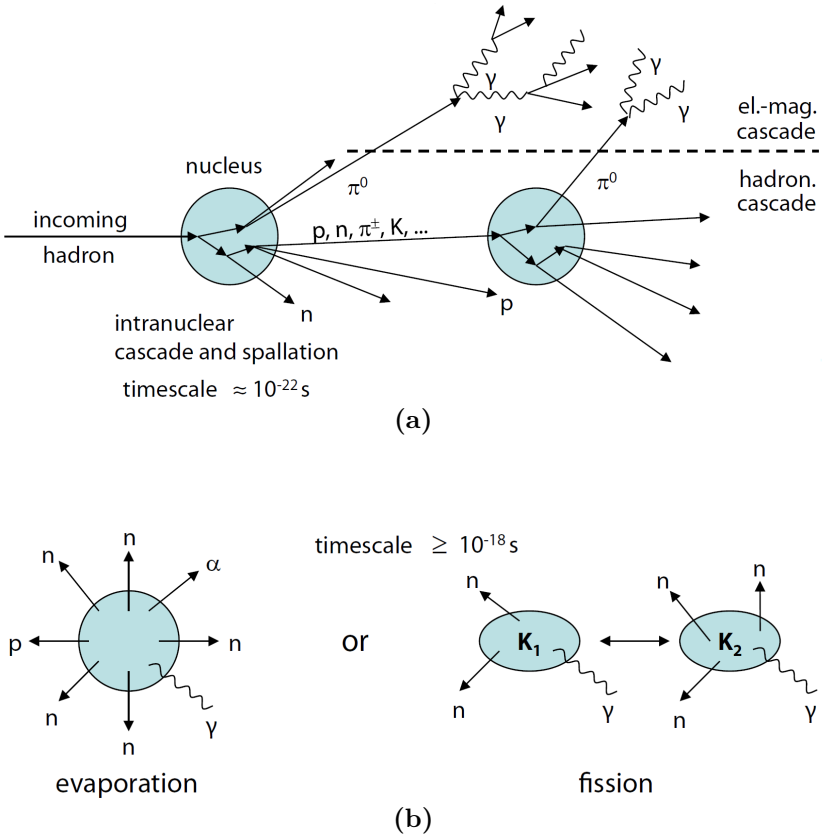


Figure 4.5: Schematic depiction of the interaction of hadrons with material. (**a**) An incoming hadron collides with a nucleus, causing spallation that emits multiple hadrons from the nucleus. These can potentially collide with other nuclei, causing an internuclear cascade. A significant amount of π^0 mesons are created in a typical cascade, which predominantly decay into photon pairs. These induce a separate, electromagnetic shower. (**b**) Nuclei are typically left in a highly excited state following spallation. After timescales in the order of 10^{-18} s, this state decays into a more stable state via evaporation of photons, nucleons, or light nuclear fragments, or occasionally via nuclear fission. Taken from Ref. [127].

slow neutrons can be captured by other nuclei, which enter an excited state with a lifetime in the order of 10^{-6} s. This is much longer than the LHC bunch crossing frequency and calorimeter integration time. Therefore, this energy is also not detected for the collision of interest.

Electromagnetic and weak processes. Short-lived hadrons produced in the processes above can decay before interacting inelastically with other nuclei. This happens, for example, for neutral pions that are created in a significant number in hadronic showers. The dominant decay $\pi^0 \to \gamma\gamma$ produces photons that result in a separate, electromagnetic shower as described above (illustrated in the upper part of Figure 4.5a). More rarely, charged pions or kaons decay weakly, resulting in neutrinos and often muons, which both contribute to the invisible energy.

The different mechanisms of energy deposition of a hadronic shower are summarized in Figure 4.6 for the case of a proton interacting with iron. The fraction of energy

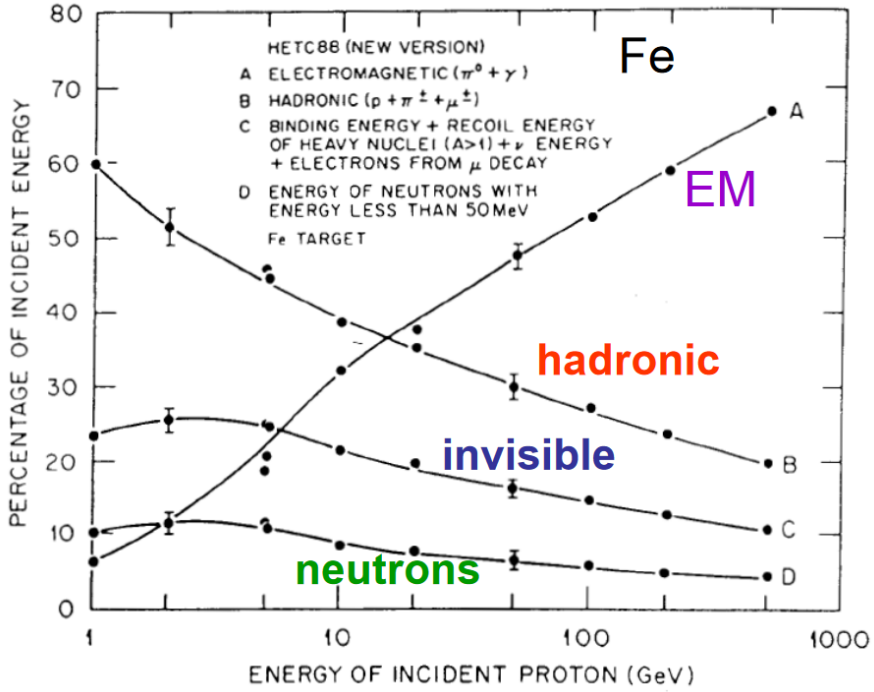


Figure 4.6: Fraction of the energy of a proton showering in iron deposited via the different interaction mechanisms as a function of proton energy. Shown are the fraction of electromagnetic showers, hadronic showers, slow neutrons, and undetected (invisible) energy deposits. Taken from Ref. [128].

deposited by each mechanism is shown as a function of the proton energy. The energy fraction deposited by slow neutrons can be counted towards the invisible energy for the time scales relevant to the ATLAS calorimeters. The electromagnetic and hadronic showers produce different responses in the calorimeter. Their energy-dependent relative contributions are accounted for in the calibration of the jet energy scale in Chapter 6.[6]

The hadronic shower also results in particle energies that exponentially decrease with the shower depth. It is characterized by the nuclear absorption length λ. It scales with the nucleon number A as $\sqrt[3]{A}$, which leads to an approximate relation to the electromagnetic radiation length in dense materials of

$$\frac{\lambda}{X_0} \approx 0.37\, Z\,. \tag{4.5}$$

Due to the significantly longer nuclear absorption length, hadronic calorimeters are designed with a larger absorber material budget. Since it does not scale as strongly with Z, lower-Z materials like iron are suitable absorber choices as well.

The hadronic calorimeter in ATLAS is based on two different technologies. The Tile Long Barrel covering $|\eta| < 1.0$ and the Tile Extended Barrels covering $0.8 < |\eta| < 1.7$ em-

[6] In principle, calorimeters can be designed to be *compensating*, i.e. to align their electromagnetic and hadronic response. This mainly depends on the material choice for the absorber (uranium is suitable, for example) and the relative thickness of absorber and active layers [127]. The ATLAS calorimeters are under-compensating, i.e. they exhibit a lower hadronic than electromagnetic response.

ploy alternating layers of more cost-effective steel absorbers and scintillating polystyrene tiles as active material, equipped with wavelength-shifting fibres and photomultipliers for the readout. They produce strictly positive signal pulses with a duration of approximately 150 ns. The Hadronic End-Cap Outer Wheels ($1.5 < |\eta| < 2.5$), Hadronic End-Cap Inner Wheels ($2.5 < |\eta| < 3.2$), and the Forward Calorimeters ($3.1 < |\eta| < 4.9$) use liquid argon as the active material like the EM layer, but use significantly thicker absorbers made of copper and tungsten. The transition region from the Barrel to the End-Cap region houses service and cryostat material that causes energy losses in the range of approximately $1.0 < |\eta| < 1.6$. This region is equipped with thin scintillators (TileGap3 scintillators, shown in pink in Figure 4.4) to sample the energy deposited in this area.

The Tile Barrel detectors and the Hadronic End-Cap Outer Wheel provide a granularity of 0.1×0.1 in the (η, ϕ) plane and longitudinally three and four layers, respectively. The Hadronic End-Cap Outer Wheel has a coarser granularity of 0.2×0.2 and four layers. The Forward Calorimeters have geometrically irregular readout channels due to space constraints very close to the beam pipe. They consist of three layers longitudinally and offer very coarse resolution in the (η, ϕ) plane. The coarser spatial resolution of the hadronic layer in comparison to the EM layer is sufficient to resolve the wider hadronic showers.

The hadronic calorimeter layer amounts to a material budget of 6 to 14 λ, depending on $|\eta|$, which contains most high-energy jets completely within the calorimeter.

Calorimeter Resolution

The energy resolution of the ATLAS calorimeters is well described by a noise term N, stochastic term S, and constant term C [129]:

$$\begin{aligned} \frac{\sigma_E}{E} &= \frac{N}{E} \oplus \frac{S}{\sqrt{E}} \oplus C \\ &= \sqrt{\left(\frac{N}{E}\right)^2 + \left(\frac{S}{\sqrt{E}}\right)^2 + C^2}\,. \end{aligned} \tag{4.6}$$

The energy-independent noise term is mainly caused by electronics and pile-up noise. It dominates at energies of a few GeV but becomes negligible at higher energies. The stochastic term is a result of Poisson statistics of the inherent statistical nature of the shower development and the interaction with the calorimeter. The constant term refers to fluctuations that correspond to a constant fraction of the particle energy, e.g. due to energy deposition in dead material, the depth at which the shower starts, and non-uniformities of the response across the calorimeter [126].

The energy resolution in the EM layer was measured in electron test beams to be approximately $\sigma_E/E = 10\%/\sqrt{E\ [\text{GeV}]} \oplus (0.2\%$ to $0.5\%)$, depending on $|\eta|$ [130, 131].

The energy measurement of hadronic showers is inherently associated with larger uncertainties, mainly due to the large effect of statistical fluctuations in the contribution of the different energy deposition mechanisms. The energy resolution for hadrons was measured to be approximately $\sigma_E/E = 50\%/\sqrt{E\ [\text{GeV}]} \oplus 5\%$ in the Tile Barrel [132], $\sigma_E/E = 70\%/\sqrt{E\ [\text{GeV}]} \oplus 6\%$ in the Hadronic End-Caps [133], and $\sigma_E/E =$

$90\%/\sqrt{E\ [\mathrm{GeV}]} \oplus 8\%$ in the Forward Calorimeters [134].

In contrast to momentum measurements from the track curvature, the relative energy resolution of the ATLAS calorimeters improves with increasing energy. This makes them the most precise way to determine the particle energy above a few tens of GeV.

4.2.4 Muon Spectrometer

Muons at energy scales in the order of GeV are minimally ionizing and therefore pass through the ID and the calorimeters without large energy losses. These can be prompt muons produced in the initial collisions or muons from decays in the hadronic shower in the calorimeters. Occasionally, hadrons of high-energy jets can also escape the calorimeters (as so-called *punch-through*).

The momenta of charged particles leaving the calorimeters are measured in the ATLAS Muon Spectrometer by tracking their trajectories. These trajectories are bent by a magnetic field of up to 4 T produced by large air-core Barrel and End-Cap toroid magnets (illustrated in Figure 4.2).

The Muon Spectrometer is designed to precisely measure muons within $|\eta| < 2.7$ and to provide muon trigger information within $|\eta| < 2.4$. It consists of three layers of Monitored Drift Tubes (MDTs) in the Barrel and up to four layers of MDTs in the End-Caps. These offer high-precision tracking with small material budgets to minimize multiple scattering that would degrade the momentum resolution. The innermost tracking layer of the End-Caps is equipped with Cathode-Strip Chambers instead of MDTs close to the beam pipe as they perform better at the high rates of the more forward regions.[7]

The Barrel layers are additionally equipped with Resistive Plate Chambers and the End-Cap layers with Thin-Gap Chambers. These provide very fast response times for the muon triggers and can measure the ϕ coordinate of hits which the MDTs can only resolve poorly.

If hits in the ID and the Muon Spectrometer can be associated with a single muon, a very good momentum resolution of up to

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.02\,\% \cdot p_\mathrm{T}\,[\mathrm{GeV}] \oplus 2\,\% \tag{4.7}$$

is achieved. The momentum resolution degrades accordingly if a track is identified in only one of the two systems.

4.2.5 Trigger and Data Acquisition System

The high LHC bunch crossing frequency of 40 MHz necessitates a trigger system for ATLAS to accommodate limited data readout bandwidth and storage capacity. With a typical event amounting to 1 MB of data, a recorded event rate of 1.2 kHz can be sustained [136]. The trigger system selects in real-time (from here on referred to as

[7] The innermost End-Cap layer has been replaced with the ATLAS New Small Wheel after Run-2 [135]. It features MicroMegas as precision trackers as they provide better performance at the high rates expected in future LHC operations.

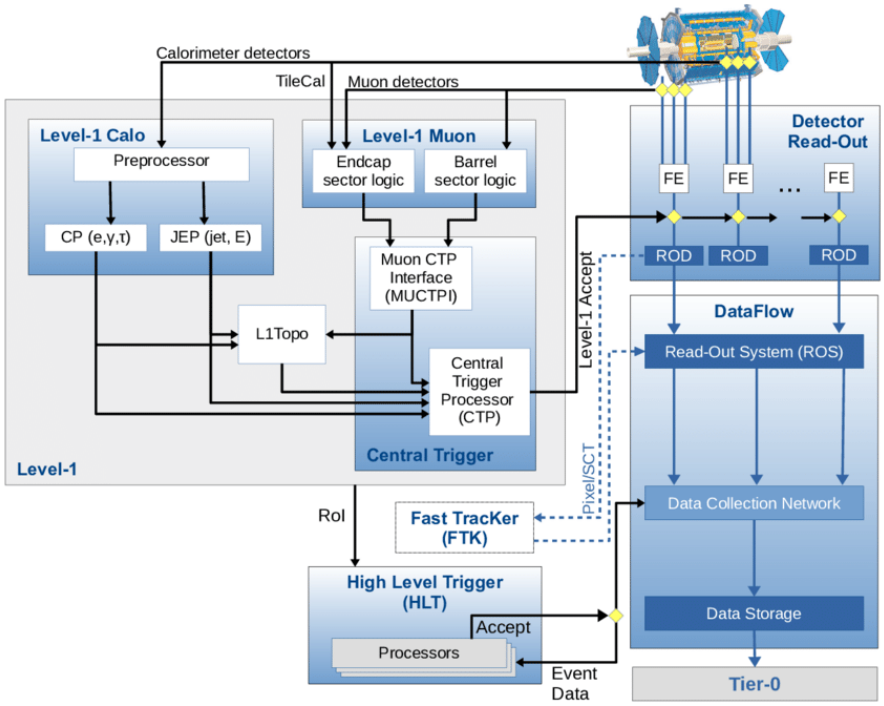


Figure 4.7: Schematic overview of the ATLAS trigger and data acquisition system during Run-2. The trigger chain consisting of Level-1 and HLT can be seen on the left side, and the data acquisition system on the right side. Note that the shown Fast TracKer was not used for data taking. Taken from Ref. [137].

online) the interesting physics events for readout while rejecting events of low momentum transfer that make up the majority of the total pp cross section.

In ATLAS, this is done in two consecutive stages: with the *Level-1* (L1) trigger implemented in custom hardware and the *High-Level Trigger* (HLT) running as software on conventional computing clusters. The interaction of the trigger stages with the data acquisition system is shown in Figure 4.7 and described in the following.

Level-1 Trigger

The L1 trigger is designed to reduce the 40 MHz bunch crossing frequency down to a rate of ≈100 kHz at which a readout of all detector systems is possible. It is implemented in custom, pipelined hardware to arrive at a trigger decision within 2.5 µs after each bunch crossing. The L1 trigger consists of three subsystems identifying potential physics signatures: The *Level-1 Calorimeter Trigger* (L1Calo), the *Level-1 Muon Trigger* (L1Muon) and the *Level-1 Topological Processor* (L1Topo). The final trigger decision is then taken by the *Central Trigger Processor* evaluating the signatures.

L1Calo [138] identifies electron, photon, τ, and jet candidates from calorimeter information. It uses reduced-granularity inputs consisting of 7168 *Trigger Towers*, facilitating fast, parallel processing. These are formed by analogue sums of the calorimeter cells in a 0.1×0.1 region in (η, ϕ) in the Barrel up to 0.4×0.4 in the Forward Calorimeter, reflecting the change in size and geometry of the calorimeter cells. The analogue signals from these Trigger Towers are digitized and calibrated to transverse energy in the *PreProcessor*.

Since the calorimeter response is slower than the bunch crossing spacing, the signals need to be assigned to the correct bunch crossing in the PreProcessor as well.

Particle signatures are reconstructed from the energy deposits using fast sliding window algorithms. The *Cluster Processor* identifies electron/photon candidates (which cannot be distinguished with the available granularity and without track reconstruction) and hadronic τ candidates. It uses the full trigger tower granularity due to the comparably narrow width of electromagnetic showers and distinguishes them from hadronic showers by the amount of energy deposited in either of the two calorimeter layers.

The candidates for the L1 jet triggers employed in this thesis are reconstructed by the *Jet/Energy-sum Processor*. It uses reduced-granularity inputs of 0.2×0.2 inclusive in the calorimeter layers because jets are expected to produce much wider showers and because they do not need to be distinguished from electromagnetic showers of much lower rates. Only the Barrel and End-Cap regions (i.e. $|\eta| < 3.2$) are considered due to the Trigger Towers in the Forward Calorimeter being too large in size and exhibiting large occupancies. A sliding window algorithm identifies jet candidates as local maxima in the energy sum in windows of 0.8×0.8, which is large enough to capture the majority of the energy of most jets. Note, that jet energy reconstructed on L1 is the sum of the calorimeter inputs calibrated for electromagnetic showers. The sizable correction to the hadronic energy scale (see Section 6.1.2) is not applied, such that the true jet energy needs to be approximately 30–50 % higher than L1 jet thresholds.

The Jet/Energy-sum Processor also reconstructs the global sum of energy in the event and the missing transverse energy[8]. For this, the Trigger Towers of the Forward Calorimeter are included as well to improve the missing transverse energy resolution.

L1Muon identifies muon candidates by finding coinciding hits in the muon trigger chambers, estimating the muon momentum from the deviation of these hits from a straight line. While muon triggers are crucial for many gauge boson or heavy flavour analyses, they are not relevant in this thesis. More information is given in Ref. [139].

Interesting physics events are not only defined by the energies and multiplicities of detected particle candidates but can also be identified by their geometric relation to each other. The L1Topo system is designed to determine higher-level variables like angular separation or invariant masses from the L1Calo and L1Muon candidates to select certain event topologies. One such topology requires dijet events with an upper threshold on the rapidity difference. The corresponding dijet trigger `L1_J50_DETA20-J50J` is used in this thesis (see Sections 5.2 and 7.2) [140]. L1Topo was commissioned during Run-2 and began operation in 2017.

The objects reconstructed by L1Calo and L1Muon, as well as the topologies identified by L1Topo, are evaluated in the Central Trigger Processor. It is configured with a set of trigger items, each containing requirements on energy thresholds or topological requirements. If the conditions of a specific trigger item are met, the event is accepted for readout and the positions of the identified object candidates are forwarded as Regions-of-Interest to the HLT as seeds for its more sophisticated reconstruction algorithms.

Most trigger items are assigned a *prescale* factor to reconcile the demand for low-

[8] For negligible particle mass, the missing transverse energy approaches the missing transverse momentum. The latter is defined as the negative sum of all momenta in the transverse plane. Any deviation from 0 (barring resolution effects) indicates a particle leaving the detector unseen.

threshold triggers of many analyses with the exponentially rising event rate as the final state energy decreases. A prescale N is defined such that only 1 in N events for which the trigger requirements are fulfilled is accepted. This allows providing triggers for objects in a wide energy spectrum without exceeding the Level-1 rate limitation.

High-Level Trigger

At the second, software-based trigger stage, the entire detector can be read out, allowing for object reconstruction using the full detector granularity. Following first, fast rejection algorithms that reduce the number of events for most trigger items, more compute-intensive reconstruction algorithms are run. These are typically similar to the sophisticated algorithms employed for offline data analyses[9] so that a high correlation between events selected in offline analyses and those accepted by the trigger is achieved [141].

Depending on the trigger item, different detector subsystem data are read out and processed. Track reconstruction, for example, is computationally intensive and not feasible to run at the total Level-1-Accept rate. Thus, only electron/photon-, muon-, τ- and b-jet triggers employ tracking information as they run at a comparably lower rate and benefit strongly from tracking. Standard jet triggers, on the other hand, amount to a large fraction of the Level-1-Accept rate and rely on calorimeter information only. The jet reconstruction algorithm is described in detail in Section 5.2. It offers close-to-offline performance even without track reconstruction.

As on L1, low-threshold HLT items are assigned a prescale to ensure that their rates remain within the bandwidth limitations. With the prescales applied, all trigger items amount to a final rate of $\approx$1.2 kHz at which events are accepted. Their event data, including the online reconstructed objects, are transmitted to the Tier-0 centre, where the raw data are catalogued, permanently stored to tape and distributed to Tier-1 centres for further processing.

[9] The term offline refers to the conventional strategy of ATLAS analyses to reconstruct objects at an arbitrary point in time from the recorded full detector readout.

5 The Trigger-Level Approach

Dijet resonance searches – like any other resonance search – face a central limitation due to the trigger and data acquisition chain: The prescales applied to low-threshold triggers both on Level-1 and in the HLT constrain the range of the invariant mass spectrum that can be studied.

Figure 5.1 shows the number of offline recorded dijet events as a function of their invariant dijet mass m_{jj} in red. The spectrum is only recorded at full statistical power above approximately $m_{jj} > 1\,\text{TeV}$ due to the lowest unprescaled HLT jet trigger during Run-2 requiring $p_\text{T} > 420\,\text{GeV}$. For lower masses, the event rate is reduced by increasingly large prescales. While changing trigger prescales show up as rather steep turn-on curves in jet p_T, the effect is washed out in m_{jj} due to the η-dependent relation between m_{jj} and p_T being averaged over (see Section 7.2). This leads to the visible wave-like features in the spectrum with each peak corresponding to a new prescaled, lower-threshold trigger.

The trigger-level analysis (TLA) presented in this thesis measures the dijet mass spectrum well below 1 TeV. It utilizes the fact that full detector information is not obligatory for a dijet resonance search. To derive the invariant mass of dijet events, only jet four-momenta are necessary – which are already reconstructed on the HLT with close-to-offline precision. Recording these *trigger-level* jets allows performing a dijet search that bypasses the HLT prescales.

The blue points in Figure 5.1 show the observed dijet mass spectrum using trigger-level jets. It is now limited only by the significantly lower L1 trigger thresholds, allowing the TLA to measure the unbiased dijet mass spectrum down to below 500 GeV.

This chapter first discusses the special event readout of trigger-level objects for the TLA, followed by a description of the HLT jet reconstruction in comparison to the offline procedure.

5.1 Event Readout

In order to record events at the full L1 rate to permanent storage, the information per event needs to be kept at a minimum. The readout is performed via the Data Scouting stream [142], introduced in Run-2 specifically for a dijet TLA.[1]

For each event, the event header (which includes global event information, error flags, and trigger decisions) and the up to 20 leading HLT jets with a transverse energy $E_\text{T} > 20\,\text{GeV}$ are recorded. For each of them, the four-momentum, the jet timing, quality variables, and structure variables like the energy by calorimeter layer or jet

[1] The success of the trigger-level dijet resonance search in Run-2 [1] as well as lessons learned from it lead to the extension of the Data Scouting stream to additional signatures for Run-3. A trigger for photon+dijet events was added and the online jet reconstruction was improved by adding tracking. Trigger items with enabled b-tagging were added to the Data Scouting stream as well to allow for trigger-level di-b-jet resonance searches [143].

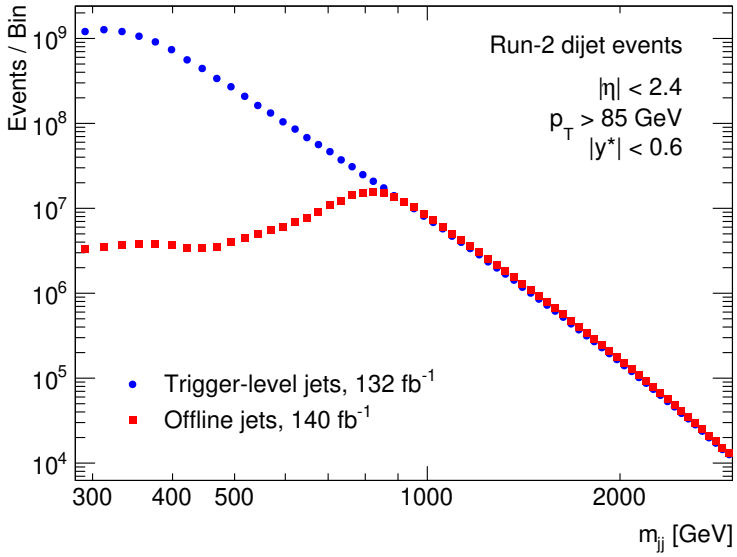


Figure 5.1: Number of dijet events recorded by the Data Scouting stream used in this thesis (blue points) in comparison to offline events recorded by any single-jet trigger (red squares), binned in the invariant dijet mass. The kinematic event selection criteria used throughout this thesis are applied: the two jets leading in p_{T} within $|\eta| < 2.4$ are selected and required to fulfil $p_{\mathrm{T}} > 85\,\mathrm{GeV}$ and $|y^*| < 0.6$, where $y^* = (y_1 - y_2)/2$ measures the difference in rapidity between the leading jets. The offline data set corresponds to the full Run-2 luminosity [118], while a slightly reduced luminosity is available for the trigger-level analysis.

width variables are saved. The latter are relevant for the jet calibration discussed in Section 6.1.3.

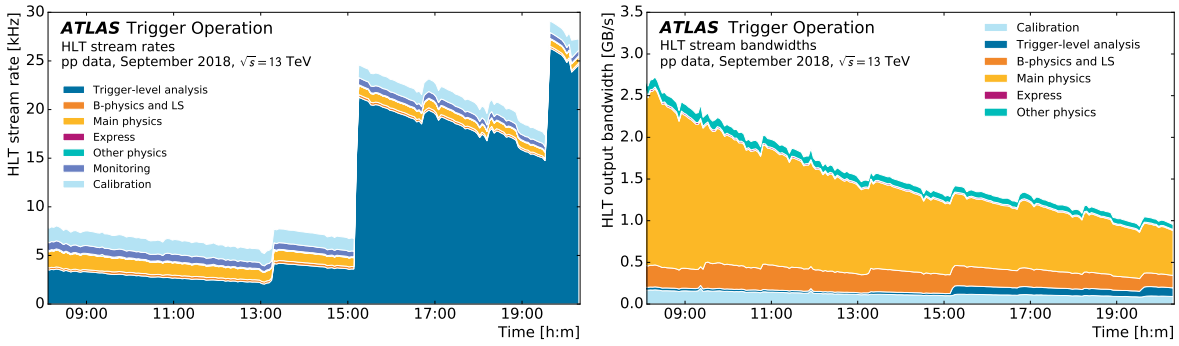
With this readout content, the average Data Scouting event amounts to 6.5 kB in size, compared to a standard event for offline analysis in the order of 1 MB [136]. This way, the Data Scouting stream consumes only a negligible part of the total ATLAS bandwidth (shown as the dark blue contribution in Figure 5.2b) while recording at the full seeding L1 rate and dominating the readout event rate (shown as the dark blue contribution in Figure 5.2a).

Jumps in the rates occur when additional triggers are activated or prescales are relaxed. During an LHC fill, the instantaneous luminosity slowly declines due to the depletion of protons per bunch and beam degradation [102]. Since trigger rates scale at least linearly with the instantaneous luminosity,[2] the bandwidth consumed by a given trigger configuration slowly declines during a fill as well. To make use of the freed bandwidth, prescales are relaxed or additional, lower-threshold triggers are activated during a run.

Several L1 jet triggers seed the Data Scouting stream. `L1_J100` requiring $E_{\mathrm{T}} > 100\,\mathrm{GeV}$ reconstructed on L1 is the main trigger used in this thesis as it ran unprescaled during most of Run-2, recording $132\,\mathrm{fb}^{-1}$ of data.

`L1_J50` was added to the Data Scouting stream in 2017 and activated at the end of

[2] L1 single object trigger rates scale approximately linearly with the instantaneous luminosity. However, triggers on missing transverse energy or multiple objects, for example, are affected stronger by pile-up such that their false-positive rate increases [136, 144]



(a) **(b)**

Figure 5.2: (**a**) Example of the large event rate recorded by the Data Scouting stream for the TLA in contrast to (**b**) the marginal bandwidth occupied by this stream in a run in September 2018. The jumps in both figures correspond to times when triggers are enabled or trigger prescales relaxed to utilize the freed bandwidth when the instantaneous luminosity declines during an LHC fill. Taken from Ref. [136].

LHC fills when bandwidth became available. Taking its prescale into account, it was active for an integrated luminosity of $5.6\,\text{fb}^{-1}$.

The L1Topo dijet trigger `L1_J50_DETA20-J50J` (from here on referred to as J50Topo) selects events with a leading L1 jet with $E_\text{T} > 50\,\text{GeV}$, a subleading jet with $E_\text{T} > 15\,\text{GeV}$ and a rapidity difference $|y^*| = \left|(y_1 - y_2)/2\right| < 1.0$ between them. It became active in 2018 and was running at the end of fills as well. Only data taken after Technical Stop 1 (TS1) [145] are considered for this thesis due to an initial misconfiguration of the trigger [146]. The data set corresponds to an integrated luminosity of $10.1\,\text{fb}^{-1}$ with a small overlap with the `L1_J50` data set.

The Data Scouting stream contains additional, lower-threshold triggers that were mostly active during special run conditions and are not considered in this thesis. A detailed overview is given in Appendix A.

5.2 Jet Reconstruction

The hadrons emerging from a quark or gluon in the final state deposit their energy in the EM and the hadronic calorimeters via hadronic showers. The charged jet constituents also leave a track in the inner detector and some particles of the shower or muons from heavy-flavour decays can escape into the muon spectrometers [147].

As discussed in Section 2.1.2, these hadronization products need to be clustered into a jet in order to define infrared-safe and collinear-safe observables that can be linked to the original partons produced in the hard interaction.

The *EMTopo* jets employed in this thesis are reconstructed primarily from the calori-

meter information.[3] This occurs in two steps: First, the energy deposits in the calorimeter are clustered into topological clusters (*topo-clusters*) corresponding approximately to single hadron showers or parts of these. Then, these topo-clusters are merged into jets via a jet clustering algorithm.

Topological Clusters

Calorimeter cells are clustered into topo-clusters with the aim to identify showers of individual particles among pile-up and other noise sources. A correspondence of one topo-cluster to a single particle can typically be achieved for compact electromagnetic showers. Hadronic showers with their large intrinsic fluctuations and long interaction lengths, on the other hand, often split up into multiple sub-showers and hence multiple topo-clusters, depending on the incoming particle type and energy and the calorimeter region [149].

The topo-clustering algorithm identifies three-dimensional, contiguous regions of calorimeter cells above a given noise threshold. It relies on the cell significances, which are defined as the ratio of the measured energy deposited in a cell over its expected noise level:

$$\varsigma_{\text{cell}} = \frac{E_{\text{cell}}}{\sigma_{\text{cell,noise}}} \,. \tag{5.1}$$

Out of all calorimeter cells, those with a significance $|\varsigma_{\text{cell}}| > 4$ are selected as seeds of topo-clusters. Next, adjacent cells with $|\varsigma_{\text{cell}}| > 2$ are added to the clusters iteratively until no more adjacent cells fulfil this criterion. During this growth stage, clusters are merged if they share cells. Finally, a last set of neighbours with a significance $|\varsigma_{\text{cell}}| > 0$ is added. This algorithm ensures keeping the fringe of the showers with deposited cell energies close to the noise level contained in the topo-clusters while retaining the noise suppression of the algorithm.

The expected noise level per cell $\sigma_{\text{cell,noise}}$ is dominated by pile-up noise in most calorimeter regions for Run-2 conditions. Thus, it must be tuned for a certain expected average number of interactions $\langle\mu\rangle$ to ensure optimal noise suppression.

For offline topo-cluster reconstruction, this noise level was retroactively set to an optimal value corresponding to $\langle\mu\rangle = 40$ throughout Run-2. It also takes out-of-time pile-up effects arising from varying LHC filling schemes into account. The online reconstructed jets available in the Data Scouting stream, however, can not be retroactively reprocessed as the underlying topo-cluster and calorimeter cell information are lost. They reflect the cell noise levels set in the HLT during the corresponding run time. This leads to a source of discrepancy of online jets with regard to offline: Out-of-time pile-up effects were only taken into account from 2016 onwards and the noise levels were optimised for $\langle\mu\rangle = 30$ in 2015 and 2016. In 2015, with the most substantial mismatch in the algorithms, the energy of online and offline reconstructed topo-clusters agreed within approximately 2%

[3] EMTopo jet reconstruction was the default in ATLAS throughout most of Run-2. More recently, *particle flow* jets are preferred in offline analyses due to their improved resolution and pile-up resilience [148]. They do not rely only on calorimeter deposits but also take tracks of the charged constituents into account. In Run-3, simplified track reconstruction is performed for events selected by the L1 jet triggers so that particle flow jets can be reconstructed on the HLT and be used in future dijet TLAs [143].

[137].

For EMTopo jets, both online and offline, the topo-clusters remain at the electromagnetic energy scale to which the calorimeter cells are calibrated. A correction for the different hadronic response only happens on the fully reconstructed jets (detailed in the following chapter).

Offline topo-clusters additionally receive an *origin correction*: Due to the finite length of the colliding proton bunches, not all jets originate at $z = 0$ but in a region of $\approx 20\,\text{cm}$ around it. Offline jets are assumed to originate from the reconstructed primary vertex with the largest sum of associated track momenta. All topo-clusters in an event receive a small correction to their four-momenta to reflect that origin [126]. Since no track and vertex reconstruction is performed for the Data Scouting stream, no origin correction can be applied to online topo-clusters.

Jet Clustering

In the ATLAS experiment, the anti-k_t algorithm [150] is used to reconstruct jets from the topo-clusters. In comparison to other popular choices like the Cambridge-Aachen [151] or k_t algorithm [152], it has the benefit of reconstructing jets with a relatively constant circular shape in the (η, ϕ) plane and of being less susceptible to pile-up [153].

The anti-k_t algorithm sequentially merges topo-clusters into jets using two distance variables:

$$d_{ij} = \min\left(p_{\mathrm{T},i}^{-2}, p_{\mathrm{T},j}^{-2}\right) \frac{\Delta R_{ij}^2}{R^2}\,, \tag{5.2}$$

$$d_{iB} = p_{\mathrm{T},i}^{-2}\,. \tag{5.3}$$

Here, d_{ij} is considered to be the distance between two input 'particles' i and j given by their distance ΔR_{ij} in the (η, ϕ) plane, normalized by a configurable radius parameter R and weighted by their momenta. The variable d_{iB} reflects the distance of particle i to the beam in momentum space. The algorithm is initiated with all topo-clusters as input particles and the minimum of all distances $\left\{d_{ij}, d_{iB}\right\}$ is determined. If it is a distance d_{ij}, particles i and j are merged into a new particle by addition of their four-momenta, all distances are recalculated, and the process is started over. If a distance d_{iB} is minimal, particle i is considered to be a fully reconstructed jet and is removed from the set.

This algorithm terminates only when all topo-clusters in an event are clustered into any jet. The exponent -2 ensures that this happens in descending order in p_{T}, leading to jets roughly circular in shape around their hardest constituents. In the ATLAS experiment, the radius parameter $R = 0.4$ is chosen for online and offline standard jet reconstruction, offering a good compromise between capturing the majority of the jet constituents and avoiding contributions from pile-up or the underlying event [154]. The missed contribution due to true jet constituents not being clustered to the jet is referred to as *out-of-cone* effect and, on average, corrected for in the jet calibration in Section 6.1.2.

Up to this point in the reconstruction chain, Run-2 online and offline EMTopo jets are almost identical, with the only differences being the topo-cluster noise thresholds before 2017 and the origin correction for offline topo-clusters. More substantial differences arise

in the subsequent jet calibration and correction steps discussed in the following chapter.

6 Jet Calibration

Jets are constructed as the sum of four-momenta of topo-clusters and as such they inherit the electromagnetic scale to which the calorimeters and, hence, the topo-clusters are calibrated. However, hadrons – and consequently jets – exhibit a lower response in the non-compensating ATLAS calorimeters. To match their particle-level energy, jets are calibrated to the *jet energy scale* (JES) in a series of corrections.

The JES calibration and its associated uncertainties are derived centrally in ATLAS for offline jets in an extensive procedure [126]. A coarse approximation of the offline procedure is applied online in the HLT considering only calorimeter information [137], which improves the correlation between offline jets used in analyses and online jets on which the trigger decision is based. Still, the jet calibration applied online only reflects the best knowledge available at the time of data-taking while refined procedures and calibration factors have been derived since [126, 155].

Thus, a custom trigger-level jet calibration is derived for this analysis that is as close as possible to the offline procedure, given the limited information available in the Data Scouting stream. This results in a better resolution of the dijet invariant mass, improving the ability to resolve narrow resonances in the spectrum and, thus, the sensitivity of this search.

The TLA calibration chain is schematically shown in Figure 6.1 and discussed in the following sections. It starts with a set of purely simulation-based calibration steps that are applied to correct the reconstructed jet energy in simulation towards the particle level. Afterwards, data-driven calibration steps are applied to align the response of jets in data with the response in simulation. These steps are adopted from the offline calibration, either following almost exactly the same procedure (green boxes) or being modified due to missing information in the Data Scouting stream (blue boxes). Finally, a custom online-to-offline correction step (yellow box) is applied to the trigger-level jets in data to correct for residual differences of their mean response with respect to offline jets. While the calibration steps are derived in terms of corrections to p_{T} or E, the resulting factors are used to scale the jet four-momenta.

An essential requirement for the jet calibration in a dijet resonance search is that it is a smooth function of the dijet invariant mass, which is fulfilled if it is a smooth function of p_{T} for individual jets. Since the QCD background is ultimately estimated by a fit of a smooth function or smooth templates to the data, any bumps, dips, or kinks in the calibration factors would propagate these features into the calibrated dijet mass spectrum and potentially show up as a false signal. With the very high statistical precision achieved by this analysis, already 10^{-4}-level features in the calibration could induce significant features in the dijet mass spectrum. Thus, all calibration steps are required to fulfil smoothness criteria in data and MC. These are discussed towards the end of this chapter.

As the derivation of each calibration step relies on MC or the offline data set, which

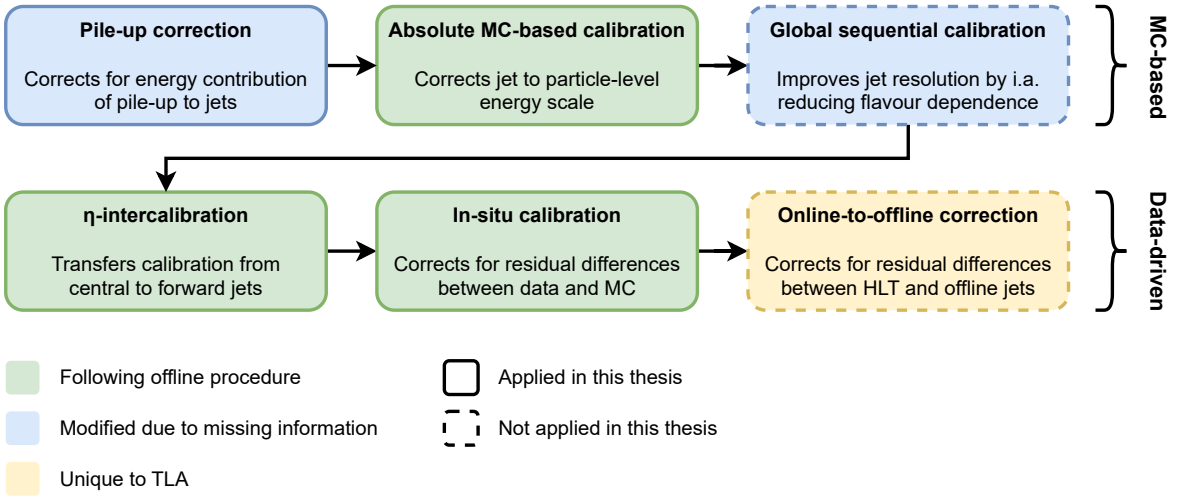


Figure 6.1: Calibration chain of the HLT jets used in this thesis. Steps directly adopted from the offline calibration or closely following its procedure are marked in green, steps with modifications from the offline procedure in blue, and steps without a counterpart in the offline chain in yellow. Dashed outlines indicate steps that are not applied in this thesis due to not fulfilling the required smoothness criteria discussed in Section 6.3.

both exhibit larger statistical uncertainties than the trigger-level data set to which the calibration will be applied, a sufficient smoothing of the calibration factors in each step is crucial.

6.1 Simulation-Based Calibration

All calibration steps described in the following are derived from the inclusive dijet Pythia MC samples described in Section 3.2. Particle-level jets, reconstructed from stable final state particles[1] with the anti-k_t algorithm with radius parameter $R = 0.4$, serve as the reference for the calibration.

6.1.1 Pile-up Correction

Pile-up collisions are dominated by low-energy QCD jet production due to its large cross section, leading to a large number of low-energy hadrons in the detector overlayed with the hard dijet event of interest. This results in an underlying energy density in each event that distorts the energy of reconstructed topo-clusters and jets.

This is accounted for in an area-based pile-up correction step. Each jet is assigned an area A by adding infinitely low-momentum *ghost* particles distributed in the (η, ϕ) plane to the anti-k_t algorithm [156]. The jet area is then determined from the spread of the ghost particles clustered into the jet. The average transverse momentum density ρ of the event is determined from the median of the transverse momenta and areas of all jets within $|\eta| < 2$. The jet p_{T} is corrected by $\Delta p_{\mathrm{T}} = -A \cdot \rho$ already online in the HLT

[1] All stable final state particles except neutrinos and muons are considered, since those deposit no or only a small amount of energy in the calorimeters.

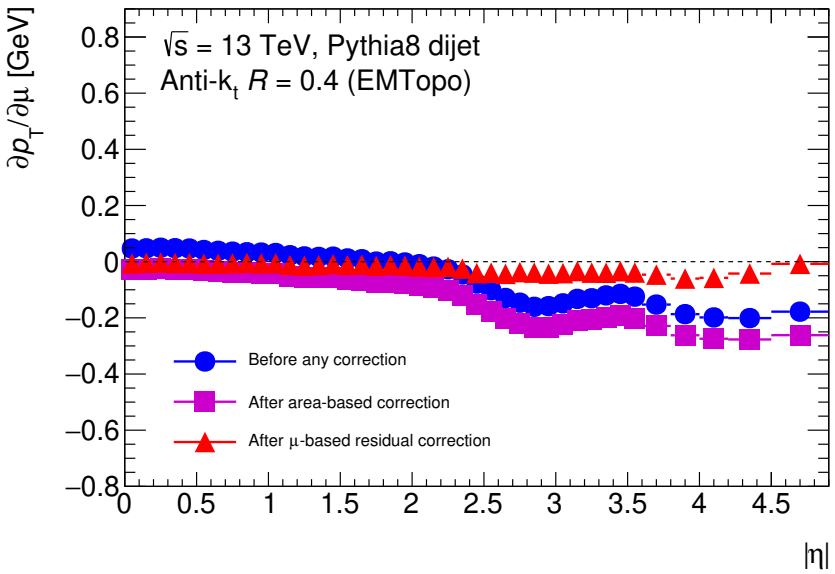


Figure 6.2: Dependence of the reconstructed trigger-level jet p_{T} on the average number of pile-up interactions μ at the individual steps of the pile-up correction in bins of $|\eta|$. Adapted from Ref. [157].

and recorded to the Data Scouting stream afterwards. The online jets with area-based pile-up correction applied are the baseline upon which the custom jet calibration in the TLA is built.

Some dependence on the pile-up activity remains after this first step, especially for large $|\eta|$ where the energy density is lower. An additional residual pile-up correction is therefore derived in MC: Reconstruction-level jets are geometrically matched to particle-level jets within $\Delta R_{\mathrm{match}} < 0.3$ and are required to be isolated from other reconstruction-level (particle-level) jets with $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ by $\Delta R_{\mathrm{iso,reco}} > 0.6$ ($\Delta R_{\mathrm{iso,truth}} > 1$). The isolation ensures that no nearby jets distort the reconstructed four-momenta on either reconstruction or particle level.

The deviation $p_{\mathrm{T}}^{\mathrm{residual}} = p_{\mathrm{T}}^{\mathrm{area\text{-}corrected}} - p_{\mathrm{T}}^{\mathrm{truth}}$ is determined in bins of $|\eta|$, $p_{\mathrm{T}}^{\mathrm{truth}}$, and the average number of interactions per bunch crossing μ. The deviation is approximately linear in μ per bin such that a linear coefficient $\beta\left(p_{\mathrm{T}}^{\mathrm{truth}}, |\eta|\right)$ can be determined from a fit. For constant $|\eta|$, β is then found to have a weak, approximately logarithmic dependence on $p_{\mathrm{T}}^{\mathrm{truth}}$. Again, a fit is performed and β is evaluated at $p_{\mathrm{T}}^{\mathrm{truth}} = 25\,\mathrm{GeV}$ as that is the scale at which pile-up effects are most dominant.

In total, the pile-up correction thus takes the form:

$$p_{\mathrm{T}}^{\mathrm{PU\text{-}corrected}} = p_{\mathrm{T}}^{\mathrm{reco}} - A \cdot \rho - \beta \cdot \mu\,. \tag{6.1}$$

The offline calibration applies an additional residual correction proportional to the number of primary vertices as a proxy for the actual number of pile-up interactions in a given event. Since tracks and primary vertices are not reconstructed for the Data Scouting stream, that correction cannot be applied in this analysis. The effect of this missing correction is found to be small for the relatively high jet energies relevant in this analysis.

Figure 6.2 shows the dependence of the reconstructed trigger-level jet p_{T} on μ at the individual steps of the pile-up correction. It can be seen that before any correction (blue points) a significant correlation exists that strongly depends on $|\eta|$. This correlation can be negative because it is dominated by the out-of-time pile-up effect of collisions in previous bunch crossings and because the liquid argon calorimeter pulse shape becomes negative after a signal peak. The area-based correction (violet squares) results in a flat shift independent of $|\eta|$ as the same energy density ρ is corrected for in the whole event. This reduces the dependence for central jets but overcorrects for jets outside the central region where the occupancy is lower. After the residual pile-up correction (red triangles), the jet p_{T} is mostly independent of the pile-up. A small residual dependence remains due to the linear approximation of the correction and the evaluation at a fixed $p_{\mathrm{T}}^{\mathrm{truth}}$.

6.1.2 Absolute Monte Carlo-Based Calibration

In the non-compensating ATLAS calorimeter, hadronic showers exhibit a different response than electromagnetic showers as detailed in Section 4.2.3. The effects of the hadronic response, energy losses due to uninstrumented regions of the detector, punch-through, and out-of-cone effects are corrected for with the absolute MC-based jet energy scale calibration (MCJES). It transfers the jets from the electromagnetic scale of the topo-clusters to the jet energy scale that matches the energy on particle level and is with 20–50 % by far the largest correction applied in the jet calibration chain.

The same matching of reconstruction-level to particle-level jets and isolation criteria as in the residual pile-up correction are applied. In bins of E^{truth} and η^{reco}, the average jet energy response $\mathcal{R}$ is defined as the mean of a Gaussian fit to the core of the $E^{\mathrm{reco}}/E^{\mathrm{truth}}$ distribution. $\mathcal{R}$ is derived in bins of E^{truth} since it does not behave Gaussian if derived in bins of E^{reco}.[2]

The determined response as a function of η is shown in Figure 6.3 for different values (bin centres) of E^{truth}. The different calorimeter regions are visible in the η dependence – most notably the End-Cap-Forward transition around $|\eta| \approx 3.1$, the Barrel-End-Cap transition around $|\eta| \approx 1.3$ where the amount of inactive material is larger (see Figure 4.4), and the transition from Outer Wheel to Inner Wheel within the End-Caps around $|\eta| \approx 2.5$ where the calorimeter granularity changes. At the same time, the response is rising with E^{truth} due to the increasing electromagnetic and decreasing invisible component of hadronic showers with increasing energy as shown in Figure 4.6.

For a given η bin, $\mathcal{R}$ is well described by a polynomial in $\log(E^{\mathrm{truth}})$:

$$\mathcal{R}(E^{\mathrm{truth}}) = \sum_{i=0}^{N_{\mathrm{max}}} a_i \log^i(E^{\mathrm{truth}}) \,, \tag{6.2}$$

where a_i are the free parameters and N_{max} is chosen between 1 and 6 based on a χ^2 goodness-of-fit test.

[2] For any calibration step, it is beneficial to find an (approximately) independent variable to derive the response in. Assuming that E_{reco} is Gaussian distributed around E_{truth}, then $E_{\mathrm{reco}}/E_{\mathrm{truth}}$ will be a Gaussian distribution if considered as a function of the independent variable E_{truth}, whereas it will follow a reciprocal normal distribution [158] if considered as a function of the dependent variable E_{reco}.

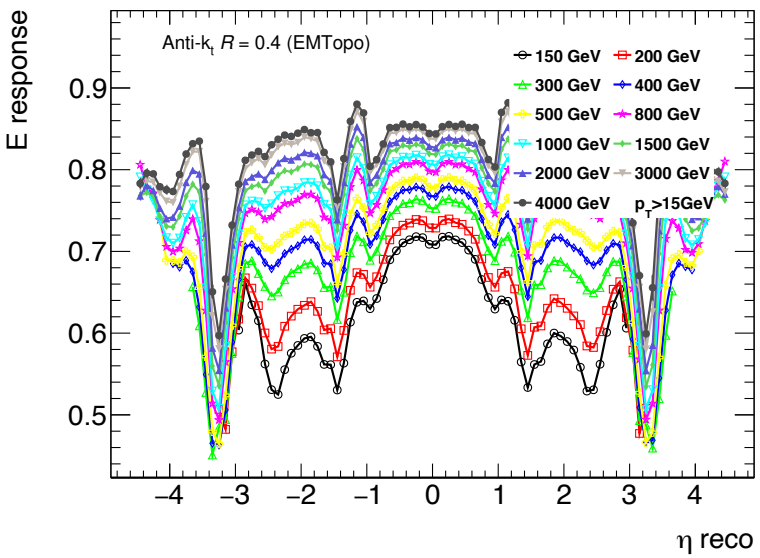


Figure 6.3: Jet energy response as a function of η for different values (bin centres) of E^{truth}. The different calorimeter regions result in the strong η dependence, most notably the End–Cap-Forward transition around $|\eta| \approx 3.1$, the Inner Wheel–Outer Wheel transition around $|\eta| \approx 2.5$, and the Barrel–End-Cap transition around $|\eta| \approx 1.3$. Adapted from Ref. [157].

To correct reconstruction-level jets by $\mathcal{R}$, its dependence on E^{truth} is translated to a dependence on E^{reco} using a numerical inversion procedure detailed in Ref. [159]. Its validity is verified in a second iteration of the response determination where a closure better than 2 ‰ for jets with $E^{\text{truth}} > 40\,\text{GeV}$ is observed. Lower-energy jets will be rejected by the $p_{\text{T}} > 85\,\text{GeV}$ requirement placed in the event selection for this thesis (detailed in Section 7.2).

The absolute MC-based calibration is the only step in which the jet four-momentum is not only linearly scaled. Due to the strong dependence of the jet response on η, jets close to detector transition regions have a bias of their measured direction towards the detector region with the higher response. This bias is determined analogously to the jet energy response in dependence of η^{reco} and E^{truth}. It is found to be small ($\Delta\eta < 0.01$ in most detector regions) and corrected for.

The procedure described here for the absolute MC-based calibration of trigger-level jets is identical to that applied to offline jets.

6.1.3 Global Sequential Calibration

Up to this point, the *average* jet response has been calibrated to the particle level in dependence of pile-up conditions, the jet energy, and η. However, within a given (p_{T}, η) bin, the response can vary jet by jet, depending on the particle content of the hadronic shower, its longitudinal and transverse shape, and fluctuations in the interaction with the calorimeter. These properties correlate with the type of parton that the jet originates from. Due to their different colour factors, quark-initiated jets often contain less, higher-energy constituents that penetrate further into the calorimeter while gluon-initiated jets

tend to have more, softer constituents – which have a lower calorimeter response and a wider transverse profile.

The global sequential calibration (GSC) aims to improve the jet energy resolution by correcting for the response dependence on certain jet structure variables – without affecting the average jet response that was calibrated in the previous step already.

Analogously to the other simulation-based calibration steps, reconstruction-level and particle-level jets are geometrically matched and the response $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}}$ is derived in bins of $p_{\mathrm{T}}^{\mathrm{truth}}$ and $|\eta|$ as a function of the jet structure variables under study.

For offline EMTopo jets, five variables are identified to have a significant impact on the resolution and are corrected for:

- f_{Tile0} , the fraction of energy measured in the first layer of the Tile calorimeter ($|\eta| < 1.7$)
- f_{EM3} , the fraction of energy measured in the last layer of the EM liquid argon calorimeter
- n_{trk} , the number of tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ that are ghost-associated to the jet
- w_{trk} , the track width: the average p_{T}-weighted distance in (η, ϕ) of all tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ that are ghost-associated to the jet
- n_{segments} , the number of reconstructed muon segments ghost-associated to the jet

Especially the track-related variables give a strong handle for offline jets to correct for the response difference between quark and gluon jets. Since tracking and muon spectrometer information are not available for trigger-level jets, calorimeter-based proxies need to be identified. In a study of different variables that can be reconstructed in the Data Scouting stream, two variables in addition to f_{Tile0} and f_{EM3} were identified to have a significant impact on the jet energy resolution:

- N_{90Const} , the minimal number of topo-clusters that make up 90 % of the jet energy
- f_{TileGap3} , the fraction of energy measured in the scintillators in the Tile gap region ($1.0 < |\eta| < 1.6$)

Since a large multiplicity of jet constituents tends to generate a large number of topo-clusters, the N_{90Const} variable exhibits a relatively strong correlation with n_{trk} and is accordingly useful to correct for quark-gluon differences. f_{TileGap3} samples the energy deposited in the Tile gap region where a relatively large amount of inactive detector material is located (see Figure 4.4).

The dependence of the trigger-level jet response on the four chosen variables f_{Tile0}, f_{EM3}, N_{90Const}, and f_{TileGap3} is determined and sequentially corrected for in this order. Sequential in this context means, that when determining the response of any of the four variables, the dependence on the previous variables has already been corrected for. This sequential approach takes the correlation between the variables into account.

The average response in bins of $p_{\mathrm{T}}^{\mathrm{truth}}$, $|\eta|$ and the jet structure variable in question is again determined by a fit to the Gaussian core of the distribution. The average response

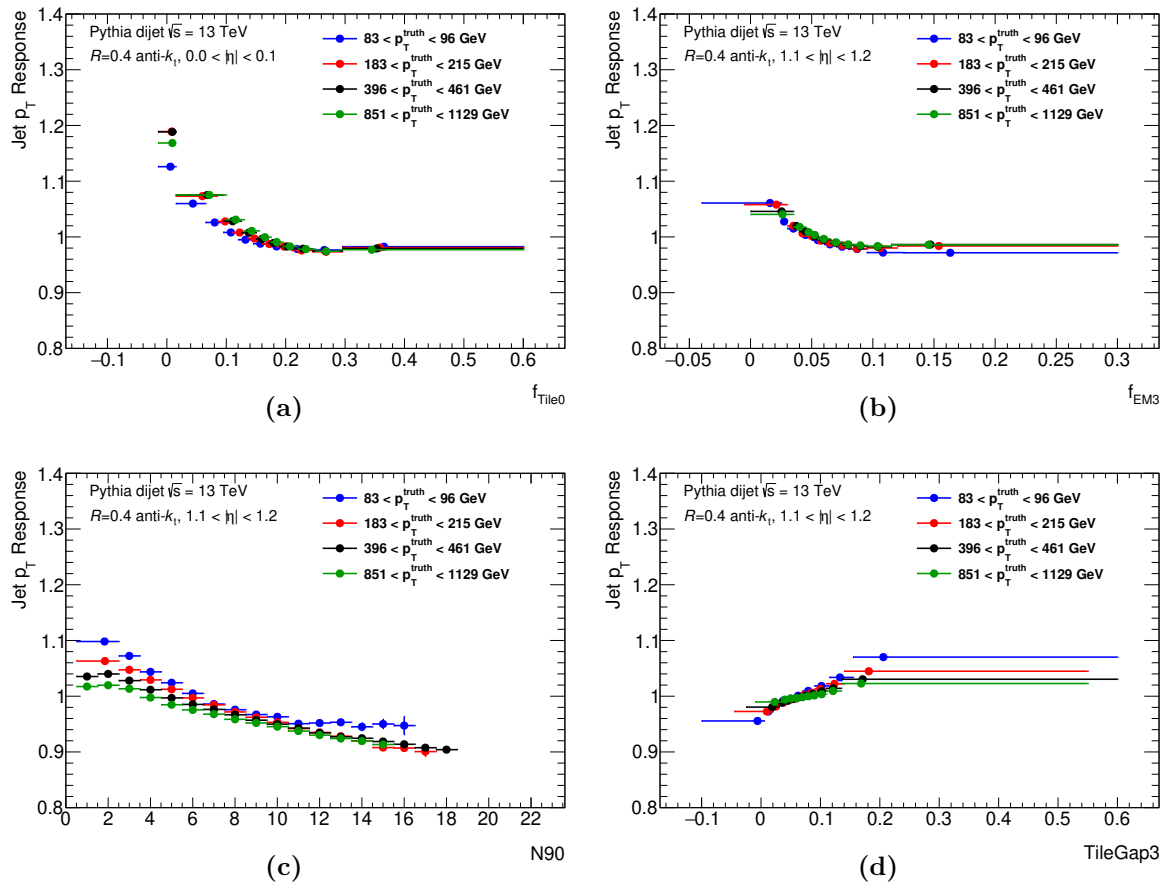


Figure 6.4: Average jet response as a function of the different GSC variables in representative $|\eta|$ bins for different p_T bins: (**a**) f_{Tile0}, (**b**) f_{EM3}, (**c**) $N_{90\text{Const}}$, and (**d**) f_{TileGap3}. The binning in p_{T} and the GSC variable is chosen to have sufficient MC statistics for the response fits. In each panel, the dependence on the previous variables has already been corrected for. Adapted from Ref. [157].

as a function of the four variables is shown in Figure 6.4 for a representative $|\eta|$ bin each. Multiple p_{T} bins are overlayed, illustrating that the p_{T} dependence of the response is low in most variables and $|\eta|$ bins.

At each step of the GSC, the jet response as a function of the jet structure variable is normalized to the response inclusive in this variable so that the average jet response in p_{T} and η is not altered by the GSC. Then, for any fixed $|\eta|$ bin, a two-dimensional Gaussian kernel smoothing in p_{T} and the structure variable is applied to remove fluctuations from the MC samples and the Gaussian response fits. The jet response is transferred from a function of $p_{\text{T}}^{\text{truth}}$ to a function of $p_{\text{T}}^{\text{reco}}$ with the same numerical inversion technique as used for the absolute MC-based calibration.

Applying the GSC to the same MC sample used to derive it validates that the residual dependence of the responses on the jet structure variables remains below 1 %, where the smoothing step prevents perfect closure. The impact of the GSC on the jet energy resolution $\sigma_{\mathcal{R}}$ can be determined from the standard deviation of the Gaussian fits to the jet response. The resolution at each of the steps is shown in Figure 6.5 for one

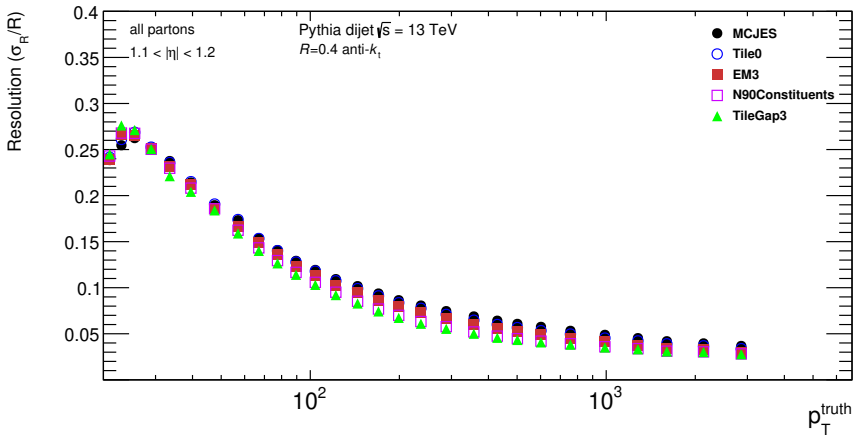


Figure 6.5: Relative jet resolution $\sigma_{\mathcal{R}}/\mathcal{R}$ as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ at the individual steps of the GSC for an exemplary $|\eta|$ bin where all GSC steps are applicable. Adapted from Ref. [157].

exemplary $|\eta|$ bin in the Tile gap region. Over the most relevant p_{T} range of 100–1500 GeV, the resolution before the application of the GSC shown in black is found to decrease from approximately 12 % at low p_{T} to 4 % at high p_{T}. Mainly due to the calorimeter granularity and the discrete nature of $N_{90\mathrm{Const}}$, the latter does not match the performance of the track-based variables in the offline GSC (compare Ref. [126]). Nevertheless, a sizable improvement of the jet energy resolution is achieved. Averaged over η, the calorimeter-based GSC improves the relative jet energy resolution by approximately 1–3 percentage points. This improvement appears much smaller than the up to 20 % dependence of the jet energy response on the individual jet structure variables shown in Figure 6.4. This seeming discrepancy is due to the majority of jets being distributed in a small range of each variable for any given p_{T} and η while the extreme ends of the distributions correspond to a small fraction of jets. For low-energy jets with $p_{\mathrm{T}} < 60$ GeV, a degradation of the jet energy resolution at individual GSC steps is observed. This is, however, well below the minimum jet p_{T} requirement of 85 GeV applied later on in this analysis.

6.2 Data-Driven Calibration

At this stage of the calibration chain, jets in MC are ensured to be calibrated to particle level. However, additional calibration steps are needed to account for differences between the jet response in data and MC. These differences can arise from an imperfect simulation of the detector or the underlying physics processes like the modelling of the hadron shower, the underlying event, or the pile-up collisions.

Therefore, in the following η intercalibration and in-situ calibration steps, the response of jets relative to other well-measured objects is determined separately in data and MC and the difference is applied as a correction to the data. Finally, the online-to-offline correction, which is unique to the trigger-level analysis, is used to correct any remaining differences between trigger-level and offline jets.

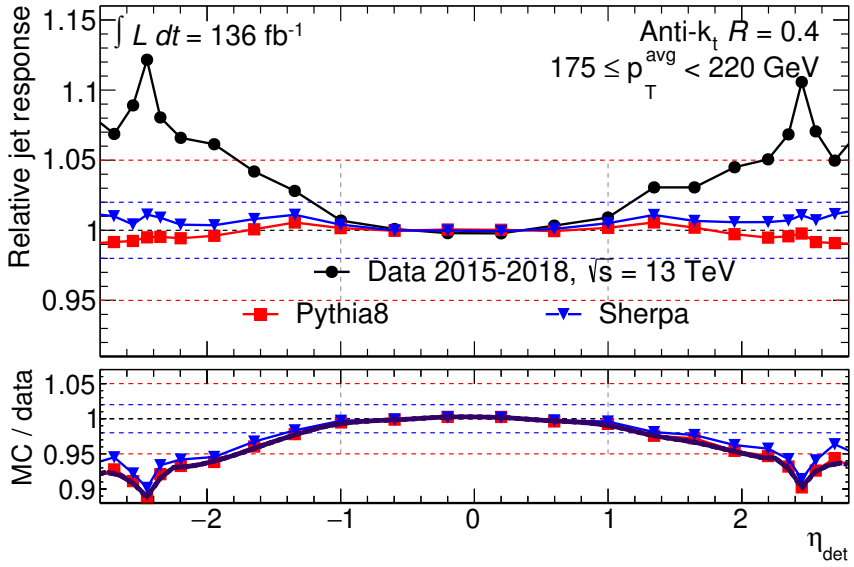


Figure 6.6: Relative jet response as a function of η as determined in the η intercalibration in data (black points) as well as in Pythia (red squares) and Sherpa (blue triangles) MC samples. The $175 < p_{\mathrm{T}}^{\mathrm{avg}} < 220$ bin is shown as an example. Adapted from Ref. [157].

6.2.1 η Intercalibration

The η intercalibration corrects for the relatively strong η dependence of the difference in the jet response between data and MC. Central jets in the region $|\eta| < 0.8$ are considered well-calibrated, while the more forward jets are corrected relative to them. To construct this relation, events with exactly two jets emerging in different η regions are selected.

Barring objects missed in the reconstruction, these two jets must be balanced in their transverse momenta. The momentum asymmetry $\mathcal{A}$ is thus a good measure for determining the η dependence of the jet response:

$$\mathcal{A} = \frac{p_{\mathrm{T}}^{\mathrm{left}} - p_{\mathrm{T}}^{\mathrm{right}}}{p_{\mathrm{T}}^{\mathrm{avg}}}\,, \tag{6.3}$$

where 'left' and 'right' denote the jet lower or higher in the order of their signed pseudo-rapidity and $p_{\mathrm{T}}^{\mathrm{avg}}$ is the average of the transverse momenta of the two balancing jets. The momentum asymmetry is determined as function of $p_{\mathrm{T}}^{\mathrm{avg}}$, since it yields approximately Gaussian distributions when binned in this variable.

Correction factors are derived in bins of $p_{\mathrm{T}}^{\mathrm{avg}}$ and η by solving a linear system of equations that relates the means of the momentum asymmetry distributions $\langle\mathcal{A}\rangle$ between different η bins to each other and to the central reference region. The resulting trigger-level jet response as a function of η is shown in Figure 6.6 for an exemplary $p_{\mathrm{T}}^{\mathrm{avg}}$ bin. The response in the two MC samples studied – Pythia (red) and Sherpa (blue) – is approximately flat due to the MC-based calibration already correctly calibrating the jet response in simulation. The response in data (black) exhibits a discrepancy of up to 12 % between the central and the forward region. The largest deviation is observed in the $2.4 < |\eta| < 2.6$ region, where the transition between the calorimeter End-Cap's Inner

Wheel and Outer Wheel enhances mismodelling in the detector simulation [149]. This transition region is ultimately not part of the selected phase space for this analysis (see Chapter 7).

One notable difference of the η intercalibration applied to the trigger-level jets in this analysis in comparison to the procedure for offline jets occurs at the smoothing of the correction factors. While a two-dimensional Gaussian kernel smoothing in p_{T} and η is applied in the offline calibration, the trigger-level calibration relies on a fit with a polynomial in $\log(p_{\mathrm{T}})$ per η bin analogous to Equation (6.2) with the exponents ranging from -3 to 3. This ensures the smoothness of the calibration, which is the foremost priority for this analysis.

6.2.2 In-Situ Calibration

The in-situ calibration is the final step to correct for differences in the jet response between data and MC. It relies on a momentum balance technique, selecting events with one or more jets recoiling against a well-calibrated photon, a Z boson decaying into electrons or muons, or other, lower-momentum jets. The uncertainties on the electron, photon and muon energy scales are small in comparison to jets, making them a suitable choice for reference objects [160, 161]. Only events with a jet occurring in the central $|\eta| < 0.8$ region are selected, as the more forward regions are calibrated relative to it in the η intercalibration. The determined in-situ calibration factors are finally applied as a function of jet p_{T} but independent of jet η to maintain the uniform response achieved by the η intercalibration.

The Z+jet balance covers the lowest jet p_{T} range from 17 GeV to 1 TeV with increasing uncertainties at higher momenta due to limited available statistics. This region is better constrained by the γ+jet selection that covers a jet p_{T} from 25 GeV to 1.2 TeV, offering higher statistics overall while being limited by photon trigger prescales at low momenta. Jets of even higher momenta are calibrated using a multijet balance, i.e. they are balanced against jets of lower momenta that can be calibrated by the Z+jet or γ+jet balance.

The derivation of the in-situ calibration factors and especially the corresponding uncertainties is an intricate process detailed in Ref. [126]. It is not rederived specifically for this analysis; instead the calibration factors for offline jets are applied to trigger-level jets as well. At this stage of the calibration chain, the trigger-level jets are sufficiently close to the offline scale already to justify using the same in-situ calibration and uncertainties. Any residual difference between the trigger-level and offline jet energy scale is corrected for in the final calibration step, further supporting the choice of applying the offline in-situ calibration.

The same approach was chosen in a previous publication of the trigger-level analysis [1]. One notable difference is that the in-situ calibration available at an earlier stage of Run-2 exhibited some tension between the results of the Z+jet and γ+jet balance, resulting in calibration factors that were not sufficiently smooth as function of p_{T}. Instead of the offline spline-based interpolation and Gaussian kernel smoothing, a fit with a function of constrained smoothness had to be used for the trigger-level analysis. In the in-situ calibration derived from the full Run-2 data, this tension is gone. At the same time, the interpolation with splines has been improved to be less susceptible to data points causing tension [126]. Thus, no alteration of the offline calibration factors is necessary in

this analysis.

6.2.3 Online-to-Offline Correction

The final step of the trigger-level jet calibration chain is the online-to-offline correction. This correction is applied to account for any residual differences between trigger-level and offline jets that remain at the end of the calibration chain. Ideally, it would be applied before the in-situ calibration so that the trigger-level jets are already exactly at the offline scale when applying the offline in-situ calibration. This is impeded by the technical implementation of the jet calibration chain. But since the online-to-offline correction is small, it is expected to approximately factorize and be applicable after the in-situ calibration without a significant change of the result.

The correction is derived from the offline Run-2 data set[3] by geometrically matching offline jets to trigger-level jets with a distance in the (η, ϕ) plane of $R < 0.2$. The kinematic selection utilized throughout this analysis (i.e. $|\eta| < 2.4$, $p_{\mathrm{T}} > 85\,\mathrm{GeV}$, and $y^* = (y_1 - y_2)/2 < 0.6$ as discussed in Chapter 7) is applied and the leading and subleading jets are considered. The response $\mathcal{R} = p_{\mathrm{T}}^{\mathrm{online}}/p_{\mathrm{T}}^{\mathrm{offline}}$ is determined in bins of offline $p_{\mathrm{T}}^{\mathrm{avg}}$ and η. The offline $p_{\mathrm{T}}^{\mathrm{avg}}$ is chosen as it yields the closest to Gaussian response distributions.

With these responses, a procedure similar to the MCJES calibration is performed: The mean response is determined by a Gaussian fit to the core of the distribution and a numerical inversion transforms it from a function of offline $p_{\mathrm{T}}^{\mathrm{avg}}$ and η to a function of trigger-level $p_{\mathrm{T}}^{\mathrm{avg}}$ and η. To reduce the effect of statistical fluctuations, a two-dimensional Gaussian kernel smoothing is applied with the kernel width chosen such that a reduced $\chi^2 \approx 1$ is achieved.

Since the trigger-level jet reconstruction is year-dependent and the pile-up conditions differ between the unprescaled J100 trigger and the end-of-fill J50 and J50Topo triggers, the correction is derived separately for each year and trigger. The correction factors are kept constant below the later applied minimum p_{T} threshold for jets to avoid large migrations from below this threshold. The same is done at high p_{T} when the available statistics in the offline jets drop below a required threshold and the response fits become unstable.

Figure 6.7 shows the derived online-to-offline correction factors in the 12 η bins, exemplary for 2017 data and the unprescaled J100 selection. They are found to be on the order of 2 % or below, which is small enough for the online-to-offline correction and in-situ calibration to approximately factorize. A permille-level asymmetry between the correction for positive η (solid lines) and negative η (dashed lines) is observed. Such an asymmetry can, for example, arise due to a small number of Tile calorimeter modules that were inactive during certain periods of Run-2 but are only partially disabled in the MC simulations [162]. If the effect is not perfectly corrected for in the previous calibration steps, a residual difference is expected here.

[3] While the offline data set is smaller than the trigger-level data set analysed in this thesis, the statistical precision of the determined responses is still sufficient to apply as a correction to all trigger-level jets after smoothing.

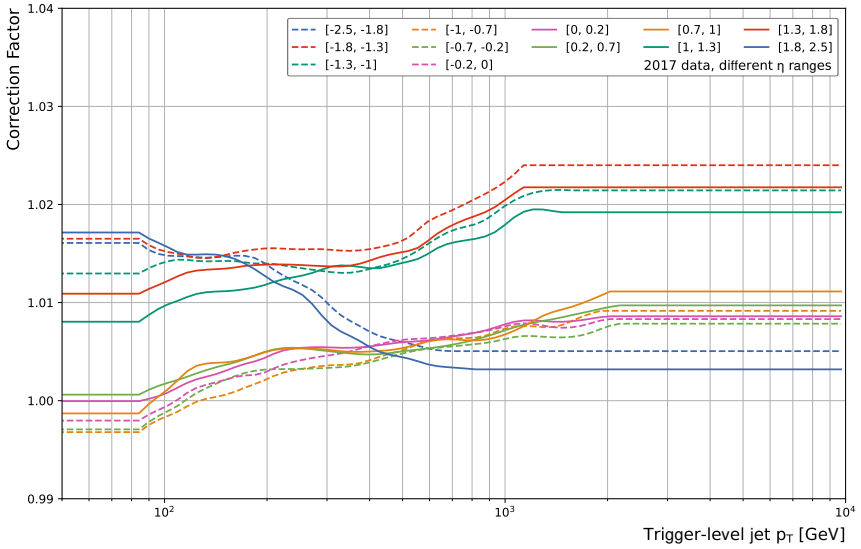


Figure 6.7: Online-to-offline correction factors applied to trigger-level jets recorded in 2017 with the unprescaled J100 trigger. The different colours correspond to the 12 η bins in which the correction is derived, with the solid lines corresponding to positive η and the dashed lines to negative η. Adapted from Ref. [157].

6.3 Calibration Smoothness

Given the large statistical power of this analysis, it is sensitive to very small signals with an amplitude in the order $\mathcal{O}(10^{-4})$ relative to the QCD background. As any localized structures in the calibration curves could induce an excess or deficit in the observed dijet mass spectrum, the smoothness of the calibration is controlled carefully. This needs to be done without directly testing the calibrated observed spectrum for excesses above the background fit to not introduce a bias in the analysis strategy by potential BSM signatures present in the data.

Instead, the correction factors applied to trigger-level jets, averaged over the full trigger-level data set, are studied. While ultimately the smoothness in m_{jj} is crucial, all calibration steps are defined as functions of p_{T} (or E), which is why potential localized structures are more easily identified in p_{T}.

Figure 6.8 shows the average correction factor applied at each of the calibration steps to the subleading jet, determined in the J100 data set. The subleading jet is chosen, because it is less biased towards a large correction by the trigger's p_{T} requirement. Still, this bias is significant for jets with $p_{\mathrm{T}} < 200\,\mathrm{GeV}$ which is why only jets above this threshold are considered here. The average correction factors shown in black are overlayed with the statistical uncertainty of the observed p_{T} spectrum to identify if structures in the calibration could lead to significant features in the observed spectrum. This overlay only serves as an approximate reference, as the size of structures in the jet response and in

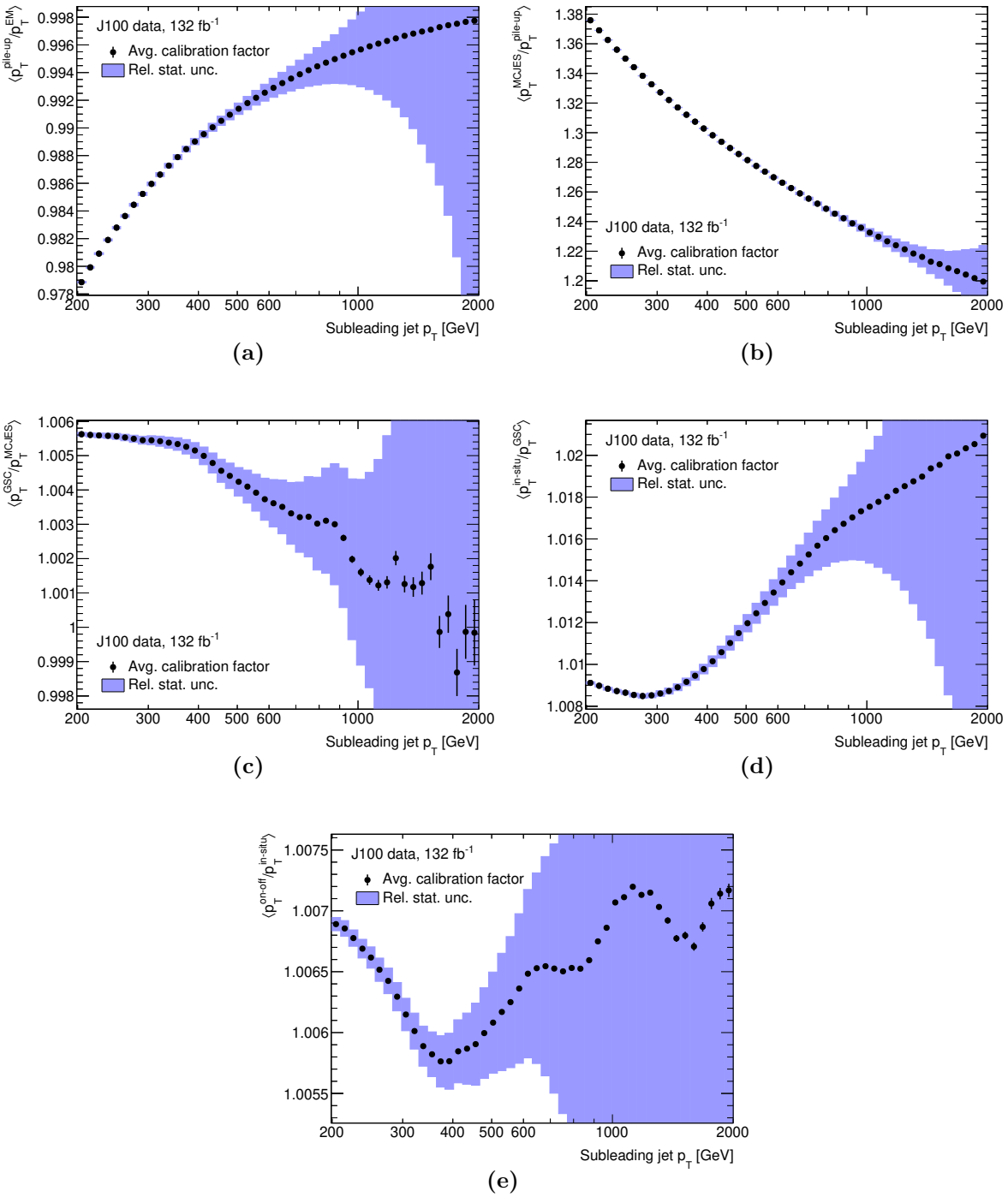


Figure 6.8: Average correction factor applied to the subleading jet in each calibration step as a function of p_{T}, determined in the J100 data set: (**a**) Pile-up correction, (**b**) MCJES, (**c**) GSC, (**d**) η intercalibration and in-situ calibration combined, and (**e**) online-to-offline correction. Black bars correspond to the uncertainty of the mean of the factors. The relative statistical uncertainty of the measured p_{T} distribution is overlayed in blue to identify statistically significant features in the calibration curves.

the final p_{T} or m_{jj} spectrum cannot be directly compared.[4] Structures in the response must be unable to mimic potential resonance signals, meaning they must be wider than the widest signals tested in this search: Gaussian resonances of 15 % width.

The pile-up correction and MCJES (Figures 6.8a and 6.8b) are smooth functions over the tested p_{T} range and only exhibit a small curvature, making them suitable for this analysis. The same is true for the η intercalibration and in-situ calibration which are for technical reasons applied as a single step in the jet calibration chain (Figure 6.8d). While the correction does exhibit a local minimum, the curvature is low enough to not mimic a signal.

The GSC, however, introduces a kink in the average jet response at $p_{\mathrm{T}} \approx 400\,\mathrm{GeV}$ and a seemingly insignificant bump at 900 GeV as shown in Figure 6.8c. While the GSC is designed to not alter the average response, this only holds perfectly true in MC and without the applied smoothing. Differences in the distribution of the jet structure variables and their correlation to another between data and MC prevent perfect closure in data. While this non-closure is well below the percent level, it can still induce features in the observed dijet mass spectrum, as will be shown later in this section. The online-to-offline correction fluctuates only at a permille level. Nevertheless, a narrow, statistically significant dip around 400 GeV and a small, seemingly insignificant dip around 800 GeV are observed. Both of these structures are potentially a result of the structures induced by the GSC response at these positions.

Similar behaviour of the calibration response is observed in the J50 data set, which corresponds to different pile-up conditions and extends to lower jet energies. The effect of the individual calibration steps on the reconstructed invariant dijet mass is determined as well. A propagation of the structures in the calibration response into m_{jj} is observed, albeit less pronounced due to the η-dependent relation between m_{jj} and p_{T} being averaged over. The results are summarized in Appendix B.

To determine the impact of the structures in the GSC and online-to-offline correction, the full calibration chain is applied to jets in the Pythia dijet MC sample and dijet events are selected as described in Chapter 7. After each stage of the calibration chain, the MC dijet mass spectrum is fitted with the background estimation technique described in Section 9.1 – in this case with the 6-parameter functional form fit. The fit residuals normalized by the statistical uncertainty in each bin, in the following labelled significances, can be seen in Figure 6.9. Only the significances are shown because the permille-level differences between the MC spectrum and fit are not visible when overlaying them. The pile-up correction stage is omitted here, since only after the large MCJES correction are the jets approximately at the true energy scale and correctly selected with the chosen m_{jj} range. Although the data-driven in-situ calibration and online-to-offline correction should only be used to correct jets in data, they are also applied to MC here to assess whether the correction behaves smoothly in m_{jj} or potentially induces features. Also note, that the MC sample is generated to offer approximately uniform statistical precision throughout the m_{jj} spectrum while the available trigger-level data set has a statistical precision rising towards low m_{jj}, exceeding that of the MC sample for $m_{jj} \lesssim 1800\,\mathrm{GeV}$.

[4] A small shift δ in the jet response at position p_{T} will migrate jets from p_{T} to $p'_{\mathrm{T}} = p_{\mathrm{T}} \cdot (1 + \delta)$. Due to the steeply falling cross section, this can migrate jets 'down the slope' to a less populated region of the p_{T} spectrum. Depending on the width and shape of such a feature in the response, the relative change in the observed p_{T} spectrum can be in the order of $1 - 10$ times δ as found in toy studies.

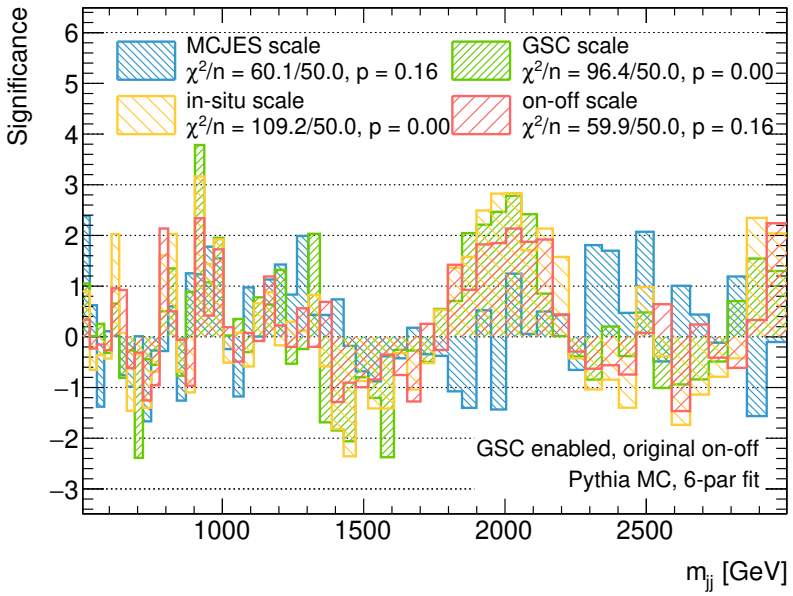


Figure 6.9: Significance of the residuals of a fit with the 6-parameter dijet function to the Pythia dijet mass spectrum at individual stages of the jet calibration. A significant excess at 2 TeV is induced by the GSC.

Thus, potential features at lower masses could be statistically insignificant in MC but significant in data. Nevertheless, the MC fits are a useful tool for assessing the overall smoothness of the calibration chain.

It can be seen, that a good fit with a p-value of 0.16 based on the χ^2-statistic[5] is achieved after the MCJES calibration (blue histogram). The GSC, however, induces a significant, broad excess around a dijet mass of 2 TeV and a single excessive bin at 900 GeV as shown in green. Since $m_{jj} \approx 2p_{\mathrm{T}}$ for back-to-back jets, these excesses are potentially caused by the apparently small feature at $p_{\mathrm{T}} \approx 900\,\mathrm{GeV}$ and the kink at 400 GeV in the average jet response shown in Figure 6.8c. These GSC-induced features remain or can only be partially resolved by the following in-situ calibration (yellow) and online-to-offline correction (red). The partial recovery of the fit quality by the online-to-offline correction reinforces the assumption that the structures observed in the average response of the online-to-offline correction mirror those in the GSC – but they cannot completely restore the smoothness.

With studies analogous to the ones shown in this section, it is found that the features arise at the $N_{90\mathrm{Const}}$ step of the GSC – potentially due to its discrete nature or its large sensitivity to quark-gluon differences. While further studies concerning the smoothness of the individual GSC steps are ongoing, results in this thesis are presented with the GSC *entirely disabled* to ensure the smoothness of the calibration chain and allow for a functioning background estimate.

The effect of disabling the GSC is, again, tested in fits of the invariant mass spectrum of Pythia dijet events. The results are shown in Figure 6.10. The GSC scale is now

[5] The determination of p-values from the χ^2-statistic is discussed in Section 9.2.3, see Equations (9.15) and (9.17).

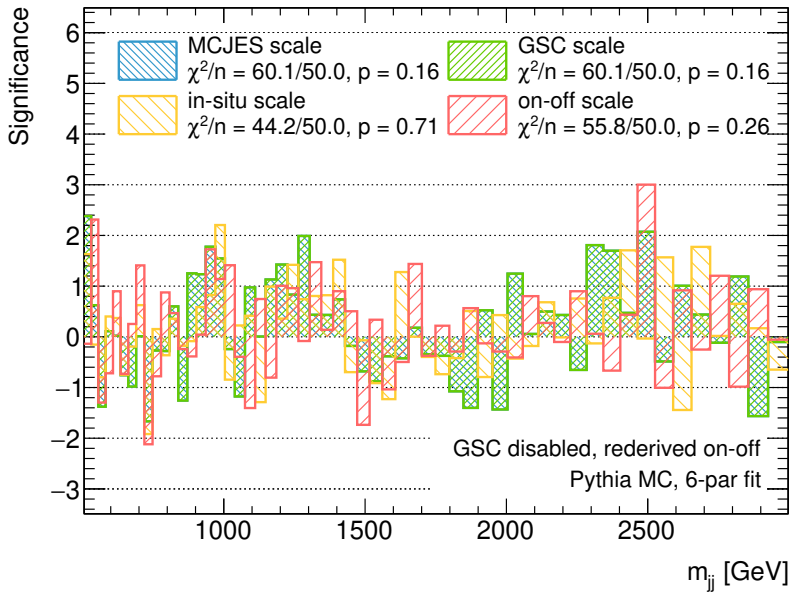


Figure 6.10: Significance of the residuals of a fit with the 6-parameter dijet function to the Pythia dijet mass spectrum at individual stages of the jet calibration. The GSC has been disabled here and the online-to-offline correction has been rederived with the GSC disabled. No significant excesses are observed.

identical to the MCJES scale since no correction is applied. The fit at the in-situ scale in yellow shows that the calibration, even though it was derived with the GSC active, is still applicable without inducing significant features.

Any residual error due to disabling the GSC is evaluated by rederiving the last step of the calibration chain, the online-to-offline calibration, with the GSC disabled for the online jets. The online-to-offline responses behave less Gaussian this way due to it being composed of multiple sub sets of slightly different responses. As a result, the mean responses determined by Gaussian fits are more prone to fluctuations. Nonetheless, updated online-to-offline correction factors are derived and tested on the Pythia MC, resulting in the fit residuals shown in red in Figure 6.10.

Since the MC does not provide sufficient statistics to identify low-energy structures in the calibration, the same tests are performed on a randomly selected partial data set corresponding to approximately 20 % of the full Run-2 J100 luminosity and 10 % of the J50 luminosity. These fractions were chosen to not exceed the sensitivity of the previous trigger-level dijet search [1] to perform calibration studies on a data set that has already been found to be signal-free with the given statistical precision.

In this partial data set, the rederived online-to-offline correction is found to induce significant features in the fit residuals at low m_{jj}, probably due to fit instabilities of the non-Gaussian responses.

Due to this non-smoothness, the online-to-offline correction is also *disabled* in this thesis.[6] Instead, its correction is treated as an uncertainty as discussed in the next

[6] At the time of writing this thesis, further studies on the interplay between GSC and online-to-offline correction smoothness as well as the stability of Gaussian fits to non-Gaussian responses are ongoing.

section.

6.4 Jet Energy Scale and Resolution Uncertainties

The calibration chain applied to trigger-level jets is associated with systematic uncertainties, which can affect the jet energy scale (JES) and resolution (JER). While the offline-derived uncertainties on the pile-up correction, in-situ calibration and JER are adopted here, uncertainties on the flavour composition and response, the η intercalibration and the online-to-offline difference are derived specifically for the TLA.

The uncertainties are parametrized in terms of nuisance parameters (NPs). They allow the propagation of the effect of individual sources of uncertainty to higher-level observables like the invariant mass spectrum of the considered signal models. This is done by independently varying each of the NPs up or down by one standard deviation of the uncertainty and repeating the calibration chain.

6.4.1 Offline-Derived Uncertainties

The pile-up uncertainties are adopted from the offline derivation. They are determined in data using a Z+jet balance as well as tracking information associated with jets and are expressed as 3 NPs influencing the energy density ρ, the μ-correction and the residual p_{T} dependence.

The in-situ measurements entail the largest number of individual uncertainty sources. MC modelling, energy scale and resolution, and statistical uncertainties from the employed electrons, muons, photons, and the multijet-balance technique are propagated to the correction factors. Since this analysis is not sensitive to the correlations between the large number of available NPs (see Section 8.3), an eigenvector decomposition is employed to combine the different sources into eight effective NPs.

The jet energy resolution is determined via an in-situ multi-jet balance technique and a noise measurement using randomly selected jet cones. A similar NP reduction scheme is employed to parameterize its uncertainty as seven effective JER NPs. They are not relevant for data but for the production of MC samples where jets are artificially fluctuated according to the uncertainty on the jet energy resolution.

6.4.2 Trigger-Level Uncertainties

Due to the different response of gluon-initiated and quark-initiated jets – and to a lesser degree also light-flavour and heavy-flavour quark-initiated jets – the calibration chain only correctly calibrates the average response for the jet composition in the chosen calibration processes. Since a potential signal is likely to entirely consist of either quark- or gluon-initiated jets, its average jet response will be different. It is thus important to determine the difference in response.

This is done by geometrically matching trigger-level jets in MC to truth jets containing the information of the initiating parton. The quark response $\mathcal{R}_q$ and the gluon response $\mathcal{R}_g$ are determined from Gaussian fits to the $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}}$ ratios. For a given fraction f_g of

gluon jets in a sample, the jet flavour composition and jet flavour response uncertainties are determined as follows:

$$\sigma_{\text{composition}} = (\mathcal{R}_g - \mathcal{R}_q) \cdot \sigma_{f_g} \,, \tag{6.4}$$

$$\sigma_{\text{response}} = f_g \cdot \sigma_{\mathcal{R}_g} \,. \tag{6.5}$$

f_g is set to 0.5 with an uncertainty σ_{f_g} of $\pm 100\,\%$ to not make assumptions on the composition of signal. This means that the flavour composition uncertainty is entirely driven by the response difference $(\mathcal{R}_g - \mathcal{R}_q)$, which is aimed to be reduced by the GSC. The jet flavour response uncertainty is approximated by the difference between the nominal gluon response determined in Pythia and alternative Sherpa samples using different hadronization models. The uncertainty on the quark response is small in comparison.

Both uncertainties have been derived specifically for the TLA, with the GSC enabled [163].[7] With the GSC disabled, the $(\mathcal{R}_g - \mathcal{R}_q)$ term is expected to increase significantly and the generator difference $\sigma_{\mathcal{R}_g}$ may be affected as well. Hence, both flavour uncertainty components are increased by a conservative factor of 100 % for this thesis. Although the flavour uncertainties are the dominant sources of systematic uncertainty for all considered signal masses, their effect is small in comparison to the statistical uncertainties of the data. This encourages the conservative approximation here.

Uncertainties on the η intercalibration are estimated from variations of the event selection criteria, the MC generator differences, a small non-closure of the polynomial fit in p_{T}, and the propagation of the statistical uncertainty. These sources are combined into six NPs.

Since the online-to-offline correction is not applied in this analysis due to its non-smoothness, the residual difference between the trigger-level and offline jet energy scale is treated as an uncertainty instead. The average correction applied by the online-to-offline correction that was derived with the GSC disabled is applied as nuisance parameter to the jet energy scale. The online-to-offline difference becomes the subdominant systematic uncertainty throughout the considered dijet mass range.

The different uncertainty sources discussed above amount to in total 27 NPs which are summarized in Table 6.1. The impact of each of these NPs on the different signal models is determined and discussed in Chapter 8.

[7] Since this uncertainty is determined purely in MC, it is not influenced by the data-driven calibration steps that were derived afterwards.

Derivation	Uncertainty Group	Nuisance Parameters
Offline	Pile-up correction	3
	In-situ calibration	8
	Jet energy resolution	7
Trigger level	Flavour uncertainties	2
	η intercalibration	6
	Online-to-offline difference	1

Table 6.1: Systematic uncertainties on the jet calibration considered in this analysis.

6.5 Calibration Performance

The performance of the calibration chain applied to trigger-level jets can be characterized by the jet energy resolution and, more importantly, by the resolution of the reconstructed invariant dijet mass. The latter is directly linked to the ability of this search to resolve narrow resonances and, thus, the search sensitivity.

The resolution of the reconstructed jet p_{T} and m_{jj} is determined analogously to the MC-based calibration steps by constructing the responses $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}}$ and $m_{jj}^{\mathrm{reco}}/m_{jj}^{\mathrm{truth}}$. This is done in the Pythia MC sample with the kinematic event selection of this search (detailed in the next chapter) applied. The resolution is then given by the standard deviation of the Gaussian distribution fit to the core of the response.

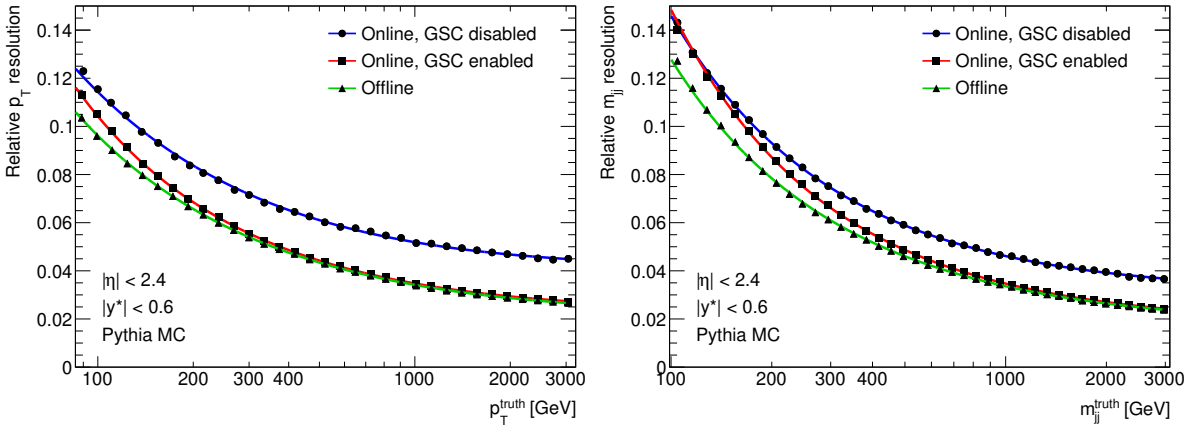
Figure 6.11 shows the determined resolution of jet p_{T} and m_{jj}, where points represent the trigger-level jets with the calibration chain applied in this thesis. A relative m_{jj} resolution between 7 % and 4 % is observed in the relevant m_{jj} range above 300 GeV. For comparison, the trigger-level jet resolution with the GSC applied and the offline jet resolution are shown as squares and triangles, respectively. The determined resolutions of both variables are well described by a function composed of a noise term, stochastic term, and constant term as given in Equation (4.6).

The m_{jj} resolution achieved in this thesis is approximately 1–1.5 percentage points worse than what could be achieved by applying the calorimeter-based GSC, which reduces the search sensitivity as will be discussed in Section 10.2.4. The fluctuations of the determined resolutions are also larger when the GSC is disabled because of a response distribution that is less Gaussian in shape.

At high energies, the resolution of trigger-level jets with applied GSC approaches that of offline jets, which validates the calorimeter-based GSC as a good approximation of the GSC applied to offline jets. The additional tracking and muon-spectrometer information available in offline jets only provide a significant resolution benefit at comparatively low energies.

6.5.1 Binning of the Invariant Dijet Mass

When measuring the m_{jj} spectrum for this search, a binning must be chosen. Narrow bins are advantageous in resolving resonances, but a binning narrower than the detector resolution does not provide any additional benefit. Therefore, the m_{jj} binning is chosen



(a) **(b)**

Figure 6.11: Resolution of the reconstructed (**a**) jet p_{T} and (**b**) m_{jj} relative to the particle level, determined in Pythia MC. The circular markers show the resolution of trigger-level jets with the calibration chain of this thesis applied, i.e. with the GSC disabled. For comparison, the resolution of trigger-level jets with the GSC enabled is shown as squares and the resolution of offline reconstructed jets as triangles. The solid coloured lines each show a fit with the function in Equation (4.6) to the respective resolution.

such that each bin width is equal to the m_{jj} resolution at that specific value of m_{jj}.

The binning used in this analysis was derived from the trigger-level m_{jj} resolution with the GSC enabled. The use of bins narrower than the true resolution in this analysis is expected to have a negligible effect.

7 Event Selection

For this analysis, collision events containing pairs of jets are selected from the Run-2 Data Scouting stream. Events with known detector malfunctions or with falsely reconstructed jets from noise sources are excluded. Selection criteria on the jet kinematics are applied to optimize the sensitivity to BSM signatures. These criteria are discussed in the following sections.

7.1 Data Quality Requirements

For standard ATLAS analyses, event selection criteria are centrally derived to select only data-taking periods where all detector subsystems were working as intended [164]. Since this analysis uses only calorimeter information, an adaptation of these criteria is applied. Only events with global errors or calorimeter malfunctions are discarded, while periods with tracking or toroid magnet failures are retained, as they are not expected to affect trigger-level jets. In addition to these central criteria, a short data-taking period with a malfunction in the online area-based jet pile-up correction was identified and discarded for this analysis.

Energy deposits in the calorimeter arising from sources other than jets can be falsely reconstructed as jets. This non-collision background includes calorimeter noise bursts, cosmic rays, or beam-induced background, which consists mainly of high-energy muons produced from interactions of the beam with residual gas or accelerator components far away from the interaction point [165].

Jets reconstructed from these noise sources do not traverse the detector from the interaction point outwards and are often asynchronous to the expected bunch crossings. Thus, they can be efficiently distinguished from real jets by the fraction of energy deposited in individual calorimeter layers and calorimeter pulse shapes. Another distinguishing variable is the charged fraction f_{ch}, defined as the sum of the transverse momenta of tracks originating from the primary vertex of the jet, relative to the transverse momentum of the jet. However, f_{ch} is not available in the Data Scouting stream as it relies on track reconstruction. The other calorimeter-based criteria derived for offline jets are applied to the trigger-level jets in this analysis. If a jet is identified as *BadLoose* [164] according to these criteria, the entire event is discarded.

The effectiveness of this calorimeter-based rejection method has been determined in events with a similar dijet topology as selected in this analysis [166]. The remaining fraction of dijet events potentially originating from noise sources has been found to be below 0.3 % and to not affect the shape of the invariant mass spectrum, which makes it a negligible background contribution.

7.2 Kinematic Selection

It is necessary to define a signal region in which the dijet topology can be measured precisely. This definition is based on jet kinematic requirements, which are discussed in this section.

7.2.1 Jet Selection

Events with at least two jets within $|\eta| < 2.4$ are selected. This excludes the Forward Calorimeters and the transition region in the End-Caps where the jet energy resolution decreases and the calibration uncertainties are largest, as discussed in Section 6.2.1.

The two jets leading in p_{T} in this η range are required to have transverse momenta of $p_{\mathrm{T}} > 85\,\mathrm{GeV}$. This achieves a strong suppression of jets produced in pile-up collisions and thus of the coincidental background of selecting pairs of jets originating from different collisions [167].[1] The p_{T} requirement also ensures that only jets are selected for which the trigger-level calibration chain is found to achieve good closure.

7.2.2 Selection Based on the Rapidity Difference

To reduce the dominant QCD background in this analysis, a selection based on half the rapidity difference between the two leading jets, $y^* = (y_1 - y_2)/2$, is applied. t-channel processes, which make up a large fraction of the QCD background, are enhanced for small deflections from the beam direction and accordingly large y^*, as discussed in Section 3.1. Conversely, the resonant s-channel signal processes favour smaller $|y^*|$ [3].

Therefore, the $|y^*|$ requirement influences the signal-to-background ratio. This dependence was determined in Ref. [10] for the case of Z' signals. Upper $|y^*|$ thresholds between 0.6 and 0.8 were found to result in the highest expected sensitivity across a wide range of Z' masses. To facilitate the comparison and combination with the offline dijet resonance search [3], $|y^*| < 0.6$ is chosen for this analysis.

The y^* selection has an additional benefit: For any given m_{jj}, requiring small $|y^*|$ selects dijet events with higher jet p_{T}. This can be seen by expressing the invariant mass of two (massless) jets as

$$m_{jj}^2 = 2\, p_{\mathrm{T}}^{j1}\, p_{\mathrm{T}}^{j2} (\cosh \Delta y - \cos \Delta\phi)\,. \tag{7.1}$$

For back-to-back dijet events ($\Delta\phi \approx \pi$, $p_{\mathrm{T}}^{j1} \approx p_{\mathrm{T}}^{j2}$), this means $m_{jj} \approx 2\, p_{\mathrm{T}}$ for small y^*, but only $m_{jj} \lesssim 2\, p_{\mathrm{T}}$ for arbitrary y^*. The former behaviour is useful for this analysis because it lowers the m_{jj} threshold above which the recorded mass spectrum is unbiased by the E_{T} requirement of the L1 trigger and the p_{T} requirement on the jets [10, 11]. This effect is discussed in the following section.

[1] In offline analyses, the pile-up rejection of low-energy jets is achieved by *jet vertex tagging*, i.e. using the reconstructed tracks associated with reconstructed jets to identify their primary vertex [168]. While this information is not available for the trigger-level jets considered in this thesis, vertex reconstruction is implemented for the Data Scouting stream in Run-3, potentially allowing probing lower-energy jets in future TLAs [143].

7.2.3 Selection Based on the Invariant Dijet Mass

The observed m_{jj} spectrum is influenced by the p_{T} requirements on individual jets. These stem either from the explicit $p_{\mathrm{T}} > 85\,\mathrm{GeV}$ selection placed in this analysis or from the Level-1 trigger selection. Both selections primarily affect the lower part of the m_{jj} spectrum. Therefore, a minimal m_{jj} threshold can be determined above which the observed m_{jj} spectrum is mostly unbiased by the requirements. Above this minimum threshold, the m_{jj} spectrum is smoothly falling, which is a requirement for the background fitting methods employed in this analysis. The optimization of this threshold is discussed in this section.

Trigger Efficiency

The L1 trigger efficiency is measured in offline events recorded by HLT single-jet triggers. These reference triggers are chosen to have a sufficiently low threshold to not bias the measured efficiency.[2] The same event selection criteria discussed above are applied to determine the trigger efficiency for the phase space selected by this analysis. Especially the $|y^*|$ requirement has a significant impact on the efficiency as a function of m_{jj}.[3] Here, m_{jj} refers to the invariant mass determined from the trigger-level jets with the full TLA calibration chain applied.

Figures 7.1a to 7.1c show the determined efficiencies of the J50, J50Topo, and J100 triggers used in this thesis. This considers data accumulated over the full Run-2 period for the J50 and J100 triggers and over the period after TS1 in 2018 for the J50Topo trigger. The J50 and J100 single-jet triggers are found to approach 100 % efficiency for large m_{jj}.

The J50Topo dijet trigger plateaus at around 99.5 % efficiency. The remaining inefficiency is due to the $|y^*| < 1.0$ requirement applied to the two leading L1 jet candidates. Due to the coarse energy resolution of the L1 jet reconstruction, the two leading L1 jet candidates do not necessarily correspond to the two leading HLT jets considered in this analysis. Consequently, the $|y^*|$ requirement can be applied to a different jet pair, resulting in proper dijet events being falsely discarded. This inefficiency is approximately constant as a function of m_{jj} at a level of 0.5 %. As a result, it mainly scales the recorded m_{jj} spectrum, while a remaining small, linear slope is absorbed by the background fit. Therefore, its effect is tolerable for this analysis.

The efficiency curves are fit with a logistic sigmoid function to smooth out statistical fluctuations:

$$\varepsilon(m_{jj}) = \frac{a}{1 + e^{-b(m_{jj}-c)}}\,. \tag{7.2}$$

[2] While the HLT trigger name, e.g. j85, suggests a higher transverse momentum threshold than required on L1, e.g. by J50, these names are misleading. L1 jet candidates are reconstructed at the electromagnetic scale, meaning that the large positive correction towards their true energy scale is missing. Additionally, the L1 jet energy resolution is significantly lower, such that 99.9 % efficiency is only reached far above the required L1 threshold.

[3] In the previous publication of the TLA, a minimal requirement on $p_{\mathrm{T}}^{\mathrm{lead}}$ was placed to ensure full trigger efficiency and then the threshold above which the m_{jj} spectrum is unbiased by the $p_{\mathrm{T}}^{\mathrm{lead}}$ requirement was identified [1]. The selection applied in the presented analysis is more efficient, because it allows both the leading and the subleading jet to have fulfilled the L1 trigger requirement. This allows for lowering the minimal m_{jj} requirement [11].

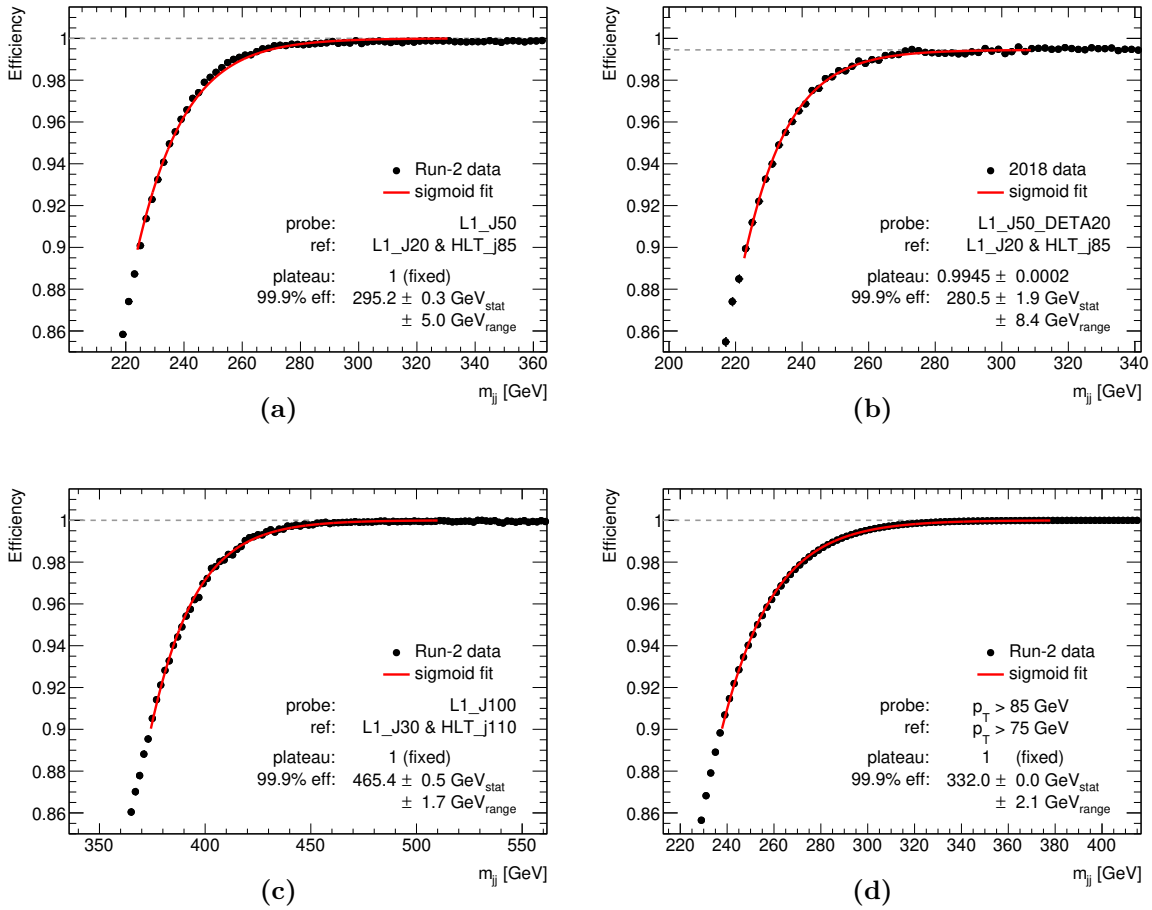


Figure 7.1: Efficiency of the trigger selection of the individual L1 triggers used in this analysis: (**a**) J50, (**b**) J50Topo, and (**c**) J100, as determined in the offline Run-2 data set recorded with jet triggers of lower threshold than the trigger to be probed. (**d**) Efficiency of the $p_{\mathrm{T}} > 85\,\mathrm{GeV}$ requirement for leading and subleading jets, relative to events selected with a $p_{\mathrm{T}} > 75\,\mathrm{GeV}$ requirement. This is measured in the J50Topo triggered Data Scouting data set.

The parameter a, which corresponds to the plateau of the efficiency, is fixed to 1 for all single-jet triggers since they reach full efficiency. However, a is left as a free parameter for J50Topo since it does not reach full efficiency. For all triggers, b and c are free parameters. The fit is performed in the m_{jj} range corresponding to an efficiency in the range $0.9 \leq \varepsilon \leq a - 10^{-4}$, which is determined iteratively from the fit. In this range, the fit describes the measured efficiency well.

The point of 99.9 % efficiency, denoted $m_{jj}^{99.9\%}$, is determined from these fits. This efficiency point was chosen because previous studies showed that the remaining 0.1 % inefficiency behaves sufficiently smoothly in m_{jj} to be absorbed by the background fit [166]. This was again validated for this analysis using the partial data set.

The systematic uncertainty on $m_{jj}^{99.9\%}$ determined with this method is estimated by a 4-point variation of the fit range corresponding to lower (upper) ε thresholds of 0.85 or 0.95 ($a - 5 \cdot 10^{-4}$ or $a - 2 \cdot 10^{-5}$). The ± 0.05 variation of the lower threshold corresponds to the range over which the sigmoid can be considered a good description of the trigger efficiency. The upper threshold has a smaller effect on the determined $m_{jj}^{99.9\%}$.

L1 Trigger	HLT Reference	Year	$m_{jj}^{99.9\%}$ [GeV]	Chosen Threshold [GeV]
J50	j85	2017	301.8 ± 5.6	344
		2018	291.8 ± 5.0	
J50Topo	j85	2018 post TS1	280.5 ± 8.4	344
J100	j110	2016	463.3 ± 1.7	481
		2017	467.8 ± 2.2	
		2018	463.4 ± 2.4	

Table 7.1: Obtained 99.9% efficiency point for all used triggers year-by-year

The upwards and downwards variation of the distance to the plateau by a factor of 5 results in an approximately symmetric variation of the fit range in m_{jj} and is considered a conservative estimate of the fit uncertainty. The statistical uncertainty on $m_{jj}^{99.9\%}$ propagated from the covariance matrix of the fit parameters is small in comparison to the range variation.

Since the trigger conditions change with the data-taking period, $m_{jj}^{99.9\%}$ is determined separately for each year and trigger. The results are presented in Table 7.1. The systematic uncertainty for J50Topo is found to be larger than for the other triggers, mainly due to the plateau parameter c being a free variable in the fit.

Efficiency of the p_{T} Selection

To determine the minimal m_{jj} thresholds for an unbiased spectrum, a second source of selection inefficiency must be considered: the p_{T} requirement on the (sub-)leading jet. Its effect is not apparent in the trigger efficiencies determined above, because both the reference and the probe selection apply the same $p_{\mathrm{T}} > 85\,\mathrm{GeV}$ selection.

The p_{T} requirement on the leading jet does not cause inefficiency in the relevant m_{jj} range. This can be understood qualitatively: Following Equation (7.1), the invariant mass of back-to-back dijet events with small y^* is approximately given by $m_{jj}^2 \approx 4\, p_{\mathrm{T}}^{\mathrm{lead}}\, p_{\mathrm{T}}^{\mathrm{sublead}}$ and therefore $m_{jj} \lesssim 2\, p_{\mathrm{T}}^{\mathrm{lead}}$. Accordingly, a leading jet p_{T} cut of 85 GeV only affects the m_{jj} spectrum below approximately 170 GeV. While this threshold is increased by events that are not pure back-to-back dijets, it does not extend to 280 GeV and beyond, where the considered triggers become efficient.

However, the same consideration does not hold true for the subleading jet. It can be significantly lower in p_{T} than the leading jet if additional objects (likely other jets) in the final state influence the momentum balance. Therefore, the $p_{\mathrm{T}} > 85\,\mathrm{GeV}$ requirement on the subleading jet has an effect up to higher m_{jj} values.

The p_{T} selection efficiency in the considered phase space is determined by selecting reference events with a looser $p_{\mathrm{T}} > 75\,\mathrm{GeV}$ requirement and then probing the fraction of events that meet the $p_{\mathrm{T}} > 85\,\mathrm{GeV}$ requirement. Figure 7.1d shows the result as a function of m_{jj}. Since this efficiency is a purely kinematic effect, it is independent of the trigger used to select the events. The J50Topo-triggered events in the Data Scouting stream are used here, as they offer a large statistical precision in the m_{jj} range of interest.

As with the trigger efficiencies, $m_{jj}^{99.9\%}$ is also determined from a fit with a logistic

sigmoid function. The value obtained is 332.0 ± 2.1GeV, which is significantly higher than the efficiency points of the J50 and J50Topo. Thus, for events selected with the J50 or J50Topo trigger, the p_{T} selection efficiency is the relevant criterion for the minimal m_{jj} threshold, while for events selected with the J100 trigger, the trigger efficiency is the relevant criterion.

Due to the limitation by the p_{T} selection efficiency, even lower-threshold Data Scouting triggers, such as J40, do not offer any additional benefit in probing lower masses and are therefore not considered in this thesis.

Thresholds for m_{jj}

The lower m_{jj} threshold for each trigger is chosen to be $m_{jj}^{99.9\%} + 2\sigma_{\mathrm{range}}$, rounded up to the next edge of the m_{jj} resolution bins used in this analysis. This choice conservatively limits the residual inefficiency to less than 0.1 %. The thresholds are found to be 344 GeV for J50 and J50Topo and 481 GeV for J100.

Due to their identical m_{jj} threshold, J50- and J50Topo-triggered events are *combined* into a single signal region (SR), hereafter referred to as J50 for simplicity. This J50 SR provides sensitivity to lower-mass signals than the J100 SR but at a reduced integrated luminosity.

In addition to a lower threshold for m_{jj}, an upper threshold is introduced to limit the fit range for the background estimate. While extending the fit range beyond the highest mass signal of interest helps to constrain the fit and increases the signal sensitivity, this improvement diminishes at some point. In a scan of possible fit ranges, the point is identified at which extending the range further has only a negligible effect on the expected limits. This point is found to be 1516 GeV for the J50 SR and 2997 GeV for the J100 SR. Limiting the considered fit range with these upper thresholds facilitates a good description of the observed m_{jj} spectrum by a single function. Additionally, it ensures large event counts per m_{jj} bin, which is relevant for an analytic approximation of the likelihood discussed in Section 9.1.1.

Table 7.2 summarizes the event selection criteria discussed in this section for both signal regions.

	J50 SR	J100 SR
(sub-)leading jet $\lvert\eta\rvert$	< 2.4	
(sub-)leading jet p_{T}	> 85 GeV	
$\lvert y^* \rvert$	< 0.6	
m_{jj}	$\in [344, 1516]$ GeV	$\in [481, 2997]$ GeV

Table 7.2: Requirements placed on the leading and subleading jet for the J50 and J100 signal regions.

8 Parametrization of Signal Models and Uncertainties

This search targets two different types of dijet resonance signals: the leptophobic Z' and generic, Gaussian-shaped resonances. To perform hypothesis tests for these signals, their expected m_{jj} spectra must be known. These predicted signal shapes are parametrized as functions depending on the signal mass to later include them in signal-plus-background fits for arbitrary masses. The theoretical and experimental systematic uncertainties are included in this parametrization to evaluate them at arbitrary signal masses as well.

8.1 Signal Parametrization

The Z' signal shape is determined from MadGraph5 MC samples (see Section 3.2), applying all event selection criteria except the m_{jj} requirement. The samples are generated at Z' masses of 350 GeV, 600 GeV, 1 TeV, and 2 TeV with a nominal coupling of $g_q = 0.1$. These values are chosen as representative mass points within the sensitive mass range for which the signal prediction is determined. For selected mass points, alternative samples are generated with $g_q = 0.02$ and $g_q = 0.2$ to confirm that the g_q dependence of the m_{jj} spectrum is negligible (see discussion in Section 2.2.2).

Interpolation between the generated mass points is required to test for arbitrary Z' masses. For this purpose, the m_{jj} spectrum of a Z' resonance is parametrized as a double-sided Crystal Ball (DSCB) function [169]. It consists of a Gaussian core with parameters μ and σ, which smoothly transitions into power-law tails with the exponents n_L, n_R and the transition points α_L, α_R being free parameters. Figure 8.1 shows the m_{jj} spectra, normalized to unity, for the four generated Z' masses overlaid with a DSCB fit. The region around and above the peak is well described by a Gaussian distribution, as it is dominated by the approximately Gaussian resolution of the invariant mass reconstruction. The tail below the peak is elongated due to potential energy losses in the dijet selection, e.g. due to jet splitting or the omission of jets outside the η acceptance. It is described by the power-law part of the DSCB function with sufficient accuracy. A small deviation from the power-law tail is observed around $0.7 m_{Z'}$ for all four Z' masses. A similar effect has been encountered in previous ATLAS dijet searches [170]. Since this does not deteriorate the description of the resonance peak, which is the most relevant for the sensitivity of this search, the effect is expected to be negligible.

The signal shape for arbitrary $m_{Z'}$ is approximated by a linear interpolation of the best-fit DSCB parameters of the four reference spectra.

For the normalization, the total cross section determined in the MC simulation is interpolated using the function

$$f(x) = p_1 \cdot x^{p_2} \cdot (1 - x)^{p_3} \,, \tag{8.1}$$

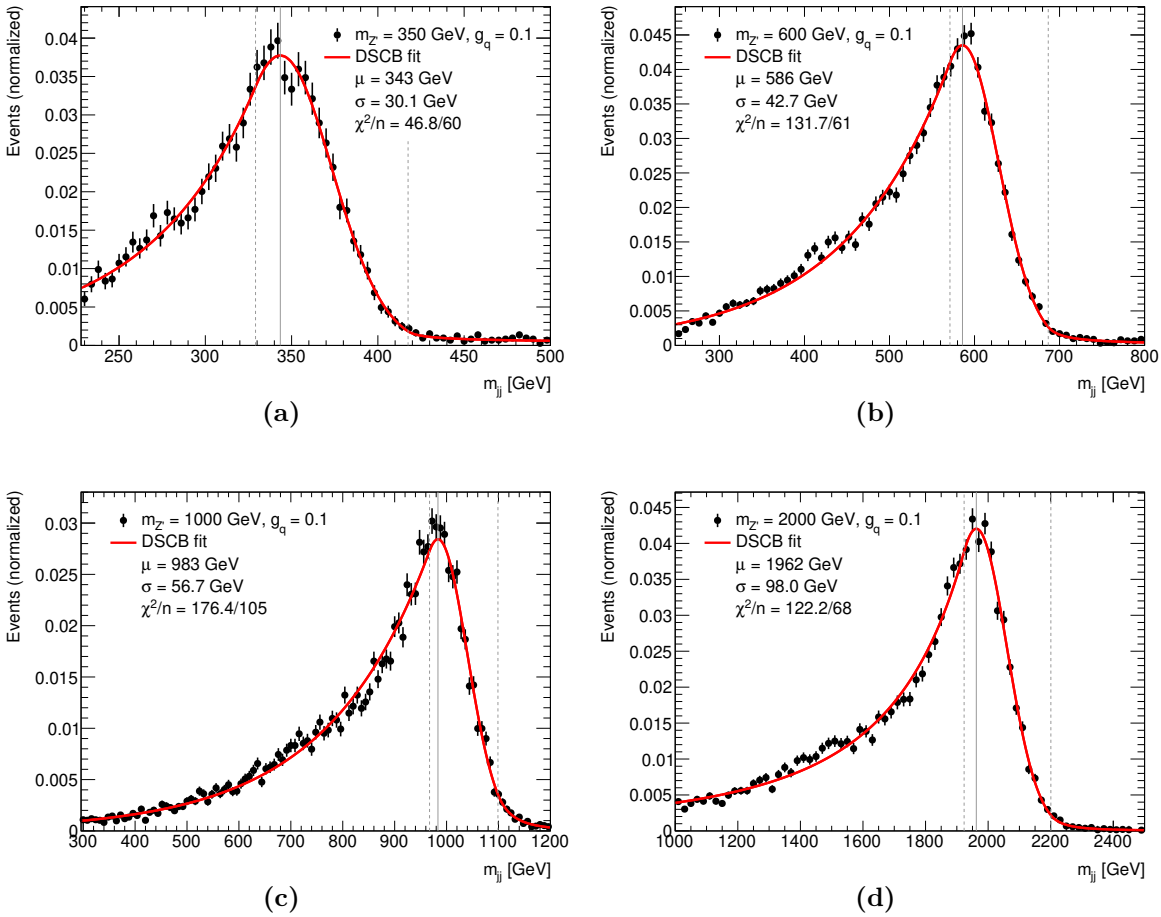


Figure 8.1: The m_{jj} distribution of the Z' signal samples and a double-sided Crystal Ball fit to the line shape. The vertical solid line indicates the best-fit mean μ, while the dashed lines indicate the transition from the Gaussian core to the power-law tails at $\mu - \alpha_L \cdot \sigma$ and $\mu + \alpha_R \cdot \sigma$.

where $x \equiv m_{Z'}/\sqrt{s}$. This function has been found to provide a good description of the Z' total cross section for a large number of tested masses [166]. The parameter p_2 encapsulates the expected $m_{Z'}^{-4}$ dependence of the partonic cross section of the resonant Z' production, combined with a potential correction from the PDFs. The parameter p_3 accounts for an additional decrease at high $m_{Z'}$ due to the PDFs decreasing towards large momentum fractions. This function describes the determined total cross sections well, as is shown in Figure 8.2 for the nominal coupling $g_q = 0.1$.

The total Z' production cross section is corrected to the visible cross section by multiplying it by $A \cdot \varepsilon$, where A is the probability of the event selection criteria to accept Z' signal events (the *acceptance*), and ε is the efficiency of the jet reconstruction. The $A \cdot \varepsilon$ factor is determined for each of the four MC samples corresponding to different $m_{Z'}$ as the ratio of events entering the SR on reconstruction level to the total number of generated events. The results are summarized in Table 8.1. The $A \cdot \varepsilon$ factor is dominated by the acceptance since the reconstruction efficiency of trigger-level jets is above 99.9 % in the considered phase space [166]. The acceptance has a small dependence on $m_{Z'}$ because the p_{T} requirement on the subleading jet and, to a lesser extent, the $|\eta| < 2.4$

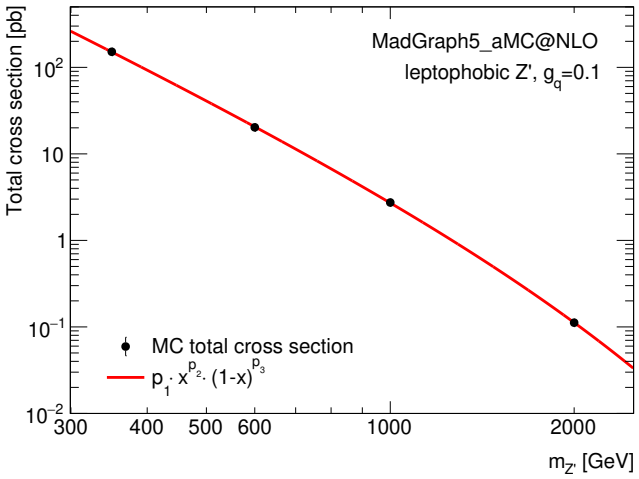


Figure 8.2: Total cross section of the Z' signals tested in this analysis. The indicated uncertainties include statistical and systematic uncertainties, where the latter are dominant. An interpolation in $m_{Z'}$ between the generated MC samples is provided by a fit with the function in Equation (8.1).

selection reduce the acceptance of low-mass Z' signals. The $A \cdot \varepsilon$ factor without the m_{jj} requirement applied is linearly interpolated in $m_{Z'}$. The lower m_{jj} threshold however has a highly non-linear influence on the acceptance: For sufficiently large $m_{Z'}$, the fraction of events with m_{jj} below the lower threshold is negligible. As $m_{Z'}$ approaches and falls below the lower m_{jj} threshold, a relatively steep transition towards an acceptance of zero occurs due to the relatively narrow resonance width. The effect is thus not interpolated linearly, but instead determined by integrating the (interpolated) DSCB function above the applied m_{jj} threshold – either 344 GeV for signals tested with the J50 SR or 481 GeV for signals tested with the J100 SR.

	$A \cdot \varepsilon$		
$m_{Z'}$ [GeV]	Excluding m_{jj} Selection	$m_{jj} > 344$ GeV	$m_{jj} > 481$ GeV
350	36 %	14 %	
600	46 %		33 %
1000	48 %		46 %
2000	48 %		47 %

Table 8.1: Acceptance of the event selection for different Z' masses. The efficiency ε of the jet reconstruction is nearly 1, such that $A \cdot \varepsilon$ is dominated by the acceptance A. Separate values for $A \cdot \varepsilon$ excluding the m_{jj} requirement of the event selection are given to allow for a potential reinterpretation with other signal models producing different m_{jj} distributions. A different m_{jj} threshold is applied based on whether a resonance of a given mass is tested in the J50 or J100 SR.

8.2 Theoretical Uncertainties

The Z' cross sections determined from MC simulations are subject to theoretical uncertainties. Three sources of uncertainty are evaluated following the PDF4LHC recommendations [171].

The strong coupling constant α_S is known to limited precision [17]. To cover its uncertainty, the cross sections are calculated for a nominal value of $\alpha_S(M_{Z^2}) = 0.118$ and for shifts of ± 0.001. A symmetrized uncertainty is then constructed from these values:

$$\delta^{\alpha_S}(\sigma) = \frac{\sigma(\alpha_S = 0.119) - \sigma(\alpha_S = 0.117)}{2}\,. \tag{8.2}$$

The effect of missing higher orders in the calculation of the hard interaction is estimated by varying the factorization and renormalization scales up and down by a factor of 2 in a seven-point variation:

$$(\mu_F, \mu_R) \times \left\{(1/2, 1/2), (1, 1/2), (1/2, 1), (1, 1), (2, 1), (1, 2), (2, 2)\right\}. \tag{8.3}$$

The scale uncertainty is then approximated by the maximal envelope of the cross section determined with each of these variations. This approach has been found to yield a reasonable estimate for the effect of missing terms at NNLO and higher [89].

The PDF uncertainty is estimated using 100 alternative replicas of the PDF. These are provided by the NNPDF group [84] and encode uncertainties on the measurements entering the PDF fit and the fitting method. The root-mean-square of the cross sections determined with each of those replicas is used as uncertainty.

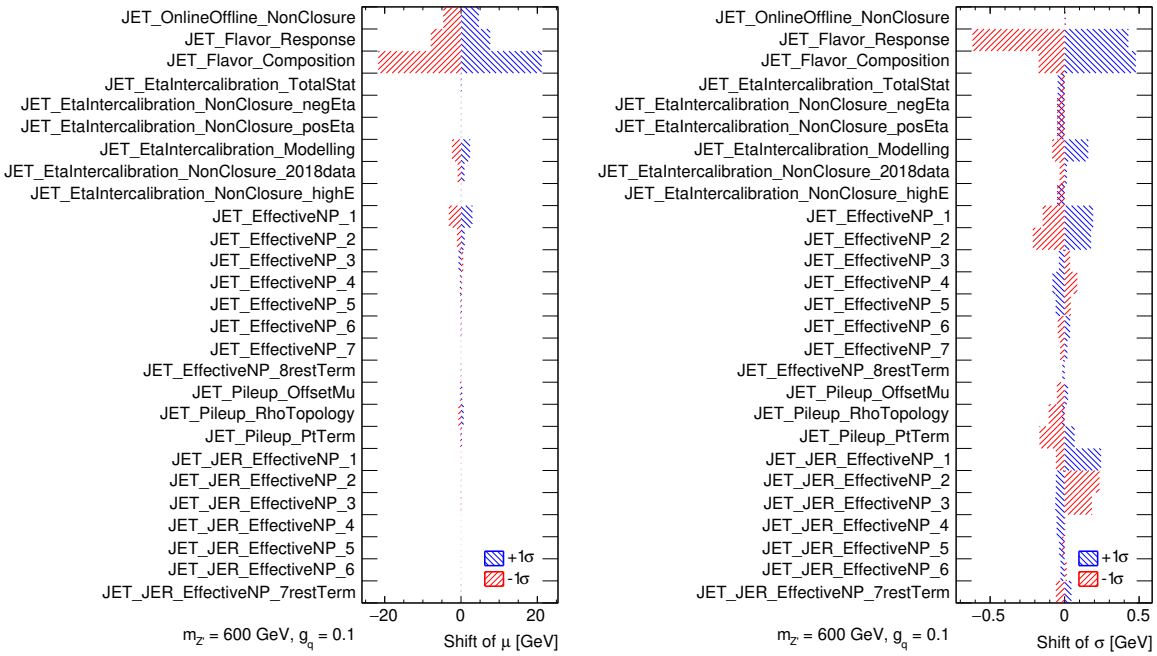
The effect of these systematic uncertainties on the total Z' cross section is evaluated with the parton-level MC generator MCFM 8.0 [172–174] at NLO.[1] As only narrow resonances are studied, the m_{jj} dependence of these uncertainties is sufficiently weak for the effect on the signal shape to be negligible.

Adding the individual theoretical uncertainties in quadrature amounts to a total uncertainty on the Z' total cross sections between $\pm 4\,\%$ at $m_{Z'} = 350\,\text{GeV}$ and $\pm 15\,\%$ at $m_{Z'} = 2000\,\text{GeV}$.

8.3 Experimental Uncertainties

Systematic uncertainties on the jet energy scale and resolution affect the position and shape of a potential resonance. Their effects on the signal templates are evaluated by independently varying the NP for each source of uncertainty up or down by 1σ. For each variation, the DSCB fit is repeated to quantify the dependence of the fit parameters on the NP. The (symmetrized) difference of μ and σ of the Gaussian core from their nominal values is treated as uncertainty. The parameters $\alpha_{\{L,R\}}$ and $n_{\{L,R\}}$ describing the power-law tails are kept at their nominal values, as the shape of the tails is less

[1] MCFM does not provide an implementation of the process $q\bar{q} \to Z' \to q\bar{q}$. Instead, the process $q\bar{q} \to Z' \to \chi\bar{\chi}$+jet is calculated. This is expected to provide a sufficient (over-)estimate of the relative effect of each of the systematic uncertainties on the total Z' production cross section. The effect of this uncertainty on the search sensitivity is found to be small.



(a) **(b)**

Figure 8.3: Impact of the individual systematic uncertainties on (**a**) the mean μ and (**b**) the width σ of the double-sided Crystal Ball fits to the Z' m_{jj} spectrum for $m'_Z = 600\,\text{GeV}$.

significant for the sensitivity and their parameters are more susceptible to fluctuations in the DSCB fits.

Figure 8.3 shows the obtained uncertainties on μ (panel (**a**)) and σ (panel (**b**)), exemplarily for $m_{Z'} = 600\,\text{GeV}$. The results for the other processed Z' masses are given in Appendix C. While the absolute size of each uncertainty increases with $m_{Z'}$, their relative importance remains largely unchanged. The JES uncertainty is dominated by the inflated flavour uncertainties and the residual online-offline jet differences – both of which are large due to the deactivated GSC (see Section 6.4). While the flavour uncertainties decrease from approximately 4.5 % to 2 % over the considered Z' mass range, the uncertainty due to online-offline differences increases approximately linear with $m_{Z'}$, with a relative uncertainty of 0.8–1.1 %. The uncertainty on σ is small, below 1 % for all tested resonance masses. It is also dominated by the flavour uncertainties, followed by the in-situ JES and JER uncertainties. The latter correspond almost entirely to shifts towards larger σ. This is due to the design of the JER NPs to apply additional fluctuations to the jets in MC. Since they can only increase the resolution, the uncertainty in the negative direction is assumed to be symmetric. The determined uncertainties for the other Z' masses are shown in Appendix C, where similar relative compositions are observed.

The correlations between the NPs are found to be below 4 %, and thus negligible, in fits to pseudo-data with injected signals. This means that the ranking of the uncertainties as shown in Figure 8.3 directly corresponds to the ranking of their effect on the sensitivity of this search. Overall, their effect on the search sensitivity is found to be small in comparison to statistical uncertainties, as discussed in Section 10.2.4.

The systematic uncertainties on μ and σ determined this way for Z' signals are applied identically to the generic Gaussian resonances of the respective masses.

9 Background Estimate

The estimation of the QCD background at the very high statistical precision achieved in the selected dijet phase space presents one of the main challenges of this analysis.

The QCD dijet background is a complex composition of many different processes (compare Figures 3.1 and 3.2) whose relative contributions change as a function of m_{jj}, primarily due to the varying quark-gluon ratio in the PDF. The particle-level spectrum is affected by the parton shower and hadronization processes and then additionally convolved with the detector response and effects of the jet reconstruction and calibration.

Due to limited computational resources, it is currently not feasible to simulate these complex processes to the very high statistical precision achieved in this analysis. Using typical background estimation methods that rely on MC would thus introduce a significant systematic limitation to the sensitivity of this search.

Instead, the smoothly falling QCD background is estimated by fitting the observed m_{jj} spectrum. The fit must be flexible enough to describe the QCD background but rigid enough to avoid accommodating narrow resonances in the spectrum. This analysis uses two complementary fit methods: a functional form fit and a novel, MC-based template fit called NLOFit.

The development of two complementary methods has significant advantages:

Blinding resilience. This analysis is performed *blinded* to reduce experimenter's bias [175]. In this context, blinded means that the procedure for deriving the background estimate is defined before testing the fit performance on the analyzed data set. The availability of two methods reduces the risk of a failed background estimate at the *unblinding* stage, as it cannot be guaranteed a priori whether an empirically developed fit method also describes a data set of previously unachieved statistical precision.

Cross-validation. Since the fit methods are of empirical nature, a possible observation of an excess could be interpreted as either a genuine BSM signal or a failure of the fit to describe the QCD background. Achieving two consistent background estimates with complementary methods reduces this ambiguity in favour of a true signal.

9.1 Functional Form Fit

The term *functional form* refers to the algebraic form of a relationship between a dependent variable – in this case, the expected event density ν – and independent variable(s) – in this case, m_{jj}. For this analysis, the N-parameter dijet function f_N is used to describe this relationship:

$$\nu\big(m_{jj}\big) = f_N(x) = p_1(1-x)^{p_2} x^{\sum_{k=3}^{N} p_k \log^{k-3}(x)}\,, \tag{9.1}$$

where $x \equiv m_{jj}/\sqrt{s}$ and $p_1, ..., p_N$ are free parameters. While this functional form is loosely inspired by terms with powers of x and $(1-x)$ appearing in PDF parameterizations

[85, 176, 177], it is ultimately a heuristic description. With $N = 4$ or $N = 5$, it has been empirically found to provide a good description of the dijet mass spectrum measured in different phase spaces and/or at lower statistical precision – either in *global fits* to the entire analyzed m_{jj} spectrum [2, 4, 178] or in *sliding-window fits*, where only a limited range of the m_{jj} spectrum is fitted at a time [1, 3].

In this analysis, a global fit to the m_{jj} spectrum is performed in the range 344–1516 GeV for the J50 SR and 481–2997 GeV for the J100 SR. While this method imposes stricter requirements on the smoothness of the m_{jj} spectrum compared to the sliding-window fit used in the previous TLA publication [1], it enhances the sensitivity to wide resonances for which previously the window width was a limiting factor [166, 167].

9.1.1 Likelihood Maximization

The fit is performed by maximizing the binned likelihood

$$\mathcal{L}_{\text{func}}(\boldsymbol{n} \,|\, \boldsymbol{p}) = \prod_{\text{bins } i} P\big(n_i \,\big|\, \nu_i(\boldsymbol{p})\big)\,, \tag{9.2}$$

$$\nu_i(\boldsymbol{p}) = \int_{\text{bin } i} f_N(x; \boldsymbol{p})\,\mathrm{d}x\,, \tag{9.3}$$

where n_i is the number of observed events in m_{jj} bin i, $\boldsymbol{p}$ is the vector of the free parameters of the dijet function $f_N(x; \boldsymbol{p})$, and P denotes the Poisson distribution. The expected number of events ν_i in bin i is given by the integral of f_N over the bin.

In this analysis, the RooFit framework [179] is used, which approximates the integral of f_N over a bin i by the value of f_N at the bin centre multiplied by the bin width. To minimize the error this approximation introduces, all functional form fits in this thesis are performed in narrow m_{jj} bins of 1 GeV within which the function behaves approximately linearly.

Computationally, it is advantageous to minimize the negative logarithm of the likelihood rather than maximizing the likelihood directly. However, the evaluation of the logarithm of the Poissonian distribution for the very high event counts present in this analysis is limited by the numerical accuracy when summed over $\mathcal{O}(1000)$ bins.[1] Therefore, the Poissonian distribution $P(n_i \,|\, \nu_i)$ is approximated by a Gaussian distribution $G\big(n_i \,\big|\, \nu_i, \sqrt{\nu_i}\big)$ with mean ν_i and standard deviation $\sqrt{\nu_i}$, where the difference is negligible for the large values of ν_i in this analysis. With this approximation, minimizing the negative log-likelihood is analytically identical to minimizing χ^2, as constant terms can be omitted in the

[1] Following the author's initiative, recent versions of RooFit [180] introduce a numerically more robust implementation of the log-likelihood evaluation. In future analyses, it is expected that the approximation of the Poisson distribution as a Gaussian will no longer be necessary, allowing the fit of spectra with both high and low event counts.

minimization:

$$
\begin{aligned}
-\log \mathcal{L}_{\text{func}}(\boldsymbol{n} \,|\, \boldsymbol{p}) &= -\sum_{\text{bins}\, i} \log P\big(n_i \,\big|\, \nu_i(\boldsymbol{p})\big) \\
&\approx \frac{1}{2} \sum_i \frac{\big(n_i - \nu_i(\boldsymbol{p})\big)^2}{\nu_i(\boldsymbol{p})} + \text{const} \\
&= \frac{1}{2}\chi^2(\boldsymbol{n}, \boldsymbol{p}) + \text{const}\,. \qquad (9.4)
\end{aligned}
$$

9.2 Monte Carlo-Based Fit

The MC-based NLOFit background estimate is developed as an alternative to the functional form fit with a heuristic function. The underlying idea is to describe the observed m_{jj} spectrum by an MC prediction of the QCD background and its dependence on uncertain simulation parameters. A constellation of these parameters that provides a good description of the data is determined in a fit. Since the QCD prediction must have a statistical precision better than the data, the computationally intensive step of the parton showering, hadronization, and detector simulation are omitted. However, their effects on the m_{jj} spectrum can be absorbed into the freedom of the simulation parameters and a good description of the data is still achieved.

Fit Method

The NLOFit considers a nominal differential cross section calculated at NLO parton-level and a set of variations generated by changing α_s, the renormalization and factorization scales, and the PDF eigenvectors within their uncertainties. The QCD background is estimated from these cross section templates in a procedure that is illustrated in Figure 9.1. The nominal spectrum, shown in blue, is scaled with a free parameter p_0. Each variation's deviation from the nominal prediction is considered a template, shown in red. Each of these templates is scaled with a parameter $p_{k>0}$ and added to the estimate. While p_0 is a truly free parameter in the fit, $\{p_{k>0}\}$ are constrained in the fit and treated like nuisance parameters of systematic uncertainties. Ultimately, a binned likelihood-based fit to the data optimizes the parameters $\{p_k\}$.

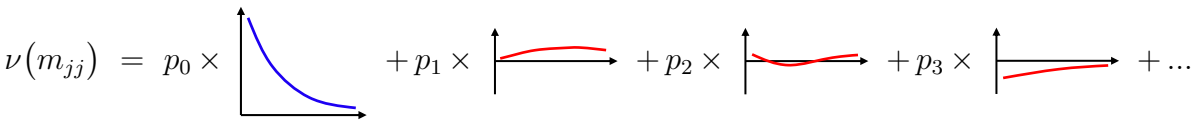
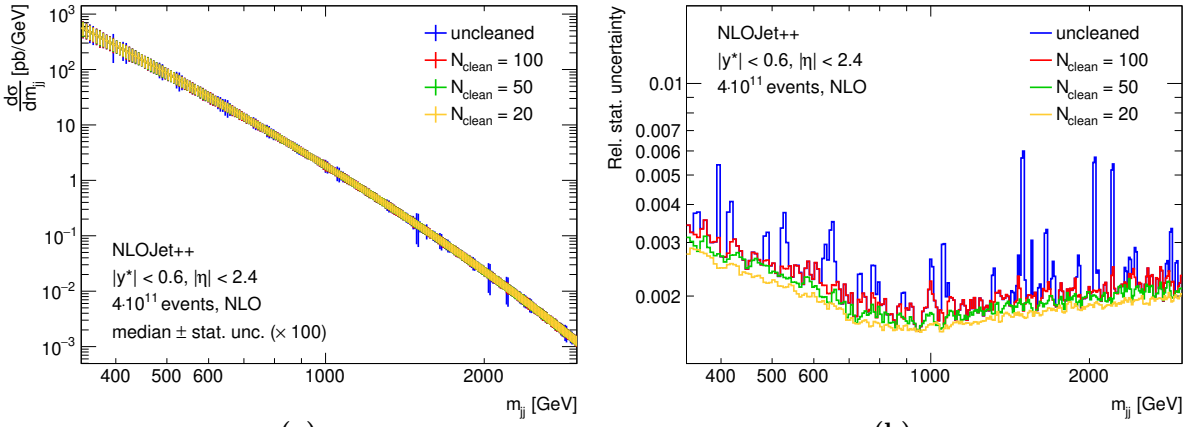


Figure 9.1: NLOFit method: The dijet invariant mass spectrum $\nu(m_{jj})$ is described by a nominal template (blue) scaled by a free parameter p_0 and a sum of relative variations (red) of the nominal template, each scaled by a constrained nuisance parameter $p_{k>0}$ (red).



(a) (b)

Figure 9.2: Effect of different choices of N_{clean} in the cleaning applied to the NLOJET++ results, determined in ≈1 % of the final statistics at NLO. (**a**) Median of the differential cross section per m_{jj} bin and its statistical uncertainty, inflated by a factor of 100 for visibility. (**b**) Relative statistical uncertainty of the differential cross section.

9.2.1 Template Generation

Parton-level dijet cross sections are calculated at LO and NLO with NLOJET++ as detailed in Section 3.3. The kinematic selection of this search is applied to the two leading jets: $|\eta| < 2.4$, $p_{\text{T}} > 85\,\text{GeV}$ and $|y^*| < 0.6$.

Event Generation

$5.2 \cdot 10^{13}$ events are generated to achieve the desired statistical precision, split into individual runs of NLOJET++ with $2 \cdot 10^7$ events each. In any one of these runs, it can rarely occur that an event with excessively large positive or negative weight is generated, leading to a very large fluctuation in one or two bins of the resulting differential cross section. This could cause a non-smoothness in the final templates that is too large given the requirements of this search, even with the very large number of generated events. Thus, a cleaning procedure described below is applied.

In a small-scale generation, the median and RMS_{99}[2] of the cross section prediction from individual runs of NLOJET++ are determined for each m_{jj} bin. Both the median and RMS_{99} are almost unaffected by potential large-weight outliers in this first, small-scale generation. Subsequent runs are discarded if, in any of the 300 m_{jj} bins, the calculated cross section deviates from the predetermined median by more than $N_{\text{clean}} \cdot \text{RMS}_{99}$. Figure 9.2 shows the effect of different choices for N_{clean}. Fluctuations induced by large weights are significantly reduced already by a loose cleaning with $N_{\text{clean}} = 100$, and the remaining non-smoothness is mostly removed at $N_{\text{clean}} = 50$. Stricter cleaning reduces the RMS further by removing events from the tails of the weight distribution at the expense of a larger fraction of discarded runs and a potentially stronger bias on the final shape of the m_{jj} spectrum. For the templates used in this thesis, $N_{\text{clean}} = 50$ is chosen

[2] To not be affected by potential large-weight fluctuations in this bin, the RMS is evaluated only using the central 99% of the distribution. This quantity is referred to as RMS_{99} from here on.

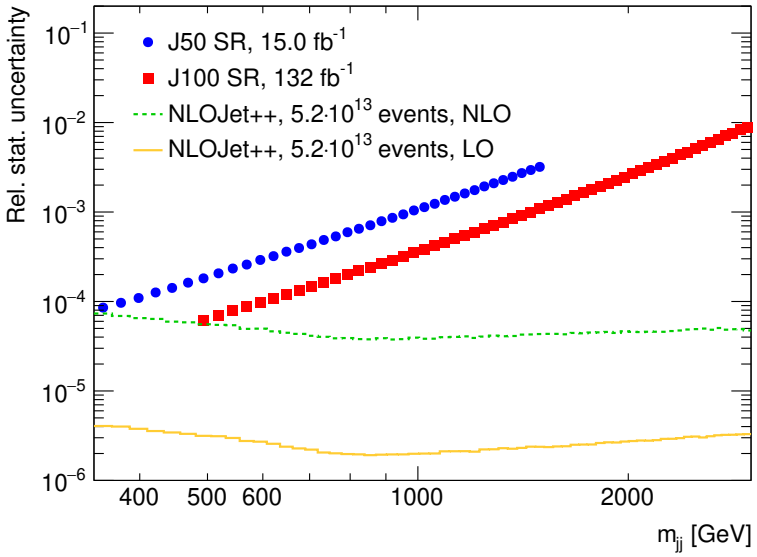


Figure 9.3: Relative statistical uncertainty of the differential dijet cross section determined with NLOJET++ at LO (solid yellow) and NLO (dashed green) accuracy. It is compared to the statistical uncertainty of the observed dijet mass spectrum in the J50 SR (blue circles) and J100 SR (red squares).

as the cleaning parameter to sufficiently remove large-weight fluctuations while keeping the bias of the predicted cross section minimal. This cleaning leads to a rejection of approximately 1.5 % of the NLOJET++ runs.

The achieved relative statistical precision of the NLOJET++ cross section is shown in Figure 9.3, where the solid yellow line depicts the LO precision and the dashed green line the NLO precision. It is approximately constant in m_{jj} due to the importance sampling applied in NLOJET++. The NLO statistical precision of the NLO calculation is a factor of $15 - 20$ lower than that of the LO calculation due to large opposite sign virtual and real corrections cancelling each other out. The NLOJET++ precision is compared to the statistical precision achieved in data in the J50 SR (blue) and J100 SR (red). This shows that the choice of $5.2 \cdot 10^{13}$ generated events achieves a statistical precision of the NLO prediction at least as good as in data throughout the m_{jj} spectrum.

NLO Reweighting

Templates with statistical fluctuations comparable to those of the analyzed data set are not sufficiently smooth for a well-performing fit as they could still result in data-fit deviations of a few standard deviations. A reweighting is applied to utilize the superior statistical precision of the LO template and the more accurate shape of the stronger fluctuating NLO template.

The ratio of the NLO and the LO cross section is fitted with a fifth-order polynomial in the m_{jj} range of interest. The fit describes this ratio well, as it is a smooth, monotonic function of m_{jj} that deviates from 1 by 30 % at most. The fit is then used to reweight the LO template bin-by-bin in m_{jj} to the more accurate NLO shape, which ensures that

the smoothness of the LO template is retained.

Additionally, the templates are scaled by the bin widths to transfer them from a differential cross section $\frac{\mathrm{d}\sigma}{\mathrm{d}m_{jj}}$ to the expected event count ν_i in an m_{jj} bin i per unit luminosity.

Systematic Variations

Systematic variations of the LO and NLO predictions are generated with APPLGRID as discussed in Section 3.3. The LO-to-NLO reweighting is applied individually to all of the varied templates. Each source of theoretical uncertainty is considered as a nuisance parameter with potentially asymmetric upward and downward variations $\delta_i^{\pm}$. These are expressed relative to the nominal LO-to-NLO reweighted prediction ν_i^0, where i iterates over the m_{jj} bins.

α_S uncertainty. While the nominal prediction is calculated with $\alpha_S = 0.118$, varied templates are produced for shifts of ± 0.001 to α_S:

$$\delta_i^{\alpha_S +} = \nu_i(\alpha_S = 0.119) - \nu_i^0 \,, \tag{9.5}$$

$$\delta_i^{\alpha_S -} = \nu_i(\alpha_S = 0.117) - \nu_i^0 \,. \tag{9.6}$$

To first order, α_S linearly scales the total cross section according to Equation (2.11). Consequently, the α_S uncertainty is approximately constant in m_{jj} with a magnitude of $\pm 1\,\%$.

Scale uncertainty. In the nominal cross section calculation, the factorization and renormalization scales are both set to $\mu = m_{jj}$ on an event-by-event basis to reflect the typical energy scale of each event. Alternative predictions are produced by varying μ_F and μ_R by factors of 2 around this nominal scale according to the seven-point variation in Equation (8.3). These seven predictions are merged into a single, symmetric nuisance parameter by creating their envelope of the maximal deviation from the nominal prediction:

$$\delta_i^{scale} = \max_{\substack{(\mu_F,\mu_R) \in \\ 7\text{-point}}} \left\{ \left| \nu_i(\mu_F, \mu_R) - \nu_i^0 \right| \right\} . \tag{9.7}$$

The symmetric approach is chosen here to prevent a non-smoothness in the template that could arise from using an asymmetric envelope. The scale variation is the largest of the studied variations, reaching up to $\pm 9\,\%$ at high m_{jj}.

PDF uncertainty. In this analysis, the CTEQ14NNLO PDF [85] is used for the NLOJET++ cross section calculation. This PDF is derived from a fit with 28 free parameters to a large number of fixed target, lepton-proton, and proton-(anti-)proton measurements. Uncertainties both from the measurements and theory predictions entering the PDF fit are propagated to correlated uncertainties on these 28 parameters. The result is diagonalized using the Hessian method [101] so that the resulting 28 eigenvectors can be varied independently and the total PDF uncertainty corresponds to the squared

sum of these variations.[3]

A set of 56 alternatives to the nominal CTEQ14NNLO PDF is published, each corresponding to an eigenvector's upward or downward variation. The cross section calculation is repeated for each of these alternative PDFs, resulting in 28 asymmetric upwards and downwards varied templates:

$$\delta_i^{\mathrm{PDF},k+} = \nu_i(\text{PDF EV}_k+) - \nu_i^0\,, \tag{9.8}$$

$$\delta_i^{\mathrm{PDF},k-} = \nu_i(\text{PDF EV}_k-) - \nu_i^0\,, \tag{9.9}$$

where $k \in \{1..28\}$ denotes the index of a PDF eigenvector EV.

The variation of the PDF parameters changes the parton luminosities [183] and, accordingly, the contribution of the different partonic subprocesses to the dijet cross section as a function of the momentum fraction x. This can be a non-linear effect, translating into asymmetric upward and downward variations and a variety of different shapes in m_{jj}. Due to the diagonalization into eigenvectors, the individual variations no longer correspond to physically meaningful parameters.

Overview. Ultimately, 30 mostly asymmetric upward and downward variations are considered as templates in the NLOFit in addition to the nominal prediction. They are summarized in Table 9.1. An examplary selection of four of these varied templates $\{\delta^k\}$ is illustrated in Figure 9.4 relative to the nominal prediction ν^0.

Uncertainty Group	Nuisance Parameters	
α_S	1	asymmetric
scale	1	symmetric
PDF	28	asymmetric

Table 9.1: Overview of the asymmetric nuisance parameters considered for the different sources of theoretical uncertainty in the NLOFit templates.

The magnitude of the variations is significantly larger than the statistical uncertainties. However, these should not be compared directly, because the variations are not to be understood as a systematic uncertainty envelope around the true prediction. Instead, they serve as templates for a fit. As such, their absolute size is less important than their smoothness, which means that their bin-to-bin fluctuations must be small in comparison to the statistical uncertainty in data.

[3] The ABMP16NNLO [181] and MMHT2014NLO [177] PDFs, which also provide Hessian eigenvector variations, were tested as alternatives. The background estimate achieved with them is compatible. PDFs that use a large number of toy replicas to parameterize their uncertainties, like NNPDF [182], cannot be used for the NLOFit due to their lack of statistically independent fit parameters.

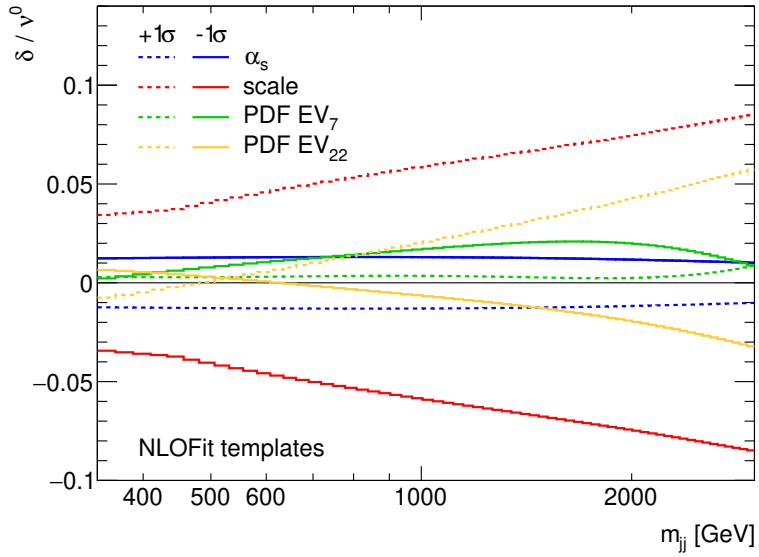


Figure 9.4: Illustration of the systematically varied templates δ used in the NLOFit relative to the nominal template ν^0. The upward (solid) and downward (dashed) variation of α_s and the scale choice are shown in blue and red. Out of the 28 PDF variations, two eigenvectors are selected exemplary and shown in green and yellow.

9.2.2 Likelihood Maximization

The NLOFit is performed by maximizing the binned likelihood, similar to the functional form fit:

$$\mathcal{L}_{\text{NLOFit}}(\boldsymbol{n} \,|\, \boldsymbol{p}) = \prod_i P\big(n_i \,\big|\, \nu_i(\boldsymbol{p})\big) \prod_{\text{vars } k>0} g(p_k)\,, \tag{9.10}$$

$$\nu_i(\boldsymbol{p}) = p_0 \cdot \nu_i^0 + \sum_{\text{vars } k>0} p_k \cdot \delta_i^k\,, \tag{9.11}$$

where the expected event count ν_i in bin i is constructed from the nominal template ν_i^0 and the sum of all variations $\{\delta_i^k\}$, as illustrated in Figure 9.1. These templates are scaled by fit parameters $\{p_k\}$, where the normalization p_0 is a free parameter and the nuisance parameters $\{p_{k>0}\}$ are constrained to be close to zero by a constraint term g.[4]

The likelihood $\mathcal{L}_{\text{NLOFit}}$ is evaluated in the m_{jj} resolution binning derived for this analysis as no integration of a function is involved in its evaluation. The same Gaussian approximation of the Poissonian distributions as used for the functional form fit in

[4] The term $p_k \cdot \delta_i^k$ in Equation (9.11) refers to the simplest case of symmetric upwards and downwards varied templates of a variation k. For asymmetric variations, an exponential extrapolation and a polynomial interpolation between the three known points is applied to ensure continuous first and second derivatives at $p_k = 0$, facilitating the numerical maximization of the likelihood. The extrapolation and interpolation are implemented with HistFactory [184].

Equation (9.4) is applied to the negative log-likelihood of the NLOFit:

$$\begin{aligned} -\log \mathcal{L}_{\text{NLOFit}}(\boldsymbol{n} \,|\, \boldsymbol{p}) &= -\sum_{\text{bins } i} \log P\big(n_i \,\big|\, \nu_i(\boldsymbol{p})\big) - \sum_{\text{vars } k>0} \log g(p_k) \\ &\approx \frac{1}{2}\chi^2(\boldsymbol{n}, \boldsymbol{p}) - \sum_{\text{vars } k>0} \log g(p_k) + \text{const} \end{aligned} \tag{9.12}$$

9.2.3 Constraint Terms

In a likelihood, the constraint term $g(p_k)$ encodes the external knowledge about the behaviour of a systematic uncertainty that is described by the nuisance parameter p_k. Typically, this knowledge stems from external measurements or theoretical considerations that provide a nominal value and uncertainty for p_k. The constraint term g is typically assumed to be a Gaussian distribution for experimental uncertainties or a uniform distribution for theoretical uncertainties.

In the NLOFit, external knowledge enters the generation of the varied templates in the form of the uncertainties on α_S and the PDF as well as the envelope of the scale choices, which estimates the uncertainty from missing higher order terms. These correspond to the uncertainties on the determined parton-level dijet cross section.

However, the NLOFit as a background estimate is not used to describe a parton-level spectrum but to describe the observed dijet mass spectrum that includes parton showering and detector effects. The nominal NLOFit prediction and its $\pm 1\sigma$ bands are not expected to describe the observed spectrum well. Instead, the generated variations are merely understood as templates that have the potential to describe a smoothly falling dijet mass spectrum. Additional freedom is granted to the considered parameter space of the $\{p_k\}$ to compensate for the missing effects in the generation of the templates. The NLOFit can be considered as a physically motivated background estimation technique for QCD spectra, but it is not to be confused with PDF fits where the best-fit values offer a physical interpretation.

In this thesis, uniform distributions are used for the constraint terms $g(p)$ of the NLOFit:

$$g(p_k) = \begin{cases} \frac{1}{2\sigma} & \text{if } -\sigma \leq p_k \leq \sigma\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{9.13}$$

where σ is a tunable parameter that describes the allowed parameter space for the $\{p_k\}$ and thus determines the number of degrees of freedom (ndf) of the NLOFit.[5] In this way, it is analogous to the number of free parameters N of the functional form fit in its ability to tune the flexibility of the fit.

With the choice of constant constraint terms, Equation (9.12) simplifies to:

$$-\log \mathcal{L}_{\text{NLOFit}}(\boldsymbol{n} \,|\, \boldsymbol{p}) \approx \frac{1}{2}\chi^2(\boldsymbol{n}, \boldsymbol{p}) + \text{const} \;\Big|\; p_k \in [-\sigma, \sigma] \;\; \forall \; k > 0\,. \tag{9.14}$$

[5] Gaussian constraint terms were tested as well, where σ could be tuned to produce fit results very similar to those achieved with the uniform constraint terms. Uniform constraints are used in this thesis as they produce a stronger dependence of ndf on σ (discussed in the following section).

Degrees of Freedom

To evaluate the goodness of fit from χ^2, ndf of the fit needs to be known. In general, it is given by the difference between independent observations – in this case, the number of m_{jj} bins considered in the fit – and the number of free fit parameters. While this relation is well defined for the functional form fit, the same does not hold true for the NLOFit due to its fit parameters being constrained.

Instead, ndf of the NLOFit is determined from the distribution of χ^2-values achieved in toy experiments for different choices of the constraint strength σ. The following procedure is used:

- An NLOFit with constraint σ is performed to the observed dijet mass spectrum separately in each SR to obtain the best-fit description of the data. This step does not introduce experimenter's bias since the fit residuals and goodness of fit remain blinded at this stage.

- 10^4 pseudo-data sets per SR are generated by drawing Poisson-fluctuated distributions from the best fit as the underlying true distribution.

- NLOFits with constraint σ are performed to all pseudo-data sets. The same choice of σ in the generation of the reference distribution and in the fit to the fluctuated pseudo-data set ensures that fit residuals are purely due to statistical fluctuations and not due to an inability to describe the underlying shape of the m_{jj} spectrum.

For both SRs and all tested values of σ, the distribution of the χ^2-values determined from the toy experiments is well described by the χ^2-distribution, which describes the probability density of observing a value of χ^2 if only statistical fluctuations are present:

$$f_{\chi^2}(\chi^2; \mathrm{ndf}) = \frac{\chi^{\mathrm{ndf}-2}\, e^{-\chi^2/2}}{2^{\mathrm{ndf}/2}\, \Gamma(\mathrm{ndf}/2)}\,. \tag{9.15}$$

Fitting $f_{\chi^2}(\chi^2; \mathrm{ndf})$ to the distribution allows for the extraction of ndf for the NLOFit. Figure 9.5 shows the toy distributions corresponding to the J100 SR, with $\sigma = 1$ and $\sigma = 10$ chosen as examples, and the fit of $f_{\chi^2}(\chi^2; \mathrm{ndf})$ to them.[6]

As expected, increasing σ, i.e. increasing the available parameter space for the NLOFit, leads to a higher number of effective parameters of the NLOFit, which manifests as a lower $\mathrm{ndf} = N_{\mathrm{bins}} - N_{\mathrm{pars}}$. The dependence of ndf extracted from the χ^2-distribution fits on σ is shown in Figure 9.6a for the toy experiments corresponding to the J50 SR and in Figure 9.6b corresponding to the J100 SR. A similar behaviour is observed in both cases: ndf decreases approximately logarithmically with σ. Fitting a logarithmic function

$$\mathrm{ndf}(\sigma) = a \cdot \log(\sigma) + b \tag{9.16}$$

reveals a similar scaling factor a for both SRs, while the offset b differs as expected due to the different number of m_{jj} bins considered in both SRs. Fluctuations of the

[6] Following the same procedure with toy experiments for the functional form fit validates that its results are well described by a χ^2-distribution with $\mathrm{ndf} = N_{\mathrm{bins}} - N_{\mathrm{pars}}$.

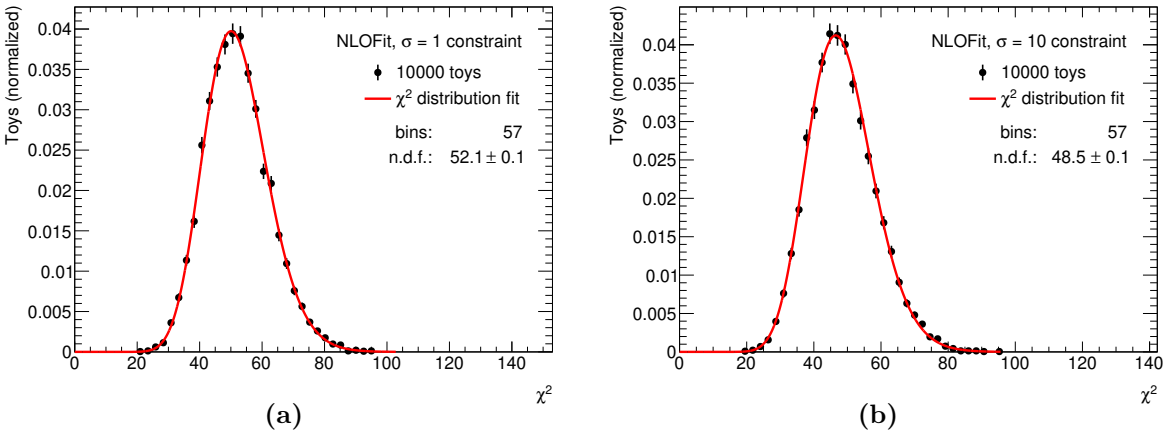


Figure 9.5: The distribution of NLOFit χ^2-values in 10^4 toy experiments of the J100 SR for a constraint of (**a**) $\sigma = 1$ and (**b**) $\sigma = 10$. A χ^2-distribution with free parameter ndf is fitted to the distributions.

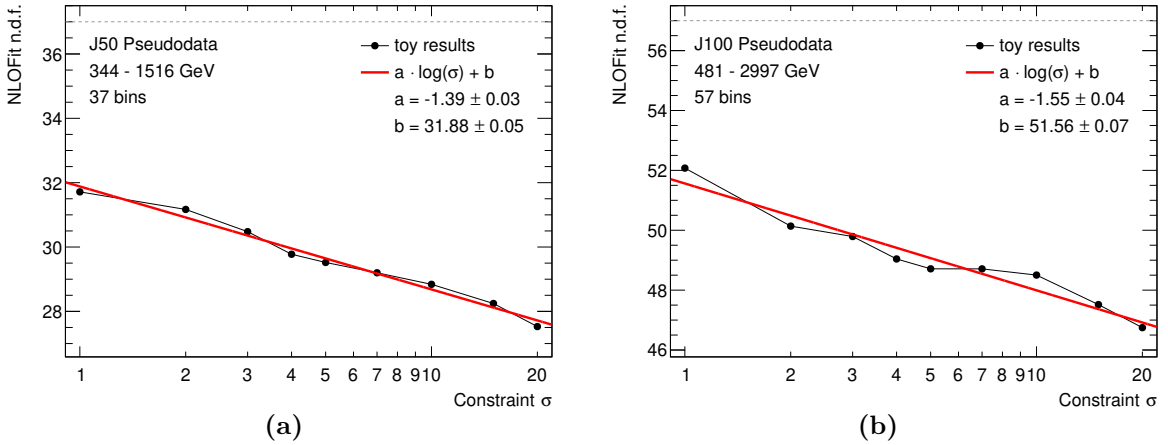


Figure 9.6: The NLOFit ndf as a function of the constraint σ for fits in the (**a**) J50 SR and (**b**) J100 SR. An approximately logarithmic dependence is observed, which is described by a logarithmic fit.

determined ndf around this logarithmic dependence could be attributed to instabilities in the minimization of the high-dimensional negative log-likelihood of the NLOFit.

The determination of ndf for the NLOFit allows calculating p-values to evaluate the goodness of fit by integrating the χ^2-distribution:

$$p(\chi^2) = \int_{\chi^2}^{\infty} f_{\chi^2}(z; \mathrm{ndf})\, \mathrm{d}z\,, \tag{9.17}$$

where for ndf the functional description in Equation (9.16) is used.

The relation ndf(σ) is also used to determine the values of σ that are considered for the NLOFit. Since σ can be any positive rational number, arbitrarily small steps could be considered as distinct fit strategies, which would be unfeasible for the valid-

ation of each strategy. Instead, in analogy to the function form fit, values of σ are chosen such that their spacing corresponds to a difference in ndf of approximately 1, i.e. $\mathrm{ndf}(\sigma') - \mathrm{ndf}(\sigma) = a \cdot \log\big(\sigma'/\sigma\big) \approx -1$. With the values of a obtained from the fits, this would correspond to a spacing of $\sigma'/\sigma \approx 2.05$ for the J50 SR and $\sigma'/\sigma \approx 1.91$ for the J100 SR. For simplicity, the same spacing of 2 is employed for both SRs. Consequently, powers of 2 are selected as values for σ that are considered as possible fit strategies in this thesis.

9.3 Signal-Plus-Background Fits

The previously described fit methods describe the case of fitting a background-only description to data. In the following validation of the background estimate, fits under the signal-plus-background (s+b) hypothesis are performed. The s+b hypothesis corresponds to adding the number of signal events S in the spectrum as a free parameter to the expected number of background events ν^b given in Equations (9.3) and (9.11):

$$\nu_i^{\mathrm{s+b}}(S, \boldsymbol{p}, \boldsymbol{\theta}) = \nu_i^{\mathrm{b}}(\boldsymbol{p}) + S \int\limits_{\mathrm{bin}\, i} s(x; \boldsymbol{\theta})\, \mathrm{d}x\,, \tag{9.18}$$

where i is an m_{jj} bin and $s(m_{jj}; \boldsymbol{\theta})$ is the m_{jj} distribution of the considered signal normalized to unity. Thus, it is either a double-sided Crystal Ball or a Gaussian distribution:

$$s(m_{jj}; \boldsymbol{\theta}) = \mathrm{DSCB}\big(m_{jj}; \mu(\boldsymbol{\theta}), \sigma(\boldsymbol{\theta}), \alpha_L, \alpha_R, n_L, n_R\big) \tag{9.19}$$

or

$$s(m_{jj}; \boldsymbol{\theta}) = G\big(m_{jj}; \mu(\boldsymbol{\theta}), \sigma(\boldsymbol{\theta})\big)\,. \tag{9.20}$$

The mean μ and width σ of either signal are subject to systematic uncertainties, parameterized by the vector of nuisance parameters $\boldsymbol{\theta}$:

$$\mu(\boldsymbol{\theta}) = \mu^0 + \sum_{\mathrm{systs}\, l} \theta_l \cdot \delta_\mu^l \tag{9.21}$$

$$\sigma(\boldsymbol{\theta}) = \sigma^0 + \sum_{\mathrm{systs}\, l} \theta_l \cdot \delta_\sigma^l\,, \tag{9.22}$$

where μ^0 and σ^0 refer to the nominal mean and width of a resonance of a certain mass, and δ_μ^l and δ_σ^l are the (symmetrized) effect of a systematic uncertainty l on the mean and the width of the signal as determined in Chapter 8. The DSCB tail parameters $\alpha_{\{L,R\}}$ and $n_{\{L,R\}}$ are not systematically varied as their influence is small.

For each systematic l, a unit Gaussian distribution $G(\theta_l; 0, 1)$ is multiplied to the likelihood to constrain its nuisance parameter θ_l. The negative log-likelihood for both the functional form fit and the NLOFit then becomes:

$$-\log \mathcal{L}(\boldsymbol{n} \,|\, S, \boldsymbol{p}, \boldsymbol{\theta}) \approx \frac{1}{2}\chi^2(\boldsymbol{n}, S, \boldsymbol{p}, \boldsymbol{\theta}) + \frac{1}{2}\sum_{\mathrm{systs}\, l} \theta_l^2 + \mathrm{const}\,, \tag{9.23}$$

where in the case of the NLOFit, the constrained parameter space on $\boldsymbol{p}$ still needs to be considered as specified in Equation (9.12). With the signal prediction incorporated, the likelihood and the χ^2 statistic of s+b fits now also depend on the expected number of signal events S and the NPs $\boldsymbol{\theta}$ parametrizing the systematic uncertainties.

In this analysis – as in most other di-X resonance searches [3, 4, 185, 186] – systematic uncertainties on the jet energy scale and resolution are only applied to the signal templates and not to the background. As the jet calibration is required to be a smooth function, the effect of all jet uncertainties on the m_{jj} spectrum itself has to be a smooth modulation as well. The background fit absorbs any such smooth modulation. Thus, the sensitivity to resonances atop the QCD background is not impacted, provided the modulation is small enough to not significantly alter the statistical uncertainty of the observed m_{jj} spectrum.

9.4 Unblinding Strategy

It cannot be guaranteed a priori whether either method of the background estimate provides a good description of the observed dijet mass spectrum and, if so, which degree of freedom is necessary for the fit. To a certain extent, fit strategies can be tested without the use of the analyzed data set by applying them to MC samples or to a partial data set corresponding to the sensitivity of the previous TLA publication – both of which were done to study the smoothness of the jet calibration in Section 6.3. However, the m_{jj} spectra constructed from these MC samples or partial data sets have larger statistical uncertainties than the spectrum analysed in this thesis. Successful fits to the former do not guarantee that the N-parameter dijet function or the NLOFit templates also describe the latter.

Ultimately, the background estimate must be validated using the observed m_{jj} spectrum. The procedure to validate and decide on a fit strategy must be defined while the analysis is still blinded to avoid introducing experimenter's bias. The term *fit strategy* here refers to a specific combination of either the functional form fit with a certain number of free parameters N or the NLOFit with a certain constraint strength σ.

The decision on a fit strategy is a three-step procedure:

- The fit strategy must provide a good description of the observed m_{jj} spectrum. This is defined to be the case if the observed $p\big(\chi^2\big)$-value exceeds 0.01.

- Pseudo-data is generated from the observed m_{jj} spectrum. The fit strategy must pass a series of validation tests on the pseudo-data.

- If multiple fit strategies are validated, the one with the lowest number of (effective) parameters is selected for the functional form fit and for the NLOFit.

This decision is performed independently for both SRs since their different m_{jj} range and statistical precision can result in different preferred fit strategies. This way, one functional form fit and one NLOFit strategy for each SR will be chosen, which can then be used for cross-validation.

9.4.1 Application of the Fit Strategies

The first step of the unblinding procedure is to apply the different fit strategies to the observed m_{jj} spectrum in each SR. The $p(\chi^2)$-value (Equation (9.17)) is used as metric for the goodness of fit, where a fit with $p(\chi^2) > 0.01$ is considered to be a good description of the data.

The potential presence of a resonance in the dijet mass spectrum adds additional complexity to this definition: If a resonance is present, the background fit is not expected to provide a good description of the observed m_{jj} spectrum but rather show an excess at the resonance mass, leading to a low $p(\chi^2)$. This case needs to be distinguished from the background fit not properly describing the QCD background, which will also result in a significant disagreement between the fit and the data and, consequently, a low $p(\chi^2)$.

Such a distinction can be made under the assumption that a potential resonance is a narrow feature in the m_{jj} spectrum, whereas the inability to describe the QCD background leads to significant deviations not only in a localized region in the spectrum. The effect of a resonance can thus be mitigated by masking a region of the m_{jj} spectrum from the fit. The masking algorithm is illustrated schematically in Figure 9.7, where panels (**a**) and (**b**) show the case of a resonance and panels (**c**) and (**d**) show the case of a bad description of the QCD background.[7] It is defined as follows:

- If $p(\chi^2) > 0.01$ is found in the initial fit, nothing needs to be done. The fit is valid.
- If $p(\chi^2) < 0.01$ is found in the initial fit, the m_{jj} range corresponding to the most significant excess is identified with the BumpHunter [187, 188] algorithm. This range is masked for a subsequent fit.
  - If $p(\chi^2) > 0.01$ is found in the masked fit, the masked region likely contains a localized excess (Figure 9.7b). The masked fit is valid.
  - If $p(\chi^2) < 0.01$ is found in the masked fit, the bad description of the data is not due to a localized effect (Figure 9.7d). The fit strategy is invalid.

The masked region is not included in the χ^2 evaluation and ndf is corrected by the number of masked bins for the determination of $p(\chi^2)$.

The same masking algorithm will also be applied for s+b fits to ensure that if the spectrum shows a true resonance, signal hypotheses at other masses can still be tested.

The chosen threshold of 0.01 for $p(\chi^2)$ has been previously determined as a suitable choice for the masking decision in toy studies with injected signals in the order of the search sensitivity [166].

Fit Results

Performing the functional form fit to the observed m_{jj} spectra results in a good fit with $p(\chi^2) > 0.01$ without masking in both SRs for a number of parameters $N \geq 4$. The fit results for various choices of $N \geq 4$ are shown in Figure 9.8a for the J50 SR and in Figure 9.8b for the J100 SR. Only the significance of the fit residuals, i.e. the fit residuals in units of the statistical uncertainty per bin, is shown since the residuals are too small to

[7] This algorithm fails if more than one significant resonance is present in the observed m_{jj} spectrum.

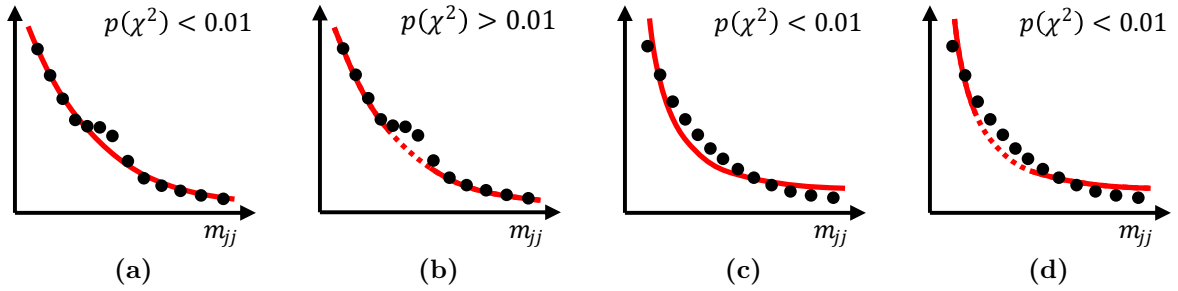


Figure 9.7: Schematic depiction of the applied masking procedure to allow for a successful background fit in the case of a resonance in the m_{jj} spectrum. (**a**) A resonance is present, such that the background fit cannot describe the data well and a low $p(\chi^2)$-value is observed. (**b**) A resonance is present and the area of the resonance has been masked for the fit. In the remaining spectrum, the description of the data is good and a higher $p(\chi^2)$-value is observed. (**c**) No resonance is present, but the fit cannot describe the shape of the QCD background in data. A low $p(\chi^2)$-value is observed. (**d**) No resonance is present and the area of the most significant excess has been masked for the fit. The description of the remaining spectrum is still poor, resulting in a low observed $p(\chi^2)$-value.

be visible when overlaying the observed m_{jj} spectrum and the fit. $p(\chi^2)$-values between 0.62 and 0.93 are observed in the J50 SR, depending on the chosen number of parameters with a notable improvement in the fit quality by introducing the p_5 to the fit. In the J100 SR, $p(\chi^2)$-values between 0.029 and 0.044 are observed, with a small improvement by adding p_5 and negligible effect of p_6 and beyond.

The NLOFit in the J50 SR yields a similar result, with a constraint strength $\sigma \geq 2$ achieving good fits, as shown in Figure 9.8c. $p(\chi^2)$-values between 0.91 and 0.98 are observed, where loosening the constraint beyond $\sigma = 4$ has a negligible effect on the fit result. In the J100 SR, a good fit is achieved for $\sigma = 4$ – but only after the most significant excess in the bin $2403\,\text{GeV} < m_{jj} < 2464\,\text{GeV}$ is masked according to the algorithm discussed above. For $\sigma \geq 8$, a $p(\chi^2) > 0.01$ is observed already in the initial fit without a masking applied. Figure 9.8d shows the residuals of these different fits. The NLOFit with $\sigma = 8$ and the same m_{jj} bin masked is included for comparison despite not being selected by the unblinding algorithm. For the unmasked strategies, a $p(\chi^2)$-value of 0.010 or 0.013 is observed, while the masked $\sigma = 4$ strategy achieves 0.017. Ultimately, the unmasked fits are chosen, as will be discussed in Section 9.5.5.

Table 9.2 summarizes the fit strategies that achieve a good fit according to the unblinding algorithm.

9.5 Fit Strategy Validation

The functional form fit and the NLOFit can be performed with arbitrary flexibility within the previously found constraints of providing a sufficiently good description of the observed m_{jj} spectra. While a method with a higher number of (effective) parameters generally provides a better fit to data, a too-flexible method could lead to a (partial) absorption of a genuine signal into the background estimate. Conversely, a too-rigid method could provide a suboptimal description of the data and thus lead to a bias in

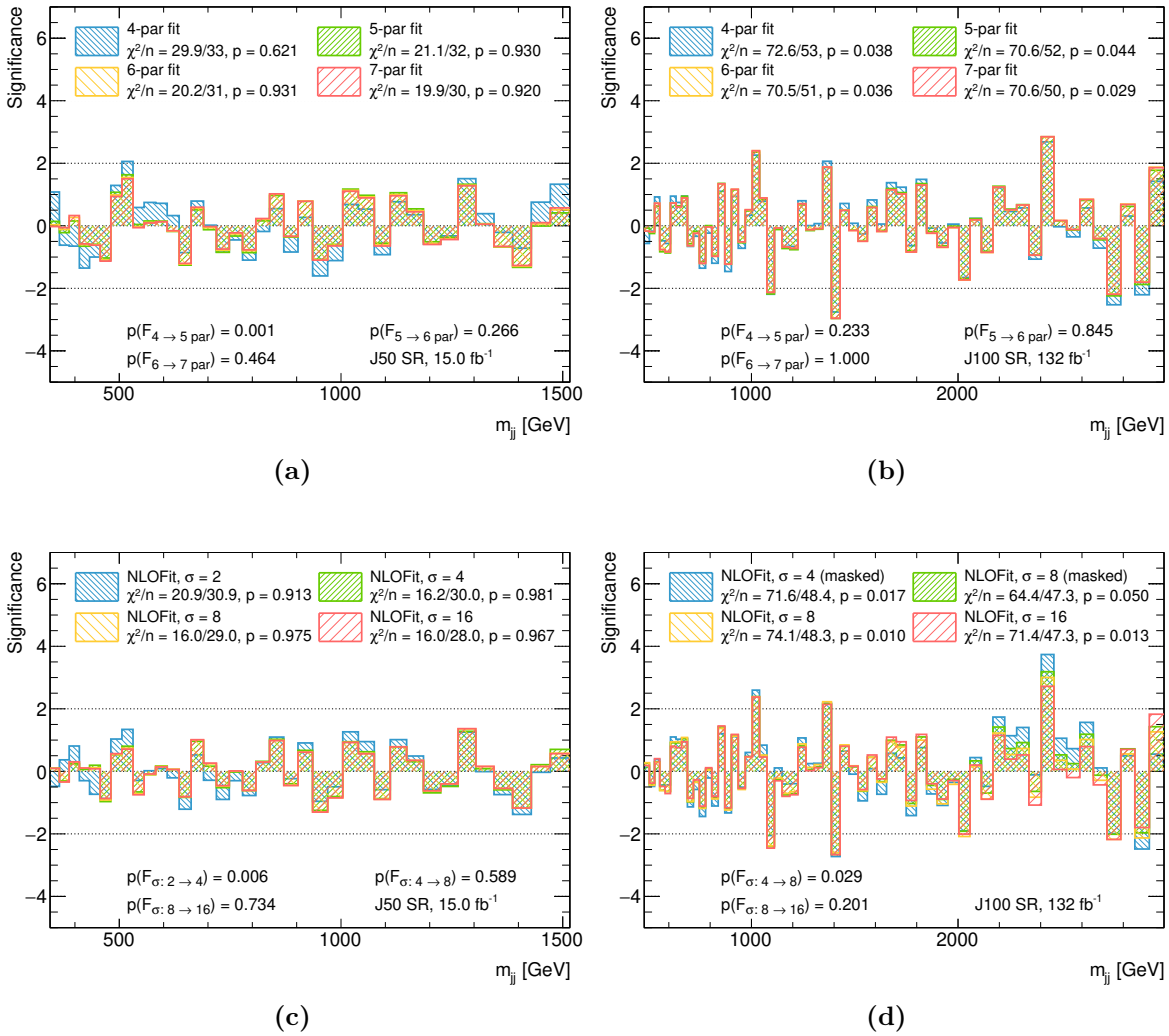


Figure 9.8: Significance of the fit residuals observed by applying various fit strategies to the observed m_{jj} spectra in both SRs. Shown are residuals of the functional form fit (**a**) in the J50 SR and (**b**) in the J100 SR for a number of parameters N between 4 and 7 in different colours and of the NLOFit (**c**) in the J50 SR and (**d**) in the J100 SR for the constraint strength σ between 2 and 16 in different colours. χ^2/ndf and the $p(\chi^2)$-value are indicated for each fit. The $p(F)$-values are discussed in Section 9.5.5.

	Fit Strategies with $p(\chi^2) > 0.01$	
Signal Region	Functional Form Fit	NLOFit
J50	$N \geq 4$	$\sigma \geq 2$
J100	$N \geq 4$	$\sigma \geq 4$ (masked) $\sigma \geq 8$ (not masked)

Table 9.2: Fit strategies that achieve a good description of the observed m_{jj} spectrum in each signal region with $p(\chi^2) > 0.01$ achieved either in the initial fit or after masking a localized excess. The number of parameters N is shown for the functional form fit and the constraint strength σ for the NLOFit.

the amount of signal identified by a fit.

Therefore, the fit bias must be estimated to validate a certain fit strategy. The methods applied in this thesis expand upon the central ATLAS recommendations [189].

9.5.1 Pseudo-Data Generation

The validation of a fit strategy requires a data set that fulfils the following criteria:

- **Statistical precision:** The statistical uncertainties of the validation data set must not be larger than those of the true data set.
- **Accurate m_{jj} spectrum:** The m_{jj} spectrum of the validation data set must accurately represent the shape of the spectrum in the true data set.
- **Absence of signal:** The validation data set must not contain a true dijet resonance to estimate the bias of the fit strategy on a background-like data set.

These criteria can be met by generating pseudo-data (PD) based on a background-like m_{jj} template extracted from data.

To extract such a template, statistical fluctuations must be smoothed out while retaining the underlying shape of the m_{jj} spectrum. This smoothing is achieved by performing a fit with the same method as the fit to be validated: A functional form fit to the observed m_{jj} spectrum serves as the PD template for the validation of the functional form fit strategies. A PD template from an NLOFit is used for the NLOFit validation, respectively. Using the background-only fit (with potential masking applied) ensures that no significant signal contamination is present in the template.

The template must be extracted with a fit of a higher number of (effective) parameters than the fit to be validated to prevent circular reasoning. In this analysis, an increase of the (effective) number of parameters of 1 is chosen. For example, the 4-parameter functional form fit is validated on PD generated from a 5-parameter template and the

$\sigma = 2$ NLOFit is validated with a template from a $\sigma = 4$ NLOFit.[8] If the strategy with a lower number of (effective) parameters can be validated using PD generated from the more flexible template, that means the additional (effective) parameter is not necessary for a sufficient description of the data.

Toy experiments are generated by adding Poissonian fluctuations to the templates extracted in both SRs so that the statistical uncertainties match those in the true data set. From here on, these toys are referred to as PD.

9.5.2 Spurious Signal

The bias of the individual fit strategies is estimated using the background-like PD. s+b fits[9] are performed on the toy experiments to find the distribution of the amount of signal S_{fit} determined in the fits.

The mean of that distribution is labelled *spurious signal* (S_{spur}) and is a direct measure of the fit bias:

$$S_{\text{spur}} \equiv \langle S_{\text{fit}} \rangle_{\text{toys}} \,. \tag{9.24}$$

The standard deviation σ_{fit} of the distribution is an estimator for the uncertainty of the s+b fit.

The fit is expected to perform nearly unbiased, i.e. the spurious signal is supposed to be small compared to the fit uncertainty. The acceptable threshold is set to be

$$\left|S_{\text{spur}}\right| < 0.5\sigma_{\text{fit}} \,. \tag{9.25}$$

This threshold ensures that the additional spurious signal uncertainty will be small in comparison to the inherent fit uncertainty when added in quadrature ($\sqrt{\sigma_{\text{fit}}^2 + (0.5\sigma_{\text{fit}})^2} \approx 1.12\sigma_{\text{fit}}$). Additionally, a larger fit bias could lead to a potential undercoverage of the exclusion limits placed by this analysis due to the likelihood penalty introduced by the Gaussian constraint term of the systematic uncertainty, as discussed in [189].

For each fit strategy, the spurious signal distributions are determined in 1000 toy experiments for all signal hypotheses (Gaussian and Z' signals, all resonance masses and Gaussian widths) considered in this analysis. Figure 9.9 shows the resulting distributions for two exemplary signal models: (**a**) a narrow, 5 %-width Gaussian resonance at 550 GeV in the J50 SR and (**b**) a wide, 15 %-width Gaussian resonance at 1600 GeV. The fit results of the 5-parameter functional form fit (blue) and the $\sigma = 8$ NLOFit (red) are included. Approximately Gaussian distributions are observed in these examples and for

[8] Different smoothing methods, like a narrow sliding-window fit, were tested. Those provide additional hyper-parameters, leading to an ambiguity between a validation of the fit strategy and the smoothing method. Extracting the template from a partial data set was also tested to potentially validate the fit methods before unblinding the analysis. However, this leads to the inflation of fluctuations of the partial data set if the template is to have a higher number of (effective) parameters than the fit to be validated.

[9] These s+b fits include the systematic uncertainties on the signal as described in Section 9.3. Thus, their effect is properly propagated in the evaluation of the fit bias. Additionally, the same masking algorithm as described in Section 9.4 is applied in all toy fits here and in the following validations to test the same fit conditions that are also applied to data.

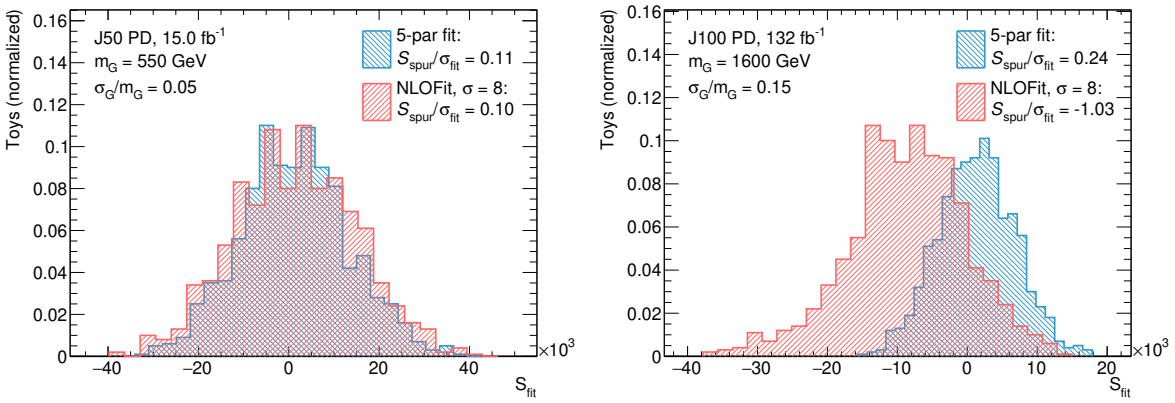


Figure 9.9: Exemplary distributions of the number of signal events determined by s+b fits in background-only PD. Functional form fit results with a 5-parameter dijet function, performed on PD from a 6-parameter template, are shown in blue and NLOFit results with $\sigma = 8$, performed on PD from a $\sigma = 16$ template, are shown in red. The chosen signal hypotheses are: (**a**) Gaussian resonance with mass 550 GeV and 5 % width and (**b**) Gaussian resonance with mass 1600 GeV and 15 % width.

all other tested signal hypotheses, indicating good fit stability. The examples highlight a case of good agreement between the functional form fit and the NLOFit with a small bias in Figure 9.9a. However, Figure 9.9b shows a case of small bias for the functional form fit but a significantly higher fit uncertainty and bias observed with the NLOFit, where the requirement in Equation (9.25) is not met.

Figure 9.10 summarizes S_{spur} and σ_{fit} determined for the functional form fit with various choices of the number of parameters N in the J50 SR (left column) and in the J100 SR (right column). The bias and uncertainty in terms of absolute event numbers decrease with the resonance mass due to the QCD background falling with m_{jj}. Simultaneously, narrower resonances exhibit smaller uncertainties than wider resonances. This is primarily due to a resonance spreading across a wider m_{jj} range, resulting in a larger QCD background in the relevant signal window. To a lesser extent, the tendency of background fit to adapt to wide resonances better than to narrow resonances leads to an increased ambiguity between signal and background component for wide resonances and, hence, an increased uncertainty.

The ratio $S_{\text{spur}}/\sigma_{\text{fit}}$ is shown in each figure to judge the compliance with the spurious signal criterion in Equation (9.25). Significant spurious signals that exceed the 50 % threshold are observed for the 4-parameter functional form fit in both SRs (top row of Figure 9.10). The mass dependence of the fit bias corresponds to the difference between the 4-parameter and 5-parameter fits to the observed m_{jj} spectra, which is especially large in the J50 SR (see Figures 9.8a and 9.8b). When generating the 5-parameter PD template, this template is not entirely describable by a 4-parameter fit, with the difference showing up as spurious signal. Although this difference is seemingly small in the J100 SR, it can still result in a non-negligible spurious signal.

The 5- (middle row) and 6-parameter (bottom row) functional form fits fulfil the spurious signal requirement, which is consistent with the negligible difference between the

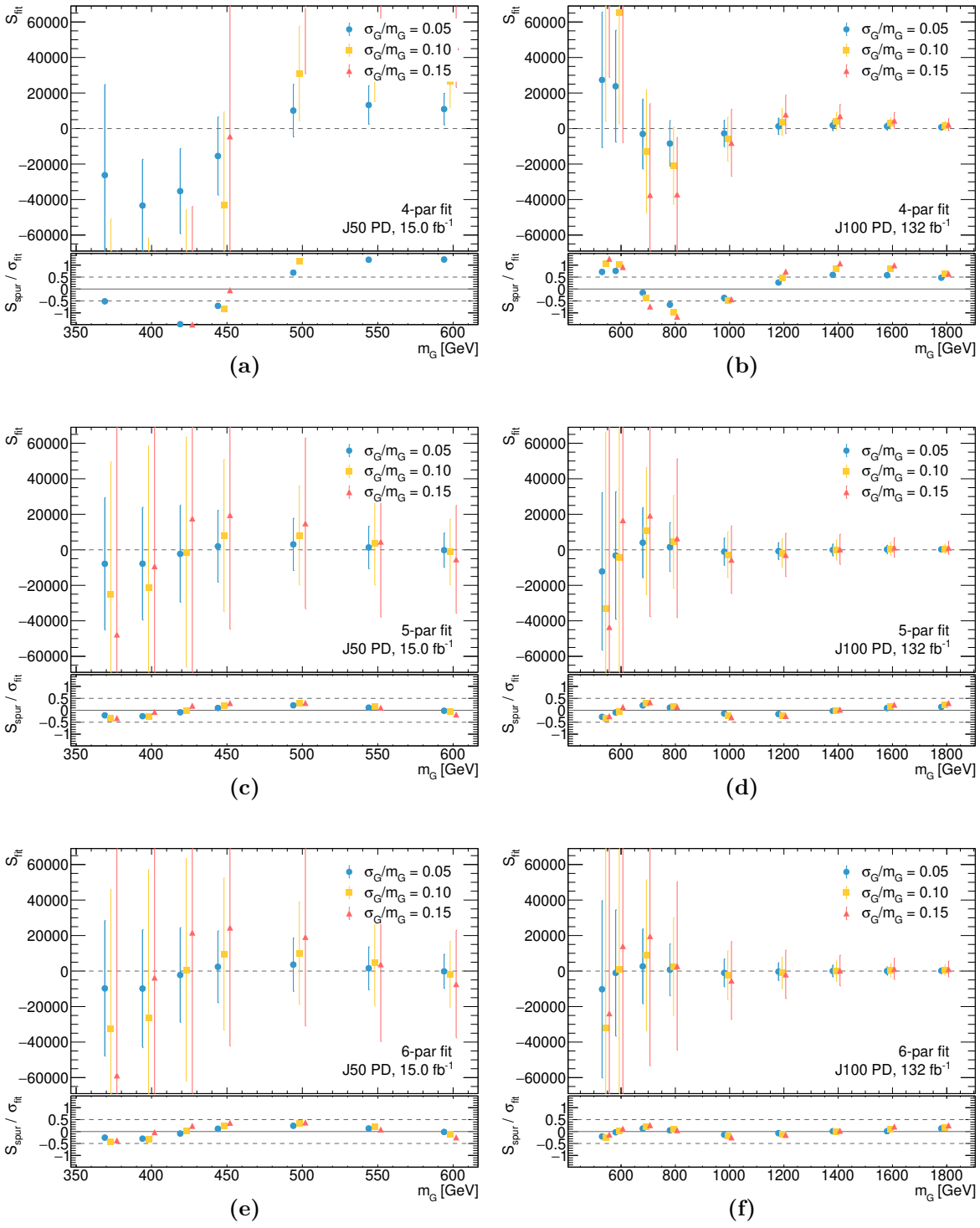


Figure 9.10: Summary of S_{spur} (points) and σ_{fit} (error bars) of the spurious signal tests for the functional form fit (**a**,**c**,**e**) to J50 PD and (**b**,**d**,**f**) to J100 PD for different choices of the number of parameters N. the results are shown as a function of resonance mass for resonance widths of 5 % (blue), 10 % (yellow), and 15 % (red). The bottom panel of each figure shows the ratio $S_{\text{spur}}/\sigma_{\text{fit}}$ with the dashed lines indicating the required threshold of $|S_{\text{spur}}| < 0.5\sigma_{\text{fit}}$. Small horizontal shifts are added to improve the readability of the different signal widths.

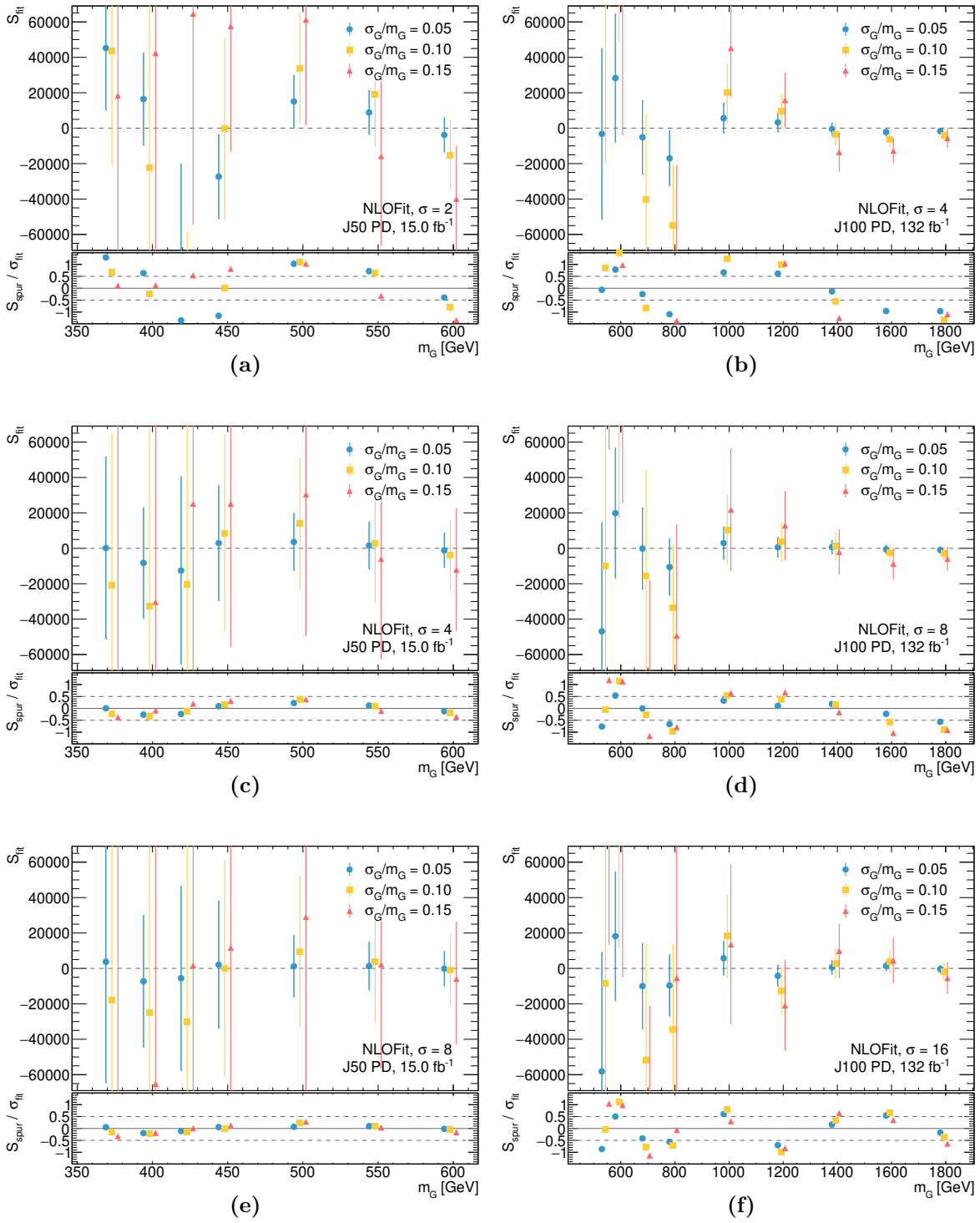


Figure 9.11: Summary of S_{spur} (points) and σ_{fit} (error bars) of the spurious signal tests for the NLOFit (**a**,**c**,**e**) to J50 PD and (**b**,**d**,**f**) to J100 PD for different choices of the constraint strength σ. the results are shown as a function of resonance mass for resonance widths of 5 % (blue), 10 % (yellow), and 15 % (red). The bottom panel of each figure shows the ratio $S_{\text{spur}}/\sigma_{\text{fit}}$ with the dashed lines indicating the required threshold of $|S_{\text{spur}}| < 0.5\sigma_{\text{fit}}$. Small horizontal shifts are added to improve the readability of the different signal widths.

corresponding fit results to the observed m_{jj} spectrum shown in Figures 9.8a and 9.8b.

The spurious signal determination of the functional form fit with Z' signals is documented in Appendix D.1. The results are consistent with the Gaussian signals found here: 5- and 6-parameter functional form fits fulfil the criterion for nearly unbiased fits.

Figure 9.11 summarizes the corresponding results of the NLOFit for both SRs. The uncertainties σ_{fit} are significantly larger than those of the functional form fit for wide signals of low masses in the J50 SR. The former indicates that the NLOFit flexibility is so high that the sensitivity to wide signals is reduced. The latter is potentially a result of multiple PDF variation templates having a zero-crossing at 400–500 GeV, which accordingly increases the fit flexibility in that region.

While the J50 SR fulfils the spurious signal requirement for constraints $\sigma \geq 4$ within the comparatively large fit uncertainties, the same does not hold true for the J100 SR. A significant fit bias is observed for all tested values of the constraint strength. This fit bias is also observed in a validation test where the PD is generated from a template with the same flexibility as the NLOFit to be validated. This indicates that the observed fit bias is not a result of the inability of the to-be-validated fit to describe the shape of the PD template. Instead, the bias in s+b fits appears to be inherent to the fit method. This bias is not expected to be a result of the constraint terms $g(p_k)$ for the fit parameters introduced in the likelihood (Equation (9.10)), since uniform constraints should not bias the fit parameters towards specific values. Instead, the bias is potentially a result of a failure to identify the global minimum of the high-dimensional likelihood.

The NLOFit is accordingly *not validated* for s+b fits and instead only used for the background-only interpretation in this analysis. While the NLOFit does fulfil the spurious signal criterion in the J50 SR, its increased uncertainties at low masses make it non-competitive with the sensitivity of the functional form fit in this SR. Hence, no s+b interpretation is performed with the NLOFit in either SR.

Spurious Signal Uncertainty

To account for the bias observed in the functional form fit, a spurious signal uncertainty is added to s+b fits. This uncertainty is set to a constant $0.5\sigma_{\text{fit}}$ for all signals as the fit bias is ensured to be smaller than this upper limit. It is included in the likelihood of s+b fits by identifying

$$S_{\text{fit}} = S + S_{\text{spur}} \tag{9.26}$$

in Equation (9.18), where S_{fit} is the best-fit number of signal events, S is the number of true signal events and S_{spur} is the number of spurious signal events. S_{spur} is then constrained to stay within its uncertainty by an additional nuisance parameter and Gaussian constraint term in the likelihood. Spurious signal is thus considered as a second source of signal that is indistinguishable from a true signal.

The spurious signal uncertainty simultaneously accounts for the inherent fit bias and the effect of not including one additional parameter in the N-parameter dijet function.

9.5.3 Signal Injections

A signal injection test probes the ability of a fit strategy to correctly determine the amount of signal present in a data set. It is performed similarly to the spurious signal

test, with the difference that a signal of the expected hypothesis is injected with a certain amplitude on top of the background-only PD.

The amplitude of the injection is denoted in units of $\sqrt{B}$, which takes into account the square root of the number of background events in the full-width-at-half-maximum (FWHM) range of the signal in question:

$$\sqrt{B} \equiv \frac{\sqrt{\sum_{\text{bins } i \text{ in FWHM}} \nu_i^{\text{b}}}}{\int_{\text{FWHM}} s(x)\,\mathrm{d}x}\,. \tag{9.27}$$

It is corrected by the integral of the normalized signal distribution $s\big(m_{jj}\big)$ over this range so that amplitudes expressed in units of $\sqrt{B}$ always correspond to the signal-to-square-root-of-background ratio in the signal's FWHM window. This definition ensures that signals at the sensitivity limit of this search correspond to an amplitude in the order of $3\sqrt{B}$ regardless of signal mass and width.[10]

With injected signal, the desired behaviour is a linear relationship between the expectation value of S_{fit} and the number of injected signal events S_{inj}. The definition of spurious signal in Equation (9.24) is generalized accordingly such that S_{spur} is now the deviation of the mean of distribution from S_{inj}:

$$S_{\text{spur}} \equiv \langle S_{\text{fit}} \rangle_{\text{toys}} - S_{\text{inj}}\,. \tag{9.28}$$

By requiring the same threshold for the spurious signal in Equation (9.25) regardless of the amplitude of the injected signal, a good linearity between S_{inj} and $\langle S_{\text{fit}} \rangle$ is ensured.

For each signal hypothesis and injection amplitude, 100 toy experiments are performed. This number of toys is sufficient as the Gaussian distribution of S_{fit} was already validated in the spurious signal test without injection.

Figure 9.12 shows the validation of the 5-parameter functional form fit for Gaussian resonances using PD generated from a 6-parameter template. Each row corresponds to Gaussian widths of either 5 %, 10 %, or 15 %. The left column shows the validation of the J50 SR and the right column shows the J100 SR. The determined values of $\langle S_{\text{fit}} \rangle \pm \sigma_{\text{fit}}$ are shown as a function of the amount of injected signal for different tested signal masses. In general, very good linearity and good agreement with the optimal behaviour indicated by the dashed diagonal line are observed for each tested signal hypothesis.

Accordingly, the 5- and 6-parameter functional form fits pass the signal injection test in both SRs. The injection of Z' signals is shown in Appendix D.2, where the same result is found and both strategies are validated.

For completeness, these signal injection tests were also performed with the NLOFit despite it not being validated for s+b fits. While the absolute amount of spurious signal still exceeds the maximal allowed threshold, it is nearly independent of the amount of injected signal. This indicates that the fit bias of the NLOFit method is most likely not a result of a tendency to partially absorb signals.

[10] An exception are wide signals towards the edge of the fit range where the search sensitivity is reduced.

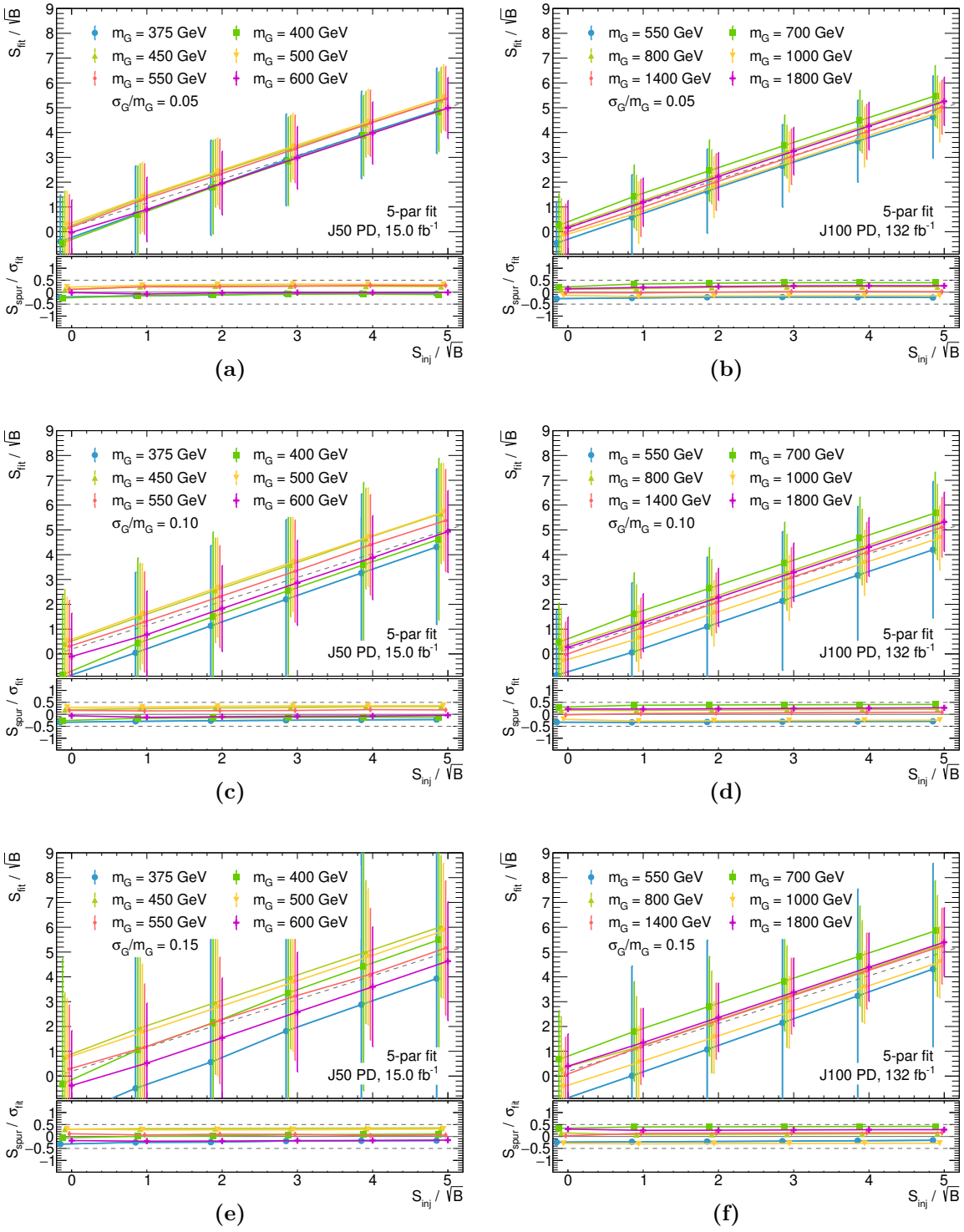


Figure 9.12: Number of signal events ($\langle S_{\text{fit}} \rangle \pm \sigma_{\text{fit}}$) determined by a 5-parameter functional form fit in PD as a function of the amount of injected signal under different signal hypotheses. The absolute value of the spurious signal in units of the fit uncertainty is shown in the bottom panels. The dashed lines at ±0.5 indicate the required maximal limit on the spurious signal. Small horizontal shifts around the integer injection amplitudes are added to improve the readability of the different signal masses.

9.5.4 Background Stability

The reliable extraction of the number of signal events from (pseudo-)data necessitates that the background component of an s+b fit is sufficiently robust to not be strongly influenced by the presence of signal. If this were not the case, it would most likely result in a non-linear relationship between $\langle S_{\text{fit}} \rangle$ and S_{inj} in the previous section.

Nonetheless, the stability of the background component of s+b fits is verified with a dedicated test. For this, the same N-parameter s+b fits to 100 toys generated from an $(N+1)$-parameter template with injected signals are considered. The background component of the s+b fits is compared to the background-like PD template. Any difference between the two then arises due to

- a difference between the N- and $(N+1)$-parameter function shapes,
- statistical fluctuations in the toys, or
- the adaptation of the background component of the s+b fit to the presence of signal.

The latter case is expected to show as a distortion of the background component around the mass of the injected signal.

For each considered signal hypothesis, an injection amplitude corresponding to the expected sensitivity is chosen – for consistency $3\sqrt{B}$ in the definition of Equation (9.27). The fit distortion is required to stay below three times the statistical uncertainty for any individual toy experiment and the ensemble of toy fits is required to show no significant systematic bias.

Figure 9.13 shows the toy fit results of the 5-parameter functional form fit relative to the reference spectra for a few selected Gaussian signal hypotheses. These show that the toy fit results exhibit a certain variance around the reference spectrum due to the Poissonian fluctuations applied to the PD toys. This variance increases around the tested resonance mass due to the ambiguity between the signal and background component in this region of the m_{jj} spectrum. The increase in variance is largest in the J50 SR and towards the edge of the m_{jj} fit range. This is potentially due to the J50 SR covering a shorter m_{jj} range, which provides a weaker constraint on the fit. The increased variance towards the edge of the fit range is most likely due to fewer data points constraining the fit in this region.

The mean of the ensemble of toy fits is consistent with the reference spectrum for all tested signal hypotheses, validating the absence of a significant bias. This test is also well suited to verify a continuous spread of the fit results arising from the statistical fluctuations of the toys and to assure the absence of rare outliers.

The 5- and 6-parameter functional form fit are validated in both SRs for Gaussian and Z' signals. The results of the latter are shown in Appendix D.2.

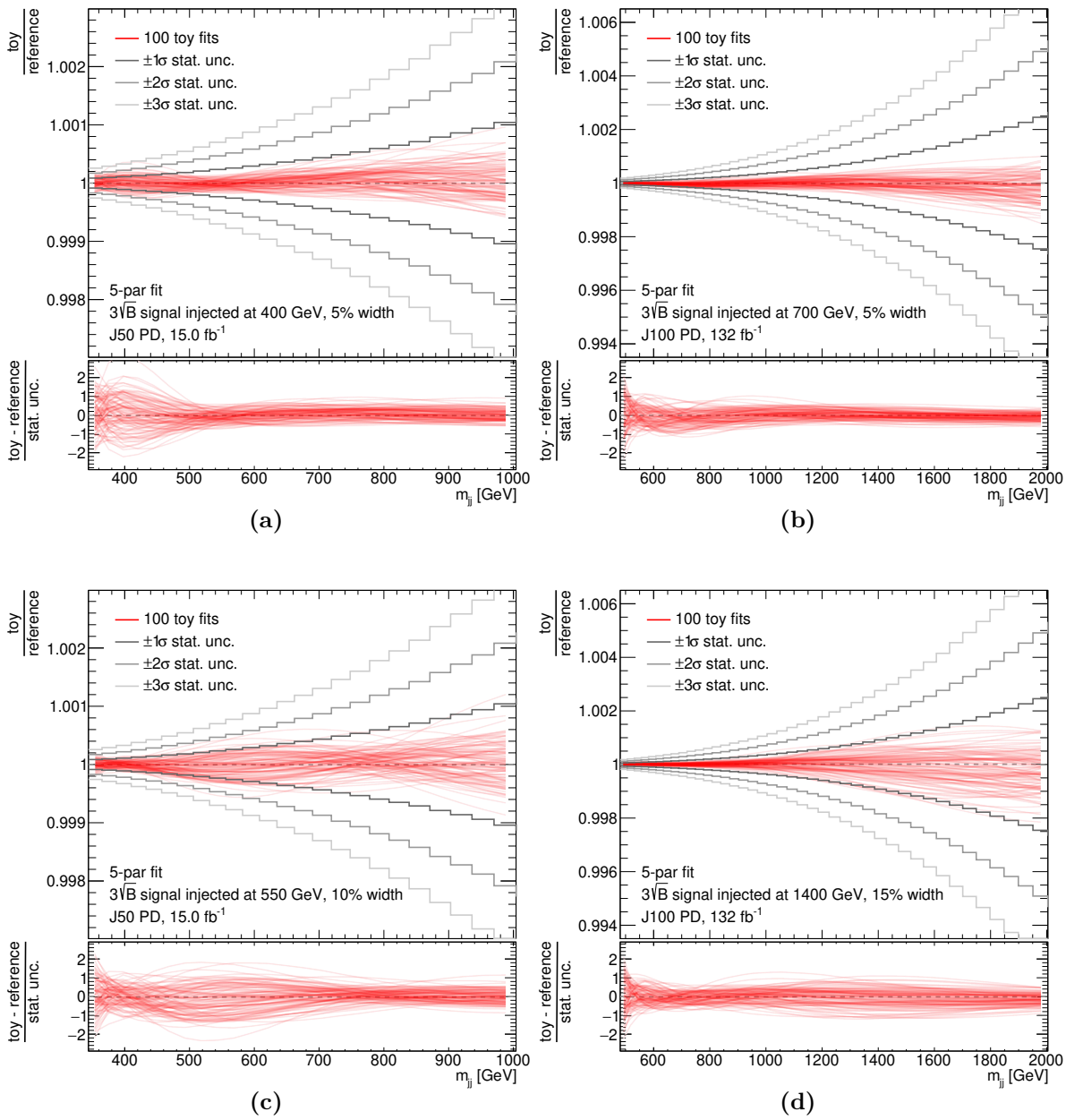


Figure 9.13: Background stability test of a 5-parameter functional form fit (**a**,**c**) to J50 PD and (**b**,**d**) to J100 PD for selected Gaussian signal hypotheses. The background component of s+b fits to PD toys with injected signal is compared to the background-only component of the PD (reference).

9.5.5 Choice of the Fit Strategy

A final fit strategy to be used in this analysis must be chosen for each SR from the fit strategies validated above. This choice is made separately per SR and for the functional form fit and the NLOFit, where the latter is only considered for background-only fits. In each case, the fit strategy with the lowest number of (practical) parameters is chosen if increasing this number further does not significantly improve the fit result.

F-Test

When increasing the complexity of a fit model, its description of the data will in general improve – barring instabilities of the fit in higher dimensions. This improvement of χ^2 could result from a better modelling of the underlying physics or simply from a better adaptation to statistical fluctuations.

The significance of the improvement due to an additional (effective) fit parameter can be evaluated by an F-test. While this test is, in general, defined for the comparison of the variance of two samples, it can be applied to compare the χ^2 of two nested fit methods performed on the same data set [189, 190]. Nested fit methods refer to a nominal and an alternative fit strategy of different complexity where the nominal is a special case of the alternative - e.g. the functional form fit with different choices of N or the NLOFit with different choices of σ.

For this application, the F-test uses the test statistic

$$F = \frac{\left(\chi^2_{\text{nom}} - \chi^2_{\text{alt}}\right) / \left(\text{ndf}_{\text{nom}} - \text{ndf}_{\text{alt}}\right)}{\chi^2_{\text{alt}} / \text{ndf}_{\text{alt}}} \,. \tag{9.29}$$

If the alternative method does not provide a better description of the true behaviour, such that the difference in χ^2 is due to statistical fluctuations alone, the expected probability density of F is given by the Fisher–Snedecor distribution f_F, for which an analytic expression exists [190].

Thus, a p-value can be determined by integrating the Fisher–Snedecor distribution:

$$p(F) = \int_F^\infty f_F(z; \text{ndf}_{\text{nom}} - \text{ndf}_{\text{alt}}, \text{ndf}_{\text{alt}})\, \mathrm{d}z \,. \tag{9.30}$$

A low $p(F)$-value indicates a significantly better description of the data by the alternative (higher-complexity) model, while a high $p(F)$-value indicates that the difference is compatible with statistical fluctuations alone.

For this analysis, an F-test is performed for each fit strategy relative to the one with (approximately) one additional (effective) parameter. A value $p(F) > 0.05$ is then required to select the lower-complexity fit. The $p(F)$-values determined for the functional form fit and the NLOFit are shown in Figure 9.8.

Functional Form Fit. For the functional form fit, $N = 4$ is not considered as it has not passed the validation tests above. In both SRs, $p\left(F_{5\to 6\,\text{pars}}\right) > 0.05$ is observed, indicating that the 5-parameter function provides a sufficient description of the observed data. Accordingly, the 5-parameter functional form fit is selected for both SRs.

NLOFit. The NLOFit is not validated for s+b fits. Nonetheless, the F-test can be used to find the proper degree of complexity for the background-only fits. In the J50 SR, $p(F_{\sigma:2\to4}) = 0.006$ is observed, indicating that the $\sigma = 4$ NLOFit provides a significantly better description of the data. The difference with $\sigma = 8$ is then found to be insignificant. Therefore, the $\sigma = 4$ NLOFit is selected for the J50 SR.

In the J100 SR, the applied masking must be taken into account. The test statistic F in Equation (9.29) can only be meaningfully constructed if both fit models describe the same data with the same binning. A fit with a masked excess can thus not be compared to an unmasked fit. Instead, a comparison between $\sigma = 4$ and $\sigma = 8$ is performed where both fits have the same m_{jj} region masked. Here, $p(F_{\sigma:4\to8}) = 0.029$ is observed, meaning that the masked $\sigma = 4$ NLOFit provides a suboptimal description of the data. For $\sigma \geq 8$, the unblinding algorithm found no necessity for masking. Thus, the unmasked $\sigma = 8$ and $\sigma = 16$ NLOFit are compared, where $p(F_{\sigma:8\to16}) > 0.05$. Therefore, the unmasked $\sigma = 8$ NLOFit is selected for the J100 SR.

Summary of the Fit Strategies

For each SR, one fit strategy for the functional form fit and one for the NLOFit have been selected with the tests and algorithms described in the previous sections. The final choice for each is summarized in Table 9.3.

	Selected Fit Strategy	
Signal Region	Functional Form Fit	NLOFit
J50	$N = 5$	$\sigma = 4$
J100	$N = 5$	$\sigma = 8$

Table 9.3: Fit strategies selected for the background estimate in each SR. The number of parameters N is shown for the functional form fit and the constraint strength σ for the NLOFit. No masking of a localized m_{jj} region is applied for any fit strategy.

10 Results

Having defined the fit strategies in the previous chapter, they are now applied to the observed m_{jj} spectra in both SRs to estimate the SM background. This allows searching for BSM resonances in the observed spectra as excesses over the estimated background. This chapter presents two interpretations of the search:

Background-only interpretation. A background-only fit is performed to test the compatibility of the data with a smoothly falling spectrum. Potential excesses are quantified with minimal assumptions on the resonance shape to search for BSM physics in a nearly model-agnostic way.

Signal interpretation. The compatibility of the data with specifically selected Gaussian or Z' resonances is tested. With these assumed signal shapes, s+b fits to the data are performed to derive exclusion limits on the signal cross section. This interpretation is more sensitive to the chosen signal models at the expense of having to assume a resonance shape.

10.1 Background-Only Interpretation

The consistency of the observed dijet mass spectrum with a smoothly falling SM background can be quantified with the $p(\chi^2)$-value alone. However, this does not provide the optimal sensitivity to a localized resonance since it considers all m_{jj} bins simultaneously and independently of each other. A true resonance, on the other hand, would likely result in a correlated excess in one or several adjacent m_{jj} bins.

10.1.1 The BumpHunter Algorithm

The BumpHunter algorithm [187, 188] is a suitable choice to identify such a localized excess. It is designed to identify the window of adjacent bins in a spectrum, in which the observed number of events most significantly deviates from the background expectation. No assumption on the resonance width or shape is made.

The BumpHunter algorithm constitutes a hypertest that combines the result of many individual hypothesis tests into a single test statistic. It iterates over all possible windows of adjacent bins in the m_{jj} spectrum, starting with windows one bin wide and increasing in width up to a configurable upper threshold for the window size (typically half of the fit range). In each of these windows, the observed number of events n and the estimated background events ν are summed as if the window were a single bin:

$$n = \sum_{\substack{\text{bins } i \\ \text{in window}}} n_i \,, \qquad \nu = \sum_{\substack{\text{bins } i \\ \text{in window}}} \nu_i \,. \tag{10.1}$$

Each of these windows is assigned a local p-value based on the Poissonian probability to observe at least as many events[1] as seen in data:

$$p_{\text{local}}(n,\nu) = \begin{cases} \Gamma(n,\nu) & \text{if } n > \nu\,, \\ 1 & \text{if } n \leq \nu\,, \end{cases} \tag{10.2}$$

where

$$\Gamma(n,\nu) = \sum_{k=n}^{\infty} \frac{\nu^k}{n!} e^{-\nu} \tag{10.3}$$

is the lower incomplete gamma function.

Since many statistically independent windows are tested for excesses, the look-elsewhere effect [191] must be accounted for. It describes that if e.g. 100 independent tests are performed, on average one of them will find a p-value below 0.01 due to statistical fluctuations alone. The *local* p-value for an excess in a specific window must hence be translated into a *global* p-value to find an excess in any of the windows.

For that purpose, the BumpHunter algorithm defines the test statistic

$$t = \min_{\text{windows}}(-\log p_{\text{local}})\,, \tag{10.4}$$

which identifies the most significant excess of all considered windows. The expected distribution of t is determined numerically to high accuracy by drawing 10^4 Poissonian fluctuated toy distributions from the background expectation and computing t for each of these toys. The global BumpHunter p-value is then given by the probability for a toy distribution to exhibit a more significant excess than the data. This corresponds to the fraction of toys for which t exceeds the observed value t_{obs}:

$$p(\text{BH}) = \frac{\#\text{toys with } t > t_{\text{obs}}}{\#\text{toys}}\,. \tag{10.5}$$

This way, $p(\text{BH})$ by definition follows a uniform distribution between 0 and 1 if excesses over the background estimate are due to statistical fluctuations alone. Observing, for example, $p(\text{BH}) < 0.05$ then means that less than 5 % of background-like toys exhibit a larger excess than the most significant one in data. Such an observation can be interpreted as evidence for a dijet resonance causing a deviation of the data from the background expectation.

10.1.2 The Sensitivity of BumpHunter

The determination of $p(\text{BH})$ assumes the ideal case that the background estimate is fixed and does not depend on the specific (toy) data set. In a realistic scenario, fluctuations in the data will influence (pull) the background fit, such that large excesses are rarer than estimated by the global BumpHunter p-value. The amount of pull depends on the flexibility of the background fit.

[1] An analogous p-value for deficits instead of excesses is also defined but not used in this search.

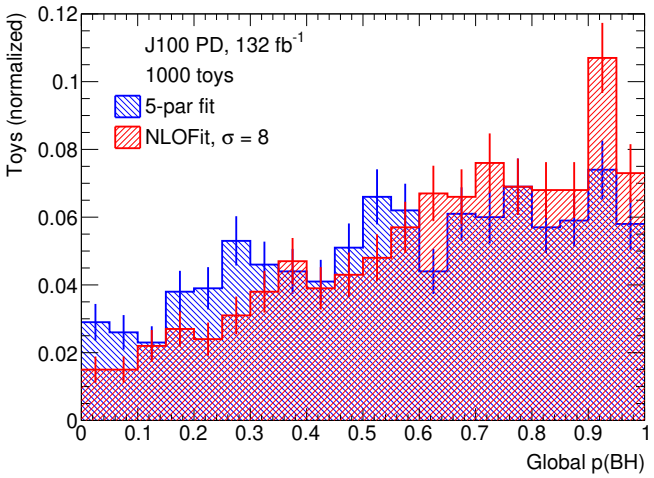


Figure 10.1: Distribution of the global BumpHunter p-values as determined in functional form fits (blue) or NLOFits (red) to 1000 PD toy data sets corresponding to the J100 SR. The statistical fluctuations due to the limited number of toys are indicated as coloured error bars. The PD is drawn from a template with one additional (effective) fit parameter. A bias towards high p-values is visible which indicates an adaptation of the fit to statistically induced excesses. This bias is stronger for the NLOFit.

This effect can be seen by performing background-only fits to PD.[2] Figure 10.1 shows the distribution of the global BumpHunter p-values as determined in background-only fits to 1000 toy data sets representing the J100 SR.

A bias towards high p-values is observed for both the functional form fit (blue) and the NLOFit (red), which indicates that both fit methods adapt to excesses that are caused by statistical fluctuations. This effect is stronger for the NLOFit, which aligns with the larger fit uncertainties observed for the NLOFit in Section 9.5.2. Such an over-flexibility can result in a reduced sensitivity of the BumpHunter search to resonances as parts of the signal are absorbed by the background-only fit.

In order to estimate the sensitivity to signals, Gaussian resonances are injected into the PD. For each combination of resonance mass, width, and amplitude, the distribution of the global BumpHunter p-values is determined in background-only fits to 200 toy data sets. The features of those distributions are summarized in Figure 10.2 for the functional form fit and in Figure 10.3 for the NLOFit. The coloured boxes represent the median and the 16 % and 84 % quantiles such that 68 % of the distribution is contained within them. The coloured dashed lines represent the 5 % and 95 % quantiles. The bottom panels show the ratio of toys for which a global BumpHunter p-value below 0.05 is found, which would indicate a significant excess.

Expectedly, a trend towards lower p(BH) with increasing amount of injected signal is observed. This means that it becomes more likely for a toy data set to have a significant excess identified by the BumpHunter algorithm. In units of $\sqrt{B}$, a signal is more likely to be identified if it is higher in mass due to a larger number of data points below

[2] The same method for the generation of the PD as outlined in the previous chapter is used, with the template corresponding to a fit with (approximately) on additional (effective) fit parameter.

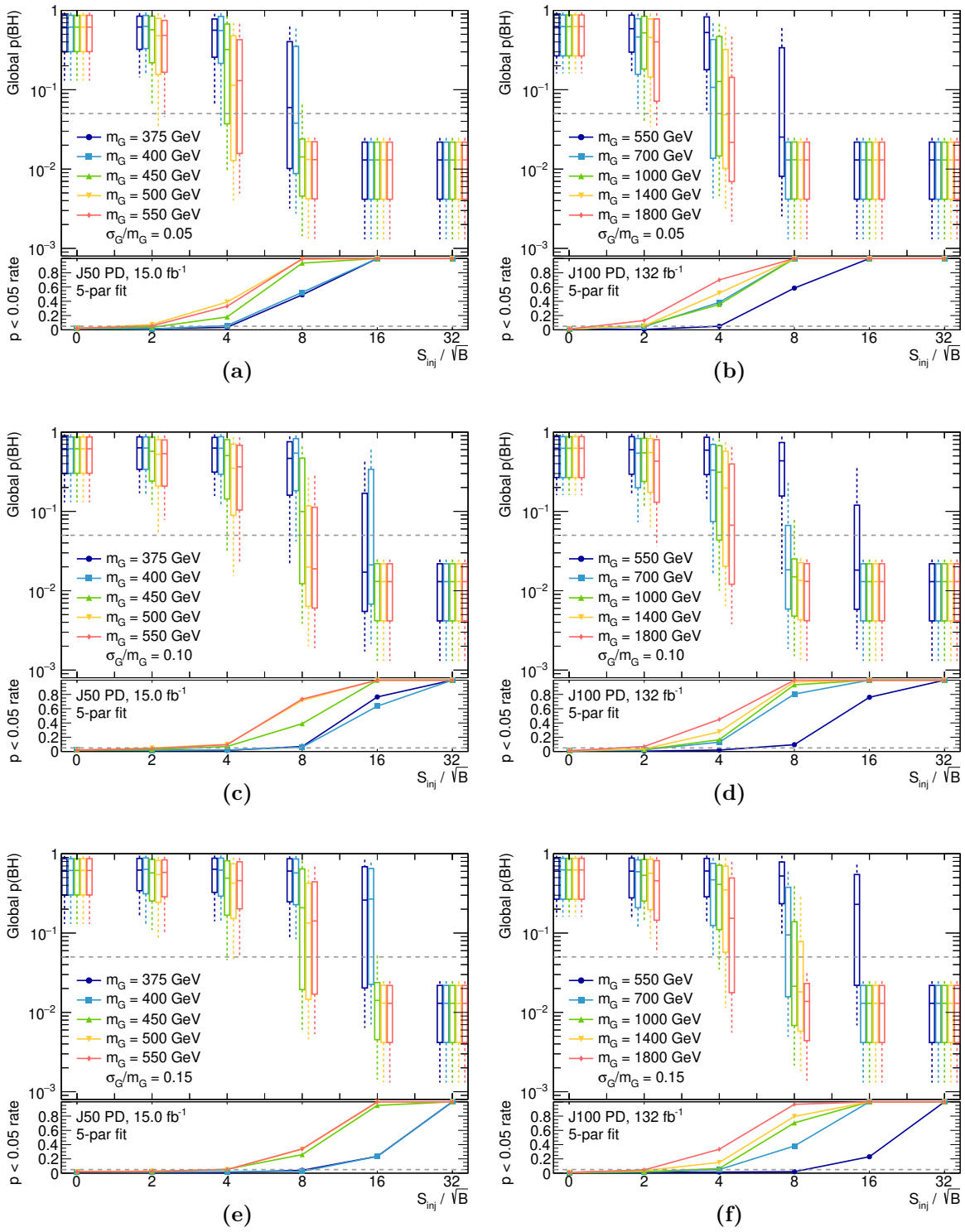


Figure 10.2: Median and spread of the global BumpHunter p-value determined in background-only functional form fits to 200 toys of (**a**,**c**,**e**) J50 PD and (**b**,**d**,**f**) J100 PD. Gaussian signal of (**a**,**b**) 5 %, (**c**,**d**) 10 %, or (**e**,**f**) 15 % width at various masses are injected. The coloured boxes indicate the median and the central 68 % of the distribution, while the coloured dashed lines correspond to the central 90 % of the distribution. The fraction of toys for which a p-value below 0.05 is found is indicated in the bottom panels as a function of the injected signal amplitude.

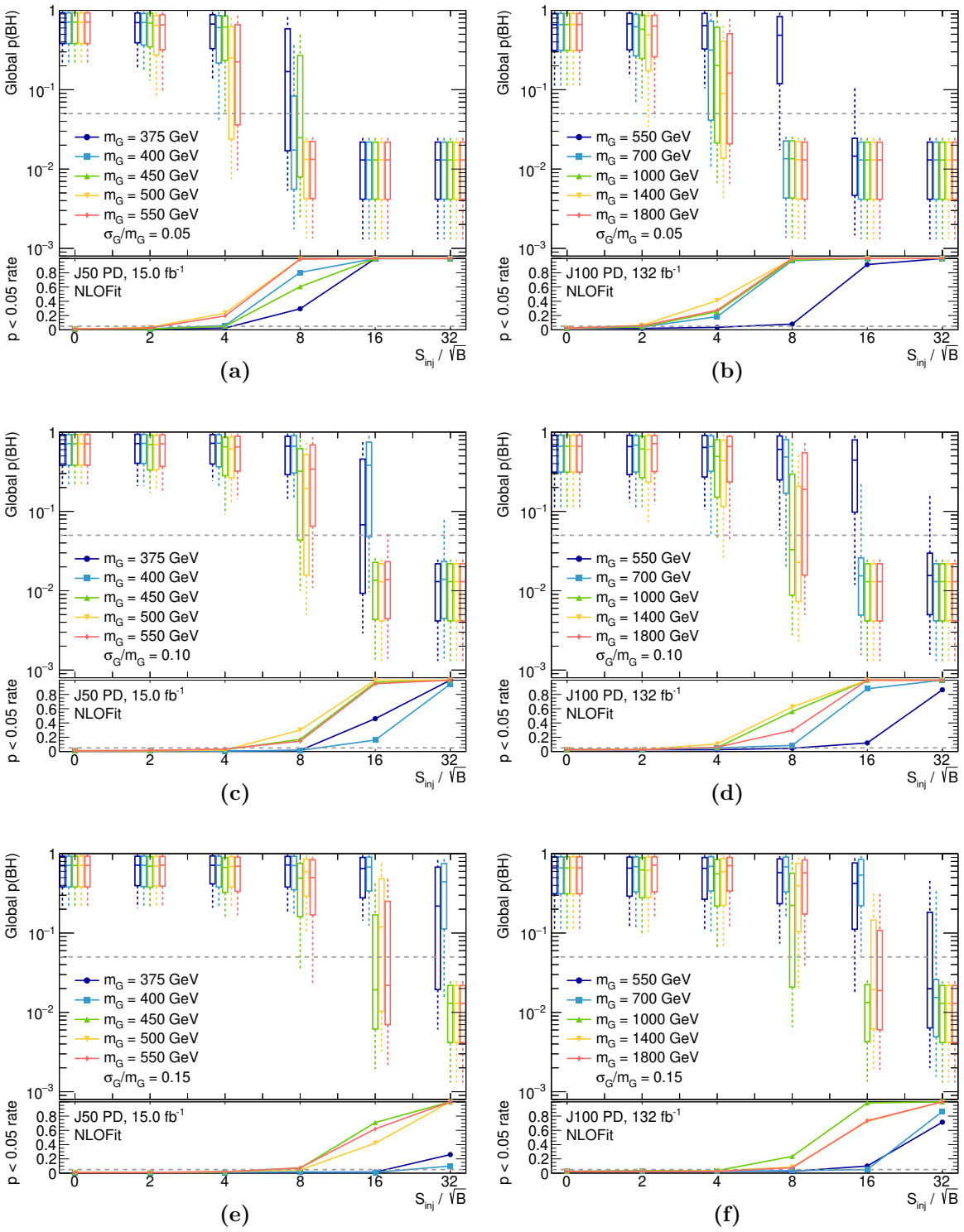


Figure 10.3: Median and spread of the global BumpHunter p-value determined in background-only NLOFits to 200 toys of (**a**,**c**,**e**) J50 PD and (**b**,**d**,**f**) J100 PD. Gaussian signal of (**a**,**b**) 5 %, (**c**,**d**) 10 %, or (**e**,**f**) 15 % width at various masses are injected. The coloured boxes indicate the median and the central 68 % of the distribution, while the coloured dashed lines correspond to the central 90 % of the distribution. The fraction of toys for which a p-value below 0.05 is found is indicated in the bottom panels as a function of the injected signal amplitude.

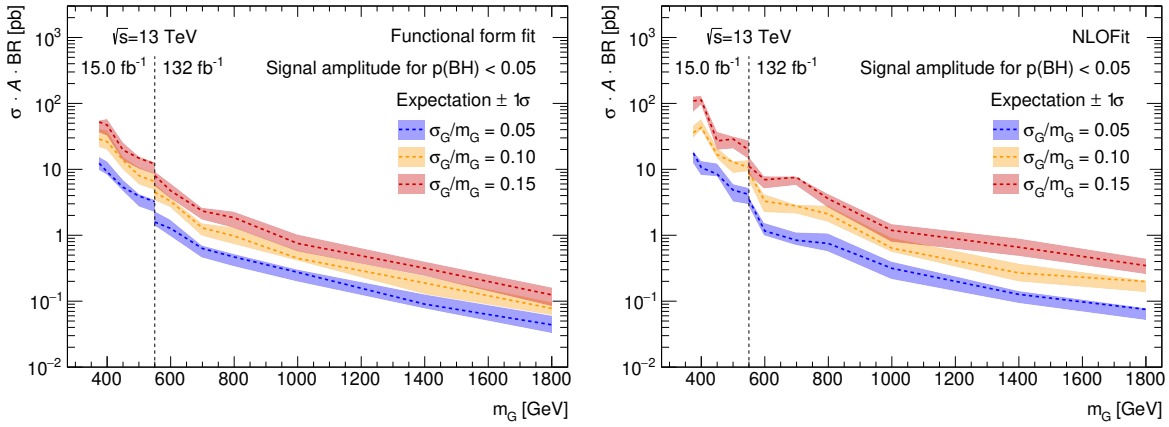


Figure 10.4: Amplitude of a Gaussian resonance that is necessary to reach a BumpHunter p-value below 0.05 using (**a**) the functional form fit and (**b**) the NLOFit as background estimate. The median and the $\pm 1\sigma$ quantiles of the amplitudes determined in 200 toys are shown. Gaussian resonances of 5 % (blue), 10 % (orange), and 15 % width (red) are tested.

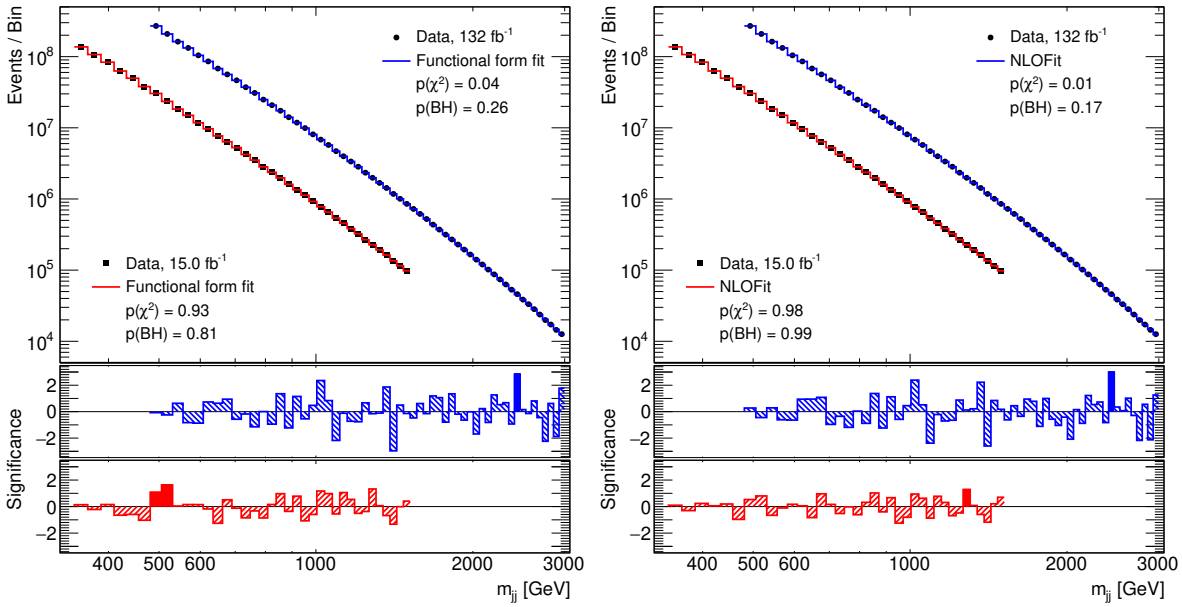
the signal that constrain the background-only fit, reducing its ability to absorb parts of the signal. Simultaneously, narrow signals are much more likely to be identified by BumpHunter since the background-only fit adapts less easily to narrow signals.

For an easier interpretation of the BumpHunter sensitivity in terms of resonance cross sections, the determined $p(BH)$-values as a function of signal amplitude are interpolated for each toy experiment to find the necessary signal amplitude to reach $p(BH) = 0.05$.[3] Figure 10.4 shows the median and the uncertainty band corresponding to the 16 % and 84 % quantiles of the necessary Gaussian resonance amplitude, using the functional form fit background estimate in panel (**a**) and the NLOFit in panel (**b**). The amplitudes are expressed in terms of the visible signal cross section times the branching ratio into a dijet signature in order to facilitate a comparison with the dedicated signal interpretation discussed in Section 10.2.3.

The comparison of the functional form fit result shows, that the search sensitivity of BumpHunter to signals is lower than that of a dedicated search for Gaussian resonances by a factor of 1.5–2 for narrow resonances. This factor increases to 3 for wider resonances, due to the stronger adaptation of the background-only fit to wide resonances. The NLOFit provides comparable results for narrow resonances, but a stronger loss in sensitivity to wide resonances – a factor of up to 5. This reflects the stronger adaptation of the NLOFit to wide signals that was already observed in s+b fits in Section 9.5.2.

In general, the BumpHunter search provides lower sensitivity to potential resonances than a dedicated s+b fit with the specific signal under test. In return, it is a relatively model-agnostic search for which no assumption on the signal shape needs to be made, other than it being a localized excess in the m_{jj} spectrum.

[3] Additional injection amplitudes are included to provide more points for the linear interpolation and to reliably reach $p(BH) < 0.05$ for each toy experiment.



(a) **(b)**

Figure 10.5: Background-only fits to the observed m_{jj} spectra in the J50 SR (red) and the J100 SR (blue), using (**a**) the functional form fit and (**b**) the NLOFit as background estimate. The significance of the fit residuals is shown in the bottom panels, where the region identified as the most significant excess by the BumpHunter algorithm is shown as solid-filled bins.

10.1.3 Results

Figure 10.5a shows the results of the functional form fit to the observed m_{jj} spectrum in the J50 SR (red) and in the J100 SR (blue). The significance of the fit residuals is shown in the bottom panel. The background-only fit describes the data well, achieving a $p(\chi^2)$-value of 0.93 (0.04) in the J50 SR (J100 SR). The most significant excess is observed in the J100 SR in the bin $2403\,\text{GeV} < m_{jj} < 2464\,\text{GeV}$ with a global BumpHunter p-value of 0.26. This corresponds to a global significance of 0.7σ.

Figure 10.5b shows the corresponding results with the NLOFit background estimate. A similarly good description of the data is achieved, with an observed $p(\chi^2)$-value of 0.98 (0.01) in the J50 SR (J100 SR). The same m_{jj} bin as for the functional form fit corresponds to the most significant excess. It is assigned a global BumpHunter p-value of 0.17, which corresponds to a global significance of 1σ.

In the J50 SR, different m_{jj} regions are identified to be the most significant excess with the two background estimates. This is the result of these two windows providing similar local p-values, such that a small difference in the background estimate can result in either of the two being the most significant excess. Global BumpHunter p-values of 0.81 and 0.99 are observed for the functional form fit and the NLOFit, respectively, which both correspond to a global significance of ~ 0.

Overall, the functional form fit and the NLOFit provide compatible background estimates, resulting in only minor differences in the fit residuals between them. This serves as cross-validation that either method achieves a reliable background estimate.

10.2 Signal Interpretation

No significant excess is observed in the dijet mass spectrum as it exhibits good compatibility with a smoothly falling background. Therefore, signal models for which an excess would be expected – generic Gaussian resonances and a Z' resonance in this thesis – can be constrained. The goal here is to determine the maximal amount of potential signal for which the signal-plus-background hypothesis H_1 still provides an adequate description of the data in comparison to the background-only hypothesis H_0.

The signal interpretation is solely based on the functional form fit as background estimate, since the NLOFit failed the validation for s+b fits.

10.2.1 Derivation of Limits

Signal models are constrained by deriving confidence intervals for model parameters. This process is reduced to a one-dimensional problem by assuming fixed values for model parameters like resonance mass and decay width and then constraining a model parameter of interest under this assumption – typically the signal cross section or a variable proportional to it. In the absence of an excess, the confidence interval for this parameter of interest becomes an upper exclusion limit. The process is repeated for all considered combinations of the other model parameters to constrain the model's parameter space.

In this thesis, the $\mathrm{CL_s}$ method is used to derive frequentist-motivated exclusion limits [192]. It provides a modified confidence level (CL) taking both hypotheses H_0 and H_1 into account. The following description is based on the $\mathrm{CL_s}$ calculation typically used in ATLAS [193].

The derivation of confidence levels relies on the likelihood of the observed data under the hypothesis H_1 with a number S of expected signal events, which is given in Equation (9.23). The parameter S is considered as the parameter of interest under which the hypothesis test is performed. The test statistic $\tilde{q}_S$ is based on the profile likelihood ratio λ, which provides the most powerful hypothesis test according to the Neyman-Pearson lemma [194]:

$$\tilde{q}_S = \begin{cases} -2\log\lambda(S,0) & \text{if } \widehat{S} < 0\,, \\ -2\log\lambda(S,\widehat{S}) & \text{if } 0 \le \widehat{S} \le S\,, \\ 0 & \text{if } \widehat{S} > S\,, \end{cases} \tag{10.6}$$

$$\lambda\big(S,\widehat{S}\big) = \frac{\mathcal{L}\Big(\boldsymbol{n}\,\Big|\,S,\widehat{\boldsymbol{p}}_S,\widehat{\boldsymbol{\theta}}_S\Big)}{\mathcal{L}\Big(\boldsymbol{n}\,\Big|\,\widehat{S},\widehat{\boldsymbol{p}}_{\widehat{S}},\widehat{\boldsymbol{\theta}}_{\widehat{S}}\Big)}\,, \tag{10.7}$$

where $\widehat{\boldsymbol{p}}_S$ and $\widehat{\boldsymbol{\theta}}_S$ are the vectors of fit and nuisance parameters that maximize $\mathcal{L}(\boldsymbol{n}\,|\,S,\boldsymbol{p},\boldsymbol{\theta})$ for a given S. $\widehat{S}$ is the best-fit signal amplitude, which corresponds to the global maximum of $\mathcal{L}\big(\boldsymbol{n}|\widehat{S},\widehat{\boldsymbol{p}}_{\widehat{S}},\widehat{\boldsymbol{\theta}}_{\widehat{S}}\big)$. The test statistic $\tilde{q}_S$ is thus always positive and, for a fixed signal hypothesis, increases monotonically with increasingly background-like observations. $\widehat{S} \ge 0$ is required to consider only positive signal cross sections as underlying truth.

$S \geq \widehat{S}$ ensures that a one-sided confidence interval is derived, which prevents the exclusion of a signal hypothesis if an upward fluctuation in the data causes the best-fit $\widehat{S}$ to exceed the tested S.

To derive a confidence level from an observed value $\tilde{q}_S^{\text{obs}}$, the expected distribution $f(\tilde{q}_S \,|\, S')$ for an assumed true value of the signal amplitude S' must be known. This distribution can be estimated with toy experiments. However, this is computationally intensive for high-dimensional likelihood fits as used in this analysis. Instead, the *asymptotic approximation* with the Asimov approach detailed in Ref. [195] is used. It provides an analytic description for $f(\tilde{q}_S \,|\, S')$ under the assumption of large event counts and of the best-fit value $\widehat{S}$ being Gaussian distributed around the true value S'. Both assumptions hold for this analysis: Large event counts of at least 100 are observed even in the lowest-statistics m_{jj} bins of 1 GeV width, and the approximately Gaussian distribution of $\widehat{S}$ has been validated in Section 9.5.2.[4]

The confidence level $\text{CL}_{\text{s+b}}$ is defined as the probability of an experiment to find a test statistic $\tilde{q}_S$ that is less signal-like than the observed value $\tilde{q}_S^{\text{obs}}$ if H_1 with signal amplitude S were true:

$$\text{CL}_{s+b}(S) \equiv \text{P}\Big(\tilde{q}_S \geq \tilde{q}_S^{\text{obs}} \,\Big|\, H_1\Big) = \int_{\tilde{q}_S^{\text{obs}}}^{\infty} f\big(\tilde{q}_S \,|\, S\big)\, \text{d}\tilde{q}_S \,. \tag{10.8}$$

This probability is translated into a confidence interval by defining a threshold α and rejecting H_1 for values of S for which $\text{CL}_{\text{s+b}}(S) < \alpha$. This confidence interval becomes an upper limit on S with the chosen definition of the test statistic $\tilde{q}_S$ in Equation (10.6). Signal amplitudes above this upper limit are then excluded at a $1 - \alpha$ confidence level. Typically, $\alpha = 0.05$ is chosen such that 95 % CL limits are derived.

The use of $\text{CL}_{\text{s+b}}$ to derive exclusion limits has the unphysical feature that it can potentially exclude signals it is not sensitive to. This can be seen for a hypothetical experiment with low sensitivity, i.e. low power to distinguish H_0 and H_1. The latter is the case if $f(\tilde{q}_S \,|\, S)$ and $f(\tilde{q}_S \,|\, 0)$ are similar distributions. If a downward fluctuation in the data results in a value of $\tilde{q}_S^{\text{obs}}$ in the upper tail of the expected distribution, this can result in H_1 being rejected. This is unphysical since the data are (almost) equally incompatible with H_0 and no statement about a preferred hypothesis should be made.

This unphysical behaviour is addressed in the CL_{s} method by penalizing $\text{CL}_{\text{s+b}}$ by the incompatibility of the data with H_0:

$$\text{CL}_{\text{s}}(S) \equiv \frac{\text{CL}_{\text{s+b}}(S)}{1 - \text{CL}_{\text{b}}(S)} \,, \tag{10.9}$$

$$1 - \text{CL}_b(S) \equiv \text{P}\Big(\tilde{q}_S \geq \tilde{q}_S^{\text{obs}} \,\Big|\, H_0\Big) = \int_{\tilde{q}_S^{\text{obs}}}^{\infty} f\big(\tilde{q}_S \,|\, 0\big)\, \text{d}\tilde{q}_S \,. \tag{10.10}$$

While $\text{CL}_{\text{s}}(S)$ cannot be strictly interpreted as a probability or a p-value, it is used analogously to derive confidence intervals by rejecting H_1 if $\text{CL}_{\text{s}}(S) < \alpha$. If the observation is consistent with H_0, the result is similar to a $\text{CL}_{\text{s+b}}$ confidence interval because

[4] The assumption that the mean of that distribution is the true signal amplitude S' is one of the reasons that the spurious signal is required to be small in comparison to the standard deviation of this distribution.

$1 - \mathrm{CL_b}(S)$ is of order 1. Conversely, the unphysical rejection of H_1 by an insensitive experiment is suppressed because $\mathrm{CL_{s+b}}(S)$ and $1 - \mathrm{CL_b}(S)$ are of similar magnitude in this case. Confidence intervals derived with the $\mathrm{CL_s}$ method are always more conservative than those derived with the $\mathrm{CL_{s+b}}$ method because $0 \leq \mathrm{CL_b}(S) < 1$ and consequently $\mathrm{CL_s}(S) > \mathrm{CL_{s+b}}(S)$.

The procedure described above results in the *observed limit* on the signal amplitude S. The *expected limit* is derived analogously by using the median value $\tilde{q}_S^A$ of the test statistic expected under H_0 in place of $\tilde{q}_S^{\mathrm{obs}}$. Uncertainty bands of $\pm 1, 2\sigma$ on the expected limit are derived by using the corresponding quantiles of $f(\tilde{q}_S \,|\, 0)$.

10.2.2 Limit Coverage

In a frequentist interpretation, confidence intervals have an associated *coverage*. It describes the expected fraction of experiments for which the derived confidence interval contains the true value S' of the parameter of interest [196]. While $\mathrm{CL_{s+b}}$ limits provide perfect coverage, their use is discouraged due to the unphysical feature described above.

$\mathrm{CL_s}$ exclusion limits are deliberately designed to *overcover*, i.e. strictly less than 5 % of experiments are expected to find 95 % $\mathrm{CL_s}$ limits that falsely exclude a true signal. However, an improper propagation of fit uncertainties or free parameters within the likelihood could potentially lead to *undercoverage*, i.e. more than 5 % of experiments excluding a true signal. This was observed for the implementation of uncertainties in the previous TLA publication [1].[5]

A coverage test is performed to ensure the validity of the $\mathrm{CL_s}$ limits derived in this analysis. For this test, signals are injected into toy experiments and for each of them the derived upper limit is compared to the true signal amplitude S'. This is done for 200 toys and various signal amplitudes as discussed in Section 9.5.3.

Figure 10.6 shows the results of the coverage test for Gaussian signals as a function of the signal amplitude. The top panels of each figure show the distribution of the $\mathrm{CL_s}$ upper limits derived in the toy experiments. The coloured boxes represent the median and the 16 % and 84 % quantiles such that 68 % of the distribution is contained within them. The coloured dashed lines represent the 5 % and 95 % quantiles. The dashed grey line in the top panels indicates the amount of injected signal below which at most 5 % of toy experiments should observe their upper limit. The bottom panels of each figure show the false exclusion rate, i.e. the fraction of toys for which the 95 % $\mathrm{CL_s}$ upper exclusion limit is below the true amplitude S'.

In both SRs and for all signals under test, proper coverage is observed. The false exclusion rate is expectedly at 0 for vanishing signal amplitude, as there is no true signal to falsely exclude. The false exclusion rate then approaches 5 % (within the statistical uncertainty due to the limited number of toy experiments) with increasing amplitude.

[5] In previous ATLAS dijet searches (trigger-level and offline), the fit parameters $\boldsymbol{p}$ were not considered as free parameters when evaluating the s+b likelihood (compare Equation (9.23) for this analysis) during the derivation of exclusion limits. Instead, a nominal background estimate and an uncertainty envelope, parametrized by one nuisance parameter, were used. This way, the possibility of the background component of the fit adapting its shape to the presence of signal was not completely accounted for. This can result in too-strict limits for wide signals and signals of low masses by a factor of up to 2 in the cross section [197, 198].

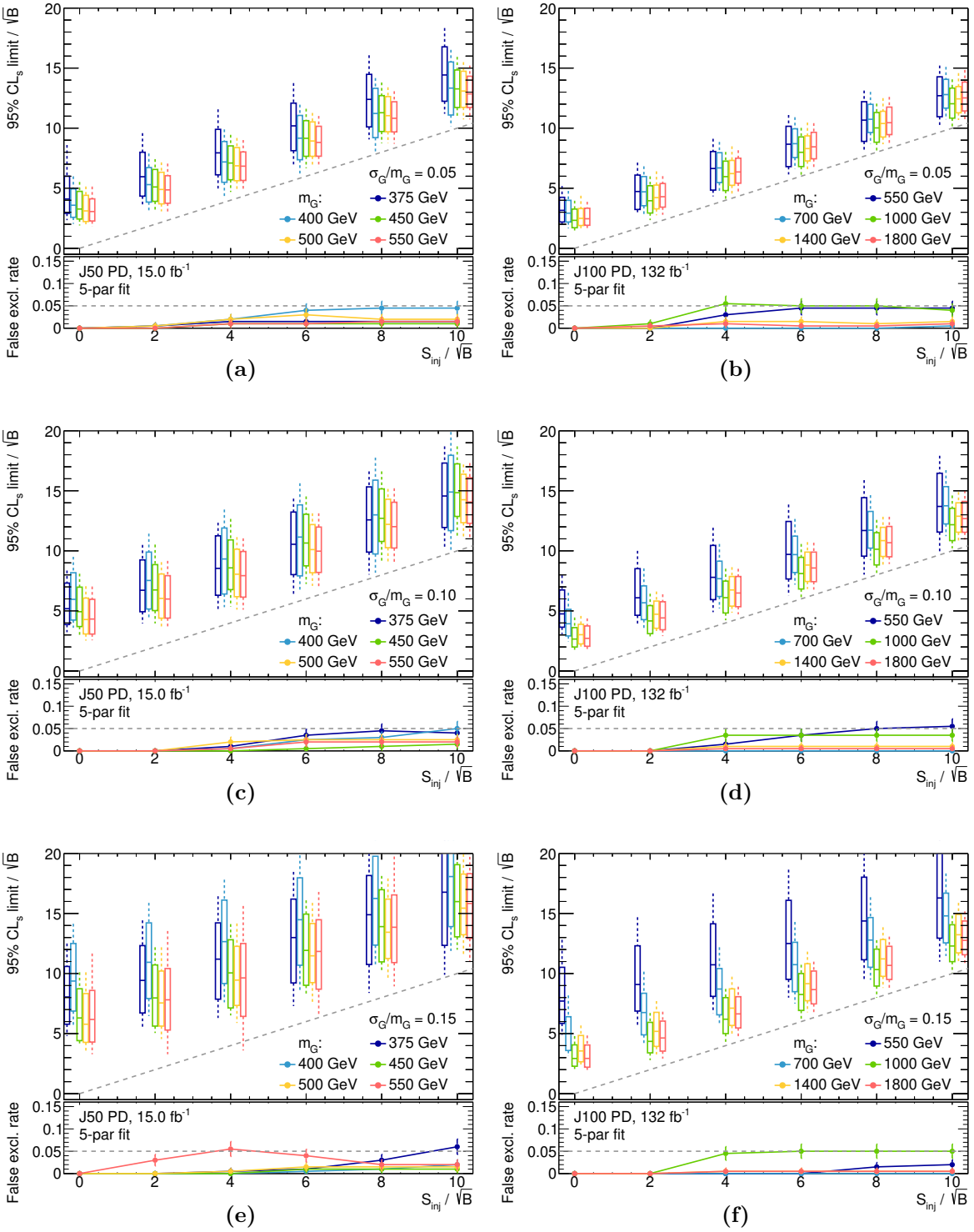


Figure 10.6: Median and spread of the 95 % CL_s upper exclusion limits determined in 200 toys of (**a**,**c**,**e**) J50 PD and (**b**,**d**,**f**) J100 PD. Gaussian signal of 5 %, 10 %, or 15 % width at various masses are injected. The coloured boxes indicate the median and the central 68 % of the distribution, while the coloured dashed lines correspond to the central 90 % of the distribution. The dashed grey line in the top panels indicates the amount of injected signal below which at most 5 % of toy experiments should fall. The bottom panels of each figure show the false exclusion rate, i.e. the fraction of toys for which the 95 % CL_s upper exclusion limit is below the true amplitude S'. Small horizontal shifts around the integer injection amplitudes are added to improve the readability of the different signal masses.

This validates the proper coverage of the CL_s method for this analysis. Consequently, it is applied to derive exclusion limits on generic Gaussian resonances and Z' resonances in the next sections.

10.2.3 Gaussian Resonances

In the derivation of exclusion limits on resonances, the two SRs in this analysis must be accounted for. While the J50 SR provides sensitivity to resonances of lower mass than the J100 SR, there is an overlap region in which both SRs are sensitive to signals. The straight-forward approach of deriving the observed limits from both SRs and using the stricter of the two results in the overlap region would be subject to the look-elsewhere effect. This would need to be corrected for with a *trials factor* [187, 199] that takes into account that two hypothesis tests for the same hypothesis were performed. Additionally, the J50 SR mostly containing events that are also part of the J100 SR in the overlapping m_{jj} region would need to be accounted for.

To avoid such a trials factor, only one SR is used to derive limits for any given signal hypothesis. The decision for a SR is based on the *expected* upper limit while the analysis is still blinded to not introduce any bias.

In general, the sensitivity is expected to be greater in the J100 SR due to its larger recorded luminosity, but at low masses, this could hypothetically be outweighed by the loss of sensitivity towards the edge of the m_{jj} fit range as observed in Section 9.5. A comparison of the expected limits shows that the sensitivity in the J100 SR is stronger than in the J50 SR already at the lowest signal mass of 550 GeV for which the J100 SR s+b fits are validated. Thus, the J100 SR is used to derive limits on resonances with a mass of 550 GeV or above and the J50 SR for masses below 550 GeV.

Exclusion limits on the signal cross section σ instead of the number of reconstructed signal events S are derived by identifying

$$S = \sigma \cdot A \cdot \varepsilon \cdot \mathcal{B} \cdot L \tag{10.11}$$

in Equation (9.18), where L is the integrated recorded luminosity of the respective SR. The ATLAS Run-2 luminosity measurement is associated with a 0.8 % uncertainty [118], which has a negligible effect on the exclusion limits. $\mathcal{B}$ is the branching ratio of the hypothetical resonance to decay as a dijet signature and $A \cdot \varepsilon$ describes the acceptance times efficiency of the applied event selection for signal events. As this section presents the search for a generic resonance instead of a specific model, limits are derived in terms of $\sigma \cdot A \cdot \varepsilon \cdot \mathcal{B}$. The results can then be interpreted for arbitrary signal models if they produce dijet resonances of approximately Gaussian shape in m_{jj}. Exemplary values for $A \cdot \varepsilon$ in Z' decays are given in Table 8.1.

Figure 10.7 shows the observed (solid lines) and expected (dashed lines) 95 % CL_s exclusion limits on Gaussian dijet resonances of 5 % width (blue), 10 % width (orange), and 15 % width (red). While the uncertainty bands of the 10 % width and 15 % width resonances are omitted for visibility, they are of similar relative size as the shown $\pm 1, 2\sigma$ bands of the 5 % width resonance. The observed limits for all signal hypotheses are in good agreement with the background-only expectation. This is consistent with the absence of an observed excess in the background-only interpretation in Section 10.1.

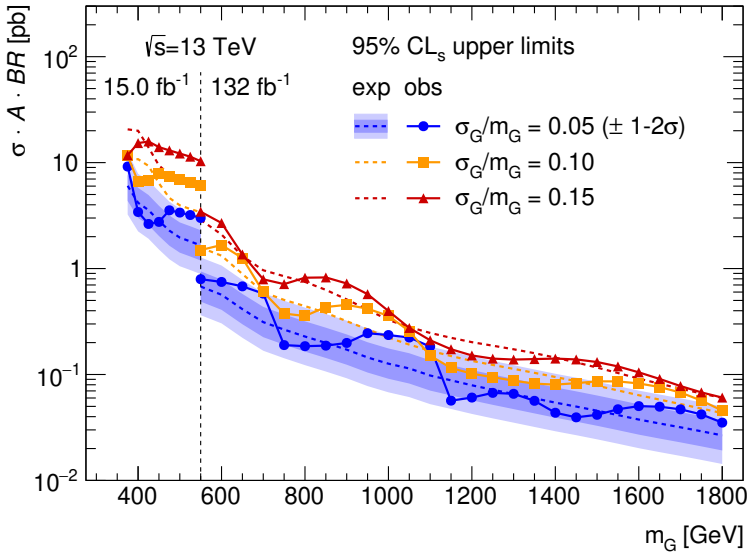


Figure 10.7: 95 % CL_s exclusion limits on Gaussian dijet resonances of 5 % width (blue), 10 % width (orange), and 15 % width (red). The observed limits are shown as solid lines and the expected limits as dashed lines. The $\pm 1, 2\sigma$ uncertainty bands on the expected limits are shown as shaded bands for the 5 % width resonances. The size of the uncertainty bands for the wider signals is similar on the logarithmic scale.

Gaussian resonances of 5 % width with $\sigma \cdot A \cdot \varepsilon \cdot \mathcal{B}$ above 9 pb at $m_G = 375$ GeV (decreasing to 35 fb at $m_G = 1800$ GeV) can be excluded at 95 % CL. For resonances of 10 % (15 %) width, the exclusion limits are weaker by a factor 1.5–2 (2–4).

The sensitivity to narrow resonances is better than the sensitivity to wide resonances for two reasons: For wide resonances, the same number of signal events is distributed across a larger m_{jj} range, which lowers the signal-to-background ratio in this range. Additionally, the flexibility of the background fit allows it to better adapt to wide signals than to narrow signals, resulting in a larger ambiguity between signal and background events.

The observed limits exhibit plateau-like structures at the largest excesses around masses of 500 GeV and 1000 GeV. These are a result of the relatively large JES uncertainty, which is dominated by the flavour composition uncertainty of 15–40 GeV (see Section 8.3). As a result of the JES uncertainty, an upward fluctuation at any position in the m_{jj} spectrum could be caused by a resonance within a certain mass range around the excess. This causes a plateau of similar likelihood for a signal around excesses.

Some mitigation of this behaviour is possible by the reduced flavour composition uncertainty that is achievable with the GSC in the calibration chain. Alternatively or additionally, exclusion limits could be derived separately for different assumed flavour compositions of the resonance, as, for example, demonstrated in Ref. [4].

Figure 10.8 compares the exclusion limits on Gaussian resonances of 5 % width achieved in this analysis (blue) to the previous TLA publication [1] (grey) and to the ATLAS offline dijet resonance search [3] (black).

The observed limits exhibit similar structures in the previous result and this analysis

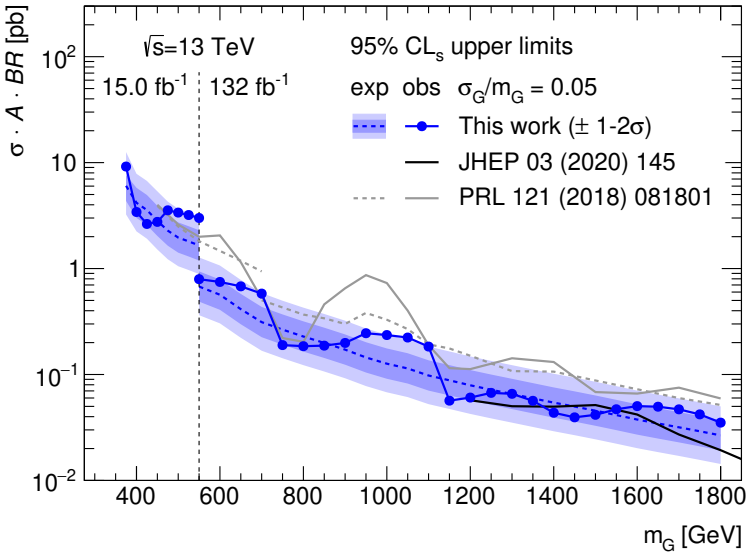


Figure 10.8: Comparison of the 95 % CL_s exclusion limits on Gaussian dijet resonances of 5 % width derived in this analysis (blue) to the previous TLA result [1] (grey) and the offline dijet resonance search [3] (black). The observed limits are shown as solid lines and the expected limits as dashed lines. The $\pm 1, 2\sigma$ uncertainty bands on the expected limits of this analysis are shown as shaded bands.

since the data used in the former are mostly a subset of the data used in the latter.

In the J100 SR, the limits of this analysis are stronger than the previous result by a factor of approximately 2, which is compatible with a $1/\sqrt{L}$ scaling of the sensitivity with the luminosity. Additionally, the J100 SR constrains resonances starting at a mass of 550 GeV, compared to 700 GeV previously. This is achieved for mainly two reasons: Firstly, the minimal m_{jj} threshold of the fit range is lowered by ensuring full trigger efficiency in terms of an m_{jj} requirement instead of a p_{T} requirement (see Section 7.2.3). Secondly, a global fit is used for the background estimate instead of a sliding-window fit, which has a stronger loss of sensitivity at the edges of the fit range due to fewer bins constraining the fit. The J50 SR also reaches lower resonance masses than probed before, due to using the lower-threshold trigger `L1_J50` instead of the previously used `L1_J75`, in addition to the other reasons listed for the J100 SR.

Another key advantage of the global fit in comparison to the previously used sliding-window fit is the sensitivity to wide signals. Previously, the relatively narrow fit window size reduced the constraint on the background estimate to prevent an adaptation to wide signals [166, 167]. Thus, only Gaussian resonances up to widths of 10 % in the low-mass SR and 7 % in the high-mass SR could be tested. In this analysis, exclusion limits on Gaussian resonances up to a width of 15 % are presented, regardless of their mass.

For the offline dijet resonance search, no expected limits on Gaussian resonances of 5 % width are published. However, the observed limits are compatible with the expectation of this analysis within its uncertainties. They are not identical to the observed limits of this search due to the difference between the trigger-level and offline jet reconstruction and calibration and, accordingly, the difference between trigger-level and offline m_{jj}.

Although both analyses are likely to consider the same collision events in the overlapping mass region, small differences in their reconstructed m_{jj} result in these collision events being mapped to different bins. This causes different statistical fluctuations in the considered m_{jj} spectra.

While the results for Gaussian resonances with a width of 5 % are compatible, the offline dijet resonance search exhibits a smaller loss in sensitivity to wide signals. This is likely due to the different treatment of fit uncertainties in the sliding-window background estimate, as discussed in the following section.

10.2.4 Z' Resonances

Given the absence of an observed dijet excess, the parameter space of the leptophobic Z' DM model can be constrained. 95 % $\mathrm{CL_s}$ exclusion limits are derived analogously to the Gaussian resonances.

Since $A \cdot \varepsilon$ was determined for this model (see Table 8.1), $\mathcal{B}$ and σ are the only free parameters in Equation (10.11). Assuming a branching ratio of 100 %, this can be translated into the coupling g_q as the free parameter using the quadratic scaling of the cross section with g_q (Equation (2.11)) and the predicted total cross section for the nominal value of $g_q = 0.1$. The uncertainty on the cross section calculation is added as uncertainty on S in Equation (9.18). The derived cross section limits can easily be interpreted for the parameter space where a decay of the Z' mediator into DM fermions is kinematically accessible by correcting for the reduced branching ratio – as long as the Z' decay width (see Equation (2.9)) does not increase significantly due to the additional decay channel.

Figure 10.9 shows the obtained exclusion limits on Z' resonances in terms of the coupling g_q. The previous TLA publication (grey) and the offline dijet resonance search (black) are included for comparison. The observed limits are compatible with the expectation under the background-only hypothesis, as is the case for the Gaussian resonance interpretation. A similar plateau effect around the most significant excesses is observed as well.

The results improve on the previous TLA publication in a similar way as for the Gaussian resonances: Lower masses are probed with both SRs and the increased luminosity improves the sensitivity. However, this improvement is not uniform in $m_{Z'}$: At high masses, the limits on g_q obtained in this analysis are stricter by approximately 30–50 %, which is consistent with the expected $1/\sqrt[4]{L}$ scaling.[6] With decreasing mass, the improvement in this analysis vanishes. This can be partially attributed to the omission of the GSC in this thesis. As the previously used GSC was less performant at high jet momenta [166], potentially due to not correcting for f_{TileGap3}, the loss in sensitivity by omitting the GSC in this thesis is lower at high masses. However, the probably more dominant factor is the different treatment of fit uncertainties in this thesis (see the discussion in Section 10.2.2). Previously, undercoverage was observed for wider signals below 1.5 TeV, probably due to not treating the fit parameters $\boldsymbol{p}$ as free parameters in the likelihood [198]. The exclusion limits presented in this thesis provide proper coverage

[6] An additional power of 1/2 needs to be considered when comparing limits on g_q due to its quadratic relation to the cross section.

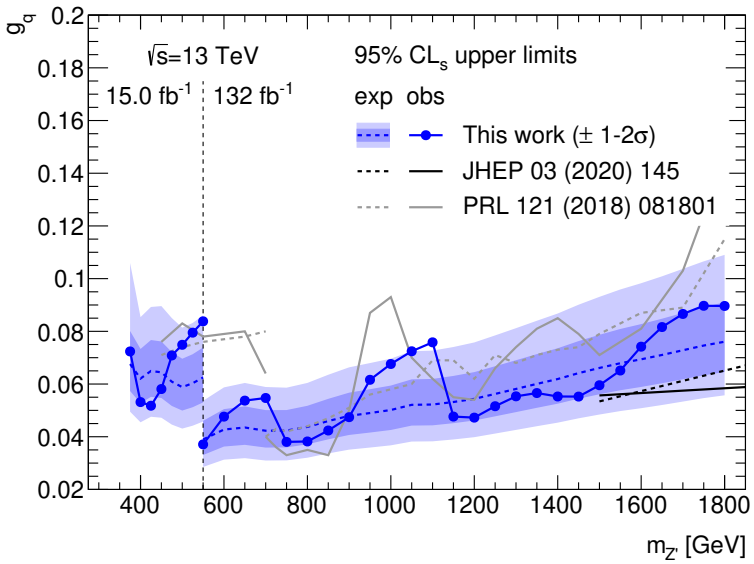


Figure 10.9: 95 % CL_s exclusion limits on the coupling g_q of a potential leptophobic Z' resonance decaying exclusively into dijets. The results of this analysis (blue) are compared to the previous TLA result [1] (grey) and the offline dijet resonance search [3] (black). The observed limits are shown as solid lines and the expected limits as dashed lines. The $\pm 1, 2\sigma$ uncertainty bands on the expected limits of this analysis are shown as shaded bands.

and are thus not more stringent than the previous results for low masses.

The J50 SR exhibits a lower sensitivity penalty in comparison to the J100 SR than the low-mass SR in the previous result. This is due to the higher acceptance with a uniform requirement of $|y^*| < 0.6$ in both SRs for this analysis, while previously $|y^*| < 0.3$ was used to reduce the minimal m_{jj} threshold of the low-mass SR.

The comparison to the offline dijet resonance search shows approximately 25 % weaker limits obtained in this analysis, which corresponds approximately to a 1σ difference in terms of the uncertainty on the expected limits. This can be attributed to the superior m_{jj} resolution of the offline jet reconstruction where a GSC is applied. Additionally, the different treatment of fit uncertainties and, consequently, a potentially overestimated sensitivity to wide signals also apply here. These two effects could be distinguished in following studies by deriving Z' exclusion limits with the calorimeter-based GSC enabled. The uncertainty treatment should then be the dominant difference between the trigger-level and the offline dijet resonance search, such that its effect can be quantified.

Overall, this analysis presents a direct extension of the offline dijet resonance search to significantly lower resonance masses. While reconstructed tracks enhance the resolution for offline jets, a good approximation is achieved by the trigger-level jets using only calorimeter information. With only little loss in sensitivity, the trigger-level dijet resonance search explores an otherwise hardly accessible region of the parameter space at very high statistical precision. The remaining gap to offline analyses can be further reduced with future ATLAS Run-3 analyses, for which track reconstruction is included on trigger level.

11 Conclusion and Outlook

Searches for dijet resonances are sensitive to a variety of models for physics beyond the Standard Model, such as a new mediator coupling to both Dark Matter and Standard Model particles. With conventional analysis approaches at the Large Hadron Collider, these searches have reduced statistical power in the sub-TeV range due to limitations in the trigger and readout bandwidth.

This thesis presents a search for dijet resonances of low mass in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ corresponding to an integrated luminosity of up to $132\,\text{fb}^{-1}$. The bandwidth limitation is overcome by using a partial event readout consisting solely of the jets reconstructed by the ATLAS High-Level Trigger.

The usage of trigger-level jets requires a custom jet calibration for this analysis to achieve a resolution and uncertainties comparable to the conventional approach. This calibration must be a smooth function of the jet energy to ensure that no resonance-like structures are induced in the recorded dijet mass spectrum. The trigger-level jet calibration chain consists of six steps, two of which are identified to not meet the required smoothness constraints. With the omission of these steps, the dijet mass spectrum resulting from Standard Model processes is smoothly falling. This allows for a search for resonances indicating physics beyond the Standard Model atop the smooth background.

The large cross section of dijet production at low masses results in a very high statistical precision of the recorded dijet mass spectrum. Two complementary methods for the background estimate are presented that can match this level of precision: a fit with a heuristic function and the novel Monte Carlo-based NLOFit. Both provide a good description of the observed spectrum and are compatible with each other, which serves as a cross-validation of the two methods.

Two signal regions, each covering different dijet mass ranges, are analyzed. Excesses over the background are quantified with the BumpHunter algorithm, which yields global p-values between 0.17 and 0.99. Given that no significant excess is observed, the data set is used to constrain physics beyond the Standard Model by deriving 95 % confidence level exclusion limits on resonance cross sections. To the author's knowledge, this thesis presents the most stringent constraints on dijet resonances in the range of 375–1200 GeV for generic, Gaussian-shaped dijet resonances and in the range of 375–1500 GeV specifically for a leptophobic axial-vector Z' Dark Matter mediator. It improves on previous cross section exclusion limits by a factor of up to 2 and extends the probed mass range by 75 GeV towards lower masses.

This search is primarily limited by statistical uncertainties, although a better jet resolution could enhance the sensitivity to narrow resonances. This is expected to be achieved in the forthcoming ATLAS publication for this search, for which the two omitted jet calibration steps have been revised. Further improvements are to be expected from subsequent ATLAS trigger-level analyses performed with Run-3 data, which have been expanded in scope following the success of the presented search. The reconstruction of

tracks for trigger-level jets improves the jet resolution and allows for the identification of jets containing b hadrons for which the Standard Model background is lower. Furthermore, the inclusion of the signature of a dijet in association with a photon from initial state radiation into the trigger-level readout allows for the probing of even lower resonance masses.

A Data Scouting Triggers

Table A.1 lists the integrated luminosity recorded by each of the Data Scouting triggers per year after applying standard data quality requirements. At the start of Run-2, in 2015, `L1_J75` was the lowest unprescaled L1 single jet trigger. With increased pile-up and accordingly larger trigger rates in 2016, `L1_J75` received a prescale and `L1_J100` became the new lowest unprescaled L1 single jet trigger. The `L1_J75` dataset was already fully analysed in a previous publication [1] and is thus not considered in this thesis.

Additional prescaled single jet L1 triggers were added to the Data Scouting stream in 2017. `L1_J50` was activated towards the end of LHC fills when bandwidth became available. `L1_J15`, `L1_J20`, and `L1_J30`, on the other hand, were almost exclusively active during special run conditions with lower trigger rates: dedicated low-$\langle\mu\rangle$ runs and during $\sqrt{s} = 5\,\text{TeV}$ collisions.

In 2018, `L1_J40` was added as a second end-of-fill trigger to the Data Scouting stream. But as discussed in Section 7.2, it does not offer additional mass range coverage in comparison to `L1_J50` due to the required minimum jet p_T threshold and is thus not analysed further. The same holds true for the lower-threshold triggers. The L1Topo dijet trigger `L1_J50_DETA20-J50J` became active in 2018 as well. It selects events with a leading Level-1 jet with $E_\text{T} > 50\,\text{GeV}$ and a subleading jet with $E_\text{T} > 15\,\text{GeV}$ and a rapidity difference $|y^*| < 1.0$. An error in the jet ordering caused an inefficiency of this trigger prior to the Technical Stop 1 (TS1) in 2018 [146]. Accordingly, only the data taken after TS1 are used in this thesis and listed in Table A.1.

The total luminosity collected by the different triggers seeding the Data Scouting stream is slightly lower than the full Run-2 luminosity available for offline analyses since the stream was not active in the early stages of 2015 and 2016 data-taking.

Trigger	Integrated Luminosity [pb^{-1}]				
	2015	2016	2017	2018	Total
J100	-	29 378	44 307	58 450	132 135
J75*	3220	3580	-	-	6800
J50	-	-	4120	1515	5635
J50Topo	-	-	-	10 142	10 142
J40*	-	-	-	3175	3175
J30*†	-	-	263	275	538
J20*†	-	-	259	174	433
J15*†	-	-	140	-	140

Table A.1: Integrated luminosity (after applying the data quality selection) recorded by each Data Scouting trigger in Run-2. Triggers not considered in this thesis are marked by asterisks. † symbols mark triggers that were mainly active during special run conditions, i.e. low-μ runs or $\sqrt{s} = 5\,\text{TeV}$ collisions.

B Additional Calibration Smoothness Results

In addition to the results discussed in Section 6.3, Figure B.1 shows the smoothness of the individual trigger-level calibration steps in the J50 data set as a function of subleading jet p_{T}. Similar behaviour of a smooth pile-up correction, MCJES and in-situ calibration is observed, while the GSC response exhibits a sharp kink at 350 GeV. Features in the online-to-offline correction are visible, but less pronounced than in the J100 data set.

Figure B.2 shows the relative effect of each calibration step on the reconstructed dijet invariant mass in the J100 data set. As a function of m_{jj}, the calibration features of the GSC and the online-to-offline correction are more washed out because their calibration factors are defined as functions of p_{T}. Due to the η-dependent relation between m_{jj} and p_{T} being averaged over, this leads to a smoother response as a function of m_{jj}. Nonetheless, the features in the average responses of the GSC and the online-to-offline correction are still visible.

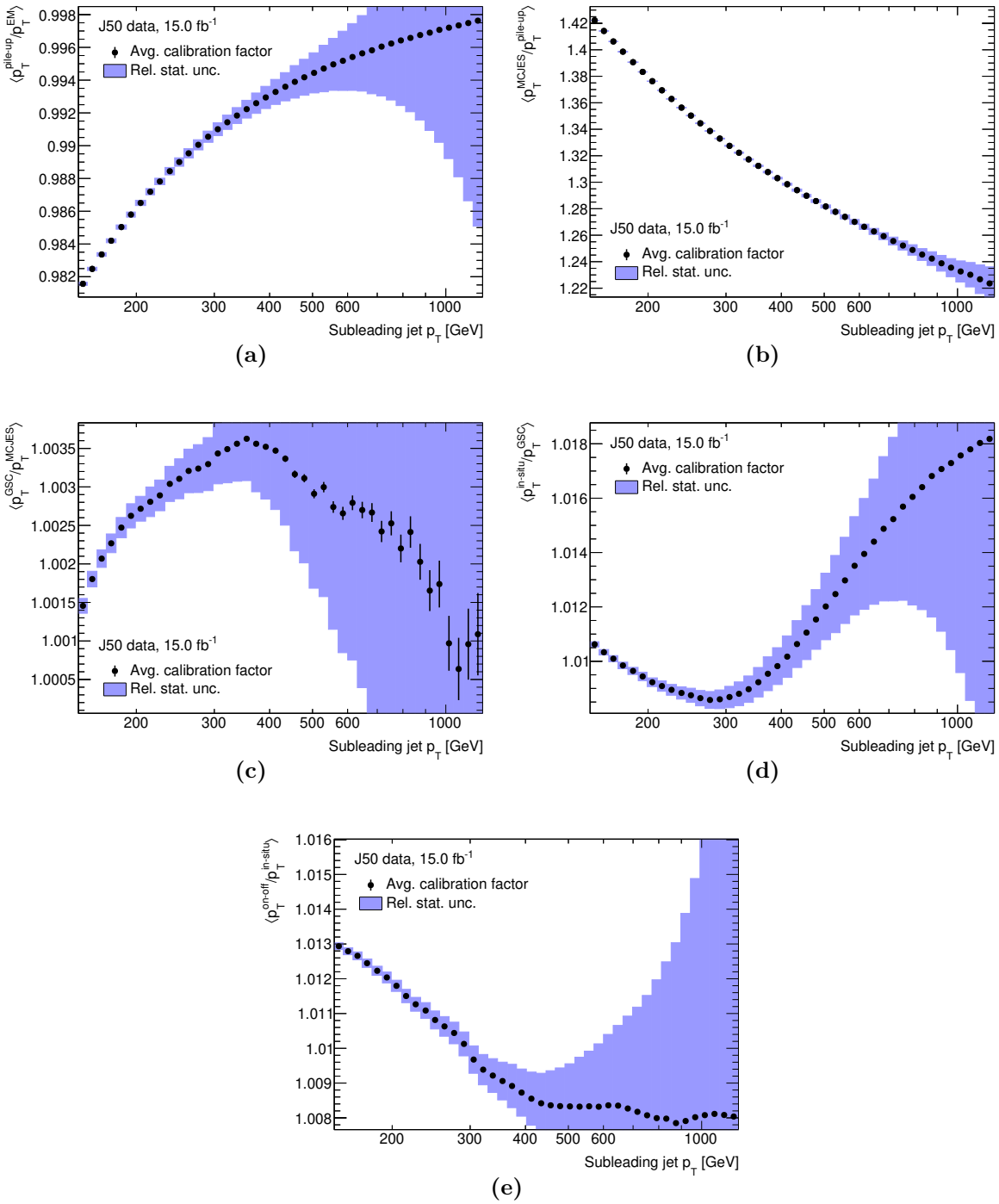


Figure B.1: Average correction factor applied to the subleading jet in each calibration step as a function of p_{T}, determined in the J50 data set: (**a**) Pile-up correction, (**b**) MCJES, (**c**) GSC, (**d**) η intercalibration and in-situ calibration combined, and (**e**) online-to-offline correction. Black bars correspond to the uncertainty of the mean of the factors. The relative statistical uncertainty of the measured p_{T} distribution is overlayed in blue to identify statistically significant features in the calibration curves.

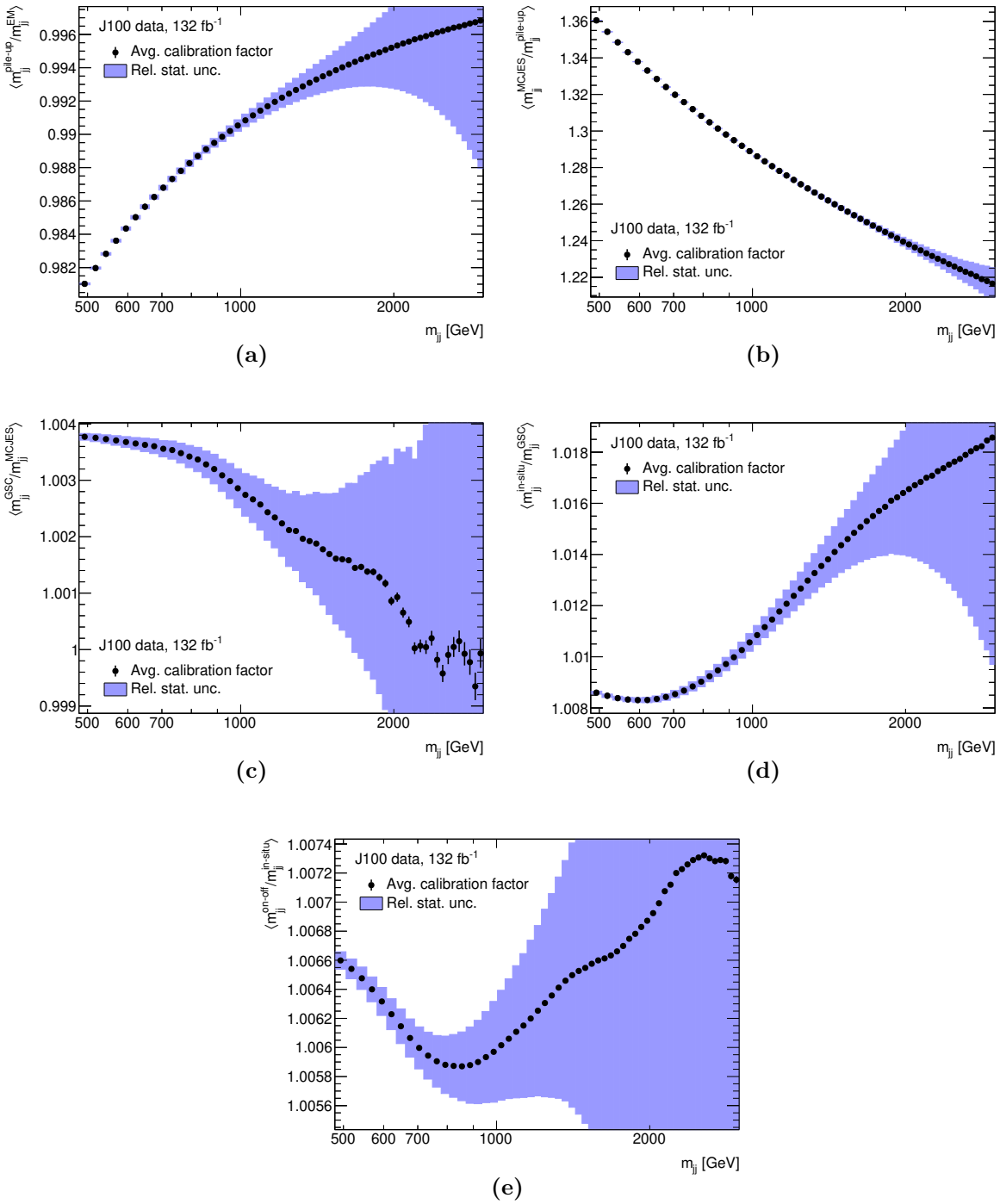


Figure B.2: Average relative change of m_{jj} at each calibration step as a function of m_{jj}, determined in the J100 data set: (**a**) Pile-up correction, (**b**) MCJES, (**c**) GSC, (**d**) η inter-calibration and in-situ calibration combined, and (**e**) online-to-offline correction. Black bars correspond to the uncertainty of the mean of the factors. The relative statistical uncertainty of the measured p_{T} distribution is overlayed in blue to identify statistically significant features in the calibration curves.

C Signal Uncertainties

The effect of the considered jet reconstruction uncertainties on the m_{jj} distribution of the Z' signals is summarized in this appendix, following the discussion in Section 8.3.

Figure C.1 shows the effect on the mean μ of the fitted DSCB distribution for the four studied mediator masses. Throughout the mass range, the flavour uncertainties, which are inflated due to the omitted GSC, dominate. The uncertainty due to the omission of the online-to-offline correction is the subleading uncertainty throughout the mass range.

Figure C.2 shows the effect on the width σ of the fitted DSCB distribution, respectively. Here, too, the flavour uncertainties dominate. The effect of the omitted online-to-offline correction on the width is zero by the definition of its implementation as uncertainty on the average jet scale only. Instead, the in-situ uncertainties of the jet energy scale and resolution are the subdominant source of uncertainty on the signal width.

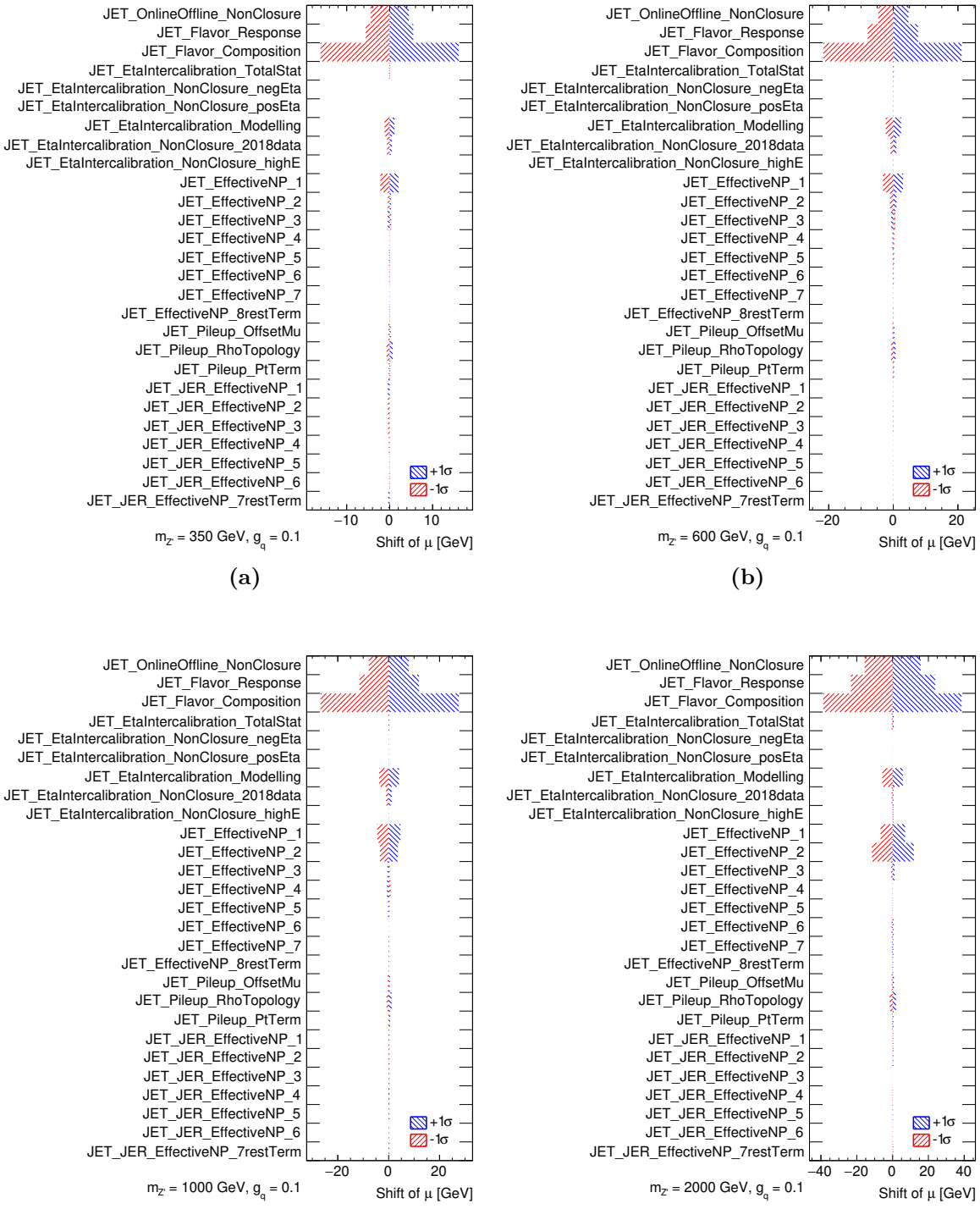


Figure C.1: Effect of the individual systematic uncertainties on the mean μ of the DSCB fits to the Z' signal templates for different values of m'_Z.

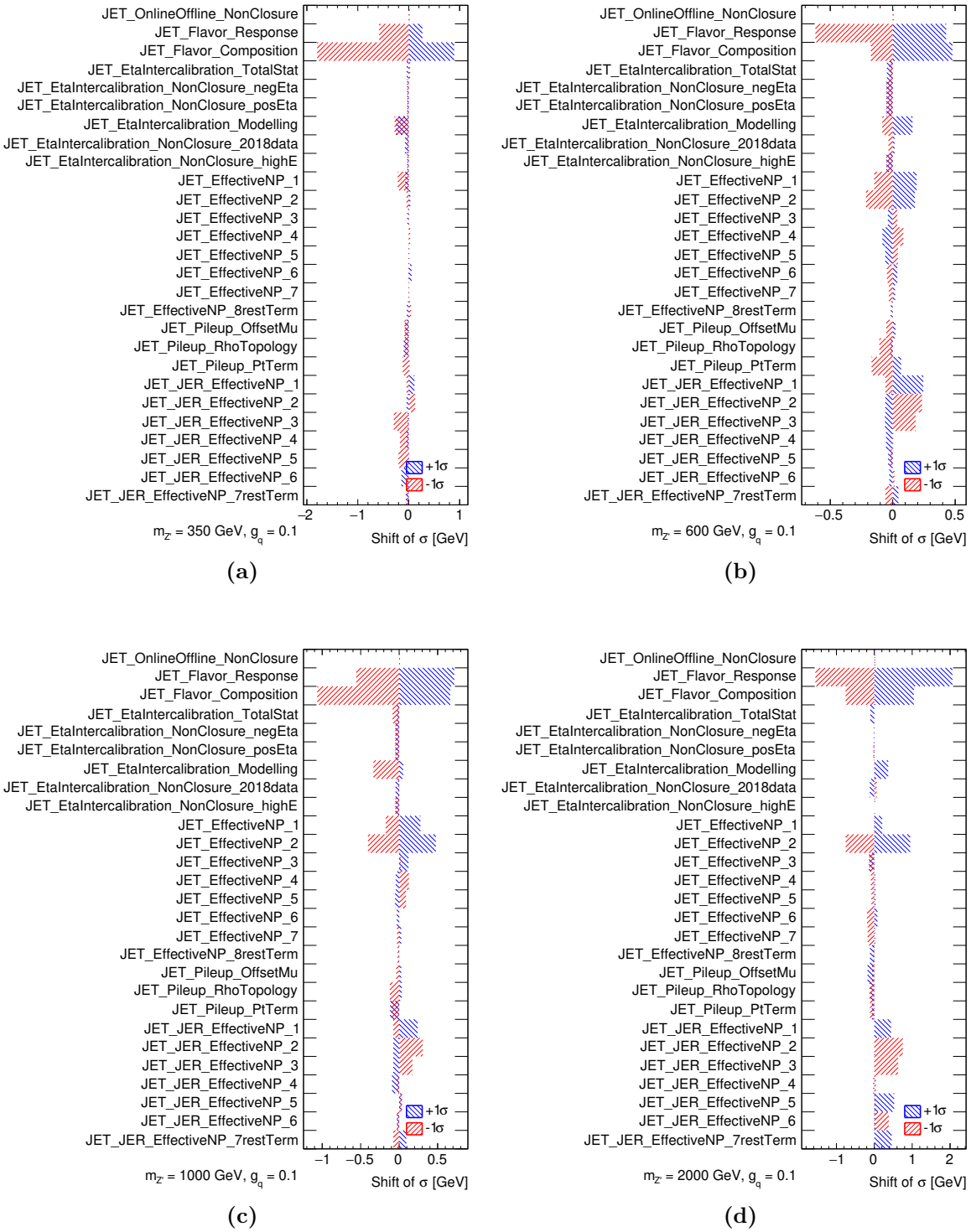
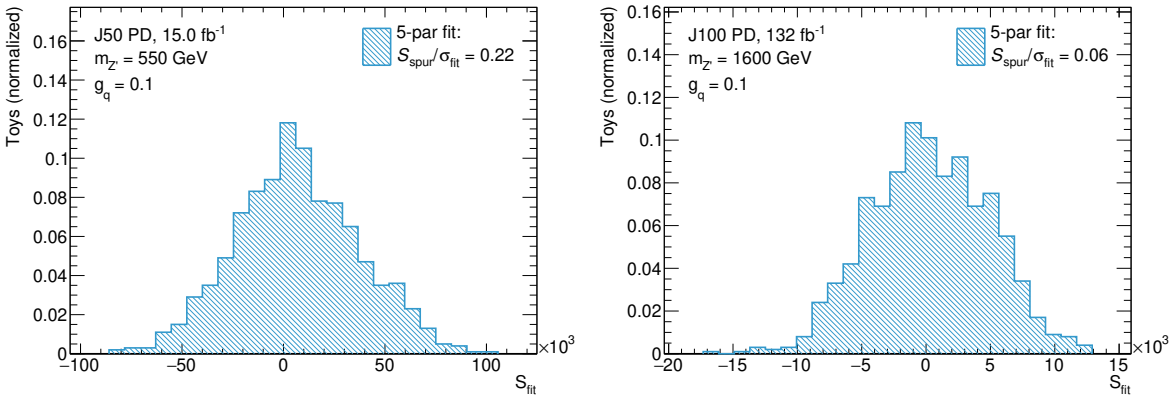


Figure C.2: Effect of the individual systematic uncertainties on the width σ of the DSCB fits to the Z' signal templates for different values of m'_Z.

D Additional Background Validation Results

D.1 Spurious Signal for Z' Resonances

The results of the spurious signal test for the functional form fit with Z' signals are very similar to those obtained for Gaussian resonances discussed in Section 9.5.2. The distribution of S_{fit} is approximately Gaussian, as shown for two exemplary Z' masses in Figure D.1 with 1000 toy experiments each. Figure D.1 shows that S_{spur} remains below the required threshold of $0.5\sigma_{\text{fit}}$ for a 5- or 6-parameter functional form fit.



(a) **(b)**

Figure D.1: Exemplary distributions of the number of signal events determined by s+b fits in background-only PD. Functional form fit results with a 5-parameter dijet function, performed on PD from a 6-parameter template, are shown. The chosen signal hypotheses are: (**a**) a Z' resonance with a mass of 550 GeV and (**b**) a Z' resonance with a mass of 1600 GeV.

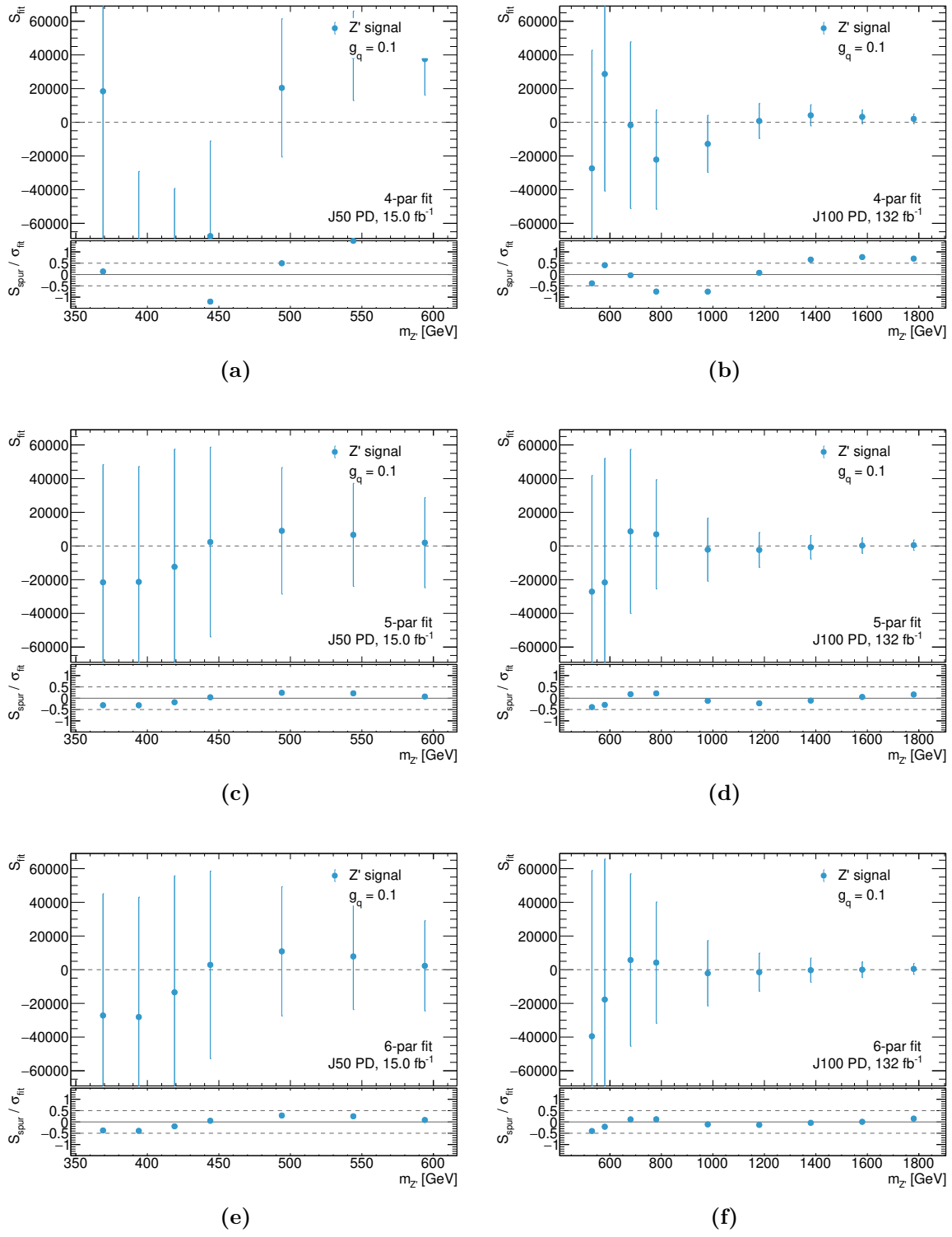
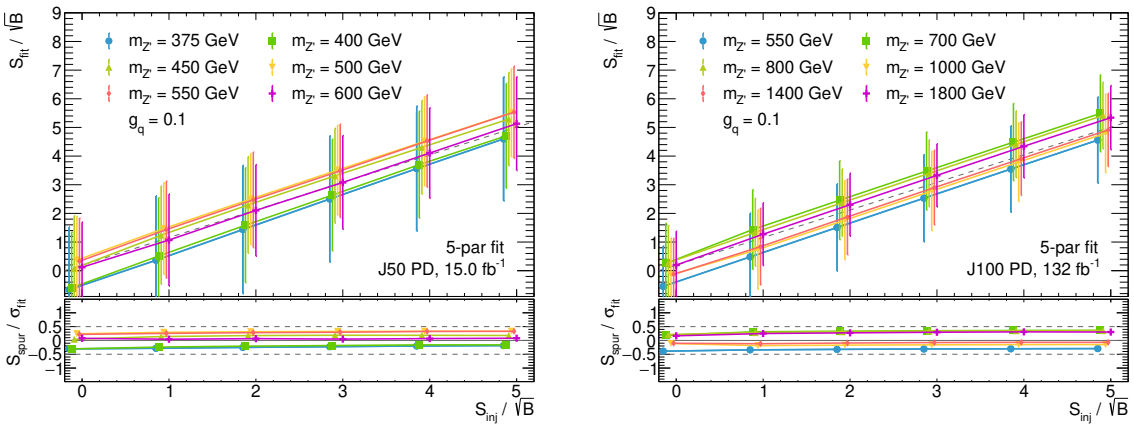


Figure D.2: Summary of S_{spur} (points) and σ_{fit} (error bars) of the spurious signal tests for the functional form fit (**a**,**c**,**e**) to J50 PD and (**b**,**d**,**f**) to J100 PD for different choices of the number of parameters N. The results for Z' signals are shown as a function of Z' mass. The bottom panel in each figure shows the ratio $S_{\text{spur}}/\sigma_{\text{fit}}$ with the dashed lines indicating the required threshold of $\left|S_{\text{spur}}\right| < 0.5\sigma_{\text{fit}}$.

D.2 Injection of Z' Resonances

The behaviour of Z' s+b functional form fits in the presence of signal exhibits good linearity, just like for Gaussian signals. The 5- and 6-parameter fits meet the required criterion $S_{\text{spur}} < 0.5\sigma_{\text{fit}}$, as is shown exemplary for the 5-parameter fit in Figure D.3.

The background component of the s+b fits exhibits a larger variance in the region of the injected signal, especially at the edge of the m_{jj} fit range, as shown for selected examples in Figure D.4. This is likely to be caused by the elongated tails of the DSCB distribution describing the Z' resonance shape (see Figure 8.1). These exponential tails could increase the ambiguity with the background component of the s+b fit. Nonetheless, no significant systematic bias of the ensemble is observed, validating the Z' s+b fits.



(a) **(b)**

Figure D.3: Number of signal events determined by a 5-parameter functional form fit $\langle S_{\text{fit}} \rangle \pm \sigma_{\text{fit}}$ in PD as a function of the amount of injected signal using the Z' signal hypothesis of different masses. The absolute value of spurious signal in units of the fit uncertainty is shown in the bottom panel. The dashed lines at ± 0.5 indicate the required maximal limit on the spurious signal. Small horizontal shifts around the integer injection amplitudes are added to improve the readability of the different signal masses.

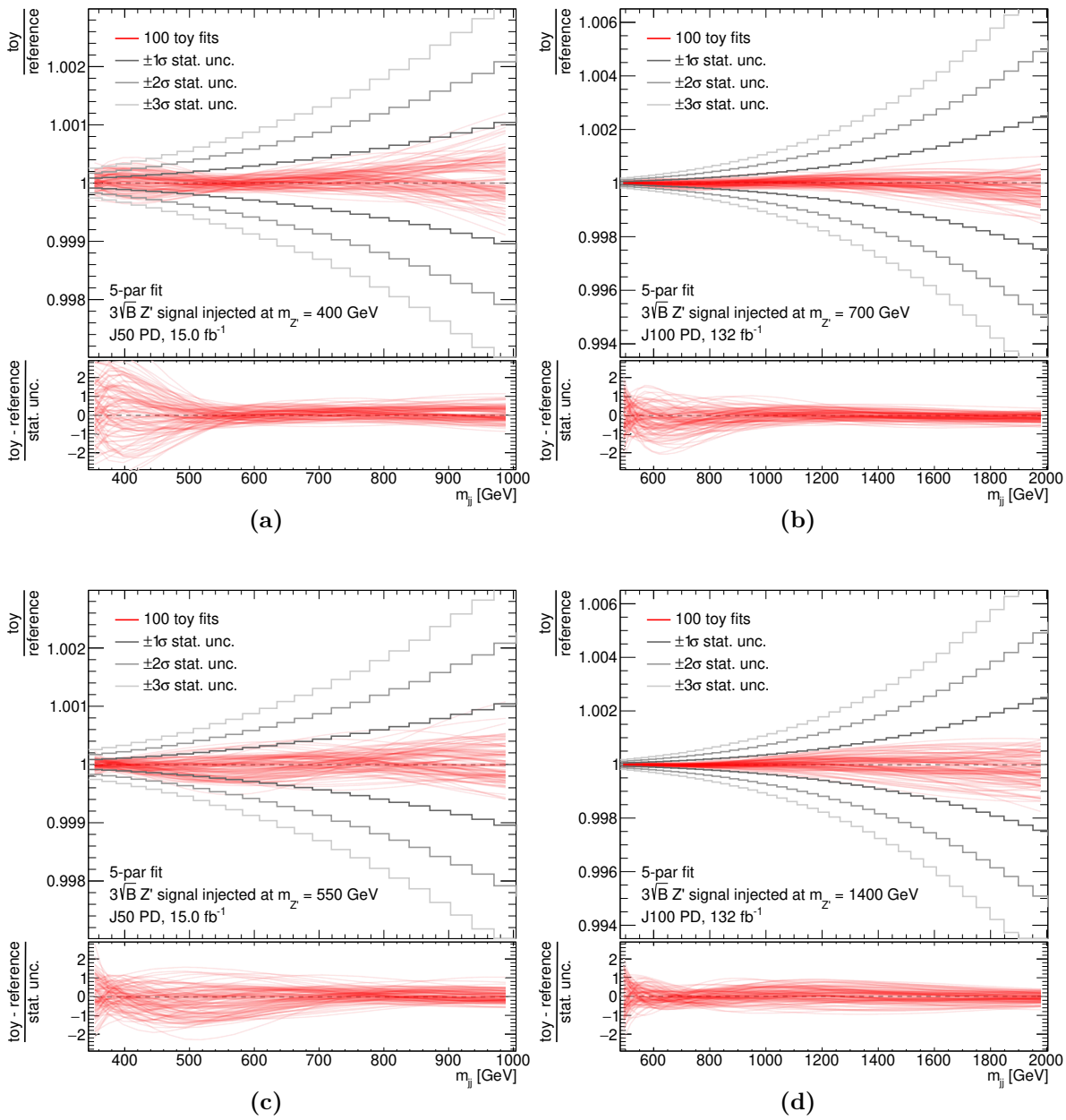


Figure D.4: Background stability test of a 5-parameter functional form fit (**a**,**c**) to J50 PD and (**b**,**d**) to J100 PD for selected signal Z' signal masses. The background component of s+b fits to PD toys with injected signal is compared to the background component of the PD (reference).

E Acronyms

BSM Beyond the Standard Model

CL Confidence level

DGLAP Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

DSCB Double-sided Crystal Ball

DM Dark Matter

EM Electromagnetic

FWHM Full-width-at-half-maximum

GSC Global sequential calibration

HLT High-Level Trigger

ID Inner Detector

JER Jet energy resolution

JES Jet energy scale

jFEX jet Feature EXtractor

L1 Level-1

L1Calo Level-1 Calorimeter Trigger

L1Muon Level-1 Muon Trigger

L1Topo Level-1 Topological Processor

LEP Large Electron-Positron Collider

LHC Large Hadron Collider

Linac Linear accelerator

LO Leading order

MC Monte Carlo

MCJES Monte Carlo-based jet energy scale calibration

MDT Monitored Drift Tube

NLO Next-to-leading order

NP Nuisance parameter

PD Pseudo-data

PDF Parton distribution function

PS Proton Synchrotron

PSB Proton Synchrotron Booster

QCD Quantum chromodynamics

s+b Signal-plus-background

SM Standard Model

SPS Super Proton Synchrotron

SR Signal region

TLA Trigger-Level Analysis

TS1 Technical Stop 1

F Bibliography

[1] ATLAS Collaboration. Search for Low-Mass Dijet Resonances Using Trigger-Level Jets with the ATLAS Detector in pp Collisions at $\sqrt{s} = 13\,\text{TeV}$. *Phys. Rev. Lett.*, 121:081801, August 2018. doi:10.1103/PhysRevLett.121.081801. (Cited on p. 2, 39, 54, 60, 67, 78, 114, 117, 118, 120, and 123)

[2] CMS Collaboration. Search for narrow and broad dijet resonances in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter mediators and other new particles. *Journal of High Energy Physics*, 2018(8):130, August 2018. ISSN 1029-8479. doi:10.1007/JHEP08(2018)130. (Cited on p. 2 and 78)

[3] ATLAS Collaboration. Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2020(3):145, March 2020. ISSN 1029-8479. doi:10.1007/JHEP03(2020)145. (Cited on p. 2, 16, 66, 78, 89, 117, 118, and 120)

[4] CMS Collaboration. Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 2005:033, 2020. doi:10.1007/JHEP05(2020)033. (Cited on p. 2, 16, 78, 89, and 117)

[5] ATLAS Collaboration. Search for low-mass resonances decaying into two jets and produced in association with a photon using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physics Letters B*, 795:56–75, 2019. ISSN 0370-2693. doi:10.1016/j.physletb.2019.03.067. (Cited on p. 2 and 3)

[6] ATLAS Collaboration. Search for light resonances decaying to boosted quark pairs and produced in association with a photon or a jet in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physics Letters B*, 788:316–335, 2019. ISSN 0370-2693. doi:https://doi.org/10.1016/j.physletb.2018.09.062. (Cited on p. 2)

[7] CMS Collaboration. Search for low mass vector resonances decaying to quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 119: 111802, Sep 2017. doi:10.1103/PhysRevLett.119.111802. (Cited on p. 2)

[8] CMS Collaboration. Search for low-mass resonances decaying into bottom quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Physical Review D*, 99(1), January 2019. ISSN 2470-0029. doi:10.1103/physrevd.99.012005. (Cited on p. 2)

[9] A. Collaboration. Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. doi:10.17181/CERN.2LBB.4IAL. (Cited on p. 3)

[10] H. Tockhorn. Improving the signal selection for trigger-level, low-mass resonance searches with ATLAS. Bachelorarbeit, Universität Heidelberg, 2021. (Cited on p. 3 and 66)

[11] E. Van der Vorst. Optimisation of Signal Selection in a Search for Low Mass Mediators in Dijet Signatures. Master's thesis, Imperial College London, 2020. (Cited on p. 3, 66, and 67)

[12] ATLAS Collaboration. Pursuit of paired dijet resonances in the Run 2 dataset with ATLAS. *Physical Review D*, 108(11), December 2023. ISSN 2470-0029. doi:10.1103/physrevd.108.112005. (Cited on p. 3)

[13] M. Thomson. *Modern Particle Physics*. Cambridge University Press, 2013. (Cited on p. 5, 8, 9, 28, and 30)

[14] J. I. Illana and A. Jimenez Cano. Quantum field theory and the structure of the Standard Model. *PoS*, CORFU2021:314, 2022. doi:10.22323/1.406.0314. (Cited on p. 5)

[15] R. M. Harris and K. Kousouris. Searches for dijet resonances at hadron colliders. *International Journal of Modern Physics A*, 26(30n31):5005–5055, December 2011. ISSN 1793-656X. doi:10.1142/s0217751x11054905. (Cited on p. 5, 20, and 22)

[16] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996. (Cited on p. 5)

[17] Particle Data Group et al. Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2022(8), August 2022. ISSN 2050-3911. doi:10.1093/ptep/ptac097. (Cited on p. 5, 6, 10, 12, and 74)

[18] S. L. Glashow. The renormalizability of vector meson interactions. *Nucl. Phys.*, 10:107–117, 1959. doi:10.1016/0029-5582(59)90196-8. (Cited on p. 5)

[19] A. Salam and J. C. Ward. Weak and electromagnetic interactions. *Il Nuovo Cimento (1955-1965)*, 11(4):568–577, Feb 1959. ISSN 1827-6121. doi:10.1007/BF02726525. (Cited on p. 5)

[20] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967. doi:10.1103/PhysRevLett.19.1264. (Cited on p. 5)

[21] H. D. Politzer. Reliable Perturbative Results for Strong Interactions? *Phys. Rev. Lett.*, 30:1346–1349, 1973. doi:10.1103/PhysRevLett.30.1346. (Cited on p. 5)

[22] D. J. Gross and F. Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys. Rev. Lett.*, 30:1343–1346, 1973. doi:10.1103/PhysRevLett.30.1343. (Cited on p. 5)

[23] Wikipedia. Standard Model of Elementary Particles, 2019. https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. (Cited on p. 6)

[24] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, August 1998. doi:10.1103/PhysRevLett.81.1562. (Cited on p. 6)

[25] Q. R. Ahmad et al. Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, July 2001. doi:10.1103/PhysRevLett.87.071301. (Cited on p. 6)

[26] F. P. An et al. Observation of Electron-Antineutrino Disappearance at Daya Bay. *Phys. Rev. Lett.*, 108:171803, April 2012. doi:10.1103/PhysRevLett.108.171803. (Cited on p. 6)

[27] Y. Cai et al. From the trees to the forest: A review of radiative neutrino mass models. *Frontiers in Physics*, 5, 2017. ISSN 2296-424X. doi:10.3389/fphy.2017.00063. (Cited on p. 6)

[28] LHCb Collaboration. Observation of structure in the J/ψ-pair mass spectrum. *Sci. Bull.*, 65:1983–1993. 11 p, July 2020. doi:10.1016/j.scib.2020.08.032. (Cited on p. 6)

[29] LHCb Collaboration. Observation of an exotic narrow doubly charmed tetraquark. *Nature Phys.*, page 24 p, September 2021. doi:10.1038/s41567-022-01614-y. (Cited on p. 6)

[30] LHCb Collaboration. Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays. *Phys. Rev. Lett.*, 115:072001, 2015. doi:10.1103/PhysRevLett.115.072001. (Cited on p. 6)

[31] LHCb Collaboration. Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$. *Phys. Rev. Lett.*, 122:222001. 11 p, April 2019. doi:10.1103/PhysRevLett.122.222001. (Cited on p. 6)

[32] C. S. Wu et al. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, February 1957. doi:10.1103/PhysRev.105.1413. (Cited on p. 7)

[33] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of Neutrinos. *Phys. Rev.*, 109:1015–1017, February 1958. doi:10.1103/PhysRev.109.1015. (Cited on p. 7)

[34] G. Bali, C. Schlichter, and K. Schilling. A Ginzburg-Landau analysis of the colour electric flux tube. *Nuclear Physics B - Proceedings Supplements*, 73(1–3):638–640, March 1999. ISSN 0920-5632. doi:10.1016/s0920-5632(99)85160-3. (Cited on p. 9)

[35] X. Fan et al. Measurement of the Electron Magnetic Moment. *Phys. Rev. Lett.*, 130:071801, February 2023. doi:10.1103/PhysRevLett.130.071801. (Cited on p. 10)

[36] E. Berti et al. Testing general relativity with present and future astrophysical observations. *Classical and Quantum Gravity*, 32(24):243001, December 2015. doi:10.1088/0264-9381/32/24/243001. (Cited on p. 10)

[37] A. Ashtekar and E. Bianchi. A short review of loop quantum gravity. *Rept. Prog. Phys.*, 84(4):042001, 2021. doi:10.1088/1361-6633/abed91. (Cited on p. 10)

[38] S. Mukhi. String theory: a perspective over the last 25 years. *Classical and Quantum Gravity*, 28(15):153001, June 2011. ISSN 1361-6382. doi:10.1088/0264-9381/28/15/153001. (Cited on p. 10)

[39] A. Bonanno et al. Critical Reflections on Asymptotically Safe Gravity. *Frontiers in Physics*, 8, 2020. ISSN 2296-424X. doi:10.3389/fphy.2020.00269. (Cited on p. 10)

[40] G. Aad et al. Exploration at the high-energy frontier: ATLAS Run 2 searches investigating the exotic jungle beyond the Standard Model. 3 2024. (Cited on p. 10 and 15)

[41] A. G. Cohen, A. D. Rújula, and S. L. Glashow. A Matter-Antimatter Universe? *The Astrophysical Journal*, 495(2):539, mar 1998. doi:10.1086/305328. (Cited on p. 10)

[42] L. Canetti, M. Drewes, and M. Shaposhnikov. Matter and antimatter in the universe. *New Journal of Physics*, 14(9):095012, September 2012. doi:10.1088/1367-2630/14/9/095012. (Cited on p. 10)

[43] A. D. Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34(5):392, May 1991. doi:10.1070/PU1991v034n05ABEH002497. (Cited on p. 10)

[44] D. S. Pereira et al. Baryogenesis: A Symmetry Breaking in the Primordial Universe Revisited. *Symmetry*, 16(1), 2024. ISSN 2073-8994. doi:10.3390/sym16010013. (Cited on p. 11)

[45] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, June 1963. doi:10.1103/PhysRevLett.10.531. (Cited on p. 11)

[46] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, February 1973. ISSN 0033-068X. doi:10.1143/PTP.49.652. (Cited on p. 11)

[47] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics*, 28(5):870–880, 11 1962. ISSN 0033-068X. doi:10.1143/PTP.28.870. (Cited on p. 11)

[48] B. Pontecorvo. Neutrino Experiments and the Problem of Conservation of Leptonic Charge. *Zh. Eksp. Teor. Fiz.*, 53:1717–1725, 1967. (Cited on p. 11)

[49] T2K Collaboration. Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580(7803):339–344, April 2020. ISSN 1476-4687. doi:10.1038/s41586-020-2177-0. (Cited on p. 11)

[50] C. Abel et al. Measurement of the Permanent Electric Dipole Moment of the Neutron. *Phys. Rev. Lett.*, 124:081803, Feb 2020. doi:10.1103/PhysRevLett.124.081803. (Cited on p. 11)

[51] R. Crewther et al. Chiral estimate of the electric dipole moment of the neutron in quantum chromodynamics. *Physics Letters B*, 88(1):123–127, 1979. ISSN 0370-2693. doi:https://doi.org/10.1016/0370-2693(79)90128-X. (Cited on p. 11)

[52] J. L. Barrow et al. Theories and Experiments for Testable Baryogenesis Mechanisms: A Snowmass White Paper, 2022. (Cited on p. 11)

[53] D. Clowe et al. A Direct Empirical Proof of the Existence of Dark Matter*. *The Astrophysical Journal*, 648(2):L109, aug 2006. doi:10.1086/508162. (Cited on p. 11)

[54] S. W. Randall et al. Constraints on the Self-Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657−56. *The Astrophysical Journal*, 679(2):1173–1180, June 2008. ISSN 1538-4357. doi:10.1086/587859. (Cited on p. 11 and 12)

[55] E. Corbelli and P. Salucci. The extended rotation curve and the dark matter halo of M33. *Monthly Notices of the Royal Astronomical Society*, 311(2):441–447, 01 2000. ISSN 0035-8711. doi:10.1046/j.1365-8711.2000.03075.x. (Cited on p. 12)

[56] V. C. Rubin and W. K. Ford, Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophysical Journal*, 159:379, February 1970. doi:10.1086/150317. (Cited on p. 11)

[57] Y. Sofue et al. Central Rotation Curves of Spiral Galaxies. *The Astrophysical Journal*, 523(1):136, sep 1999. doi:10.1086/307731. (Cited on p. 11)

[58] N. Aghanim et al. Planck 2018 results: I. Overview and the cosmological legacy of Planck. *Astronomy & Astrophysics*, 641:A1, September 2020. ISSN 1432-0746. doi:10.1051/0004-6361/201833880. (Cited on p. 11)

[59] N. Aghanim et al. Planck 2018 results: VI. Cosmological parameters. *Astronomy & Astrophysics*, 641:A6, September 2020. ISSN 1432-0746. doi:10.1051/0004-6361/201833910. (Cited on p. 12)

[60] B. Famaey and S. S. McGaugh. Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions. *Living Reviews in Relativity*, 15(1), September 2012. ISSN 1433-8351. doi:10.12942/lrr-2012-10. (Cited on p. 12)

[61] P. Tisserand et al. Limits on the Macho content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. *Astronomy & Astrophysics*, 469(2): 387–404, April 2007. ISSN 1432-0746. doi:10.1051/0004-6361:20066017. (Cited on p. 12)

[62] T. D. Brandt. Constraints on MACHO Dark Matter from Compact Stellar Systems in Ultra-Faint Dwarf Galaxies. *The Astrophysical Journal Letters*, 824(2):L31, June 2016. ISSN 2041-8213. doi:10.3847/2041-8205/824/2/l31. (Cited on p. 12)

[63] S. White, C. Frenk, and M. Davis. Clustering in a neutrino-dominated universe. *The Astrophysical Journal*, 274:L1–L5, October 1983. doi:10.1086/184139. (Cited on p. 12)

[64] V. Springel, C. S. Frenk, and S. D. M. White. The large-scale structure of the Universe. *Nature*, 440(7088):1137–1144, Apr 2006. ISSN 1476-4687. doi:10.1038/nature04805. (Cited on p. 12)

[65] S. Tremaine and J. E. Gunn. Dynamical Role of Light Neutral Leptons in Cosmology. *Phys. Rev. Lett.*, 42:407–410, Feb 1979. doi:10.1103/PhysRevLett.42.407. (Cited on p. 12)

[66] L. Randall, J. Scholtz, and J. Unwin. Cores in dwarf galaxies from Fermi repulsion. *Monthly Notices of the Royal Astronomical Society*, 467(2):1515–1525, 01 2017. ISSN 0035-8711. doi:10.1093/mnras/stx161. (Cited on p. 12)

[67] R. C. Cotta et al. Dark matter in the MSSM. *New Journal of Physics*, 11(10): 105026, oct 2009. doi:10.1088/1367-2630/11/10/105026. (Cited on p. 13)

[68] J. Abdallah et al. Simplified models for dark matter searches at the LHC. *Physics of the Dark Universe*, 9-10:8–23, 2015. ISSN 2212-6864. doi:10.1016/j.dark.2015.08.001. (Cited on p. 13)

[69] D. Abercrombie et al. Dark Matter benchmark models for early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum. *Physics of the Dark Universe*, 27:100371, 2020. ISSN 2212-6864. doi:10.1016/j.dark.2019.100371. (Cited on p. 13 and 15)

[70] M. Ackermann et al. Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data. *Phys. Rev. Lett.*, 115:231301, November 2015. doi:10.1103/PhysRevLett.115.231301. (Cited on p. 14)

[71] K. Choi et al. Search for Neutrinos from Annihilation of Captured Low-Mass Dark Matter Particles in the Sun by Super-Kamiokande. *Phys. Rev. Lett.*, 114:141301, April 2015. doi:10.1103/PhysRevLett.114.141301. (Cited on p. 14)

[72] IceCube Collaboration. Improved limits on dark matter annihilation in the Sun with the 79-string IceCube detector and implications for supersymmetry. *Journal of Cosmology and Astroparticle Physics*, 2016(04):022–022, April 2016. ISSN 1475-7516. doi:10.1088/1475-7516/2016/04/022. (Cited on p. 14)

[73] F. Calore et al. AMS-02 antiprotons and dark matter: Trimmed hints and robust bounds. *SciPost Physics*, 12(5), May 2022. ISSN 2542-4653. doi:10.21468/scipostphys.12.5.163. (Cited on p. 14)

[74] C. Pérez de los Heros. Status, Challenges and Directions in Indirect Dark Matter Searches. *Symmetry*, 12(10):1648, October 2020. ISSN 2073-8994. doi:10.3390/sym12101648. (Cited on p. 14)

[75] R. K. Leane. Indirect Detection of Dark Matter in the Galaxy, 2020. (Cited on p. 14)

[76] J. Aalbers et al. First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment. *Physical Review Letters*, 131(4), July 2023. ISSN 1079-7114. doi:10.1103/physrevlett.131.041002. (Cited on p. 14)

[77] E. Aprile et al. First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment. *Physical Review Letters*, 131(4), July 2023. ISSN 1079-7114. doi:10.1103/physrevlett.131.041003. (Cited on p. 14)

[78] O. Buchmueller et al. Characterising dark matter searches at colliders and direct detection experiments: vector mediators. *Journal of High Energy Physics*, 2015(1): 37, Jan 2015. ISSN 1029-8479. doi:10.1007/JHEP01(2015)037. (Cited on p. 14)

[79] ATLAS Collaboration. Dark matter summary plots for s-channel, 2HDM+a, Higgs portal and Dark Higgs models. Technical report, CERN, Geneva, 2023. http://cds.cern.ch/record/2865335. (Cited on p. 14 and 17)

[80] A. Collaboration. Search for new phenomena in dijet events using $37\,\mathrm{fb}^{-1}$ of pp collision data collected at $\sqrt{s} = 13\,\mathrm{TeV}$ with the ATLAS detector. *Physical Review D*, 96(5), September 2017. ISSN 2470-0029. doi:10.1103/physrevd.96.052004. (Cited on p. 16)

[81] A. Buckley et al. General-purpose event generators for LHC physics. *Physics Reports*, 504(5):145–233, 2011. ISSN 0370-1573. doi:10.1016/j.physrep.2011.03.005. (Cited on p. 19)

[82] J. Campbell et al. Event Generators for High-Energy Physics Experiments. Technical report, 2022. https://cds.cern.ch/record/2807002. (Cited on p. 19)

[83] S. Schumann. Lecture: Simulation Methods for HEP, 2023. (Cited on p. 19)

[84] R. D. Ball et al. Parton distributions with LHC data. *Nuclear Physics B*, 867(2): 244–289, 2013. ISSN 0550-3213. doi:10.1016/j.nuclphysb.2012.10.003. (Cited on p. 19, 23, and 74)

[85] S. Dulat et al. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev. D*, 93:033006, February 2016. doi:10.1103/PhysRevD.93.033006. (Cited on p. 19, 23, 78, and 82)

[86] T. Kinoshita. Mass singularities of Feynman amplitudes. *J. Math. Phys.*, 3:650–677, 1962. doi:10.1063/1.1724268. (Cited on p. 21)

[87] T. D. Lee and M. Nauenberg. Degenerate Systems and Mass Singularities. *Phys. Rev.*, 133:B1549–B1562, March 1964. doi:10.1103/PhysRev.133.B1549. (Cited on p. 21)

[88] S. Catani and M. Seymour. A general algorithm for calculating jet cross sections in NLO QCD. *Nuclear Physics B*, 485(1):291–419, 1997. ISSN 0550-3213. doi:10.1016/S0550-3213(96)00589-5. (Cited on p. 21)

[89] E. Bagnaschi et al. An extensive survey of the estimation of uncertainties from missing higher orders in perturbative calculations. *JHEP*, 02:133, 2015. doi:10.1007/JHEP02(2015)133. (Cited on p. 21 and 74)

[90] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506: 250–303, 2003. doi:10.1016/S0168-9002(03)01368-8. (Cited on p. 22)

[91] J. Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270, 2006. doi:10.1109/TNS.2006.869826. (Cited on p. 22)

[92] T. Sjöstrand, S. Mrenna, and P. Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, 2008. ISSN 0010-4655. doi:10.1016/j.cpc.2008.01.036. (Cited on p. 23)

[93] ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. Technical report, CERN, Geneva, 2014. https://cds.cern.ch/record/1966419. (Cited on p. 23)

[94] A. Ryd et al. EvtGen: A Monte Carlo Generator for B-Physics. May 2005. https://evtgen.hepforge.org/doc/EvtGenGuide.pdf. (Cited on p. 23)

[95] E. Bothmann et al. Event generation with Sherpa 2.2. *SciPost Phys.*, 7:034, 2019. doi:10.21468/SciPostPhys.7.3.034. (Cited on p. 23)

[96] J.-C. Winter, F. Krauss, and G. Soff. A modified cluster-hadronisation model. *The European Physical Journal C - Particles and Fields*, 36(3):381–395, August 2004. ISSN 1434-6052. doi:10.1140/epjc/s2004-01960-8. (Cited on p. 23)

[97] T. Sjöstrand. The Lund Monte Carlo for jet fragmentation. *Computer Physics Communications*, 27(3):243–284, 1982. ISSN 0010-4655. doi:10.1016/0010-4655(82)90175-8. (Cited on p. 23)

[98] Z. Nagy. NLOJet++. https://www.desy.de/~znagy/Site/NLOJet++.html, 2010. (Cited on p. 23)

[99] M. Cacciari, G. P. Salam, and G. Soyez. FastJet user manual. *The European Physical Journal C*, 72(3):1896, March 2012. ISSN 1434-6052. doi:10.1140/epjc/s10052-012-1896-2. (Cited on p. 23)

[100] T. Carli et al. A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: the APPLGRID project. *The European Physical Journal C*, 66(3):503–524, April 2010. ISSN 1434-6052. doi:10.1140/epjc/s10052-010-1255-0. (Cited on p. 24)

[101] J. Pumplin et al. Uncertainties of predictions from parton distribution functions. 2. The Hessian method. *Phys. Rev. D*, 65:014013, 2001. doi:10.1103/PhysRevD.65.014013. (Cited on p. 24 and 82)

[102] L. Evans and P. Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, 2008. doi:10.1088/1748-0221/3/08/S08001. (Cited on p. 25 and 40)

[103] ATLAS Collaboration et al. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008. doi:10.1088/1748-0221/3/08/S08003. (Cited on p. 25 and 26)

[104] LHC Run 3: physics at record energy starts tomorrow. https://atlas.cern/Updates/Press-Statement/LHC-Run3-Starts, 2022. (Cited on p. 25)

[105] T. Taylor and D. Treille. The Large Electron Positron Collider (LEP): Probing the Standard Model. *Adv. Ser. Direct. High Energy Phys.*, 27:217–261, 2017. doi:10.1142/9789814749145_0007. (Cited on p. 25)

[106] R. P. Walker. Synchrotron radiation. 1994. doi:10.5170/CERN-1994-001.437. (Cited on p. 25)

[107] J. Jowett. Colliding Heavy Ions in the LHC. 2018. doi:10.18429/JACoW-IPAC2018-TUXGBD2. (Cited on p. 25)

[108] M. Vretenar et al. *Linac4 design report*, volume 6 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2020. doi:10.23731/CYRM-2020-006. (Cited on p. 25)

[109] M. Zinser. *The Large Hadron Collider*, pages 47–51. Springer International Publishing, Cham, 2018. ISBN 978-3-030-00650-1. doi:10.1007/978-3-030-00650-1_4. (Cited on p. 26)

[110] B. Salvachua. Overview of Proton-Proton Physics during Run 2. In *9th LHC Operations Evian Workshop*, pages 7–14, 2019. (Cited on p. 25)

[111] Restarting the LHC: Why 13 TeV? June 2014. https://cds.cern.ch/record/1998739. (Cited on p. 26)

[112] CMS Collaboration et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008. doi:10.1088/1748-0221/3/08/S08004. (Cited on p. 26)

[113] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716 (1):1–29, September 2012. ISSN 0370-2693. doi:10.1016/j.physletb.2012.08.020. (Cited on p. 26)

[114] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, September 2012. ISSN 0370-2693. doi:10.1016/j.physletb.2012.08.021. (Cited on p. 26)

[115] ALICE Collaboration et al. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008. doi:10.1088/1748-0221/3/08/S08002. (Cited on p. 26)

[116] LHCb Collaboration et al. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, 2008. doi:10.1088/1748-0221/3/08/S08005. (Cited on p. 26)

[117] W. Herr and B. Muratori. Concept of luminosity. 2006. doi:10.5170/CERN-2006-002.361. (Cited on p. 27)

[118] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. *The European Physical Journal C*, 83(10): 982, October 2023. ISSN 1434-6052. doi:10.1140/epjc/s10052-023-11747-w. (Cited on p. 27, 40, and 116)

[119] J. Wenninger and M. Hostettler. LHC Report: colliding at an angle. https://home.cern/news/news/accelerators/lhc-report-colliding-angle. (Cited on p. 27)

[120] J. Boyd. LHC Run-2 and Future Prospects. 2021. doi:10.23730/CYRSP-2021-005.247. (Cited on p. 27)

[121] J. Pequenao. Computer generated image of the whole ATLAS detector. March 2008. https://cds.cern.ch/record/1095924. (Cited on p. 28)

[122] J. Pequenao. Computer generated image of the ATLAS inner detector. March 2008. https://cds.cern.ch/record/1095926. (Cited on p. 29)

[123] A. L. Rosa. The ATLAS Insertable B-Layer: from construction to operation, 2016. (Cited on p. 28)

[124] ATLAS Collaboration. The ATLAS Inner Detector commissioning and calibration. *The European Physical Journal C*, 70(3):787–821, Dec 2010. ISSN 1434-6052. doi:10.1140/epjc/s10052-010-1366-7. (Cited on p. 29)

[125] S. Boutle, others, and on behalf of the ATLAS Collaboration. Primary vertex reconstruction at the ATLAS experiment. *Journal of Physics: Conference Series*, 898(4):042056, October 2017. doi:10.1088/1742-6596/898/4/042056. (Cited on p. 29)

[126] ATLAS Collaboration. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *The European Physical Journal C*, 81(8):689, August 2021. ISSN 1434-6052. doi:10.1140/epjc/s10052-021-09402-3. (Cited on p. 30, 34, 43, 45, 52, and 54)

[127] H. Kolanoski and N. Wermes. *Particle Detectors*. Oxford University Press, 6 2020. ISBN 978-0-19-885836-2. (Cited on p. 30, 31, 32, and 33)

[128] R. Kogler and P. Schleper. Lecture: Detectors and analysis methods, 2018. https://www.physik.uni-hamburg.de/en/iexp/gruppe-schleper/lehre/detectors-methods-ss18. (Cited on p. 33)

[129] R. Wigmans. *Calorimetry: Energy Measurement in Particle Physics*. Oxford University Press, September 2017. ISBN 9780198786351. doi:10.1093/oso/9780198786351.001.0001. (Cited on p. 34)

[130] M. Aharrouche et al. Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(2):601 – 623, 2006. ISSN 0168-9002. doi:10.1016/j.nima.2006.07.053. (Cited on p. 34)

[131] B. Aubert et al. Performance of the ATLAS electromagnetic calorimeter end-cap module 0. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 500(1):178–201, 2003. ISSN 0168-9002. doi:https://doi.org/10.1016/S0168-9002(03)00344-9. (Cited on p. 34)

[132] J. Abdallah et al. Study of energy response and resolution of the ATLAS Tile Calorimeter to hadrons of energies from 16 to 30 GeV. *The European Physical Journal C*, 81(6):549, Jun 2021. ISSN 1434-6052. doi:10.1140/epjc/s10052-021-09292-5. (Cited on p. 34)

[133] B. Dowler et al. Performance of the ATLAS hadronic end-cap calorimeter in beam tests. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 482(1):94 – 124, 2002. ISSN 0168-9002. doi:10.1016/S0168-9002(01)01338-9. (Cited on p. 34)

[134] J. P. Archambault et al. Energy calibration of the ATLAS Liquid Argon Forward Calorimeter. *Journal of Instrumentation*, 3(02):P02002, 2008. doi:10.1088/1748-0221/3/02/P02002. (Cited on p. 35)

[135] B. Stelzer. The New Small Wheel Upgrade Project of the ATLAS Experiment. *Nuclear and Particle Physics Proceedings*, 273-275:1160–1165, 04 2016. doi:10.1016/j.nuclphysbps.2015.09.182. (Cited on p. 35)

[136] ATLAS Collaboration. Operation of the ATLAS trigger system in Run 2. *Journal of Instrumentation*, 15(10):P10004, October 2020. doi:10.1088/1748-0221/15/10/P10004. (Cited on p. 35, 40, and 41)

[137] ATLAS Collaboration. Performance of the ATLAS trigger system in 2015. *The European Physical Journal C*, 77(5):317, May 2017. ISSN 1434-6052. doi:10.1140/epjc/s10052-017-4852-3. (Cited on p. 36, 43, and 45)

[138] R. Achenbach et al. The ATLAS Level-1 Calorimeter Trigger. Technical report, CERN, Geneva, 2008. doi:10.1088/1748-0221/3/03/P03001. (Cited on p. 36)

[139] F. Anulli et al. The Level-1 Trigger Muon Barrel System of the ATLAS experiment at CERN. *Journal of Instrumentation*, 4(04):P04010, April 2009. doi:10.1088/1748-0221/4/04/P04010. (Cited on p. 37)

[140] ATLAS Collaboration. Performance of the ATLAS Level-1 topological trigger in Run 2. *The European Physical Journal C*, 82(1):7, January 2022. ISSN 1434-6052. doi:10.1140/epjc/s10052-021-09807-0. (Cited on p. 37)

[141] C. Schiavi. ATLAS High-Level Trigger algorithms for Run-2 data-taking. Technical report, CERN, Geneva, 2015. https://cds.cern.ch/record/2016651. (Cited on p. 38)

[142] R. Lince Amaral Farto Abreu et al. Data Scouting in ATLAS: Trigger, Tier0 and Analysis. Technical report, CERN, Geneva, 2014. https://cds.cern.ch/record/1957173. (Cited on p. 39)

[143] S. Alderweireldt et al. The Run-3 ATLAS Trigger System. Technical report, CERN, Geneva, 2022. https://cds.cern.ch/record/2845056. (Cited on p. 39, 42, and 66)

[144] ATLAS Collaboration. Trigger Operation Public Results: 2018 *pp* at 13 TeV, Trigger rates and bandwidth. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults#2018_pp_at_13_TeV, 2018. (Cited on p. 40)

[145] by Rende Steerenberg and A. Schaeffer. LHC Report: Back in production. 2018. https://cds.cern.ch/record/2633046. (Cited on p. 41)

[146] J. Damp. private communication. (Cited on p. 41 and 123)

[147] E. Ritsch. Fast Calorimeter Punch-Through Simulation for the ATLAS Experiment. Master's thesis, Innsbruck U., 2011. http://cds.cern.ch/record/1388275. (Cited on p. 41)

[148] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *The European Physical Journal C*, 77(7):466, July 2017. ISSN 1434-6052. doi:10.1140/epjc/s10052-017-5031-2. (Cited on p. 42)

[149] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *The European Physical Journal C*, 77(7):490, July 2017. ISSN 1434-6052. doi:10.1140/epjc/s10052-017-5004-5. (Cited on p. 42 and 54)

[150] M. Cacciari et al. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008. doi:10.1088/1126-6708/2008/04/063. (Cited on p. 43)

[151] G. C. Blazey et al. Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop, 2000. (Cited on p. 43)

[152] Y. Dokshitzer et al. Better jet clustering algorithms. *Journal of High Energy Physics*, 1997(08):001, September 1997. doi:10.1088/1126-6708/1997/08/001. (Cited on p. 43)

[153] R. Atkin. Review of jet reconstruction algorithms. *Journal of Physics: Conference Series*, 645(1):012008, September 2015. doi:10.1088/1742-6596/645/1/012008. (Cited on p. 43)

[154] R. Atkin. Review of jet reconstruction algorithms. *Journal of Physics: Conference Series*, 645(1):012008, September 2015. doi:10.1088/1742-6596/645/1/012008. (Cited on p. 43)

[155] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 96:072002, October 2017. doi:10.1103/PhysRevD.96.072002. (Cited on p. 45)

[156] M. Cacciari and G. P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(1):119–126, 2008. ISSN 0370-2693. doi:10.1016/j.physletb.2007.09.077. (Cited on p. 46)

[157] L. M. Baltes et al. Search for low-mass dijet resonances in run-2 data with Trigger-object-Level Analysis. Technical report, CERN, Geneva, 2018. https://cds.cern.ch/record/2632454. (Cited on p. 47, 49, 51, 52, 53, and 56)

[158] N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Volume 1*. Wiley Series in Probability and Statistics. Wiley, 1994. ISBN 9780471584957. (Cited on p. 48)

[159] ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{\mathrm{s}} = 7\,\mathrm{TeV}$. *The European Physical Journal C*, 73(3): 2304, March 2013. ISSN 1434-6052. doi:10.1140/epjc/s10052-013-2304-2. (Cited on p. 49)

[160] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006, December 2019. doi:10.1088/1748-0221/14/12/P12006. (Cited on p. 54)

[161] ATLA Collaboration. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 76(5):292, May 2016. ISSN 1434-6052. doi:10.1140/epjc/s10052-016-4120-y. (Cited on p. 54)

[162] E. S. Kuwertz et al. ATLAS Data Quality Operations and Performance during Run 2. Technical report, CERN, Geneva, 2019. https://cds.cern.ch/record/2663169. (Cited on p. 55)

[163] F. Del Rio. Calibración de jets a nivel del trigger en colisiones protón-protón a altas energías. Master's thesis, Universidad de Buenos Aires, 2020. (Cited on p. 62)

[164] ATLAS Collaboration. ATLAS data quality operations and performance for 2015–2018 data-taking. *Journal of Instrumentation*, 15(04):P04003–P04003, April 2020. doi:10.1088/1748-0221/15/04/p04003. (Cited on p. 65)

[165] ATLAS Collaboration. Beam-induced and cosmic-ray backgrounds observed in the ATLAS detector during the LHC 2012 proton-proton running period. *Journal of Instrumentation*, 11(05):P05013, May 2016. doi:10.1088/1748-0221/11/05/P05013. (Cited on p. 65)

[166] R. Abreu et al. Search for low-mass dijet resonances in 2016 data with Trigger-object-Level Analysis. Technical report, CERN, Geneva, 2016. https://cds.cern.ch/record/2226514. (Cited on p. 65, 68, 72, 78, 90, 118, and 119)

[167] H. Meyer zu Theenhausen. *Search for Light Dijet Resonances using Trigger Jets with the ATLAS Experiment at the Large Hadron Collider*. PhD thesis, Universität Heidelberg, 2018. (Cited on p. 66, 78, and 118)

[168] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. Technical report, CERN, Geneva, 2014. https://cds.cern.ch/record/1700870. (Cited on p. 66)

[169] T. Skwarnicki. *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986. (Cited on p. 71)

[170] N. Nishu et al. Search for New Phenomena in Dijet Events using the complete Run-2 dataset collected with the ATLAS Detector at $\sqrt{s}$ =13 TeV. Technical report, CERN, Geneva, 2018. https://cds.cern.ch/record/2646455. (Cited on p. 71)

[171] J. Butterworth et al. PDF4LHC recommendations for LHC Run II. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, January 2016. doi:10.1088/0954-3899/43/2/023001. (Cited on p. 74)

[172] J. M. Campbell and R. K. Ellis. Update on vector boson pair production at hadron colliders. *Physical Review D*, 60(11), November 1999. ISSN 1089-4918. doi:10.1103/physrevd.60.113006. (Cited on p. 74)

[173] J. M. Campbell, R. K. Ellis, and C. Williams. Vector boson pair production at the LHC. *Journal of High Energy Physics*, 2011(7), July 2011. ISSN 1029-8479. doi:10.1007/jhep07(2011)018. (Cited on p. 74)

[174] P. J. Fox and C. Williams. Next-to-leading order predictions for dark matter production at hadron colliders. *Physical Review D*, 87(5), March 2013. ISSN 1550-2368. doi:10.1103/physrevd.87.054030. (Cited on p. 74)

[175] J. R. Klein and A. Roodman. Blind analysis in nuclear and particle physics. *Annual Review of Nuclear and Particle Science*, 55(1):141–163, 2005. doi:10.1146/annurev.nucl.55.090704.151521. (Cited on p. 77)

[176] S. Alekhin, J. Blümlein, and S. Moch. Parton distribution functions and benchmark cross sections at next-to-next-to-leading order. *Phys. Rev. D*, 86:054009, September 2012. doi:10.1103/PhysRevD.86.054009. (Cited on p. 78)

[177] L. A. Harland-Lang et al. Parton distributions in the LHC era: MMHT 2014 PDFs. *The European Physical Journal C*, 75(5):204, May 2015. ISSN 1434-6052. doi:10.1140/epjc/s10052-015-3397-6. (Cited on p. 78 and 83)

[178] T. Aaltonen et al. Search for new particles decaying into dijets in proton-antiproton collisions at $\sqrt{s} = 1.96\,\text{TeV}$. *Phys. Rev. D*, 79:112002, June 2009. doi:10.1103/PhysRevD.79.112002. (Cited on p. 78)

[179] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling, 2003. (Cited on p. 78)

[180] R. Brun, F. Rademakers, et al. ROOT - An Object Oriented Data Analysis Framework, November 2023. https://root.cern/releases/release-63002. (Cited on p. 78)

[181] S. Alekhin et al. The new ABMP16 PDF. *PoS*, DIS2016:016, 2016. doi:10.22323/1.265.0016. (Cited on p. 83)

[182] R. D. Ball et al. Parton distributions for the LHC run II. *Journal of High Energy Physics*, 2015(4):40, April 2015. ISSN 1029-8479. doi:10.1007/JHEP04(2015)040. (Cited on p. 83)

[183] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard interactions of quarks and gluons: a primer for LHC physics. *Reports on Progress in Physics*, 70(1): 89–193, December 2006. ISSN 1361-6633. doi:10.1088/0034-4885/70/1/r02. (Cited on p. 83)

[184] K. Cranmer et al. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical report, New York U., New York, January 2012. https://cds.cern.ch/record/1456844. (Cited on p. 84)

[185] ATLAS Collaboration. Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physics Letters B*, 796:68–87, 2019. ISSN 0370-2693. doi:10.1016/j.physletb.2019.07.016. (Cited on p. 89)

[186] ATLAS Collaboration. Search for resonances decaying into photon pairs in 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physics Letters B*, 822:136651, 2021. ISSN 0370-2693. doi:10.1016/j.physletb.2021.136651. (Cited on p. 89)

[187] G. Choudalakis. On hypothesis testing, trials factor, hypertests and the BumpHunter. 2011. doi:10.48550/arXiv.1101.0390. (Cited on p. 90, 105, and 116)

[188] L. Vaslin et al. pyBumpHunter: A model independent bump hunting tool in Python for High Energy Physics analyses. *SciPost Phys. Codeb.*, 2023:15, 2023. doi:10.21468/scipostphyscodeb.15. (Cited on p. 90 and 105)

[189] ATLAS Collaboration. Recommendations for the Modeling of Smooth Backgrounds. ATL-PHYS-PUB-2020-028, 2020. https://cds.cern.ch/record/2743717. (Cited on p. 93, 94, and 103)

[190] P. R. Bevington, D. K. Robinson, and G. Bunce. Data Reduction and Error Analysis for the Physical Sciences, 2nd ed. *American Journal of Physics*, 61(8): 766–767, August 1993. ISSN 0002-9505. doi:10.1119/1.17439. (Cited on p. 103)

[191] L. Lyons. Open statistical issues in Particle Physics. *The Annals of Applied Statistics*, 2(3):887 – 915, 2008. doi:10.1214/08-AOAS163. (Cited on p. 106)

[192] A. L. Read. Modified frequentist analysis of search results (the CL_s method). 2000. doi:10.5170/CERN-2000-005.81. (Cited on p. 112)

[193] ATLAS Collaboration, CMS Collaboration, and LHC Higgs Combination Group. Procedure for the LHC Higgs boson search combination in summer 2011. 2011. https://cds.cern.ch/record/1379837. (Cited on p. 112)

[194] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:289–337, 1933. doi:10.1098/rsta.1933.0009. (Cited on p. 112)

[195] G. Cowan et al. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1554, February 2011. ISSN 1434-6052. doi:10.1140/epjc/s10052-011-1554-0. (Cited on p. 113)

[196] L. Lista. Practical Statistics for Particle Physicists. *CERN Yellow Reports: School Proceedings*, 2017. doi:10.23730/CYRSP-2017-005.213. (Cited on p. 114)

[197] N. Saoulidou et al. CMS-ATLAS Dijet Resonance Search Statistics. October 2020. https://indico.cern.ch/event/957652. (Cited on p. 114)

[198] C. Doglioni and M. Toscani. Internal Discussion with the ATLAS Statistics Committee. October 2020. https://indico.cern.ch/event/960806/#14-joint-discussion-with-stati. (Cited on p. 114 and 119)

[199] E. Gross and O. Vitells. Trial factors for the look elsewhere effect in high energy physics. *The European Physical Journal C*, 70(1–2):525–530, October 2010. ISSN 1434-6052. doi:10.1140/epjc/s10052-010-1470-8. (Cited on p. 116)

Acknowledgements

This thesis would not have been possible without the support of a number of great people during my PhD.

First and foremost, I want to thank Prof. Dr. Monica Dunford for giving me the opportunity to work on various interesting research topics, for her supervision since my time as a Master's student, and for changing my mind about leaving for other fields of physics twice. Despite her ever-increasing responsibilities in ATLAS, she always found time for regular chats with me to discuss physics and, most importantly, share fun anecdotes. Despite physics, she taught me how to manage an analysis team, when to be pragmatic, and – with ample support from two experts – how to play several board games.

I also want to express my gratitude to Prof. Dr. Hans-Christian Schulz-Coulon for sparking my interest in particle physics in the first place and for welcoming me to the great ATLAS group at KIP, where I have stayed ever since my Bachelor's. He always created a positive working environment for the group and ensured that we not only worked together as colleagues but also enjoyed doing so. The yearly conference travels have always been a highlight!

Special thanks also to Prof. Dr. Stephanie Hansmann-Menzemer for being my second referee for this thesis. She supported my studies with various opportunities provided in our Research Training Group, like the physics teams, my own and other people's student lectures, and the regular Pizza seminars.

I would like to thank Pavel, Rainer, and Martin for always providing expert knowledge on any physics questions I could think of and on the nitty-gritty details of the ATLAS detector and its trigger system. In general, chatting with you is always enriching and fun.

Quite a few adventures were to be had with my fellow farm admins Sebastian, Martin, and Thomas. With you, I spent long but (usually) fun hours shaping our beloved farm into what it is today. We managed to realize many cool (or not so cool during long outages of the air conditioning) projects which always gave me something fun to do as a change from the constant analysis work.

I want to express further thanks to my colleagues and friends for diligent proofreading: Pavel, Rainer, Sebastian, Martin, Anke, Lisa, Varsiha, Mathias, Hannah, Thomas, and Krissy. You managed to spot the tiniest errors (I hope) and, rightfully so, forced me to rewrite incomprehensible sections.

Special thanks go to Martin and Iris, who helped me to enjoy life outside of physics with countless board game nights – or rather days during the pandemic when I spent most of my weekends at your place. I fondly look back on our climbing sessions, and there is no better motivation than seeing one of you send a route that I could not.

Finally, I want to express my gratitude to all my friends and family for showing me unwavering support during the stressful final phases of the PhD and throughout my studies in general. I would not be in this position without all of you.