CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch 2021/06/05

Search for new particles in events with energetic jets and large missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search is presented for new particles produced in proton-proton collisions at $\sqrt{s} = 13$ TeV at the LHC, using events with energetic jets and large missing transverse momentum. The analysis is based on a data sample corresponding to an integrated luminosity of 101 fb^{-1}, collected in 2017−2018 with the CMS detector. Separate categories are defined for events with narrow jets from initial-state radiation and with large-radius jets consistent with a hadronic decay of a W or a Z boson. Novel machine learning techniques are used to identify hadronic W and Z boson decays. The analysis is combined with an earlier search based on a data sample corresponding to an integrated luminosity of 36 fb^{-1}, collected in 2016. No significant excess of events is observed with respect to the standard model background expectation, as determined from control samples in data. The results are interpreted in terms of limits on the branching fraction of an invisible decay of the Higgs boson, as well as constraints on simplified models of dark matter, on first-generation scalar leptoquarks decaying to quarks and neutrinos, and on gravitons in models with large extra dimensions. Several of the new limits are the most restrictive to date.

1 Introduction

The standard model (SM) of particle physics has been widely recognized as a very successful, yet incomplete theory. Many well-established features of the universe, such as gravity and the existence of dark matter (DM), are not described in the SM. It is therefore paramount to search for evidence of physics beyond the SM (BSM). Attempts at finding BSM physics often center around the production of new, hypothetical particles, which subsequently decay to the observable SM particles. In this search, we aim at the scenarios that are hidden from such searches, because the BSM particles do not necessarily decay into detectable species.

Scenarios with new particles that are not directly observable in collider detectors are motivated by many BSM theories. One of the strongest motivations stems from the idea of particle DM. Over the last decades, cosmological evidence for the existence of DM has been steadily accumulating [1], yet with few hints as to its nature or detailed properties. One theoretically attractive model of DM is that of a thermally produced weakly interacting massive particle (WIMP). If such particles have just the right masses and couplings, the abundance of DM in the universe, as well as many of the observed phenomena commonly ascribed to DM can be explained. In this search, multiple scenarios of DM production are considered. A Higgs portal scenario [2–4] is tested, in which DM particles are produced in decays of a recently discovered Higgs boson with a mass of 125 GeV [5–7]. Although the properties of the new boson have been already measured to an impressive precision and found to agree with those of the SM Higgs boson, its decay branching fractions are still not sufficiently constrained and allow a decay branching fraction to invisible particles of up to about 20% [8, 9]. Beyond the Higgs portal scenario, simplified models of DM production [10] via new bosonic mediators with spin zero or one are explored. Colorlessl mediators coupled to pairs of quarks and to a pair of DM particles, are considered, as well as colored mediators, which decay into a single quark together with a single DM candidate. The latter scenario is referred to as a "fermion portal" [11]. In addition to a search for DM, a scenario with large extra dimensions proposed by Arkani-Hamed, Dimopoulos, and Dvali (ADD) [12, 13] is tested. In this model, the existence of additional spatial dimensions beyond the known three could explain the large difference in strength between the gravitational and electroweak (EW) interactions. In this scenario, gravitons can be produced in proton-proton (pp) collisions via their enhanced couplings to quarks or gluons and avoid detection by escaping in the additional dimensions.

Since in these models the final-state particles are not detectable, one needs a visible detector signature to be able to identify and record such events. We use energetic hadronic jets accompanying the invisible particles to trigger on signal candidates. The experimental signature therefore comprises one or more energetic jets and a large missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$). While $p_{\mathrm{T}}^{\mathrm{miss}}$ is the intrinsic result of BSM or SM particles escaping a detector without leaving any trace, hadronic jets are produced either as a result of initial-state gluon radiation, or are due to the hadronic decays of energetic heavy SM vector bosons (V) produced in association with BSM particles. For energetic V bosons, the hadronic decay products are Lorentz boosted in the lab frame and are reconstructed as a single, large-radius jet with a characteristic substructure. Novel machine learning algorithms based on artificial neutral networks are used in order to identify such signatures and efficiently suppress the overwhelming background coming from quantum chromodynamics (QCD) production of parton jets. Separate signal categories are defined for events with and without an identified V candidate. Several control samples in data are used to constrain background contributions to the signal regions. The search is carried out with the CMS detector at the CERN LHC, in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$, using a data set collected in 2017–2018, corresponding to an integrated luminosity of $101\,\mathrm{fb}^{-1}$.

The chosen experimental signature allows us to also probe for other BSM scenarios, with new particles decaying semivisibly. One such scenario probed by the present search is production of leptoquarks (LQs). The LQs are hypothetical scalar or vector particles that carry both baryon and lepton numbers [14–16]. Here, a scenario with a single scalar LQ type is considered. This first-generation LQ decays into a u quark and an electron neutrino (ν_e), and can be either produced in pairs [17], via the strong interaction through their couplings to gluons, or singly [18, 19], in association with a ν_e, through its coupling to the u quark and ν_e. Both processes result in a jets $+p_T^{miss}$ signature.

Searches for new phenomena in events with jets (including those coming from energetic V decays) and p_T^{miss} at $\sqrt{s} = 13\,\text{TeV}$ have been previously published by the CMS [20] and ATLAS [21, 22] Collaborations, based on data recorded in 2016 (and also 2015 in the case of ATLAS), corresponding to an integrated luminosity of approximately $36\,\text{fb}^{-1}$. Compared to these results, we tripled the amount of analyzed data and enhanced the analysis sensitivity by the means of improved identification of hadronically decaying V bosons. We further extend the sensitivity by combining the new results with those from Ref. [20].

This note is organized as follows. After discussing the CMS detector in Section 2 and simulated samples in Section 3, we describe the event selection in Section 4, followed by the background estimation in Section 5. Section 6 contains the results of the analysis and their interpretation in the context of the above scenarios. We summarize the content of this note in Section 7.

2 The CMS detector and event reconstruction

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid.

The silicon tracker measures charged particles within the pseudorapidity range $|\eta| < 2.5$. During the LHC running period when the data used in this note were recorded, the silicon tracker consisted of 1856 silicon pixel and 15 148 silicon strip detector modules.

In the region $|\eta| < 1.74$, the HCAL cells have widths of 0.087 in pseudorapidity and 0.087 in azimuth (ϕ). In the η-ϕ plane, and for $|\eta| < 1.48$, the HCAL cells map on to 5×5 arrays of ECAL crystals to form calorimeter towers projecting radially outwards from close to the nominal interaction point. For $|\eta| > 1.74$, the coverage of the towers increases progressively to a maximum of 0.174 in $\Delta\eta$ and $\Delta\phi$. The hadron forward (HF) calorimeter uses steel as an absorber and quartz fibers as the sensitive material. The two halves of the HF are located 11.2 m from the interaction region, one on each end, and together they provide coverage in the range $3.0 < |\eta| < 5.2$.

Events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of about 4 μs [23]. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage [24].

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [25].

The candidate vertex with the largest value of summed physics-object p_{T}^2 is taken to be the primary vertex (PV) of the pp interaction. The physics objects are the jets, clustered using the jet finding algorithm [26, 27] with the tracks assigned to candidate vertices as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the p_{T} of those jets.

The particle-flow (PF) algorithm [28] aims to reconstruct and identify each individual particle in an event, with an optimized combination of information from the various elements of the CMS detector. In this process, the identification of the PF candidate type (photon, electron, muon, and charged and neutral hadrons) plays an important role in the determination of the particle direction and energy. The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from a combination of the electron momentum at the PV as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

For each event, hadronic jets are clustered from the PF candidates using the infrared- and collinear-safe anti-k_{T} algorithm [26, 27] with a distance parameter of 0.4 or 0.8. Depending on the respective distance parameter, these jets are referred to as "AK4" or "AK8" jets. Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the entire p_{T} spectrum and detector acceptance. Additional pp interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, charged particles identified as not originating from the PV are discarded and an offset correction is applied to correct for the remaining neutral pileup contributions. Jet energy corrections are derived from simulation to bring measured response of jets to that of particle-level jets on average. In situ measurements of the momentum balance in the dijet, γ + jet, Z + jet, and multijet events are used to account for any residual differences in the jet energy scale (JES) and jet energy resolution (JER) in data and simulation [29]. The JER amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV. Additional selection criteria [30] are applied to each jet to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures.

The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [31]. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is modified to account for corrections to the energy scale and resolution of the reconstructed jets in the event. Anomalous high-$p_{\mathrm{T}}^{\mathrm{miss}}$ events can be due to a variety of reconstruction failures, detector malfunctions, or noncollision backgrounds. Such events are rejected by dedicated filters that are designed to eliminate more than 85–90% of the spurious high-$p_{\mathrm{T}}^{\mathrm{miss}}$ events with a signal efficiency exceeding 99.9% [31].

Large-radius AK8 jets are used for the identification of hadronic decays of W and Z bosons. The pileup-per-particle identification (PUPPI) algorithm [32] is used to mitigate the effect of pileup at the reconstructed-particle level, making use of local shape information, event pileup properties, and tracking information. Charged particles identified as not originating from the

PV are discarded. For each neutral particle, a local shape variable is computed using the surrounding charged particles within the tracker acceptance ($|\eta| < 2.5$) compatible with the PV, and using both charged and neutral particles in the region outside of the tracker coverage. The momenta of the neutral particles are then rescaled according to their probability to originate from the PV deduced from the local shape variable, superseding the need for jet-based pileup corrections [30]. The modified mass drop tagger algorithm [33, 34], also known as the soft-drop (SD) algorithm, with the angular exponent $\beta = 0$, soft cutoff threshold $z_{\text{cut}} < 0.1$, and characteristic radius $R_0 = 0.8$ [35], is applied to remove soft, wide-angle radiation from the jet.

3 Simulated samples

Monte Carlo (MC) simulated event samples are used to model signal and background contributions to all the analysis regions. In all cases, parton showering, hadronization, and underlying event properties are modeled using PYTHIA [36] version 8.202 or later with the underlying event tune CP5 [37]. Simulation of interactions between particles and the CMS detector is based on GEANT4 [38] . The same reconstruction algorithms used for data are applied to simulated samples as well. The NNPDF3.1 next-to-next-to-leading order (NNLO) set of parton distribution functions (PDFs) [39] is used for generation of all samples.

For the V + jets processes, predictions with up to two partons in the final state are obtained at next-to-leading order (NLO) in QCD using MADGRAPH5_aMC@NLO version 2.4.2 [40] with the FxFx matching scheme [41] between the jets from the matrix element calculations and the parton shower. The γ + jets samples are simulated at NLO in QCD with up to one additional parton using MADGRAPH5_aMC@NLO version 2.6.5. This version is also used for all other MADGRAPH5_aMC@NLO samples, unless indicated otherwise. Samples of events with top quark pairs are generated at NLO in QCD with up to two additional partons in the matrix element calculations using MADGRAPH5_aMC@NLO and the FxFx jet matching scheme. Their cross sections are normalized to the inclusive cross section of the top quark pair production at NNLO in QCD [42]. Events with single top quarks are simulated using POWHEG 2.0 [43, 44] and normalized to the inclusive cross section calculated at NNLO in QCD [45] for single top quarks produced in association with a W boson, and NLO in QCD [46, 47] for production in association with a quark. Production of diboson events (WW, WZ, and ZZ) is simulated at leading order (LO) in QCD using PYTHIA, and normalized to the cross sections at NNLO accuracy for WW production [48] and at NLO accuracy for the others [49]. The production of Wγ and Zγ events is simulated using MADGRAPH5_aMC@NLO at NLO in QCD. Samples of QCD multijet events are generated at LO using PYTHIA.

For the Higgs portal signal model, POWHEG is used to generate separate signal samples for the different production modes of the Higgs boson: via gluon fusion [50], in association with a SM vector boson (VH) [51], and via vector boson fusion (VBF) [52]. The samples are generated by enforcing decays of the SM Higgs boson to neutrinos, and are normalized to the SM cross sections evaluated at next-to-NNLO in QCD and NLO in EW corrections for the gluon fusion production, and at NNLO in QCD and NLO in EW for the VBF and VH modes [53]. Events for the simplified model scenarios of DM production are generated using MADGRAPH5_aMC@NLO and the DMSIMP model implementation [54–56]. For the case of spin-1 mediators, events with a pair of DM particles and either one or two additional partons are generated at NLO in QCD, and FxFx jet merging is used. The couplings between the mediator and quarks, as well as between the mediator and the DM particles are set to $g_{\text{q}} = 0.25$ and $g_{\chi} = 1.0$, respectively, as recommended by the LHC Dark Matter Working Group [57]. For DM production via a spin-0 mediator, signal samples are generated at the effective LO, with one additional parton in

the matrix element calculations, and the respective couplings are set to $g_q = g_\chi = 1.0$ [57]. Separate samples are generated for different coupling types (vector, axial vector, scalar, and pseudoscalar), as well as for different mass hypotheses for the mediator and DM particles. Signal events for the fermion portal scenario are generated using MADGRAPH5_aMC@NLO and the implementation of Ref. [58]. The mediator in this case is assumed to couple to right-handed up quarks and a Dirac fermion DM candidate with a coupling of $\lambda = 1$. The single- and pair-production of scalar LQs is simulated at LO in QCD using MADGRAPH5_aMC@NLO version 2.6.0 with an implementation provided by the authors of Ref. [17]. Decays of each LQ to an up quark and an electron neutrino are enforced, and separate samples are generated for the LQ mass values between 0.5 and 2.5 TeV, as well as for the LQ-u-ν_e coupling values ranging from 0.01 to 1.5, depending on the LQ mass. Finally, events with graviton production in the ADD scenario are generated at LO using PYTHIA [59]. In this case, samples of signal events are generated for the number of extra dimensions d between 2 and 7, and the values of the fundamental Planck scale M_D between 5 and 15 TeV.

4 Event selection

The key feature of the analysis strategy is the extensive use of control data samples for the purpose of precise prediction of the background contributions in the signal regions (SRs). The leading SM background contributions originate from $Z \to \nu\overline{\nu}$ and $W \to \ell\overline{\nu}$ production, the properties of which are constrained using control regions (CRs) with charged leptons that are enriched in $Z \to \ell\ell$ and $W \to \ell\overline{\nu}$ events, respectively. Additionally, a CR enriched in γ + jets events is defined. The V + jets events in these CRs share many kinematic properties of the processes in the SRs and are used to constrain the latter. The CR and SR definitions share as many of selection criteria as possible, in order to ensure that minimal selection biases are introduced. For each SR, five CRs are defined: single-electron and single-muon CRs enriched in $W \to \ell\overline{\nu}$ events, dielectron and dimuon CRs enriched in $Z \to \ell\ell$ events, and the fifth CR enriched in γ + jets events.

The SR events are selected using a trigger with a p_T^{miss} requirement of at least 120 GeV. The trigger requirement for the SR is based on an online calculation of p_T^{miss} based on all PF candidates reconstructed at the HLT, except for muons, and is therefore also used to collect data populating the single-muon and dimuon CRs. The control samples with electrons are selected based on single-electron triggers with the thresholds of $p_T > 35$ GeV ($p_T > 32$ GeV) and $p_T > 115$ GeV, and with a single-photon trigger with a requirement of $p_T > 200$ GeV for 2017 (2018). The single-electron triggers differ in their usage of isolation requirements: while the lowest threshold trigger requires electrons to be well isolated, the higher-threshold trigger does not, which allows for an improved efficiency at high p_T. Similarly, the single-photon trigger avoids the reliance on the online track reconstruction and increases the overall efficiency at $p_T > 200$ GeV. The photon trigger is also used to select events for the photon control sample. During the 2017 data taking, a gradual shift in the timing of the inputs of the ECAL L1 trigger in the region at $|\eta| > 2.0$ caused a specific trigger inefficiency. For events containing an electron or a photon (a jet) with p_T larger than ≈50 (≈100) GeV in this region, the efficiency loss is up to ≈10–20%, depending on p_T, η, and time. Correction factors were computed from data and applied to the acceptance evaluated by simulation for the 2017 samples.

At the analysis level, a requirement of $p_T^{miss} > 250$ GeV is applied to the SR events in order to ensure the p_T^{miss} trigger efficiency of at least 95%. Events are separated into three mutually exclusive categories based on the properties of the highest p_T ("leading") jets in the event. For the mono-V categories, the leading AK8 jet is required to have $p_T > 250$ GeV and $|\eta| < 2.4$. In order

to preferentially select events where an AK8 jet originates from a hadronic decay of a W or Z boson, the jet is further required to be V tagged with the DEEPAK8 algorithm [60] and to have an SD-corrected mass of $65 < m_{\mathrm{SD}} < 120\,\mathrm{GeV}$. The mono-V category is further subdivided into a low- and high-purity regions based on the DEEPAK8 score of the jet. Events that do not pass the mono-V selection are considered for the monojet category. In this case, the leading AK4 jet in the event is required to have $p_{\mathrm{T}} > 100\,\mathrm{GeV}$, $|\eta| < 2.4$, and to pass quality criteria based on the composition of the jet in terms of different types of PF candidates, such as a minimum charged hadron energy fraction of 10% and a maximum neutral hadron energy fraction of 80% [30]. In all categories, further requirements are imposed in order to suppress reducible background processes. Events are rejected if they contain a well-reconstructed and isolated electron (photon) with $p_{\mathrm{T}} > 10$ (15) GeV and $|\eta| < 2.5$, or a muon with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.4$; or a hadronically decaying τ lepton with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$ and $|\eta| < 2.3$ [61–63]. These requirements efficiently reject events with leptonic decays of the Z and W bosons or top quarks, as well as backgrounds with photons. Contributions from top quark processes are further suppressed by rejecting events with AK4 jets that have $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, $|\eta| < 2.4$, and are identified to have originated from the hadronization of a b quark ("b-tagged jets") using the DEEPCSV algorithm [64] with a "medium" working point. Finally, topological requirements are applied in order to reject contributions from QCD multijet events. These events do not have $p_{\mathrm{T}}^{\mathrm{miss}}$ from genuine sources and require a $p_{\mathrm{T}}^{\mathrm{miss}}$ mismeasurement in order to pass the SR selection, which can happen in two main ways. In the first case, the energy of a jet in the event could be misreconstructed either as a result of an interaction between the jet with poorly instrumented or inactive parts of the detector, or because of failures in the readout of otherwise functioning detector modules. In these cases, artificial $p_{\mathrm{T}}^{\mathrm{miss}}$ is generated with a characteristically small azimuthal angle difference between the misreconstructed jet $\vec{p}_{\mathrm{T}}$ and the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ vector. Such events are rejected by requiring $\Delta\phi(\vec{p}_{\mathrm{T}}^{\mathrm{jet}}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5\,\mathrm{rad}$. In the second case, large $p_{\mathrm{T}}^{\mathrm{miss}}$ is generated due to failures of PF reconstruction, which are suppressed by considering an alternative calculation of $p_{\mathrm{T}}^{\mathrm{miss}}$ based on calorimeter energy clusters and muon candidates, rather than the full set of all PF candidates. While the calorimeter-based $p_{\mathrm{T}}^{\mathrm{miss}}$ has significantly worse resolution than PF $p_{\mathrm{T}}^{\mathrm{miss}}$, it is much simpler and more robust. To reduce the multijet background with PF reconstruction failures, events are required to have $\Delta p_{\mathrm{T}}^{\mathrm{miss}}(\text{PF–calorimeter}) = |p_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{PF}) / p_{\mathrm{T}}^{\mathrm{miss}}(\text{calorimeter}) - 1| < 0.5$. A similar criterion is constructed using an alternative $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation based exclusively on charged-particle candidates. Since charged particles are only reconstructed within the coverage of the pixel tracking detector, this $p_{\mathrm{T}}^{\mathrm{miss}}$ variant is robust against noise and PU contributions in the forward calorimeters. Events in the SR are required to have a maximum angular separation in the transverse plane between the regular and charged-candidate $p_{\mathrm{T}}^{\mathrm{miss}}$ vectors of $\Delta\phi(\mathrm{PF}, \mathrm{charged}) < 2\,\mathrm{rad}$. Finally, a section of the hadronic calorimeter was not functioning during a part of the 2018 data taking period, leading to irrecoverable mismeasurement in a localized region of the detector ($-1.57 < \phi < -0.87$, $-3.0 < \eta < -1.3$). To avoid contamination from such mismeasurement, events where any jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ is found in the corresponding $\eta - \phi$ region are rejected in the analysis of the 2018 data set. Events where the mismeasurement is so severe that a jet is fully lost in this region are found to contribute at low values of $p_{\mathrm{T}}^{\mathrm{miss}} < 470\,\mathrm{GeV}$ and to have a characteristic signature in $\phi(\vec{p}_{\mathrm{T}}^{\mathrm{miss}})$. Such events are rejected by requiring that $\phi(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}) \notin [-1.62, -0.62]$ if $p_{\mathrm{T}}^{\mathrm{miss}} < 470\,\mathrm{GeV}$.

In each of the CRs, the same selection criteria are applied as for the corresponding SR (high- or low-purity mono-V or monojet), with two exceptions: the charged-lepton and photon rejection criteria are inverted to allow the exact number of desired leptons or photons for each CR, and the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ vector used in the SR definition is replaced by the hadronic recoil vector $\vec{U}$. The hadronic recoil is defined as the vectorial sum of the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ vector and the transverse momentum vectors of the selected charged lepton(s) or the photon in each event. The

hadronic recoil therefore acts as a proxy of the momentum of the V boson or a photon in each CR, convoluted with the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution, which is equivalent to the role of $p_{\mathrm{T}}^{\mathrm{miss}}$ in the SRs. In order to enhance the purity of the CRs, specific additional selection criteria are applied. For the charged-lepton CRs, at least one of the leptons is required to pass a more strict set of quality criteria and have $p_{\mathrm{T}} > 40$ (20) GeV electrons (muons), while the photon in the photon CR is required to have $p_{\mathrm{T}} > 230\,\mathrm{GeV}$ in order to ensure high trigger efficiency. Additionally, events in the single-lepton CRs are required to have a transverse mass $m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\mathrm{miss}} p_{\mathrm{T}}^{\ell} \left(1 - \cos\left(\Delta\phi(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\ell})\right)\right)} < 160\,\mathrm{GeV}$, and events in the single-electron CR are required to have $p_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ in order to reject contributions from QCD multijet events. Finally, in order to enrich the dilepton CRs with Z events, the two leptons are required to have opposite signs and to form a dilepton candidate with an invariant mass in the range $60 < m_{\ell\ell} < 120\,\mathrm{GeV}$, consistent with the mass of the Z boson [65].

5 Background estimation

Background estimation and signal extraction are performed simultaneously, using a joint maximum-likelihood (ML) fit in each SR and all the corresponding CRs. A likelihood function is constructed to model the expected background contributions in each recoil variable bin of the SR and CRs, as well as the expected signal yield in each bin of the SR. The best fit background model, as well as the best fit signal strength, are obtained by maximizing the likelihood.

5.1 Likelihood function

The likelihood function is defined in the same way as described in Ref. [66] and previously used in Ref. [20]. Separate approaches are adopted to estimate the dominant (Z + jets, W + jets, γ + jets) and subdominant ($\mathrm{t\bar{t}}$, diboson, and QCD multijet) backgrounds.

The predictions for the dominant backgrounds are based on the yield of $\mathrm{Z} \to \nu\overline{\nu}$ events in each bin of the SR. The per-bin yields for this process are defined as free parameters of the likelihood function. The yields for the W + jets contribution to the SR, as well as the yields of the γ + jets process in the photon CR and the $\mathrm{Z} \to \ell\ell$ process in the dilepton CRs, are defined relative to the $\mathrm{Z} \to \nu\overline{\nu}$ yields by introducing a set of per-bin transfer factors. The yields of $\mathrm{W} \to \ell\overline{\nu}$ events in the single-lepton CRs are similarly related via a transfer factor to the $\mathrm{W} \to \ell\overline{\nu}$ event yield in the SRs. In all cases, the central values of the transfer factors are obtained from the ratios of the simulated recoil spectra of the respective processes in the SRs and CRs. For the minor backgrounds, such as $\mathrm{t\bar{t}}$ and QCD multijet production, the nominal expected yield per region is obtained directly from simulation (top quark and diboson backgrounds, as well as QCD multijet production in the single-lepton CRs) or by dedicated estimates based on control samples in data (QCD multijet production in the SRs and photon CRs).

Systematic uncertainties are incorporated in the likelihood function as nuisance parameters, as described in more detail below. In the case of the V + jets processes, the nuisance parameters affect the values of the transfer factors in each recoil variable bin and thus control the ratios of the contributions from different processes, as well as the ratios of the yields in the SRs and various CRs. For the subdominant background processes, the yields in each bin are directly parameterized in terms of the nuisance parameters. The final free parameter of the likelihood function is the signal strength modifier μ, which—for a given signal hypothesis—controls the signal normalization relative to the theoretical cross section.

The likelihood method relies on the accurate predictions of the ratios between the dominant backgrounds in the SRs and CRs, as well as on the absolute normalization and shape of the

recoil distributions for the subdominant backgrounds. To achieve the most accurate possible predictions for these quantities, weights are applied to each simulated event to take into account both experimental and theoretical effects not present in the MC simulated samples. The experimental corrections are related to the trigger efficiencies, identification and reconstruction efficiencies of charged leptons, photons and b-tagged jets, and the pileup distribution in simulation. Theoretical corrections are applied to the V + jets processes in order to model the effects of NLO terms in the perturbative EW corrections [67]. The corrections are parameterized as functions of the generator-level boson p_{T} and are evaluated separately for the $\mathrm{W}(\ell\overline{\nu})$ + jets, Z + jets, and γ + jets processes. For the diboson processes (WW, WZ,and ZZ), EW and QCD NLO corrections are applied differentially in the boson p_{T}. The EW corrections are obtained from Ref. [68], while the QCD corrections are derived from simulated samples generated with MADGRAPH5_aMC@NLO and POWHEG. The EW NLO corrections for the $\mathrm{W}\gamma$ and $\mathrm{Z}\gamma$ processes are similarly obtained from Refs. [69, 70].

The validity of the predictions is checked by considering the differential ratio of yields in the control regions. The yield ratio serves as a proxy for the ratios between the different V + jets processes, which the fit relies on. The yield ratios between the dilepton and single-lepton CRs, and dilepton and photon CRs are shown in Figs. 1 and 2, respectively. Good agreement is observed between prediction and data. In the monojet categories, it is found that the rate of $\mathrm{W} \to \ell\overline{\nu}$ events is initially underpredicted relative to $\mathrm{Z} \to \ell\ell$ and γ events. This underprediction is corrected in the ML fit.

5.2 Estimation of the QCD multijet background

The contributions from QCD multijet events in each SR and the corresponding photon CR are estimated from data. Multijet events do not carry large intrinsic $p_{\mathrm{T}}^{\mathrm{miss}}$, and therefore could only contribute to the SR if one of the hadronic jets in an event is significantly misreconstructed or partially lost, leading to the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ vector and the transverse momentum vector of the jet to be aligned. The contribution from such events is estimated from a CR that is enriched in multijet events by inverting the requirement on $\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{j}})$ relative to the SR. The recoil spectrum of multijet events in the SR is obtained by multiplying the spectrum in data in this CR by a transfer factor obtained from simulation. The nonmultijet background components, as predicted from simulation, are subtracted from data before applying the transfer factor. The performance of the method is tested by splitting the low-$\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{j}})$ CR into parts across different borders in $\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{j}})$, and verifying that an estimate based on the low-$\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{j}})$ part of the region can correctly predict the QCD multijet background contribution in the high-$\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{j}})$ part. The method is found to correctly predict the QCD background contribution to approximately 25% for various choices of $\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{j}})$ thresholds, which is assigned as a normalization uncertainty to the QCD multijet background estimate in the SR. Uncertainties related to the finite size of multijet samples, as well as the choice of the transfer factor binning are taken into account.

In the photon CR, multijet events can contribute if additionally a jet is misreconstructed as an isolated photon. The fraction of photons resulting from jet misreconstruction is estimated by from the distribution of the lateral shower width of the photons. The distribution of this variable shows a characteristic peak for genuine photons, while being significantly more flat for the contribution from jets misreconstructed as photons. A template fit is performed to the distribution in data in order to extract the relative contributions of the two components. Templates for genuine photons are obtained from simulation, while templates for misreconstructed jets are taken from a CR in data with an inverted photon isolated requirement, which is en-

riched in QCD multijet events. The fraction of photons originating from jet misreconstruction is found to range between 3.5% at $p_{\mathrm{T}} = 200\,\mathrm{GeV}$ and 1% at 800 GeV. A prediction for the recoil distribution in QCD multijet events in the photon CR is obtained by weighting the photon candidate spectrum in data by the misreconstructed jet fraction evaluated at the respective p_{T} of the photon candidates. A 25% uncertainty is assigned to the normalization of the QCD multijet background to account for mismodeling of the shower width in simulation. The statistical uncertainty in the determination of the differential recoil shape is taken into account, and ranges from less than 1% at low recoil values up to 10 (20)% at a recoil value of 1.4 TeV in the 2017 (2018) data set.

5.3 Systematic uncertainties

The inputs to the ML fit are subject to various experimental and theoretical uncertainties. The overall experimental uncertainty is dominated by the uncertainties in the efficiency of identifying and reconstructing lepton and photon candidates, as well as the uncertainty in the trigger efficiency. The uncertainties in the efficiencies of reconstructing and identifying electron candidates are 1.0 and 2.5%, respectively. For muons, the corresponding uncertainties are 1%, with an additional 1% uncertainty in the efficiency of the isolation criteria. Finally, for photons, the uncertainty in the reconstruction efficiency is negligible, and the uncertainty in the identification efficiency ranges between 4% at $p_{\mathrm{T}} = 200\,\mathrm{GeV}$ and 12% at 1 TeV. The uncertainties in the identification efficiency of lepton candidates are further propagated to the estimate of the contribution from background processes in the SRs, where events with identified leptons are rejected. These uncertainties predominantly affect the $\mathrm{W} \to \ell\overline{\nu}$ process, and their magnitude is taken to be 1–2% of the total $\mathrm{W} \to \ell\overline{\nu}$ yield for the identification of τ leptons, 1.5% for electrons, and less than 0.5% for muons. The uncertainty in the photon energy calibration modeling is 1% of the photon momentum, leading to an effect on the background yield in the photon control region of up to 3% at low recoil values. The uncertainty in the b tagging efficiency leads to an uncertainty of 6% in the normalization of background processes with top quarks, and 2% in the normalization of the diboson and QCD multijet processes. The uncertainties in the trigger efficiency are 2% for either the electron or photon triggers, and 1% per identified muon for the $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger for the recoil values of less than 400 GeV, and negligible above this threshold. The muon multiplicity dependence of the $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger uncertainty reflects the differences in the reconstruction of muons at the trigger and offline levels, which affect the calculation of the hadronic recoil value. Uncertainties of 50–100% are assigned to the normalization of the QCD multijet contributions, depending on the analysis region. Finally, additional uncertainties of 20% each are assigned to the rate of the Drell–Yan events entering the single-lepton CRs and of the γ + jets events entering the single-electron CRs.

The theoretical uncertainties in the transfer factors related higher-order effects in QCD and EW perturbative expansion are calculated according to the prescription given in Ref. [67] and implemented, as described in Ref. [20]. The uncertainty related to the modeling of PDFs is estimated using the replicas provided in the PDF4LHC15 PDF set [71–74]. Additionally, uncertainties of 10% each are assigned to the cross sections of the diboson and top quark processes, and a further 10% normalization uncertainty is assigned to account for the differences in the p_{T} spectrum of simulated and observed top quark events [75]. For the diboson and $\mathrm{V}\gamma$ processes, additional uncertainties related to unknown mixed QCD–EW NLO corrections are estimated based on the product of the individual EW and QCD correction terms. These uncertainties range between 1 and 10%, depending on the process and boson p_{T}.

The likelihood functions obtained for the monojet and mono-V categories, as well as for the two data taking years, are combined in order to maximize the statistical power of the analysis.

The results based on the data set analyzed here, which corresponds to an integrated luminosity of $101 \pm 2\,\text{fb}^{-1}$, is further combined with the results of an earlier analysis [20] based on a data set collected at the same center-of-mass energy in 2016 and corresponding to an integrated luminosity of $36\,\text{fb}^{-1}$. In this combination, all experimental uncertainties are assumed to be fully correlated between the categories in a given data set, and uncorrelated between the data sets. All theoretical uncertainties are assumed to be fully correlated between the categories and data sets. The parameters controlling the yield of the Z + jets process in each bin of each SR are assumed to be uncorrelated between the categories and data sets.

6 Results and interpretation

The ML fit is performed combining the analysis categories, as well as the 2017 and 2018 data sets. The $p_{\text{T}}^{\text{miss}}$ distributions in the SRs before ("prefit") and after ("postfit") the fit are shown in Fig. 3 for the monojet category and in Fig. 4 for the mono-V category. In all cases, good agreement is observed between the background-only postfit result and the data. The corresponding distributions for the CRs are provided as supplementary material.

6.1 Combination with the previous result

The results presented here are combined with those of an earlier analysis in the same final state, based on a 2016 data set of pp collisions at the same center-of-mass energy, corresponding to an integrated luminosity of $36\,\text{fb}^{-1}$ [20]. The combination is performed by defining a combined likelihood describing all the analysis regions in all data sets. For this purpose, the effects of all theoretical uncertainties are assumed to be correlated, while most experimental sources of uncertainty are uncorrelated between different data taking years. The experimental uncertainties related to the JES, as well as those related to the determination of the integrated luminosity are partially correlated between the data taking years, which is taken into account by splitting the total uncertainty into its correlated and uncorrelated components. In order to harmonize the theoretical signal treatment between the data sets, the signal templates from Ref. [20] are replaced by the templates derived from simulated samples with generator configurations identical to those used in the analysis of the more recent data sets. This procedure affects the excluded cross sections based on the 2016 data set alone by up to 13%, depending on the signal hypothesis. The effect is reduced to a few percent level in the fully combined final result.

6.2 Higgs portal interpretation

The results are also interpreted in terms of the exclusion limits at 95% confidence level (CL) on the branching fraction of an otherwise SM-like Higgs boson to particles without detectable detector interactions (invisible decays). The limits are derived assuming the SM production cross section for the Higgs boson. In the monojet category, values of $\mathcal{B}(\text{H} \to \text{inv.})$ larger than 59.6% are excluded (36.2% expected). In the mono-V case, branching fractions of more than 37.0% are excluded (31.0% expected). Finally, the combination of all categories yields an exclusion limit of $\mathcal{B}(\text{H} \to \text{inv.}) < 27.8\%$ (25.3% expected). These limits are summarized in Fig. 5.

6.3 Interpretation in a DM simplified model with a colorless mediator

The results are further interpreted in terms of simplified models of DM production. In a model with a spin-1 mediator, exclusion limits are calculated in the two-dimensional parameter space of the DM and mediator particle masses, m_{DM} and m_{med}. The coupling between the mediator and the SM quarks is set to a constant value of $g_{\text{q}} = 0.25$, and the mediator-DM coupling is set

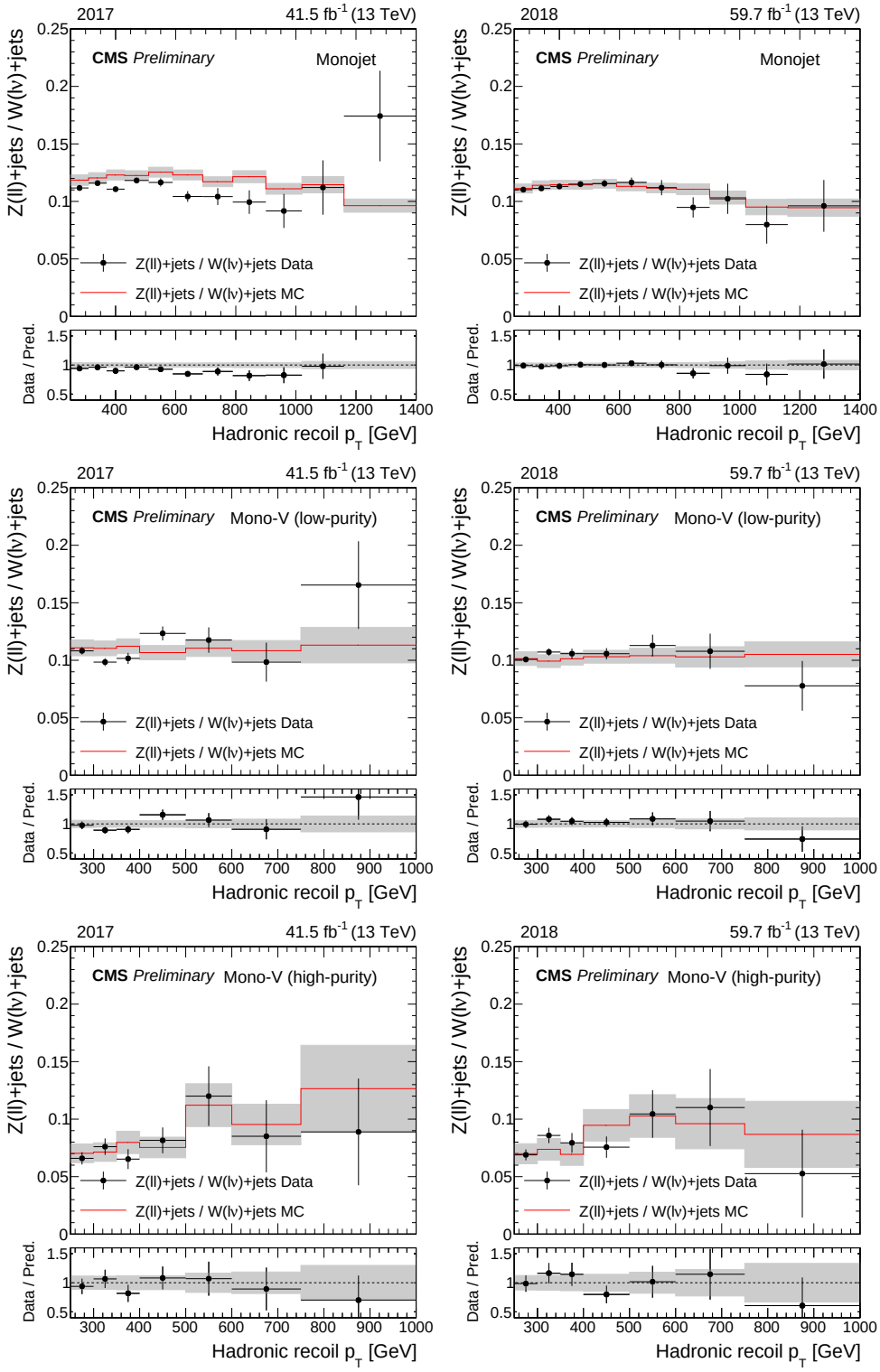

Figure 1: Ratio of the dilepton and single-lepton control region yields predicted using simulation (red solid line), and observed in data (black points). The gray band represents the total uncertainty in the ratio. In the lower panels, the ratio of data over prediction is shown. From upper to lower, the rows show the monojet, low-purity, and high-purity mono-V categories, while the left (right) column represents the 2017 (2018) data set.

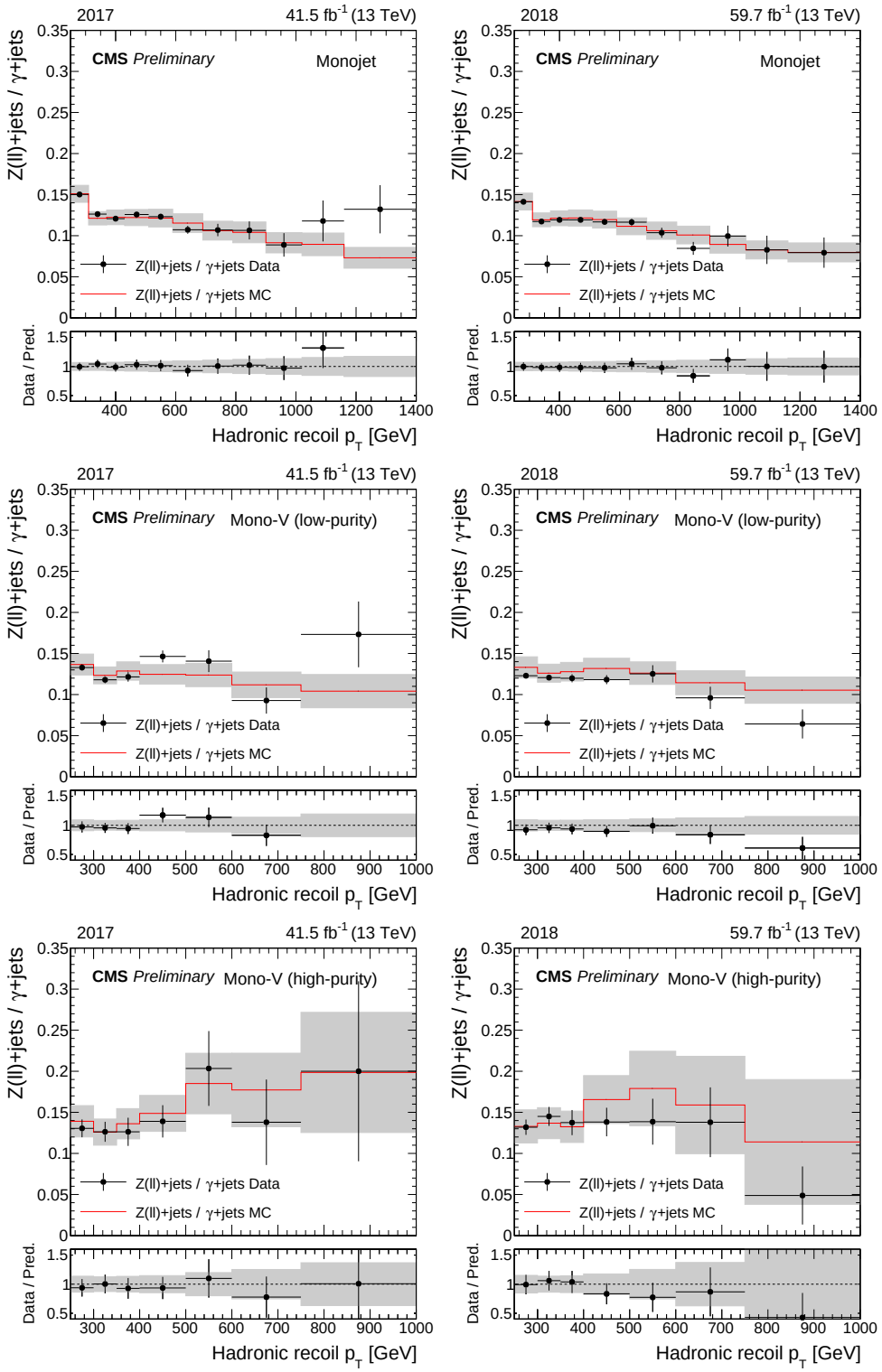

Figure 2: Same as Fig. 1, but for the ratio of the dilepton and photon control regions.

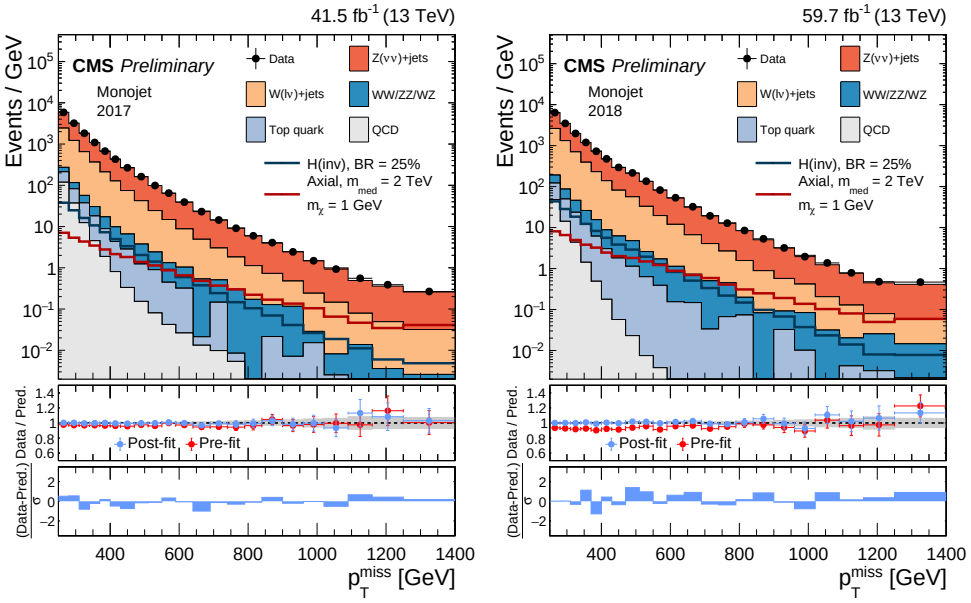

Figure 3: Comparison between data and the background prediction in the monojet signal regions before and after the simultaneous fit. The fit includes all control samples and the signal region in all categories and both data taking years, and the background-only fit model is used. The resulting distributions are shown separately for 2017 (left) and 2018 (right column). Templates for two signal hypothesis are shown overlaid as dark blue and dark red solid lines. The last bin includes the overflow. In the middle panels, ratios of data to the pre-fit background prediction (red open points) and post-fit background prediction (blue solid points) are shown. The gray band in the lower panels indicates the post-fit uncertainty after combining all the systematic uncertainties. Finally, the distribution of the pulls, defined as the difference between data and the post-fit background prediction divided by the quadratic sum of the post-fit uncertainty in the prediction and statistical uncertainty in data, is shown in the lower panels.

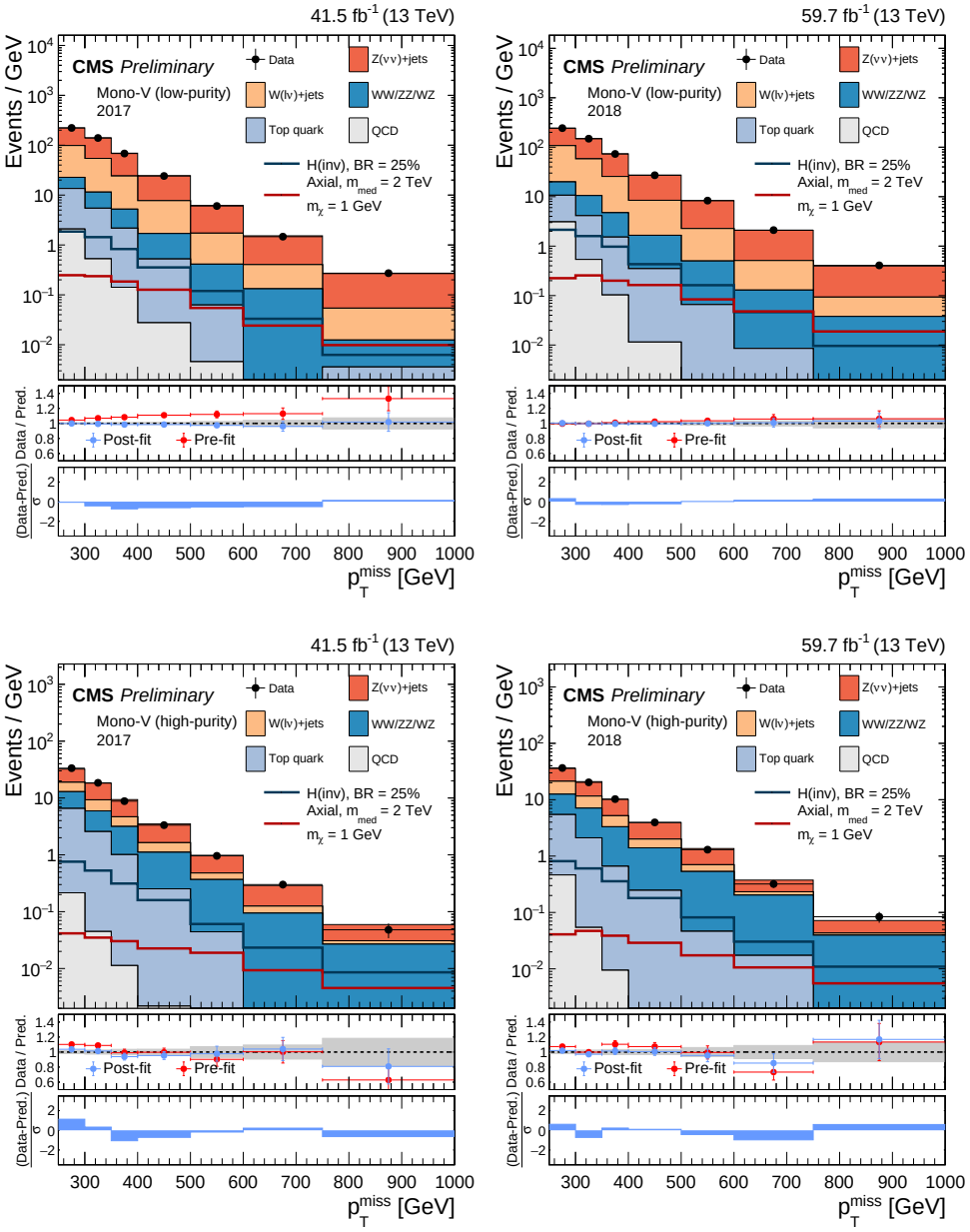

Figure 4: Comparison between data and the background prediction in the mono-V signal regions before and after the simultaneous fit. The fit includes all control samples and the signal region in all categories and both data taking years, and the background-only fit model is used. The resulting distributions are shown separately for 2017 (left column) and 2018 (right column), as well as for the low- and high-purity categories (upper and lower rows, respectively).

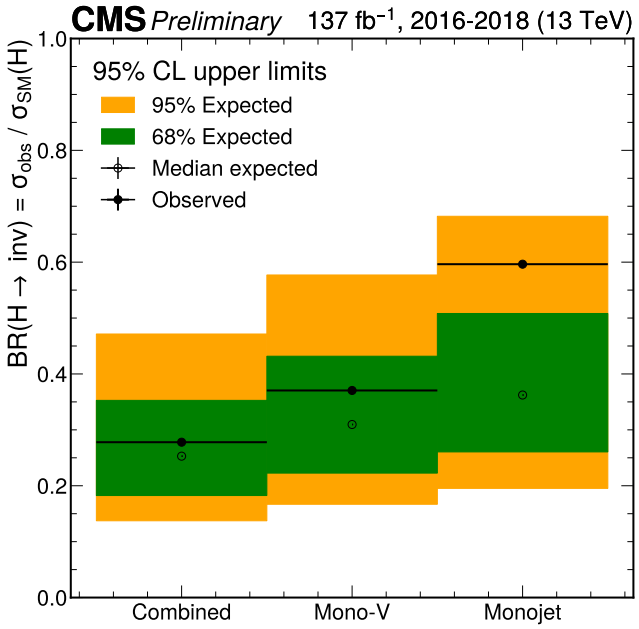

Figure 5: Upper limits at 95% CL on the branching fraction of the Higgs boson to invisible final states. The results are shown separately for the monojet and mono-V categories, as well as for their combination. The final combined limit is 27.8% (25.3% expected).

to $g_\chi = 1.0$, and both vector and axial-vector type couplings are considered in separate interpretations. The resulting exclusion limits at 95% CL on the signal strength μ are shown in Fig. 6. The m_{med} values of up to 2.0 TeV (2.2 TeV expected) are excluded for low m_{DM} values. The maximum excluded values of m_{med} decrease with the m_{DM} increase, as the branching fraction of the mediator to dark matter decays diminishes. The dependence of the branching fraction on m_{DM} is more pronounced in the case of an axial-vector mediator, leading to a reduced maximal exclusion reach in m_{DM} of 0.7 TeV compared to 1 TeV for the vector case. Compared to the results of Ref. [20], the combined limits improve the maximal exclusion in terms of the mediator mass by approximately 400 GeV, or 20%. In addition to the constraints in the m_{DM}-m_{med} plane, we also obtain exclusion limits in the planes of m_{med} and g_{q}, as well as m_{med} and g_χ, which are shown in Fig. 7 for the case of axial-vector couplings. The coupling value exclusion is derived analytically from the signal strength exclusion at the default coupling values by rescaling the signal cross section according to the production cross section and decay branching fractions of the mediator, using the formalism of Ref. [57]. The DM candidate mass m_{DM} is fixed to $m_{\text{med}}/3$. For low mediator masses, values of g_{q} (g_χ) as low as 0.02 (0.06) are excluded, providing significant additional insight into the probed parameter space, compared to the mass exclusion for fixed coupling values.

The expected upper limits on the signal strength in the case of spin-0 mediators are shown in Fig. 8. The mediator couplings are assumed to be $g_{\text{q}} = 1.0$ and $g_\chi = 1.0$, and the DM candidate mass is fixed to 1 GeV. For scalar mediators, signal strengths larger than 1.2 can be excluded at low mediator mass values of $\approx$ 50 GeV. A pseudoscalar mediator with a mass below $m_{\text{med}} = 490$ GeV is excluded (440 GeV expected). In both cases, the signal strength limits show distinctive features around the top quark decay threshold of $m_{\text{med}} = 2m_{\text{t}}$. As the mediator is produced via a top quark loop, the signal cross section is enhanced as the

mediator mass approaches the threshold from below. Above the threshold, the decay of the mediator into a pair of top quarks becomes possible, leading to a significant suppression of the branching fraction to DM candidates, and therefore the effective signal cross section.

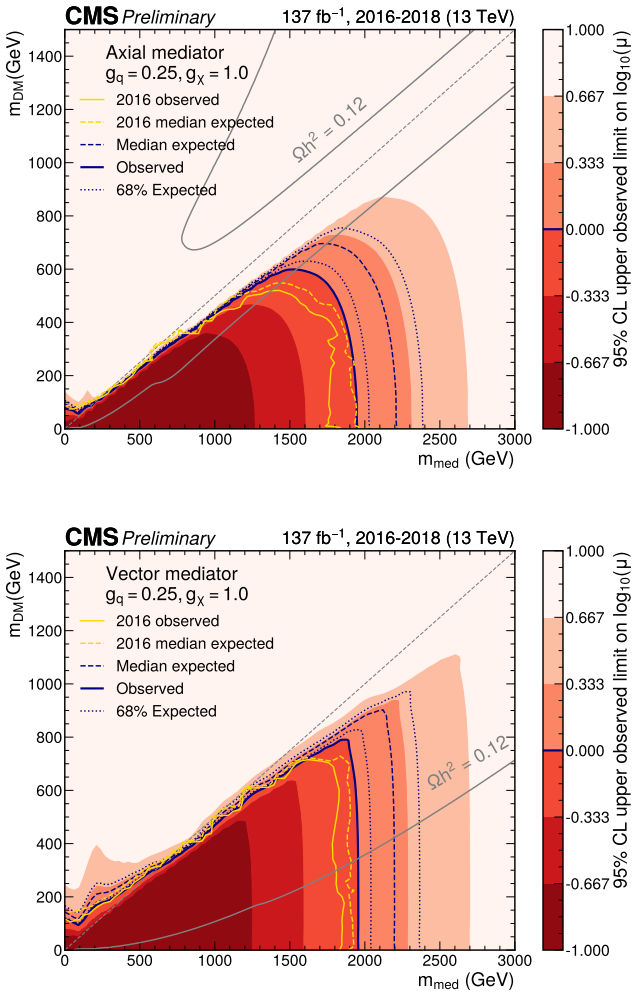

Figure 6: Exclusion limits at 95% CL on the signal strength $\mu = \sigma/\sigma_{\text{theo}}$ in the m_{med}-m_{DM} plane for the coupling values of $g_{\text{q}} = 0.25$, $g_\chi = 1.0$ for an axial-vector (upper) or vector (lower) mediator. The blue solid line indicates the observed exclusion boundary $\mu = 1$. The blue dashed and dotted lines represent the expected exclusion and the the 68% confidence level interval around the expected boundary, respectively. Parameter combinations with larger values of μ (indicated by a darker shade in the color scale) are excluded. The observed exclusion reaches up to $m_{\text{med}} = 2.0\,\text{TeV}$ for low values of $m_{\text{DM}} = 1\,\text{GeV}$ (2.2 TeV expected). Yellow solid and dashed lines represent the observed and expected exclusion boundaries from Ref. [20]. The gray dashed line indicates the diagonal $m_{\text{med}} = 2m_{\text{DM}}$, above which only off-shell mediator production contributes to the jet+$p_{\text{T}}^{\text{miss}}$ final state. The steep increase of the signal strength limit above the diagonal leads to fluctuations of the exclusion contour, which are due to finite precision in the interpolation method in this region. The gray solid lines represent parameter combinations for which the simplified model reproduces the observed DM relic density in the universe under the assumption of a thermal freeze-out mechanism [57, 76].

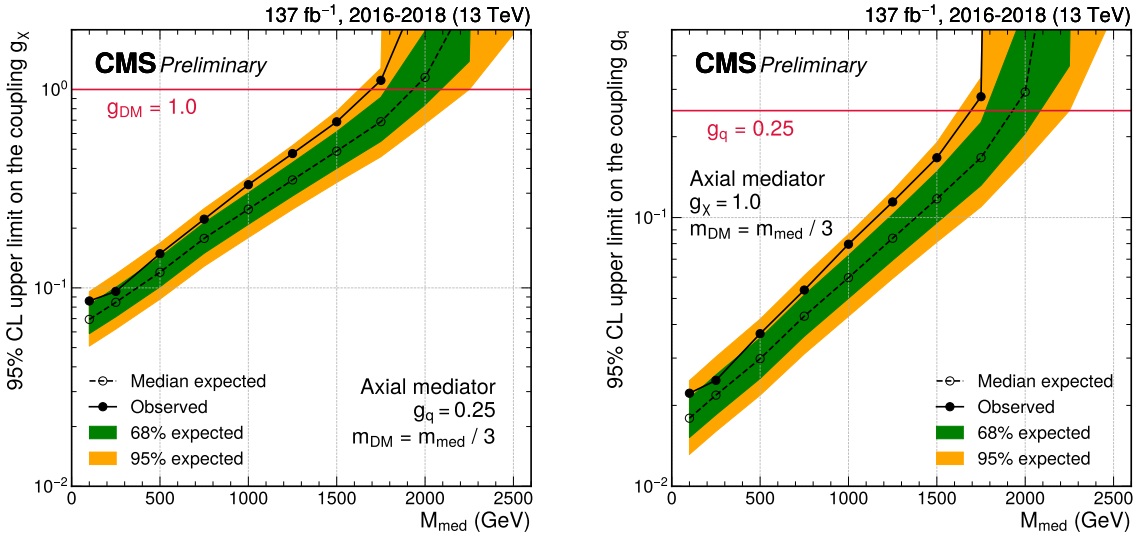

Figure 7: Exclusion limits at 95% CL on the couplings g_χ (left) and g_q (right) for an axial-vector mediator. In each panel, the result is shown as a function of the mediator mass m_{med}, with the mass of the DM candidate fixed to $m_{DM} = m_{med}/3$. In either case, only one coupling is varied, while the other coupling is fixed at its default value ($g_q = 0.25$ or $g_\chi = 1.0$).

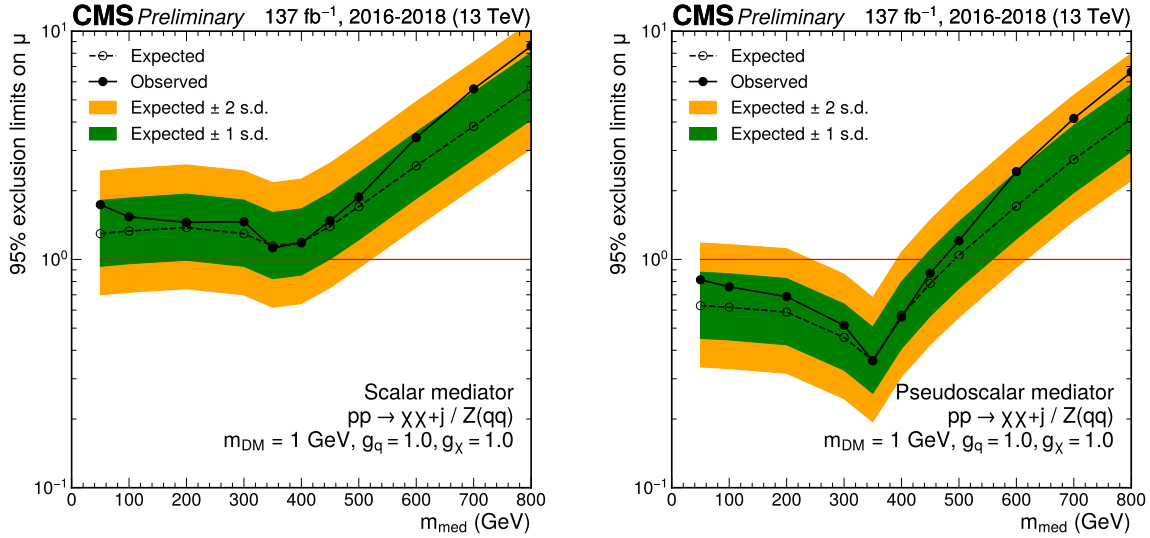

Figure 8: Upper limits at 95% CL on the signal strength $\mu = \sigma/\sigma_{theo}$ as a function of m_{med} for scenarios with scalar (left) and pseudoscalar (right) mediators and coupling values of $g_q = 1.0$, $g_\chi = 1.0$, for a constant value of $m_{DM} = 1\,\text{GeV}$. The red solid line indicates the exclusion boundary $\mu = 1$. In the case of a pseudoscalar mediator, m_{med} values up to 480 GeV are excluded (440 GeV expected).

6.4 Fermion portal interpretation

In the fermion portal model, the results of the analysis are shown in Fig. 9 in the plane of the mediator mass m_Φ and the DM candidate mass m_χ. At low m_χ values, the mediator masses of up to 1.5 TeV are excluded (1.7 TeV expected).

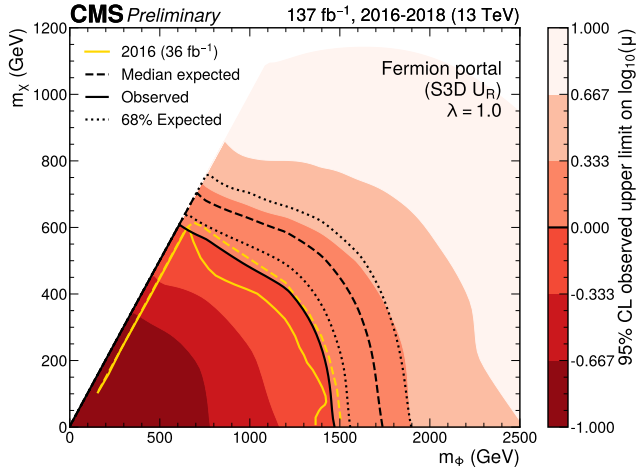

Figure 9: Exclusion limits at 95% CL in in the plane of the mediator mass m_Φ and the DM candidate mass m_χ. The black dashed and dotted lines represent the median expected exclusion and the 68% confidence level interval around it, respectively. The black solid line represents the observed exclusion. The yellow solid line shows the observed exclusion from Ref. [20] for comparison.

6.5 ADD interpretation

In the ADD scenario, lower limits on the fundamental Planck scale M_D for the number of extra dimensions d ranging from 2 to 7 are shown in Fig. 10. For the lowest number of extra dimensions considered here, $d = 2$, M_D values of up to 10.8 TeV are excluded (12.1 TeV expected). As the number of extra dimensions increases, the probed M_D value is reduced to 5.4 TeV for $d = 6$ (5.9 TeV expected). Compared to the result of Ref. [20], these limits represent an improvement of approximately 8% for low values of d (20% expected). At larger values, the relative gain in M_D sensitivity is smaller, as a result of the dependence of the signal cross section on M_D, which becomes steeper as d increases.

6.6 Leptoquark interpretation

Finally, upper limits are placed on the production cross section of LQs coupled to up quarks and neutrinos with a coupling value λ. The branching fraction for the decay of the LQ into an electron neutrino and up quark is assumed to be 100% (also referred to as to $\beta = 0$ in the literature). The limits are shown in Fig. 11. Generally, both single and pair LQ production contributes to the signal, with the coupling λ mainly influencing the single production rate. The pair production dominates at lower LQ masses of $m_{LQ} < 1$ TeV, which has already been excluded by previous searches [77]. In the higher-mass regime $m_{LQ} > 1$ TeV the contribution from single production is increased, providing additional sensitivity to the value of λ. The minimal excluded coupling λ ranges from about 0.5 at $m_{LQ} = 1$ TeV (0.4 expected) to $\lambda = 0.9$ at $m_{LQ} = 1.5$ TeV (0.7 expected) and $\lambda = 1.8$ at 2 TeV (1.25 expected).

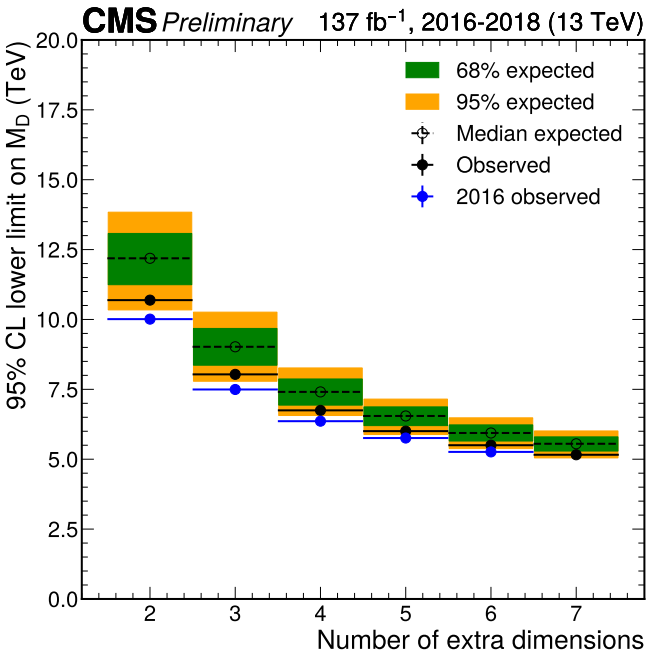

Figure 10: Exclusion limits at 95% CL on M_D in the ADD scenario for different values of the number of extra dimensions d. The blue markers show the result from Ref. [20] for comparison.

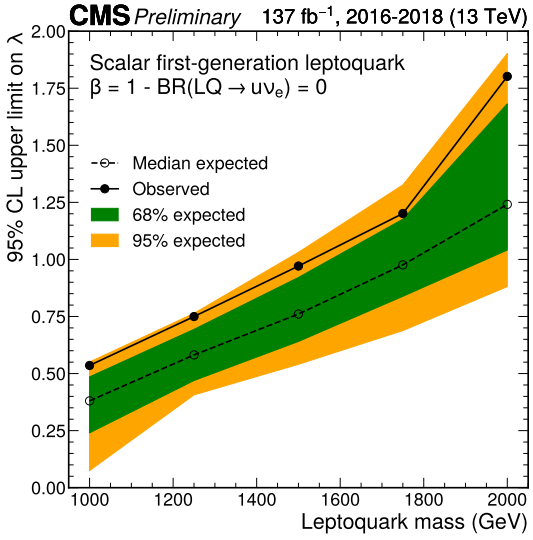

Figure 11: Upper limits at 95% CL on the leptoquark coupling λ as a function of the leptoquark mass. The branching fraction for the decay of the leptoquark into an electron neutrino and up quark is assumed to be 100% ($\beta = 0$). The dashed line indicates the median expected exclusion contour. For a leptoquark mass value of 1 TeV, coupling values as low as approximately 0.5 are excluded (0.4 expected). The upper limit increases with the leptoquark mass increase, reaching $\lambda = 0.9$ at a mass of 1.5 TeV (0.7 expected) and $\lambda = 1.8$ at 2 TeV (1.25 expected).

7 Summary

A search for physics beyond the standard model in events with energetic jets and large missing transverse momentum is presented. A data set of proton-proton collisions at a center-of-mass energy of 13 TeV, corresponding to and integrated luminosity of 101 fb^{-1} is analyzed, and the analysis results are combined with those of an earlier search using an independent data set collected at the same center-of-mass energy, corresponding to an integrated luminosity of 36 fb^{-1} [20]. Separate analysis categories are defined for events with a large-radius jet consistent with a decay of a W or a Z boson, and for events without such a jet. A joint maximum-likelihood fit over a combination of signal and control regions is used to constrain standard model (SM) background processes and to extract possible signal. The data are found to be in a good agreement with the fit results, with no evidence for a significant signal contribution found. The result is interpreted in terms of exclusion limits on the parameters of a number of models of beyond-the-SM physics. In simplified models of the production of dark matter (DM) candidates via a spin-1 s-channel mediator, values of the mediator mass of up to 1.95 TeV are excluded, assuming the couplings of $g_{\text{q}} = 0.25$ between the mediator and quarks, and $g_{\chi} = 1.0$ between the mediator and the Dirac fermion DM particles. Assuming a fixed ratio $m_{\text{DM}} = m_{\text{med}}/3$, coupling values as low as $g_{\text{q}} = 0.018$ and $g_{\chi} = 0.070$ can be excluded for a mediator mass value of $m_{\text{med}} = 100$ GeV. In a similar model with a pseudoscalar spin-0 mediator, mediator masses of up to 470 GeV are excluded. We further constrain the branching fraction of the Higgs boson decay to invisible particles to be below 27.8%. In a model of large extra dimensions, values of the fundamental Planck scale below from 10.7 to 5.2 TeV can be excluded, depending on the number of extra dimensions between 2 and 7. Finally, the production of leptoquarks decaying into the up quark and the electron neutrino is excluded for coupling values between the leptoquarks and the SM fermions larger than between 10^{-5} and 1.8 for leptoquark masses between 0.5 and 1.8 TeV.

References

[1] G. Bertone, D. Hooper, and J. Silk, "Particle dark matter: evidence, candidates and constraints", *Phys. Rept.* **405** (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:hep-ph/0404175.

[2] S. Kanemura, S. Matsumoto, T. Nabeshima, and N. Okada, "Can WIMP dark matter overcome the nightmare scenario?", *Phys. Rev. D* **82** (2010) 055026, doi:10.1103/PhysRevD.82.055026, arXiv:1005.5651.

[3] O. Lebedev, H. M. Lee, and Y. Mambrini, "Vector Higgs-portal dark matter and the invisible Higgs", *Phys. Lett. B* **707** (2012) 570, doi:10.1016/j.physletb.2012.01.029, arXiv:1111.4482.

[4] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, "Implications of LHC searches for Higgs–portal dark matter", *Phys. Lett. B* **709** (2012) 65, doi:10.1016/j.physletb.2012.01.062, arXiv:1112.3299.

[5] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[6] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[7] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[8] ATLAS Collaboration, "Combination of searches for invisible Higgs boson decays with the ATLAS experiment", *Phys. Rev. Lett.* **122** (2019) 231801, doi:10.1103/PhysRevLett.122.231801, arXiv:1904.05105.

[9] CMS Collaboration, "Search for invisible decays of a Higgs boson produced through vector boson fusion in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **793** (2019) 520, doi:10.1016/j.physletb.2019.04.025, arXiv:1809.05937.

[10] D. Abercrombie et al., "Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum", *Phys. Dark Univ.* **27** (2020) 100371, doi:10.1016/j.dark.2019.100371, arXiv:1507.00966.

[11] Y. Bai and J. Berger, "Fermion portal dark matter", *JHEP* **11** (2013) 171, doi:10.1007/JHEP11(2013)171, arXiv:1308.0612.

[12] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "The hierarchy problem and new dimensions at a millimeter", *Phys. Lett. B* **429** (1998) 263, doi:10.1016/S0370-2693(98)00466-3, arXiv:hep-ph/9803315.

[13] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity", *Phys. Rev. D* **59** (1999) 086004, doi:10.1103/PhysRevD.59.086004, arXiv:hep-ph/9807344.

[14] J. C. Pati and A. Salam, "Unified lepton-hadron symmetry and a gauge theory of the basic interactions", *Phys. Rev. D* **8** (1973) 1240, doi:10.1103/PhysRevD.8.1240.

[15] J. C. Pati and A. Salam, "Lepton number as the fourth color", *Phys. Rev. D* **10** (1974) 275, doi:10.1103/PhysRevD.10.275. [Erratum: doi:10.1103/PhysRevD.11.703.2].

[16] H. Georgi and S. L. Glashow, "Unity of all elementary particle forces", *Phys. Rev. Lett.* **32** (1974) 438, doi:10.1103/PhysRevLett.32.438.

[17] B. Diaz, M. Schmaltz, and Y.-M. Zhong, "The leptoquark hunter's guide: Pair production", *JHEP* **10** (2017) 097, doi:10.1007/JHEP10(2017)097, arXiv:1706.05033.

[18] J. L. Hewett and S. Pakvasa, "Leptoquark production in hadron colliders", *Phys. Rev. D* **37** (1988) 3165, doi:10.1103/PhysRevD.37.3165.

[19] O. J. P. Eboli and A. V. Olinto, "Composite leptoquarks in hadronic colliders", *Phys. Rev. D* **38** (1988) 3461, doi:10.1103/PhysRevD.38.3461.

[20] CMS Collaboration, "Search for new physics in final states with an energetic jet or a hadronically decaying W or Z boson and transverse momentum imbalance at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **97** (2018) 092005, doi:10.1103/PhysRevD.97.092005, arXiv:1712.02345.

[21] ATLAS Collaboration, "Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector", *JHEP* **01** (2018) 126, doi:10.1007/JHEP01(2018)126, arXiv:1711.03301.

[22] ATLAS Collaboration, "Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **10** (2018) 180, doi:10.1007/JHEP10(2018)180, arXiv:1807.11471.

[23] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

[24] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[25] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[26] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[27] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[28] CMS Collaboration, "Particle-flow reconstruction and global event description with the cms detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[29] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[30] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2017.

[31] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[32] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[33] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.

[34] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC", *Phys. Rev. Lett.* **100** (2008) 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.

[35] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[36] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[37] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[38] GEANT4 Collaboration, "GEANT4 — a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[39] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[40] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[41] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[42] M. Czakon, P. Fiedler, and A. Mitov, "Total top-quark pair-production cross section at hadron colliders through $o(\alpha_s^4)$", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[43] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].

[44] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[45] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-", *Phys. Rev. D* **82** (2010) 054018, doi:10.1103/PhysRevD.82.054018, arXiv:1005.4451.

[46] M. Aliev et al., "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR", *Comput. Phys. Commun.* **182** (2011) 1034, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.

[47] P. Kant et al., "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions", *Comput. Phys. Commun.* **191** (2015) 74, doi:10.1016/j.cpc.2015.02.001, arXiv:1406.4403.

[48] T. Gehrmann et al., "W^+W^- production at hadron colliders in next to next to leading order QCD", *Phys. Rev. Lett.* **113** (2014) 212001, doi:10.1103/PhysRevLett.113.212001, arXiv:1408.5243.

[49] J. M. Campbell and R. K. Ellis, "An update on vector boson pair production at hadron colliders", *Phys. Rev. D* **60** (1999) 113006, doi:10.1103/PhysRevD.60.113006, arXiv:hep-ph/9905386.

[50] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, "Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM", *JHEP* **02** (2012) 088, doi:10.1007/JHEP02(2012)088, arXiv:1111.2854.

[51] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, "$HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO", *JHEP* **10** (2013) 083, doi:10.1007/JHEP10(2013)083, arXiv:1306.2542.

[52] P. Nason and C. Oleari, "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG", *JHEP* **02** (2010) 037, doi:10.1007/JHEP02(2010)037, arXiv:0911.5299.

[53] L. H. C. S. W. Group, "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", 10, 2016. arXiv:1610.07922.

[54] O. Mattelaer and E. Vryonidou, "Dark matter production through loop-induced processes at the LHC: the s-channel mediator case", *Eur. Phys. J. C* **75** (2015) 436, doi:10.1140/epjc/s10052-015-3665-5, arXiv:1508.00564.

[55] M. Backović et al., "Higher-order QCD predictions for dark matter production at the lhc in simplified models with s-channel mediators", *Eur. Phys. J. C* **75** (2015) 482, doi:10.1140/epjc/s10052-015-3700-6, arXiv:1508.05327.

[56] M. Neubert, J. Wang, and C. Zhang, "Higher-order QCD predictions for dark matter production in mono-Z searches at the LHC", *JHEP* **02** (2016) 082, doi:10.1007/JHEP02(2016)082, arXiv:1509.05785.

[57] LHC Dark Matter Working Group, "Recommendations of the LHC dark matter working group: Comparing LHC searches for dark matter mediators in visible and invisible decay channels and calculations of the thermal relic density", *Phys. Dark Univ.* **26** (2019) 100377, doi:10.1016/j.dark.2019.100377, arXiv:1703.05703.

[58] C. Arina, B. Fuks, and L. Mantani, "A universal framework for t-channel dark matter models", *Eur. Phys. J. C* **80** (2020) 409, doi:10.1140/epjc/s10052-020-7933-7, arXiv:2001.05024.

[59] S. Ask et al., "Real emission and virtual exchange of gravitons and unparticles in Pythia8", *Comput. Phys. Commun.* **181** (2010) 1593, doi:10.1016/j.cpc.2010.05.013, arXiv:0912.4233.

[60] CMS Collaboration, "Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques", *JINST* **15** (2020) P06005, doi:10.1088/1748-0221/15/06/P06005, arXiv:2004.08262.

[61] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021) P05014, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

[62] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[63] CMS Collaboration, "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P10005, doi:10.1088/1748-0221/13/10/P10005, arXiv:1809.02816.

[64] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[65] Particle Data Group Collaboration, "Review of particle physics", *PTEP* **2020** (2020) 083C01, doi:10.1093/ptep/ptaa104.

[66] CMS Collaboration, "Search for dark matter produced with an energetic jet or a hadronically decaying W or Z boson at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2017) 014, doi:10.1007/JHEP07(2017)014, arXiv:1703.01651.

[67] J. M. Lindert et al., "Precise predictions for v+jets dark matter backgrounds", *Eur. Phys. J. C* **77** (2017) 829, doi:10.1140/epjc/s10052-017-5389-1, arXiv:1705.04664.

[68] J. Baglio, L. D. Ninh, and M. M. Weber, "Massive gauge boson pair production at the LHC: a next-to-leading order story", *Phys. Rev. D* **88** (2013) 113005, doi:10.1103/PhysRevD.88.113005, arXiv:1307.4331. [Erratum: doi:10.1103/PhysRevD.94.099902].

[69] A. Denner, S. Dittmaier, M. Hecht, and C. Pasold, "NLO QCD and electroweak corrections to W+γ production with leptonic W-boson decays", *JHEP* **04** (2015) 018, doi:10.1007/JHEP04(2015)018, arXiv:1412.7421.

[70] A. Denner, S. Dittmaier, M. Hecht, and C. Pasold, "NLO QCD and electroweak corrections to Z+γ production with leptonic Z-boson decays", *JHEP* **02** (2016) 057, doi:10.1007/JHEP02(2016)057, arXiv:1510.08742.

[71] J. Butterworth et al., "PDF4LHC recommendations for LHC run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[72] S. Dulat et al., "New parton distribution functions from a global analysis of quantum chromodynamics", *Phys. Rev. D* **93** (2016) 033006, doi:10.1103/PhysRevD.93.033006, arXiv:1506.07443.

[73] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs", *Eur. Phys. J. C* **75** (2015) 204, doi:10.1140/epjc/s10052-015-3397-6, arXiv:1412.3989.

[74] NNPDF Collaboration, "Parton distributions for the LHC run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[75] M. Czakon, D. Heymes, and A. Mitov, "High-precision differential predictions for top-quark pairs at the LHC", *Phys. Rev. Lett.* **116** (2016) 082003, doi:10.1103/PhysRevLett.116.082003, arXiv:1511.00549.

[76] Planck Collaboration, "Planck 2018 results. VI. cosmological parameters", *Astron. Astrophys.* **641** (2020) A6, doi:10.1051/0004-6361/201833910, arXiv:1807.06209.

[77] CMS Collaboration, "Searches for physics beyond the standard model with the m_{T2} variable in hadronic final states with and without disappearing tracks in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **80** (2020) 3, doi:10.1140/epjc/s10052-019-7493-x, arXiv:1909.03460.