2022/03/16
Archive Hash: d0622d3-D
Archive Date: 2022/03/16

# Search for a right-handed W boson and a heavy neutrino in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

The CMS Collaboration

## Abstract

A search is presented for a right-handed W boson ($\text{W}_\text{R}$) and a heavy neutrino (N), in a final state consisting of two same-flavor leptons (ee or $\mu\mu$) and two quarks. The search is performed with the CMS experiment at the CERN LHC using a data sample of proton-proton collisions at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of $138\,\text{fb}^{-1}$. The search covers two regions of phase space, one where the decay products of the heavy neutrino are merged into a single large-area jet, and one where the decay products are well separated. The expected signal is characterized by an excess in the invariant mass distribution of the final-state objects. No significant excess over the standard model background expectations is observed. The observations are interpreted as upper limits on the product of $\text{W}_\text{R}$ production cross sections and branching fractions assuming that couplings are identical to those of the standard model W boson. For N masses $m_\text{N}$ equal to half the $\text{W}_\text{R}$ mass $m_{\text{W}_\text{R}}$ ($m_\text{N} = 0.2\,\text{TeV}$), $m_{\text{W}_\text{R}}$ is excluded at 95% confidence level up to 4.7 (4.8) and 5.0 (5.4) TeV for the electron and muon channels, respectively. This analysis provides the most stringent limits on the $\text{W}_\text{R}$ mass to date.

# 1 Introduction

Left-right (LR) symmetric models [1–4], which extend the electroweak sector of the standard model (SM) by a right-handed SU(2) group, provide a possible explanation for parity violation in the SM as the consequence of spontaneous symmetry breaking at a multi-TeV mass scale. These models predict a heavy partner of the SM W boson, a heavy, right-handed gauge boson $W_R$ that is coupled to right-handed fermions. In addition, LR models also provide an explanation for the small mass of SM neutrinos through the seesaw mechanism [5–7], which requires the existence of a heavy right-handed neutrino (N) for each lepton flavor. The heavy neutrinos couple exclusively to leptons of the corresponding flavor, and have different masses.

The coupling strength of the $W_R$ boson to the SM particles ($g_R$) is a free parameter in most LR models. We assume LR symmetry, such that $g_R$ is the same as the SM coupling constant $g_L$.

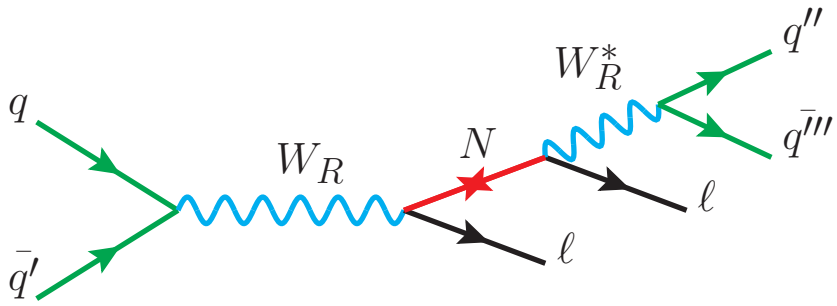


Figure 1: Feynman diagram for the production of a heavy neutrino via the decay of a $W_R$ boson.

The dominant production process for the $W_R$ boson at the CERN LHC is the Drell–Yan (DY) mechanism. The leading order Feynman diagram for this process is shown in Fig. 1. Although the potential Majorana nature of the right-handed neutrinos implies that the final-state charged leptons can have the same sign, we do not impose charge requirements on the final-state leptons, in order to remain sensitive to the widest possible range of models. Searches for $W_R$ bosons and heavy neutrinos have been performed by the ATLAS [8–10] and CMS [11–15] Collaborations using LHC proton-proton (pp) collision data at $\sqrt{s} = 8$ and 13 TeV. These searches have excluded regions of phase space with masses of the right-handed W boson and heavy neutrino ($m_{W_R}$ and $m_N$, respectively) up to several TeV.

In this paper, we extend the region of parameter space covered in the previous 2016 CMS search for $W_R$ bosons in events with two same-flavor leptons (e or $\mu$) and two jets using 2016 data [12], by including the regime where the $W_R$ boson is heavy compared to the N ($m_{W_R}/m_N \geq 10$). In this scenario, the heavy neutrinos are produced with large transverse momentum ($p_T$) and their decay products are collimated along their direction of motion. Therefore, the heavy neutrino decay is reconstructed in a single jet and identified using jet substructure techniques, similar to those discussed in Refs. [16–18] and previously applied in Ref. [10]; we refer to these as "boosted" events, in contrast to "resolved" events where the two jets from the heavy neutrino decay are reconstructed separately. The inclusion of boosted events leads to significant improvements in the search sensitivity in the region where $m_N < 0.5$ TeV. To obtain maximum sensitivity, a statistical combination of the resolved and boosted results is performed. For this analysis we use the data sample collected in 2016–2018 with the CMS detector at the LHC corresponding to an integrated luminosity of 138 $\text{fb}^{-1}$. Tabulated results are provided in HEPData [19].

# 2 The CMS detector

A detailed description of the CMS detector, together with a definition of the coordinate system and the relevant kinematic variables, can be found in Ref. [20]. 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 ($\eta$) 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.

Events of interest are selected using a two-tiered trigger system [21]. The first level, 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 time interval of less than 4 $\mu$s [22]. The second level, known as the high-level trigger, 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.

# 3 Simulated samples

All signal events are simulated using the MADGRAPH5_aMC@NLO 2.6.5 [23] Monte Carlo (MC) event generator at leading order (LO) following the prescriptions described in Refs. [17, 18] for various $m_{W_R}$ hypotheses in the range from 0.2 to 7.0 TeV, with $m_N$ ranging from 0.1 TeV up to $m_{W_R}$ where N is a Majorana neutrino. The events are generated in final states with two same-flavor (SF) leptons (ee or $\mu\mu$). The production cross sections are scaled to next-to-LO (NLO) in quantum chromodynamics (QCD) using $K$ factors obtained from the same generator.

The background samples are simulated with several MC event generators. The Z and W boson production associated with jets is simulated with the MADGRAPH5_aMC@NLO 2.2.2 (2.4.2) generator for up to four parton-level jets at LO for the 2016 (2017–2018) samples, with the MLM matching scheme between jets from matrix element calculations and parton showers [24]. The POWHEG 2.0 [25–28] generator is used to model the $t\bar{t}$ production, as well as tW and $t$-channel single top quark production at NLO. The $s$-channel single top quark production is generated with the MADGRAPH5_aMC@NLO generator at NLO precision. The inclusive decay of $t\bar{t}$ produced in association with a W and Z boson is simulated by MADGRAPH5_aMC@NLO at NLO and LO precision, respectively. The PYTHIA generator [29] is used to simulate diboson processes (WW, WZ, and ZZ) at LO. Triple vector boson (WWW, WWZ, WZZ, and ZZZ) events are generated at NLO using MADGRAPH5_aMC@NLO.

The generators used for the signal and background processes are interfaced with PYTHIA 8.226 (8.230) [29] to simulate the parton showering and hadronization in the 2016 (2017–2018) samples. The PYTHIA parameters for the underlying event description are set with either the CUETP8M1 [30] tune for the 2016 samples or the CP5 [31] tune for the 2017–2018 samples. For the description of the parton distribution functions (PDFs), the NNPDF3.0 [32] PDF sets are used to produce all simulated background samples in 2016. For the 2017–2018 background samples, the NNPDF3.1 next-to-NLO (NNLO) PDF sets [33] are used. The NNPDF3.1 NNLO PDF sets are used to simulate all the signal samples.

All simulated samples were processed through a GEANT4 simulation [34] of the CMS detector. Multiple pp collisions may occur in the same or adjacent LHC bunch crossings (pileup) and contribute to the overall event activity in the detector. This effect is included in the simulation,

| MC Generator | 사용된 물리 프로세스 | 계산 정확도 |
|---|---|---|
| MadGraph5_aMC@NLO | $Z$+jets, $W$+jets (2016: v2.2.2, 2017-2018: v2.4.2) | Leading Order (LO) |
| POWHEG 2.0 | $t\bar{t}$, $tW$, $t$-채널 단일 top 생산 | Next-to-Leading Order (NLO) |
| MadGraph5_aMC@NLO | $s$-채널 단일 top 생산 | NLO |
| MadGraph5_aMC@NLO | $t\bar{t}+W$ (NLO), $t\bar{t}+Z$ (LO) | NLO & LO |
| PYTHIA | WW, WZ, ZZ diboson 과정 | LO |
| MadGraph5_aMC@NLO | $WWW, WWZ, WZZ, ZZZ$ 트리플 보손 생산 | NLO |



and the distribution of the number of pileup interactions is adjusted to match the one observed in the data, assuming a total inelastic cross section of 69.2 mb [35].

# 4 Event reconstruction and object selection

For this search, the events were selected at the trigger level by requiring the presence of a high-momentum lepton in the event. A combination of three triggers that require an isolated electron with $p_{\mathrm{T}} > 27\,(32)$ GeV, an electron with $p_{\mathrm{T}} > 115$ GeV, or a photon with $p_{\mathrm{T}} > 175\,(200)$ GeV is used to collect events that contain at least one electron in 2016 (2018). In 2017, two triggers with the requirement of an isolated electron with $p_{\mathrm{T}} > 35$ GeV or a photon with $p_{\mathrm{T}} > 200$ GeV were used. The single-electron triggers differ in their usage of isolation requirements: while the lower-threshold trigger requires electrons to be well isolated, the higher-threshold trigger does not, which gives an improved efficiency at high $p_{\mathrm{T}}$. Similarly, the single-photon trigger avoids reliance on the online track reconstruction and increases the overall efficiency for electrons with $p_{\mathrm{T}} > 200$ GeV. Events containing at least one isolated muon are collected by triggers that require a muon with a minimum $p_{\mathrm{T}}$ of 50 GeV.

The candidate vertex with the largest value of summed physics-object $p_{\mathrm{T}}^2$ is taken to be the primary pp interaction vertex. The physics objects used for this determination are the jets, clustered using the jet finding algorithm [36, 37] with the tracks assigned to candidate vertices as inputs, as well as the remaining single tracks (including identified leptons), and the associated missing transverse momentum, taken as the negative vector sum of the $p_{\mathrm{T}}$ of those objects.

The global event reconstruction (also called particle-flow event reconstruction) [38] aims to reconstruct and identify each individual particle in an event, with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, electron, muon, charged and neutral hadron) plays an important role in the determination of the particle direction and energy. Photons (e.g., coming from $\pi^0$ decays or from electron bremsstrahlung) are identified as ECAL energy clusters not linked to the extrapolation of any charged particle trajectory to the ECAL. Each electron (e.g., coming from photon conversions in the tracker material or from B hadron semileptonic decays) is identified as a primary charged particle track and potentially many ECAL energy clusters corresponding to this track extrapolation to the ECAL and to possible bremsstrahlung photons emitted along the way through the tracker material. Muons (e.g., from B hadron semileptonic decays) are identified as tracks in the central tracker consistent with either a track or several hits in the muon system, and associated with calorimeter deposits compatible with the muon hypothesis. Charged hadrons are identified as charged particle tracks neither identified as electrons, nor as muons. Finally, neutral hadrons are identified as HCAL energy clusters not linked to any charged hadron trajectory, or as a combined ECAL and HCAL energy excess with respect to the expected charged hadron energy deposit.

The energy of each photon is obtained from the ECAL measurement. The energy of each electron is determined from a combination of the electron momentum at the primary interaction vertex 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 momentum of each muon is obtained from the trajectory of the corresponding track in the magnetic field. The energy of each charged hadron is determined from a combination of the 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 each neutral hadron is obtained from the corresponding corrected ECAL and HCAL

energies.

For each event, the jets are clustered from the reconstructed particles using the infrared and collinear safe anti-$k_T$ algorithm [36, 37] with a distance parameter of 0.4 (AK4 jets) or 0.8 (AK8 jets). The 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 5–10% of the true momentum over the whole $p_T$ spectrum and detector acceptance [39].

Pileup interactions can contribute additional tracks and calorimetric energy deposits to the reconstructed jets, increasing their apparent momentum. To mitigate this effect for AK4 jets, charged particles identified as originating from pileup vertices are discarded and an offset correction is applied to correct for the remaining contributions. Jet energy corrections are derived by comparing the average measured energies of jets in data with those of simulated jets. In situ measurements of the momentum balance in dijet, photon+jet, Z+jet, and multijet events are used to account for any residual differences in jet energy scale in data and simulation [39]. The typical jet energy resolution is 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV. Additional selection criteria are applied to each jet to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures [40]. For AK8 jets, the pileup-per-particle identification algorithm [41, 42] 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.

The momentum resolution for electrons with $p_T \approx 45$ GeV from $Z \to ee$ decays ranges from 1.7% to 4.5%. It is generally better in the barrel region than in the endcaps, and also depends on the bremsstrahlung energy emitted by the electron as it traverses the material in front of the ECAL [43].

Muons are measured in the range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive-plate chambers. The single-muon trigger efficiency exceeds 90% over the full $\eta$ range, and the subsequent efficiency to reconstruct and identify muons is greater than 96%. Matching muons to tracks measured in the silicon tracker results in a fractional $p_T$ resolution of 1% in the barrel and 3% in the endcaps for muons with $p_T$ up to 100 GeV, and of better than 7% in the barrel for muons with $p_T$ up to 1 TeV [44].

To reconstruct resolved $W_R$ candidates, events with two leptons and at least two AK4 jets are selected. Events with additional leptons are rejected, and if more than two jets exist, the two jets with the highest $p_T$ are used. The leading (subleading) lepton is required to have $p_T > 60(53)$ GeV and to be within the fiducial acceptance ($|\eta| < 2.4$). Electrons are rejected if the cluster lies in the range $1.44 < |\eta| < 1.57$, which corresponds to the transition region between the barrel and endcap sections of the ECAL, where the performance is degraded. To suppress muons originating from hadron decays or pion punch-through in jets, the $p_T$ sum of additional tracks that originate at the PV and are inside a cone of $\Delta R < 0.3$, where $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, is required to be less than 10% of the muon $p_T$. Electrons are also required to be isolated, i.e., the $p_T$ sum of all tracks in a cone of $\Delta R < 0.3$ centered on the electron candidate, not associated with the electron and originating from the PV, must be below 5 GeV. Dedicated identification algorithms, optimized for the selection of high-momentum leptons [43, 45] are used. The two jet candidates must each have $p_T > 40$ GeV and be within $|\eta| < 2.4$. To avoid overlaps between leptons and jets, they are required to be separated in $\Delta R$ by at least 0.4. An event is considered resolved if it contains four well-separated final-state objects as described above.

Events that fail the resolved selection criteria are used to reconstruct the boosted $W_R$ candidates. The $N \to \ell q\overline{q'}$ decay is reconstructed as a single AK8 jet. The boosted signature requires

an AK8 jet with $p_T > 200\,\text{GeV}$ and two leptons that have the same $p_T$ requirements as in the resolved search. The leading lepton is required to be separated in the azimuthal direction by $\Delta\phi > 2.0$ with respect to the AK8 jet. The subleading lepton, with the isolation requirement removed, is required to fall within the AK8 jet cone. The AK8 jet with the highest $p_T$ is chosen when multiple jets satisfy the conditions above, and events with additional isolated leptons are rejected. The jet is groomed [46] to remove soft and wide-angle radiation using the soft-drop algorithm [47, 48], with the soft radiation fraction parameter $z$ set to 0.1 and the angular exponent parameter $\beta$ set to 0. The groomed jet is used to compute the soft-drop jet mass ($m_{SD}$) and each AK8 jet is required to have $m_{SD} > 40\,\text{GeV}$, to optimize the signal selection and background rejection.

The lepton subjet fraction ($LSF_3$) algorithm [49] is used to determine the consistency of the jet with three subjets, where one subjet is dominated by the four-momentum of the lepton. This algorithm clusters the constituents of a jet into three subjets using the exclusive-$k_T$ algorithm, and $LSF_3$ is defined as the ratio of the lepton $p_T$ to its associated subjet $p_T$ ($p_T(\ell)/p_T(\text{subjet})$). Jets with higher $LSF_3$ values are considered to have a more isolated lepton within the jet and therefore the lepton is assumed to originate from a prompt decay. A selection of $LSF_3 > 0.75$ is required for all signal AK8 jets, which removes more than 81 (94)% of dielectron (dimuon) background events that have leptons originating from the decay of particles with non-negligible lifetimes (nonprompt leptons), while keeping 87 (92)% of signal events with $(m_{W_R}, m_N) = (5.0, 0.2)\,\text{TeV}$.

For the resolved events we reconstruct an invariant mass distribution from the two leptons and two AK4 jets ($m_{\ell\ell jj}$), and for the boosted events we consider the distribution in invariant mass of the lead lepton and AK8 jet ($m_{\ell J}$). We search for deviations from the expected SM background in these distributions in the mass regions with $m_{\ell\ell jj}$ or $m_{\ell J} > 0.8\,\text{TeV}$. To reduce the contribution from Z boson production, we also impose a requirement of $m_{\ell\ell} > 0.4\,(0.2)\,\text{TeV}$ in the resolved (boosted) search. For the signal samples with $m_{W_R} = 5.0\,\text{TeV}$ and $m_N = 0.2\,(3.0)\,\text{TeV}$, the product of acceptance and efficiency is 21–22 (49–50)% and 34–43 (54–67)% in the dielectron and dimuon boosted (resolved) signal regions (SRs), respectively, depending on the year.

# 5 Background estimation

The dominant SM processes that contribute to the background in each SR are DY production of lepton pairs with additional jets in the final state (DY+jets), leptonic decays of pair-produced top quarks, and single top production with an associated W boson. These backgrounds are estimated from simulation and the modeling is corrected using control regions (CRs), as described below. A schematic diagram presenting the CRs and SR is shown in Fig. 2.

The $t\bar{t}$ events with at least one hadronically decaying W boson, $s$- and $t$-channel production of a single top quark, and W boson production in association with jets can contribute to the SRs when nonprompt leptons from, e.g., semileptonic decays of B hadrons, are misidentified as signal-like leptons. These types of background are labeled as nonprompt, and are estimated by means of simulated samples. Multiboson and $t\bar{t}$ production in association with a gauge boson are rare SM processes that can also contribute to the SRs, and their contributions are obtained from simulation as well.

To correct the mismodeled $p_T$ distribution of the Z boson in the DY simulation [50, 51], we apply a $K$ factor as a function of the generator-level Z boson $p_T$ ("$p_T(Z)$ correction"). The NLO-to-LO $K$ factor in QCD is obtained by taking the ratio between DY samples simulated to NLO and LO precision. The uncertainties from renormalization and factorization scale, PDF

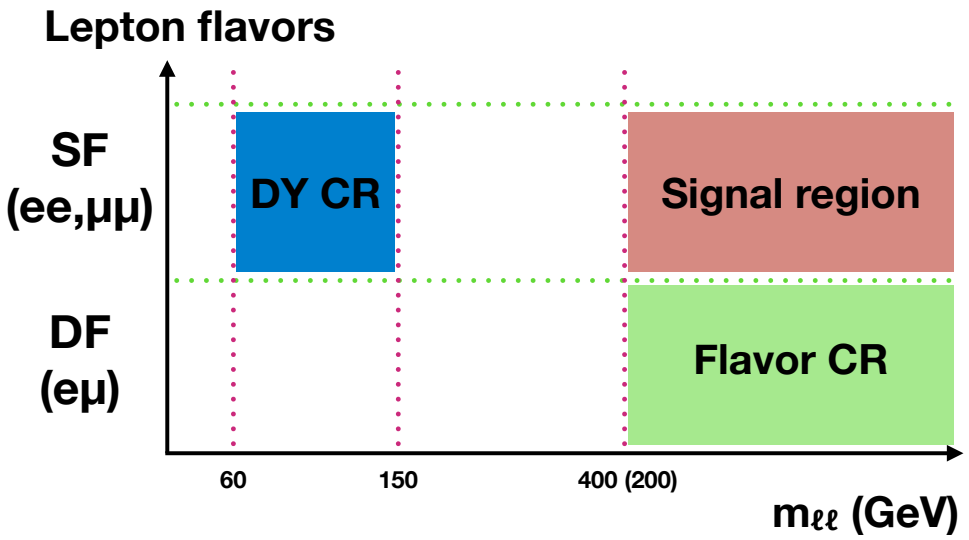


Figure 2: A schematic diagram of the analysis region. The minimum values of the dilepton mass in the SR and the flavor CR is 400 (200) GeV for the resolved (boosted) region. The backgrounds from tW and $t\bar{t}$ production are estimated from the flavor CR (green), where different flavor (DF) leptons are required. The DY background is estimated from the DY CR (blue).

variations, and the statistical uncertainties in the MC samples are taken into account. The nNLO electroweak correction (combination of the $\mathcal{O}(\alpha^2\alpha_S)$ contributions and the $\mathcal{O}(\alpha^3\alpha_S)$ electroweak Sudakov logarithms in the next-to-leading log approximation) is taken from Ref. [51]. We define DY CRs to be the same as the SR but modify the dilepton mass requirement to $60 < m_{\ell\ell} < 150$ GeV (blue region in Fig. 2). For the boosted DY CRs, a lepton is not required to be located inside an AK8 jet, and the $LSF_3$ requirement is removed when collecting the AK8 jets. Owing to a limitation of the DY MC generator, which is restricted to the LO simulation of events with a maximum of four partons in the hard scattering, a shape discrepancy is observed in the leading AK4 and AK8 jet $p_T$ distributions in the resolved and boosted DY CRs, respectively. We correct the shape of the jet $p_T$ distributions to match with those observed in data by a bin-by-bin rescaling ("DY reshape"). Since this mismodeling is due to the higher-order effect of QCD jet splitting, the dielectron and dimuon events will be affected in a similar way, thus the two event samples are summed when extracting the reweighting values, typically ranging from 0.8 to 1.2. The DY reshape correction modifies the shape of the $m_{\ell\ell jj}$ and $m_{\ell J}$ distributions and brings the distribution in simulation into good agreement with data. The uncertainty assigned to the DY reshape correction is dominated by the statistical uncertainty, and is treated as uncorrelated across bins in the invariant mass spectrum. The normalization of the DY processes is allowed to vary in the fitting procedure, which is described in Section 7. Several checks were performed to validate the extrapolation from low to high $m_{\ell\ell}$, including verifying the agreement in the intermediate region, $150 < m_{\ell\ell} < 400(200)$ GeV for the resolved (boosted) analysis.

To constrain the normalization of the $t\bar{t}$ and tW backgrounds in each SR, a corresponding control region (flavor CR) dominated by $t\bar{t}$ and tW events is defined by requiring the two leptons to have different flavors (the green region in Fig. 2). For the boosted analysis, separate control regions are used for the dielectron and dimuon searches. These require a lead muon plus a jet with an electron contained within it (e-jet), and a lead electron plus a jet with a muon contained within it ($\mu$-jet), respectively. The normalizations of these backgrounds are extracted by performing a simultaneous fit across the SRs and flavor CRs.

The invariant mass distributions in the DY CRs, flavor CRs, and the SRs are simultaneously fitted (Section 7). The $m_{\ell\ell jj}$ and $m_{\ell J}$ distributions after the fitting (post-fit) in the DY CRs and flavor CRs are shown in Figs. 3 and 4, respectively.

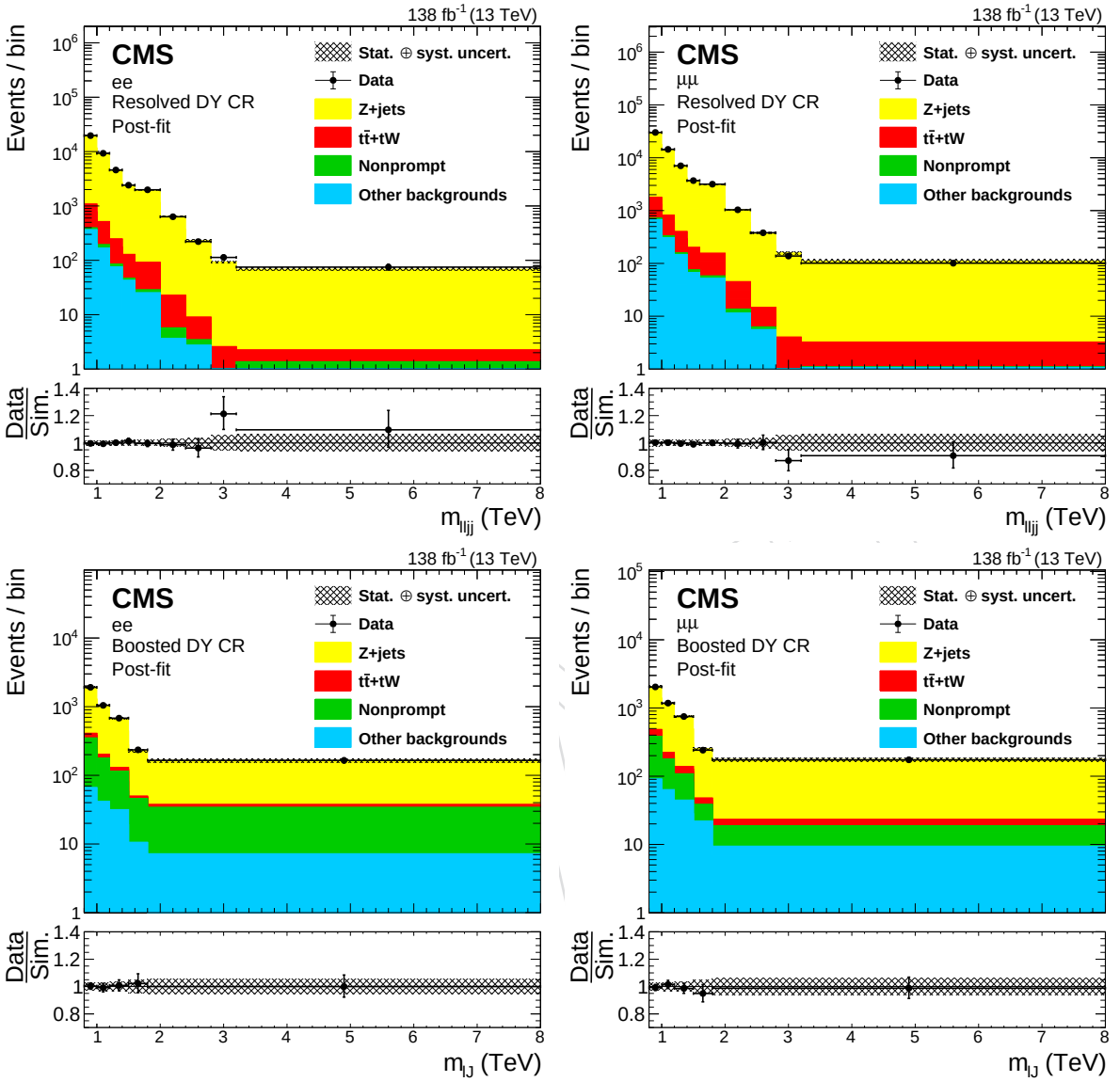


Figure 3: The $m_{\ell\ell jj}$ ($m_{\ell J}$) distributions in the resolved (boosted) DY CRs are shown in the upper (lower) row. Results in the ee ($\mu\mu$) channels are shown in the left (right) plots. The hatched uncertainty bands on the simulated background histograms include statistical and systematic components.

# 6 Systematic uncertainties

The $m_{\ell\ell jj}$ ($m_{\ell J}$) distribution is used to perform a binned maximum likelihood fit in the resolved (boosted) SR. The shapes of the invariant mass distributions are affected by a number of systematic uncertainties. Each systematic uncertainty is treated either as uncorrelated or fully correlated across the three data-taking years. A complete list of systematic uncertainties and their correlations is given in Table 1.

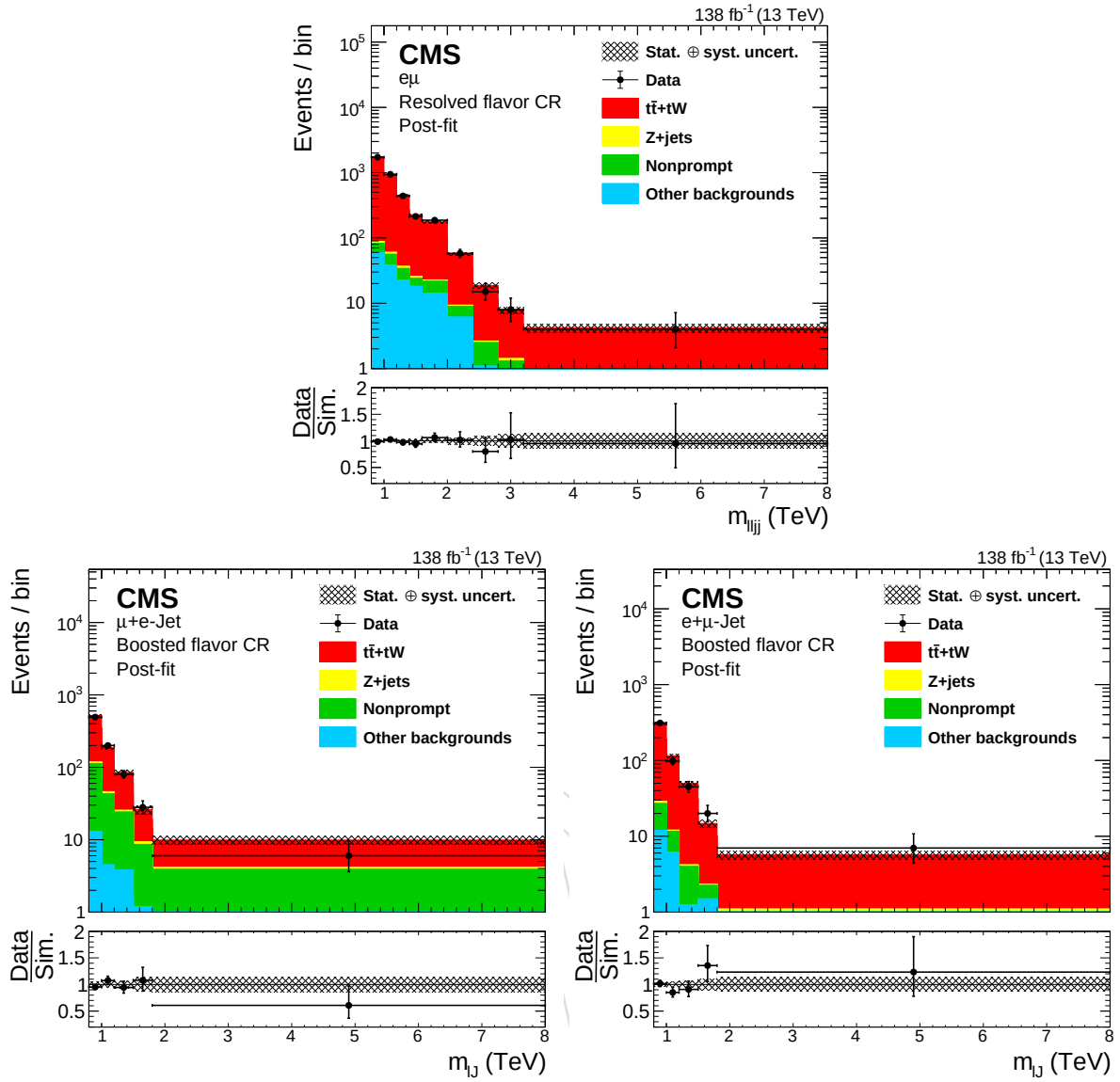


Figure 4: The reconstructed mass of the $W_R$ boson in the resolved (upper), boosted with e-jet (lower left), and boosted with $\mu$-jet (lower right) flavor CR.

The integrated luminosities for the 2016, 2017, and 2018 data-taking years have 1.2–2.5% individual uncertainties [52–54], while the overall uncertainty for the 2016–2018 period is 1.6%. The uncertainties in the expected signal and background yields associated with the pileup simulation are estimated by varying the total cross section of the inelastic scattering used in the simulation by $\pm$5%. The resultant uncertainties are 0.0–0.9% and 0.2–1.1% in signal and background yields, respectively.

The lepton trigger, reconstruction, identification, and isolation efficiencies are measured in both data and simulation using $Z \to \ell^{\pm}\ell^{\mp}$ ($\ell$ = e or $\mu$) events and data-to-simulation scale factors are applied to all simulation samples to correct any discrepancies. The uncertainties in the lepton scale factors are evaluated by changing the scale factors by $\pm$1 standard deviation from their nominal values. The momenta of leptons are varied in the simulation within their $\pm$1 standard deviation from their nominal value, and the differences in the mass distributions are taken as the uncertainties. The uncertainties in the jet energy scale and resolutions [39] affect the signal and background yields by less than 5%. The jet selection efficiency associated with

the requirement $\mathrm{LSF}_3 > 0.75$ is measured by taking a sample of AK8 jets from boosted W bosons in $\mathrm{t\bar{t}}$ events where one W boson decays leptonically and the other decays hadronically and injecting a simulated lepton (e or $\mu$) in the direction of the hadronically decaying W boson, to emulate the 3-prong jet topology. The data-to-simulation scale factor obtained for the $\mathrm{LSF}_3$ requirement, ranging from 0.95 to 1.05, is then applied to the simulated samples. The resulting uncertainty in the yields amounts to less than 10% for both the signal and background.

As described in Section 5, the two corrections applied to the DY simulation are varied within their uncertainties to estimate their impact on the invariant mass distributions.

Owing to the imperfect description of nonprompt leptons in the simulation, a conservative 100% normalization uncertainty, cross-checked with a dedicated control region with zero b jets, is assigned to the nonprompt background estimation. A 50% uncertainty is assigned to the rare SM background. The theoretical uncertainties originating from the strong coupling constant $\alpha_S$, PDFs, and renormalization/factorization scales are the dominant sources of uncertainties for the signal shape estimate. The PDF and $\alpha_S$ uncertainties are estimated from the standard deviation of the weights from the PDF replicas provided in the NNPDF3.1 PDF set [33], following the PDF4LHC procedure [55].

Table 1: Summary of the relative uncertainties in the total yields of the signal and background (bkgd) predictions. The uncertainties are given for the resolved (boosted) SR. The values for the signal correspond to $m_{\mathrm{W_R}} = 5\,\mathrm{TeV}$. The range given for each systematic uncertainty source covers the variation across the years and the year-to-year treatment is shown as either fully correlated (C) or uncorrelated (U). The EW1, EW2, and EW3 entries are the uncertainties from the $\mathcal{O}(\alpha^2)$ Sudakov terms, the NLO portion of the nNLO EW uncertainty, and the uncertainty in the Sudakov approximation at nNLO, respectively.

| Source | Process | Corr. | ee bkgd. (%) | ee signal (%) | $\mu\mu$ bkgd. (%) | $\mu\mu$ signal (%) |
|---|---|---|---|---|---|---|
| Integrated luminosity | All | C | 1.6 (1.6) | 1.6 (1.6) | 1.6 (1.6) | 1.6 (1.6) |
| Electron reconstruction | All | C | 1.0–1.6 (0.5–0.8) | 0.8–1.4 (0.4–0.8) | — | — |
| Electron energy resolution | All | C | <0.1 (<0.1) | <0.1 (<0.1) | — | — |
| Electron energy scale | All | C | 0.5–1.9 (0.5–2.4) | 0–0.2 (0–0.4) | — | — |
| Electron identification | All | C | 3.1–3.3 (1.8–1.9) | 4.1–4.4 (2.1–2.4) | — | — |
| Electron trigger | All | U | 0–0.1 (0.2–0.4) | <0.1 (0.1–0.2) | — | — |
| Muon reconstruction | All | C | — | — | 0.4–1.0 (0.3–0.7) | 4.4–36.8 (5.6–30.7) |
| Muon momentum scale | All | C | — | — | 0.4–2.5 (0.5–3.7) | 0.1–0.2 (0.1–0.3) |
| Muon identification | All | C | — | — | 0.2–1.2 (0.1–0.6) | 0.2–1.1 (0.1–0.5) |
| Muon isolation | All | C | — | — | 0.1–0.2 (0–0.1) | 0.1–0.2 (0–0.1) |
| Muon trigger | All | U | — | — | 0.1–0.2 (0.1–0.2) | 0.7–1.6 (0.5–1.3) |
| Jet energy scale | All | C | 1.9–4.2 (0.9–1.9) | 0–0.2 (0–0.2) | 2.1–3.5 (0.6–0.9) | 0–0.2 (0–0.4) |
| Jet energy resolution | All | U | 0.5–1.5 (0.7–2.0) | 0–0.3 (0–0.4) | 0.2–1.3 (0.2–1.1) | 0–0.3 (0–0.3) |
| Jet mass scale | All | C | <0.1 (1.0–1.8) | <0.1 (0.1–1.0) | <0.1 (1.4–1.7) | <0.1 (0.2–1.0) |
| LSF scale factor | All | U | — (6.7–8.7) | — (6.7–8.7) | — (5.8–7.1) | — (5.8–7.1) |
| Pileup modeling | All | C | 0.2–1.1 (0.5–1.2) | 0.1–0.8 (0.1–0.9) | 0.3–0.5 (0.2–1.2) | 0.1–0.5 (0–0.6) |
| $p_T$ (Z) correction, MC stat. | DY+jets | C | 0.6–2.0 (0.6–2.0) | — | 0.6–2.0 (0.6–2.0) | — |
| $p_T$ (Z) correction, renorm./fact. scales | DY+jets | C | 6.3–7.1 (6.7–7.3) | — | 6.1–7.1 (6.8–7.4) | — |
| $p_T$ (Z) correction, PDF replicas | DY+jets | C | 0.6–2.0 (0.6–2.0) | — | 0.6–2.0 (0.6–2.0) | — |
| $p_T$ (Z) correction, PDF $\alpha_S$ | DY+jets | C | 0.8–4.7 (0.8–5.2) | — | 0.8–4.7 (0.8–5.3) | — |
| $p_T$ (Z) correction, EW1 [51] | DY+jets | C | <0.1 (<0.1) | — | <0.1 (<0.1) | — |
| $p_T$ (Z) correction, EW2 [51] | DY+jets | C | 0.3 (0.3) | — | 0.3 (0.3) | — |
| $p_T$ (Z) correction, EW3 [51] | DY+jets | C | 0.1 (0.1) | — | 0.1 (0–0.1) | — |
| DY reshape | DY+jets | C | 8.5–9.1 (10.1–11.6) | — | 8.5–9.2 (9.7–11.6) | — |
| Nonprompt background normalization | Nonprompt | U | 100 (100) | — | 100 (100) | — |
| Rare SM background normalization | Others | C | 50 (50) | — | 50 (50) | — |
| PDF replicas | Signal | C | — | 5.9–11.9 (8.8–40.3) | — | 2.8–6.8 (17.5–40.6) |
| $\alpha_S$ | Signal | C | — | 0–0.2 (0.2–1.3) | — | 0–0.2 (0.2–1.2) |
| Renorm./fact. scales | Signal | C | — | 0–0.1 (0.3–2.3) | — | 0–0.1 (2.1–2.9) |

# 7 Results

A maximum-likelihood fit is performed on the $m_{\ell\ell\mathrm{jj}}$ ($m_{\ell\mathrm{J}}$) distributions in the resolved (boosted) SRs and CRs, with the systematic uncertainties described in Section 6 treated as nuisance parameters. The background-only post-fit invariant mass distributions are shown in Fig. 5 and no significant excess over the background expectations is observed. A simulated signal distribution corresponding to mass values at the limit of sensitivity is shown for comparison. The shape of this distribution is very similar for the resolved and the boosted analysis, with the peak falling in the last bin. The enhancement visible around 2 TeV in the resolved distribution is due to sculpting of the $\mathrm{W_R}$ off-shell mass distribution. The upper limits on the product of the cross section for $\mathrm{W_R}$ production and the branching fractions $\sigma(\mathrm{pp} \to \mathrm{W_R})\mathcal{B}(\mathrm{W_R} \to \mathrm{ee}(\mu\mu)\mathrm{q}\overline{\mathrm{q}'})$ for various $m_{\mathrm{W_R}}$ and $m_{\mathrm{N}}$ hypotheses are obtained using the distributions of the likelihood ratio calculated using the asymptotic approximation [56] and the $\mathrm{CL_s}$ criterion [57, 58]. The excluded phase space as a function of $m_{\mathrm{W_R}}$ obtained from the expected and observed upper limits at 95% confidence level (CL) is shown in Fig. 6. With $m_{\mathrm{N}} = m_{\mathrm{W_R}}/2$, the observed (expected) lower limit at 95% CL on the mass of the $\mathrm{W_R}$ is 4.7 (5.2) TeV and 5.0 (5.2) TeV for the electron and muon channels, respectively; for $m_{\mathrm{N}} = 0.2$ TeV, limits exclude the phase space up to $m_{\mathrm{W_R}} = 4.8\,(5.0)$ and 5.4 (5.3) TeV for the electron and muon channels. The local $p$-value of the signal strength, as a function of $m_{\mathrm{W_R}}$ and $m_{\mathrm{N}}$, is obtained from fits to the data with the signal stength at each point treated as a free parameter. We observe the most extreme $p$-value to occur in the electron channel for the $(m_{\mathrm{W_R}}, m_{\mathrm{N}}) = (6.0, 0.8)$ TeV mass point and to be $1.58 \times 10^{-3}$, corresponding to a local significance of $2.95\sigma$. The look-elsewhere effect [59] is taken into account by using pseudo-experiments to calculate the probability under the background-only hypothesis of observing a similar or larger excess in the electron channel across the full analysis mass range. This probability is $2.7 \times 10^{-3}$, corresponding to a global significance of $2.78\sigma$. The upper limits across the entire $(m_{\mathrm{W_R}}, m_{\mathrm{N}})$ plane are shown in Fig. 7. The results are compared with previous searches for $\mathrm{W_R}$ bosons and heavy neutrinos performed by the ATLAS Collaboration [9] using LHC pp collision data at $\sqrt{s} = 13$ TeV and the excluded regions from this analysis in $m_{\mathrm{W_R}}$, with $m_{\mathrm{N}} = m_{\mathrm{W_R}}/2$, match these previous results in the electron channel and extend them by 0.3 TeV in the muon channel. Furthermore, for values of $m_{\mathrm{N}}$ less than 0.5 TeV, the introduction of the boosted analysis provides a significant increase in sensitivity, extending the range of $m_{\mathrm{W_R}}$ excluded in this region beyond 4.8 and 5.0 TeV in the electron and muon channels, respectively.

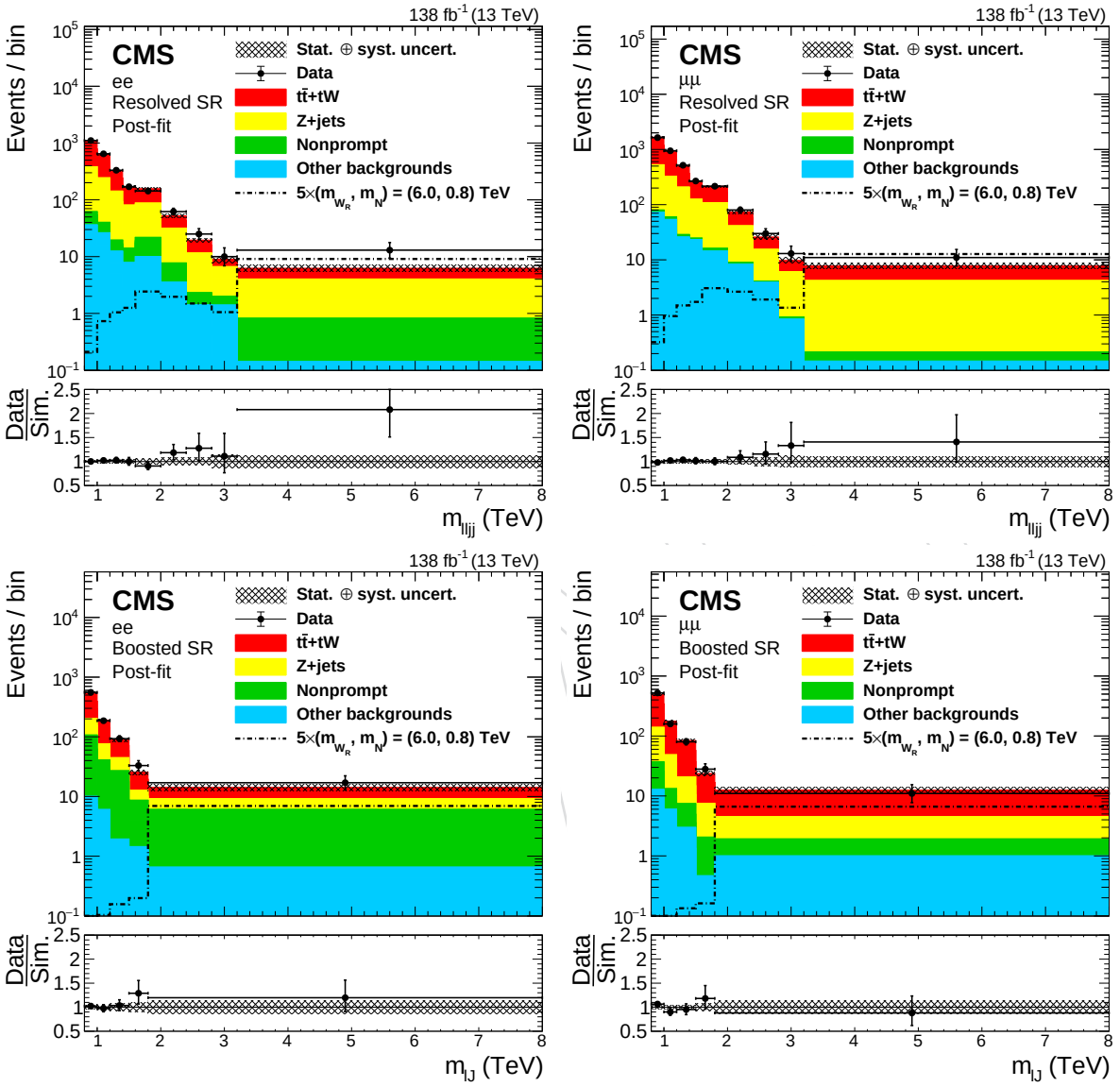


Figure 5: The background-only post-fit $m_{\ell\ell \mathrm{jj}}$ ($m_{\ell \mathrm{J}}$) distributions in the resolved (boosted) SR are shown in the upper (lower) plot. Results for the dielectron (dimuon) channel are shown on the left (right). Statistical and systematic uncertainties in the expected background yields are represented by the hatched band. The simulated signal distribution is scaled up by a factor of five to enhance visibility.

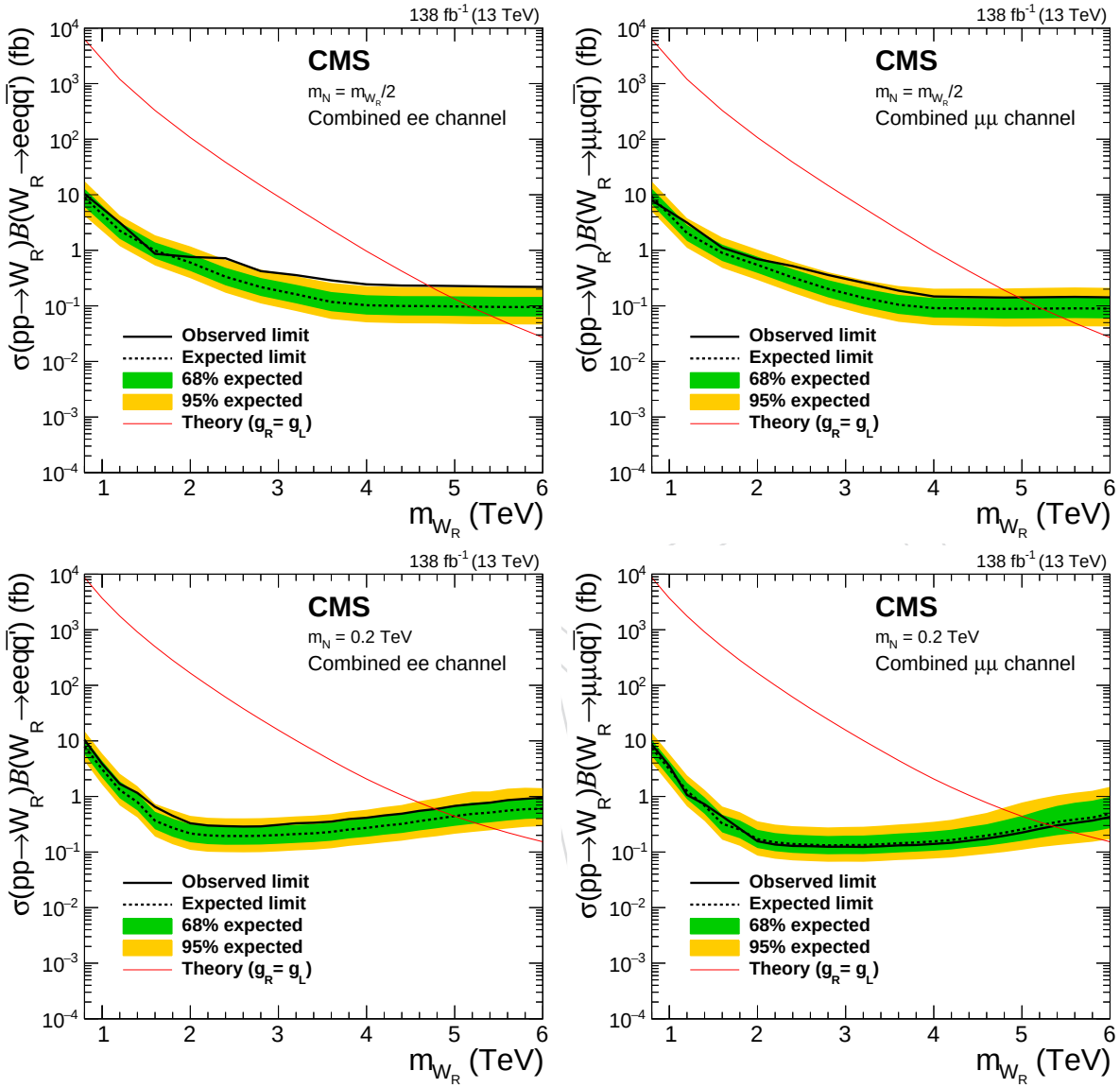


Figure 6: The expected (black dashed line) and the observed (black solid line) 95% CL upper limits on the product of the cross section for $W_R$ production and the branching fractions for the electron channel (left) and muon channel (right) from the combination of the resolved and boosted categories. The plots in the upper (lower) row are the results for the $m_N = m_{W_R}/2$ ($m_N = 0.2\,\text{TeV}$) mass point, which correspond to the resolved (boosted) $W_R$ topology. The green (inner) and yellow (outer) bands indicate the 68 and 95% coverage of the expected upper limits. The red solid lines represent the values expected from the theory [17].

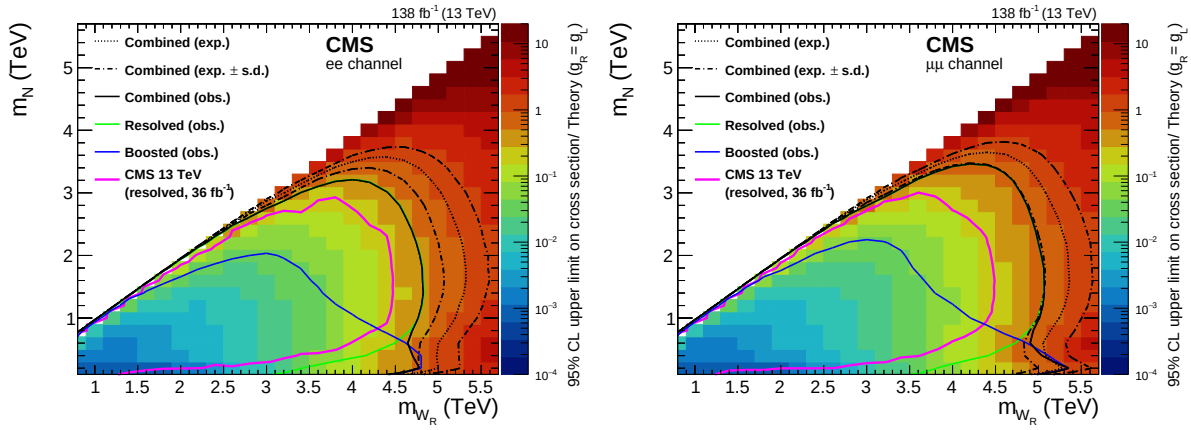


Figure 7: The observed 95% CL upper limits on the product of the production cross sections and the branching fractions of a right-handed $W_R$ boson divided by the theory expectation for a coupling constant $g_R$ equal to the SM coupling of the $W_R$ boson ($g_L$), for the electron channel (left) and muon channel (right). The observed exclusion regions are shown for the resolved (solid green), boosted (solid blue), and combined (solid black) channels, together with the expected exclusion region for the combined result (dotted black). The dash-dotted lines represent the 68% coverage of the boundaries of the expected exclusion regions. The observed exclusion regions obtained in the previous search performed by the CMS Collaboration [12] are bounded by the magenta lines. The biggest improvement can be seen in the $m_N < 0.5$ TeV region, where the new boosted category greatly improves the sensitivity over the previous result.

# 8 Summary

A search for right-handed bosons ($W_R$) and heavy right-handed neutrinos (N) in the left-right symmetric extension of the standard model has been presented. The analysis is based on proton-proton collision data collected at $\sqrt{s} = 13$ TeV by the CMS detector, corresponding to an integrated luminosity of 138 fb$^{-1}$. The final state consists of events with two same-flavor leptons (ee or $\mu\mu$) and two quarks, and is identified through two regions: the resolved region, where all four objects are well separated, and the boosted region, where the heavy neutrino decay is identified using jet substructure techniques applied to large-area jets. The addition of the boosted region greatly improves the search sensitivity in the region where $m_N < 0.5$ TeV. No significant excess over the standard model background expectations is observed in the invariant mass distributions. Upper limits are set on the products of the $W_R$ and N production cross sections and their branching fraction to two leptons and two quarks assuming that couplings are identical to those of the standard model. For N masses $m_N$ equal to half the $W_R$ mass $m_{W_R}$ ($m_N = 0.2$ TeV), $m_{W_R}$ is excluded at 95% confidence level up to 4.7 (4.8) and 5.0 (5.4) TeV for the electron and muon channels, respectively. This analysis provides the most stringent limits on the $W_R$ mass to date.

# Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid and

other centers for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC, the CMS detector, and the supporting computing infrastructure provided by the following funding agencies: BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES and BNSF (Bulgaria); CERN; CAS, MoST, and NSFC (China); MINCIENCIAS (Colombia); MSES and CSF (Croatia); RIF (Cyprus); SENESCYT (Ecuador); MoER, ERC PUT and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRI (Greece); NKFIA (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR, and NRC KI (Russia); MESTD (Serbia); MCIN/AE and PCTI (Spain); MOSTR (Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract Nos. 675440, 724704, 752730, 758316, 765710, 824093, 884104, and COST Action CA16108 (European Union); the Leventis Foundation; the Alfred P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the F.R.S.-FNRS and FWO (Belgium) under the "Excellence of Science – EOS" – be.h project n. 30820817; the Beijing Municipal Science & Technology Commission, No. Z191100007219010; the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Deutsche Forschungsgemeinschaft (DFG), under Germany's Excellence Strategy – EXC 2121 "Quantum Universe" – 390833306, and under project number 400140256 - GRK2497; the Lendület ("Momentum") Program and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the New National Excellence Program ÚNKP, the NKFIA research grants 123842, 123959, 124845, 124850, 125105, 128713, 128786, and 129058 (Hungary); the Council of Science and Industrial Research, India; the Latvian Council of Science; the Ministry of Science and Higher Education and the National Science Center, contracts Opus 2014/15/B/ST2/03998 and 2015/19/B/ST2/02861 (Poland); the Fundação para a Ciência e a Tecnologia, grant CEECIND/01334/2018 (Portugal); the National Priorities Research Program by Qatar National Research Fund; the Ministry of Science and Higher Education, projects no. 14.W03.31.0026 and no. FSWW-2020-0008, and the Russian Foundation for Basic Research, project No.19-42-703014 (Russia); MCIN/AEI/10.13039/501100011033, ERDF "a way of making Europe", and the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia María de Maeztu, grant MDM-2017-0765 and Programa Severo Ochoa del Principado de Asturias (Spain); the Stavros Niarchos Foundation (Greece); the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Kavli Foundation; the Nvidia Corporation; the SuperMicro Corporation; the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

# References

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

[2] R. N. Mohapatra and J. C. Pati, "A natural left-right symmetry", *Phys. Rev. D* **11** (1975) 2558, doi:10.1103/PhysRevD.11.2558.

[3] G. Senjanović and R. N. Mohapatra, "Exact left-right symmetry and spontaneous violation of parity", *Phys. Rev. D* **12** (1975) 1502, doi:10.1103/PhysRevD.12.1502.

[4] W.-Y. Keung and G. Senjanović, "Majorana neutrinos and the production of the right-handed charged gauge boson", *Phys. Rev. Lett.* **50** (1983) 1427, doi:10.1103/PhysRevLett.50.1427.

[5] R. N. Mohapatra and G. Senjanović, "Neutrino mass and spontaneous parity nonconservation", *Phys. Rev. Lett.* **44** (1980) 912, doi:10.1103/PhysRevLett.44.912.

[6] A. Das, P. S. B. Dev, and R. N. Mohapatra, "Same sign versus opposite sign dileptons as a probe of low scale seesaw mechanisms", *Phys. Rev. D* **97** (2018) 015018, doi:10.1103/PhysRevD.97.015018, arXiv:1709.06553.

[7] M. Gell-Mann, P. Ramond, and R. Slansky, "Complex spinors and unified theories", *Conf. Proc. C* **790927** (1979) 315, doi:10.1142/9789812836854_0018, arXiv:1306.4669.

[8] ATLAS Collaboration, "Search for heavy Majorana neutrinos with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* **07** (2015) 162, doi:10.1007/JHEP07(2015)162, arXiv:1506.06020.

[9] ATLAS Collaboration, "Search for heavy Majorana or Dirac neutrinos and right-handed W gauge bosons in final states with two charged leptons and two jets at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **01** (2019) 016, doi:10.1007/JHEP01(2019)016, arXiv:1809.11105.

[10] ATLAS Collaboration, "Search for a right-handed gauge boson decaying into a high-momentum heavy neutrino and a charged lepton in $pp$ collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **798** (2019) 134942, doi:10.1016/j.physletb.2019.134942, arXiv:1904.12679.

[11] CMS Collaboration, "Search for heavy neutrinos and W bosons with right-handed couplings in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J. C* **74** (2014) 3149, doi:10.1140/epjc/s10052-014-3149-z, arXiv:1407.3683.

[12] CMS Collaboration, "Search for a heavy right-handed W boson and a heavy neutrino in events with two same-flavor leptons and two jets at $\sqrt{s} = 13$ TeV", *JHEP* **05** (2018) 148, doi:10.1007/JHEP05(2018)148, arXiv:1803.11116.

[13] CMS Collaboration, "Search for heavy Majorana neutrinos in same-sign dilepton channels in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **01** (2019) 122, doi:10.1007/JHEP01(2019)122, arXiv:1806.10905.

[14] CMS Collaboration, "Search for heavy neutrinos or third-generation leptoquarks in final states with two hadronically decaying $\tau$ leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2017) 077, doi:10.1007/JHEP03(2017)077, arXiv:1612.01190.

[15] CMS Collaboration, "Search for third-generation scalar leptoquarks and heavy right-handed neutrinos in final states with two tau leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2017) 121, doi:10.1007/JHEP07(2017)121, arXiv:1703.03995.

[16] A. Ferrari et al., "Sensitivity study for new gauge bosons and right-handed Majorana neutrinos in pp collisions at $\sqrt{s} = 14$ TeV", *Phys. Rev. D* **62** (2000) 013001, doi:10.1103/PhysRevD.62.013001.

[17] O. Mattelaer, M. Mitra, and R. Ruiz, "Automated neutrino jet and top jet predictions at next-to-leading-order with parton shower matching in effective left-right symmetric models", 2016. arXiv:1610.08985.

[18] M. Mitra, R. Ruiz, D. J. Scott, and M. Spannowsky, "Neutrino jets from high-mass $W_R$ gauge bosons in TeV-scale left-right symmetric models", *Phys. Rev. D* **94** (2016) 095016, doi:10.1103/PhysRevD.94.095016, arXiv:1607.03504.

[19] "HEPData record for this analysis", 2021. doi:10.17182/hepdata.114866.

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

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

[22] 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.

[23] 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.

[24] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[25] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[26] 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.

[27] 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].

[28] S. Frixione, G. Ridolfi, and P. Nason, "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[29] 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.

[30] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[31] 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.

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

[33] 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.

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

[35] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, arXiv:1802.02613.

[36] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-$k_{\mathrm{T}}$ jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[37] 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.

[38] 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.

[39] 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.

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

[41] 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.

[42] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/p09018, arXiv:2003.00503.

[43] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[44] 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.

[45] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[46] G. P. Salam, "Towards jetography", *Eur. Phys. J. C* **67** (2010) 637, doi:10.1140/epjc/s10052-010-1314-6, arXiv:0906.1833.

[47] 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.

[48] 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.

[49] C. Brust et al., "Identifying boosted new physics with non-isolated leptons", *JHEP* **04** (2015) 079, doi:10.1007/JHEP04(2015)079, arXiv:1410.0362.

[50] CMS Collaboration, "Measurements of differential Z boson production cross sections in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **12** (2019) 061, doi:10.1007/JHEP12(2019)061, arXiv:1909.04133.

[51] 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.

[52] CMS Collaboration, "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS", *Eur. Phys. J. C* **81** (2021) 800, doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.

[53] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[54] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2017.

[55] 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.

[56] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[57] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[58] A. L. Read, "Presentation of search results: the $CL_s$ technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[59] E. Gross and O. Vitells, "Trial factors for the look elsewhere effect in high energy physics", *Eur. Phys. J. C* **70** (2010) 525, doi:10.1140/epjc/s10052-010-1470-8, arXiv:1005.1891.