CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2020/02/02
Archive Hash: 5b132ae
Archive Date: 2020/02/02

Search for the electroweak production of two Z bosons in the 4ℓjj final state in pp collisions at $\sqrt{s} = 13$ TeV

Y. An[1], R. Bellan[2], C. Charlot[3], R. Covarelli[2], D. Giljanovic[3,4], H. He[5], D. Lelas[4], C. Li[1], Q. Li[1], and A. Savin[5]

[1] Peking University.
[2] Università degli Studi di Torino and INFN.
[3] Laboratoire Leprince-Ringuet.
[4] University of Split.
[5] University of Wisconsin, Madison.

Abstract

A search for the electroweak production of of two Z bosons in association with two jets in pp collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV is documented in this note. The analysis uses the full Run2 data set collected with the CMS experiment at the LHC, corresponding to an integrated luminosity of 137 fb^{-1}. Leptonic decays of the Z bosons are considered, ZZ $\rightarrow \ell\ell\ell'\ell'$, where $\ell = e, \mu$. The significance of the electroweak signal is XY standard deviations (XY expected). The fiducial cross section of the electroweak production is measured to be XY, in agreement with the Standard Model prediction of XY. Limits on anomalous quartic couplings are derived in terms of the effective field theory operators T8 and T9.

Contents

1 Introduction

In this note we present a search for the electroweak production of two Z bosons produced in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV.

The search is carried out in the fully-leptonic final state $ZZ \to 4\ell$, where $\ell = e, \mu$ and the full Run2 dataset of approximately 137 fb^{-1} is used. The signal is extracted by means of discriminants based on the matrix-element description by the MFCM calculator at the leading-order of electroweak and QCD production processes.

The analysis strategy follows closely the CMS analysis of the same final state performed with 2016 data only [1, 2]. The object and ZZ candidate selections used in this analyis are those of the $H \to ZZ \to 4\ell$ analysis [3, 4], with only minor changes to better describe the two-jet phase space.

This note is organized as follows: Section 2 summarizes the dataset as well as all Monte Carlo samples used for this analysis, Section 3 describes the object and event selection, while Section 6 deals with the background estimation and the related uncertainties. Section 7 lists the several sources of systematic uncertainty. Section 9 presents the results of the VBS search, the measurement of the electroweak production fiducial cross section as well as limits on anomalous quartic couplings.

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 [5]; the key components for this analysis are summarized here. The CMS experiment is characterized by a superconducting solenoid located in the central region of the detector, providing an axial magnetic field of 3.8 T parallel to the beam direction. A silicon pixel and strip tracker, a crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter are located within the solenoid and cover the absolute pseudorapidity range $|\eta| < 3.0$, where pseudorapidity is defined as $\eta = -ln[\tan(\theta/2)]$. The Forward Hadronic calorimeters (HF) are placed outside the magnet yoke, 11 m far from the interaction point, extending the pseudorapidity coverage up to $|\eta| = 5$.

2 Data sets and Monte Carlo samples

2.1 CMS data

This analysis uses a data sample recorded by the CMS experiment during 2016, 2017 and 2018 corresponding to $\mathcal{L} = 137$ fb^{-1} of data.

Only data that passed the quality certification by all detector subsystems is used in the analysis and only luminosity sections included in the respective golden JSONs are used for further analysis. The MINIAOD format is chosen to perform the analysis. The HLT paths used for 2016, 2017 and 2018 collisions data are reported in Table **??** [3, 4].

The used data sets are listed in Table 4.The analysis relies on five different primary datasets (PD), *DoubleEG*, *DoubleMu*, *MuonEG*, *SingleElectron*, and *SingleMuon*, each of which combines a certain collections of HLT paths, whose exact requirements depend on the year of data taking. *DoubleEG* and *SingleElectron* are merged into *EGamma* in 2018. To avoid duplicate events from different primary datasets, events are taken:

- from DoubleEG if they pass the diEle (`HLT_EleXX_EleYY_CaloIdXX_TrackIdXX_IsoXX(_DZ)`) or triEle triggers (`HLT_EleXX_EleYY_EleZZ_CaloIdXX_TrackIdXX`) where XX,

HLT path	prescale	primary dataset
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	1	DoubleEG
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	1	DoubleEG
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL	1	DoubleEG
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL	1	DoubleMuon
HLT_TripleMu_12_10_5	1	DoubleMuon
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL	1	MuonEG
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL	1	MuonEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL	1	MuonEG
HLT_Ele25_eta2p1_WPTight	1	SingleElectron
HLT_Ele27_WPTight	1	SingleElectron
HLT_Ele27_eta2p1_WPLoose_Gsf	1	SingleElectron
HLT_IsoMu20 OR HLT_IsoTkMu20	1	SingleMuon
HLT_IsoMu22 OR HLT_IsoTkMu22	1	SingleMuon

Table 1: Trigger paths used in 2016 collision data.

HLT path	prescale	primary dataset
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_*	1	DoubleEG
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL	1	DoubleEG
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8	1	DoubleMuon
HLT_TripleMu_12_10_5	1	DoubleMuon
HLT_TripleMu_10_5_5_D2	1	DoubleMuon
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ	1	MuonEG
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	1	MuonEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_DZ	1	MuonEG
HLT_Ele35_WPTight_Gsf_v*	1	SingleElectron
HLT_Ele38_WPTight_Gsf_v*	1	SingleElectron
HLT_Ele40_WPTight_Gsf_v*	1	SingleElectron
HLT_IsoMu27	1	SingleMuon

Table 2: Trigger paths used in 2017 collision data.

YY and ZZ are year-dependent thresholds;

- from DoubleMuon if they pass the diMuon (HLT_MuXX_TrkIsoVVL_MuYY_TrkIsoVVL) or triMuon (HLT_TripleMu_XX_YY_ZZ) triggers and fail the diEle and triEle triggers,
- from MuEG if they pass the MuEle (HLT_MuXX_TrkIsoXX_EleYY_CaloIdYY_TrackIdYY_IsoYY) or MuDiEle (HLT_MuXX_DiEleYY_CaloIdYY_TrackIdYY) or DiMuEle (HLT_DiMuXX_EleYY_CaloIdYY_TrackIdYY) triggers and fail the diEle, triEle, diMuon and triMuon triggers,
- from SingleElectron if they pass the singleElectron trigger (HLT_EleXX_etaXX_WPLoose/Tight(_Gsf)) and fail all the above triggers.
- from SingleMuon if they pass the singleMuon trigger (HLT_IsoMuXX OR HLT_IsoTkMuXX) and fail all the above triggers.

2.2 Simulation

Several Monte Carlo (MC) event generators are used to simulate the signal and background contributions. The MC samples are employed to optimize the event selection, evaluate the signal efficiency and acceptance, and optimise the search strategy for Vector Boson Scattering. Most of these samples differ from the choice in the Higgs analysis because the samples listed below are better suited for the 2-jet phase space, rather than an inclusive 4ℓ analysis.

The signal for this analysis is the purely electroweak production of two jets and two leptonically decaying Z bosons. The hard process of the signal is simulated at leading order (LO) using

HLT path	prescale	primary dataset
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	DoubleEG
HLT_DoubleEle25_CaloIdL_MW_v*	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v*	1	DoubleMuon
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ_v*	1	MuonEG
HLT_Ele32_WPTight_Gsf_v*	1	SingleElectron
HLT_IsoMu24_v*	1	SingleMuon

Table 3: Trigger paths used in 2018 collision data.

Data stream	**Run and reconstruction version**
DoubleMuon DoubleEG MuonEG SingleMuon SingleElectron	Run2016B-17Jul2018-v1 Run2016C-17Jul2018-v1 Run2016D-17Jul2018-v1 Run2016E-17Jul2018-v1 Run2016F-17Jul2018-v1 Run2016G-17Jul2018-v1 Run2016H-17Jul2018-v1
DoubleMuon DoubleEG MuonEG SingleMuon SingleElectron	Run2017B-31Mar2018-v1 Run2017C-31Mar2018-v1 Run2017D-31Mar2018-v1 Run2017E-31Mar2018-v1 Run2017F-31Mar2018-v1
DoubleMuon MuonEG SingleMuon EGamma	Run2018A-17Sep2018-v1 Run2018B-17Sep2018-v1 Run2018C-17Sep2018-v1 Run2018D-PromptReco-v2

Table 4: List of data samples used in the analysis. All runs for each of the 5 data streams are used, for a total of 76 primary datasets in the MINIAOD format.

MadGraph5_aMCatNLO 2.4.2 [6] by explicitly reducing the number of allowed QCD vertices to zero:

```
generate p p > z z j j QCD=0, z > l+ l-
```

The last decay is limited to electrons and muons and it is performed by MadSpin, in order to preserve the spin correlations between the leptons. This includes diagrams featuring the Standard Model Higgs boson ($m_H = 125$ GeV) produced in vector boson fusion as well as the interference with the non-Higgs diagrams. The motivation of using the decay chain syntax is twofold, as it allows to speed up the generation and to populate the phase-space probed by the analysis. The latter is crucial to ensure sufficient statistics. The resulting set of diagrams includes triboson production with one hadronic W/Z decay, a contribution that is suppressed by requiring a minimum di-jet invariant mass of 100 GeV. A second signal sample generated with Phantom [7] is used to validate the MadGraph prediction.

The irreducible QCD-induced $pp \to ZZ$ processes are produced at next-to-leading-order (NLO) with up to 2 extra parton emission with MadGraph5_aMCatNLO [6], and merged using the FxFx scheme. This sample was specifically developed and requested for the ZZjj analysis; it will be the nominal sample in the statistical analysis.

```
generate p p > l+ l- l+ l- [QCD] @0
add process p p > l+ l- l+ l- j [QCD] @1
```

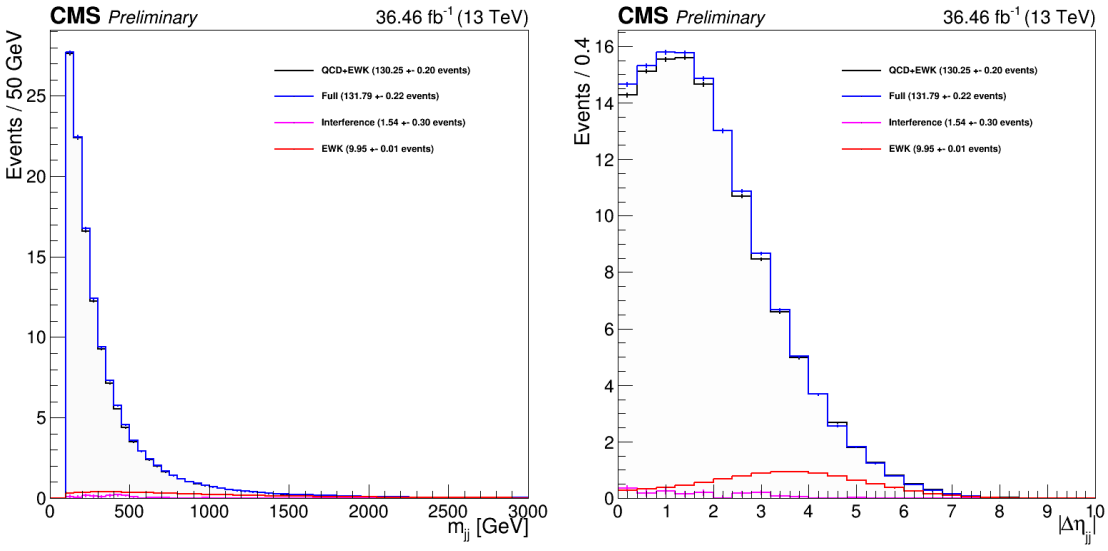
The interference between the signal and the QCD background is evaluated by generating dedicated samples (only done in 2016) that include the electroweak and QCD as well as their interference:

```
QCD only: generate p p > z z j j QCD=2 QED=2, z > l+ l-
QCD+EWK: generate p p > z z j j QCD=2 QED=4, z > l+ l-
```

The cross-sections reported by the generator with 2016 settings are reported in Table 5. It can be seen that the interference is positive and amounts to about 0.0426 fb or 10% of the electroweak signal. The opposing kinematics of the QCD and electroweak production cause the interference to be concentrated in the same phase-space as the QCD background, as is shown in the tagging-jet mass m_{jj} and $|\Delta\eta_{jj}|$ distribution in Fig. 1. Because of its small size and the background-like kinematics, the interference is neglected.

Table 5: Cross sections of the electroweak and QCD-induced production of the $4\ell jj$ final state and the interference. The phase space is that of the generation, i.e. $m_{jj} > 100$ GeV and includes the branching ratios for the Z decays to electrons or muons.

σ_{QCD}	$\sigma_{electroweak}$	$\sigma_{sum} = \sigma_{QCD} + \sigma_{electroweak}$	σ_{full}	$\sigma_{full} - \sigma_{sum}$
9.335 fb	0.4404 fb	9.7754 fb	9.818 fb	0.0426 fb



(a) (b)

Figure 1: Dijet invariant mass (left) and $|\Delta\eta|$ separation (right) distributions at GEN level for the electroweak signal, the QCD background and the interference between the two.

Aside from the dominant QCD background mediated by the tree-level processes, there is also a gluon loop-induced ZZ production process, which is a NNLO diagram and therefore is not included in the nominal QCD sample. Though suppressed by the two additional strong couplings, it nevertheless contributes to inclusive ZZ production at the 10% level.

This process is simulated at leading order (LO) with 1 jet using `MadGraph5_aMCatNLO 2.4.2` [6], by explicitly requesting a loop-induced process:

```
generate g g > z z j [noborn=QCD], z > l+ l-
```

The second jet in the event and the last decay are both simulated by `PYTHIA`, since the processing time becomes too long otherwise. Both these effects reduce the quality of the simulation, because the Z natural decay width and lepton correlations are not considered, and the second jet is produced via parton showering. Therefore the process is validated with a different MC sample generated at LO inclusive with `MCFM 7.0` [8]. The modelling of this process is discussed in more detail in the following.

A new gluon loop-induced ZZ sample simulated by `MadGraph5_aMCatNLO` [6] is also under studied, with up to 2 jets modeled from matrix-element and the MLM matching [9] implemented for the first time. The phase-space of dijet produced from the loop-induced process is expected to be more accurately modeled with this sample. The simulation details and kinematics study are described in Appendix B.

The list of MC samples and their cross sections are shown in Table 6. All cross sections used in the analysis are those returned by the generator and reported in Table 6, with no additional k-factors being used.

The `PYTHIA 8` [10] package is used for parton showering, hadronization, and the underlying event simulation for all MC samples. All samples are generated with the NNPDF 3.0 (in 2016) or 3.1 (2017-18) parton distribution functions (PDFs) [11].

The MC samples are reweighed based on the per-event true number of interactions to match the level of pileup observed in data as per general recipes.

2.3 Higher-order corrections

Higher-order cross-sections are known to be small in VBS and not precisely calculated. Calculators such as VBFNLO [12] do exist which enable NLO QCD theoretical estimations. However, recent papers for other VBS final states [13] show that this description is inadequate because NLO EWK corrections are comparable in size and, in fact, the interplay between the two corrections makes the "purely EWK" and "purely QCD" contributions ill-defined at the NLO. Therefore, we just adopt the cross sections provided by the MC generators. The same choice is done for the minor backgrounds WWZ and $t\bar{t}$ Z.

The qqZZ background is generated at NLO, while the fully differential cross section has been computed at NNLO [14], but are not yet available in a partonic level event generator. Therefore NNLO/NLO *k*-factors for the qqZZ background process are applied to the MadGraph5_aMC@NLO sample. The NNLO/NLO *k*-factors are applied in the analysis differentially as a function of $m(\mathrm{ZZ})$.

Additional NLO electroweak corrections which depend on the initial state quark flavor and kinematics are also applied to the qqZZ background process in the region $m(\mathrm{ZZ}) > 2m(\mathrm{Z})$ where the corrections have been computed [15]. The differential QCD and electroweak *k*-factors can be seen in Figure 2.

As detailed in the previous paragraph, ggZZ modelling uses a dedicated LO simulation of the 1-jet bin only. The most precise estimations of the cross-section are at the NLO (technically at the $\mathrm{N^3LO}$ with respect to the tree-level process) [16, 17]. In order to apply k-factors from this work, we verify a few things:

Table 6: List of signal and background samples used in the analysis.

Process	Generator	Cross Section [fb]	Sample name	Remarks
Signal samples 2016				
$ZZ \to 4\ell$ + 2jets	MadGraph (LO)	0.441	ZZJJTo4L_EWK_13TeV-madgraph-pythia8/[1]	$m_{jj} > 100$ GeV
$ZZ \to 4\mu$ + 2jets	Phantom (LO)	0.418	VBFToHiggs0PMContinToZZTo4muJJ_M125_GaSM_13TeV_phantom_pythia8/[1]	used to cross-check MadGraph sample
$ZZ \to 4e$ + 2jets	Phantom (LO)	0.418	VBFToHiggs0PMContinToZZTo4eJJ_M125_GaSM_13TeV_phantom_pythia8/[1]	used to cross-check MadGraph sample
$ZZ \to 2e2\mu$ + 2jets	Phantom (LO)	0.836	VBFToHiggs0PMContinToZZTo2e2muJJ_M125_GaSM_13TeV_phantom_pythia8/[1]	used to cross-check MadGraph sample
Irreducible background samples 2016				
$ZZ \to 4\ell$ + 0,1jets	MadGraph (NLO)	1218	ZZTo4L_13TeV-amcatnloFXFX-pythia8/[1]	
$gg \to ZZ \to 4\ell$ + 1 jet	MadGraph (LO)	4.45	/eos/cms/store/user/covarell/ggzz1jet/	
$gg \to ZZ \to 4\mu$	MCFM (LO)	1.59	GluGluToContinToZZTo4mu_13TeV_DefaultShower_MCFM701_pythia8/[1]	used to cross-check MadGraph sample
$gg \to ZZ \to 4e$	MCFM (LO)	1.59	GluGluToContinToZZTo4e_13TeV_DefaultShower_MCFM701_pythia8/[1]	used to cross-check MadGraph sample
$gg \to ZZ \to 2e2\mu$	MCFM (LO)	3.19	GluGluToContinToZZTo2e2mu_13TeV_DefaultShower_MCFM701_pythia8/[1]	used to cross-check MadGraph sample
Minor background samples 2016				
$t\bar{t}Z \to 4\ell\ 2\nu$	MadGraph	0.253	ttZJets_13TeV_madgraphMLM/[1]	
WWZ+ jets	MadGraph (NLO)	0.1651	WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	
Signal samples 2017-18				
$ZZ \to 4\ell$ + 2jets	MadGraph (LO)	0.427	ZZJJTo4L_EWK_TuneCP5_13TeV-madgraph-pythia8/[2,3]	$m_{jj} > 100$ GeV
$ZZ \to 4\mu$ + 2jets	Phantom (LO)	0.418	VBFToHiggs0PMContinToZZTo4muJJ_M125_GaSM_13TeV_phantom_pythia8/[2,3]	used to cross-check MadGraph sample
$ZZ \to 4e$ + 2jets	Phantom (LO)	0.418	VBFToHiggs0PMContinToZZTo4eJJ_M125_GaSM_13TeV_phantom_pythia8/[2,3]	used to cross-check MadGraph sample
$ZZ \to 2e2\mu$ + 2jets	Phantom (LO)	0.836	VBFToHiggs0PMContinToZZTo2e2muJJ_M125_GaSM_13TeV_phantom_pythia8/[2,3]	used to cross-check MadGraph sample
Irreducible background samples 2017-18				
$ZZ \to 4\ell$ + 0,1jets	MadGraph (NLO)	1218	ZZTo4L_TuneCP5_13TeV-amcatnloFXFX-pythia8/[2,3]	
$gg \to ZZ \to 4\ell$ + 1 jet	MadGraph (LO)	4.45	/eos/cms/store/user/covarell/ggzz1jet/	
$gg \to ZZ \to 4\mu$	MCFM (LO)	1.59	GluGluToContinToZZTo4mu_13TeV_TuneCP5_MCFM701_pythia8/[2,3]	used to cross-check MadGraph sample
$gg \to ZZ \to 4e$	MCFM (LO)	1.59	GluGluToContinToZZTo4e_13TeV_TuneCP5_MCFM701_pythia8/[2,3]	used to cross-check MadGraph sample
$gg \to ZZ \to 2e2\mu$	MCFM (LO)	3.19	GluGluToContinToZZTo2e2mu_13TeV_TuneCP5_MCFM701_pythia8/[2,3]	used to cross-check MadGraph sample
Minor background samples 2017-18				
$t\bar{t}Z \to 4\ell\ 2\nu$	MadGraph	0.253	ttZJets_TuneCP5_13TeV_madgraphMLM/[2,3]	
WWZ+ jets	MadGraph (NLO)	0.1651	WWZ_TuneCP5_13TeV-amcatnlo-pythia8/[2,3]	

[1] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6-v*
[2] RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14-v*
[3] RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v*

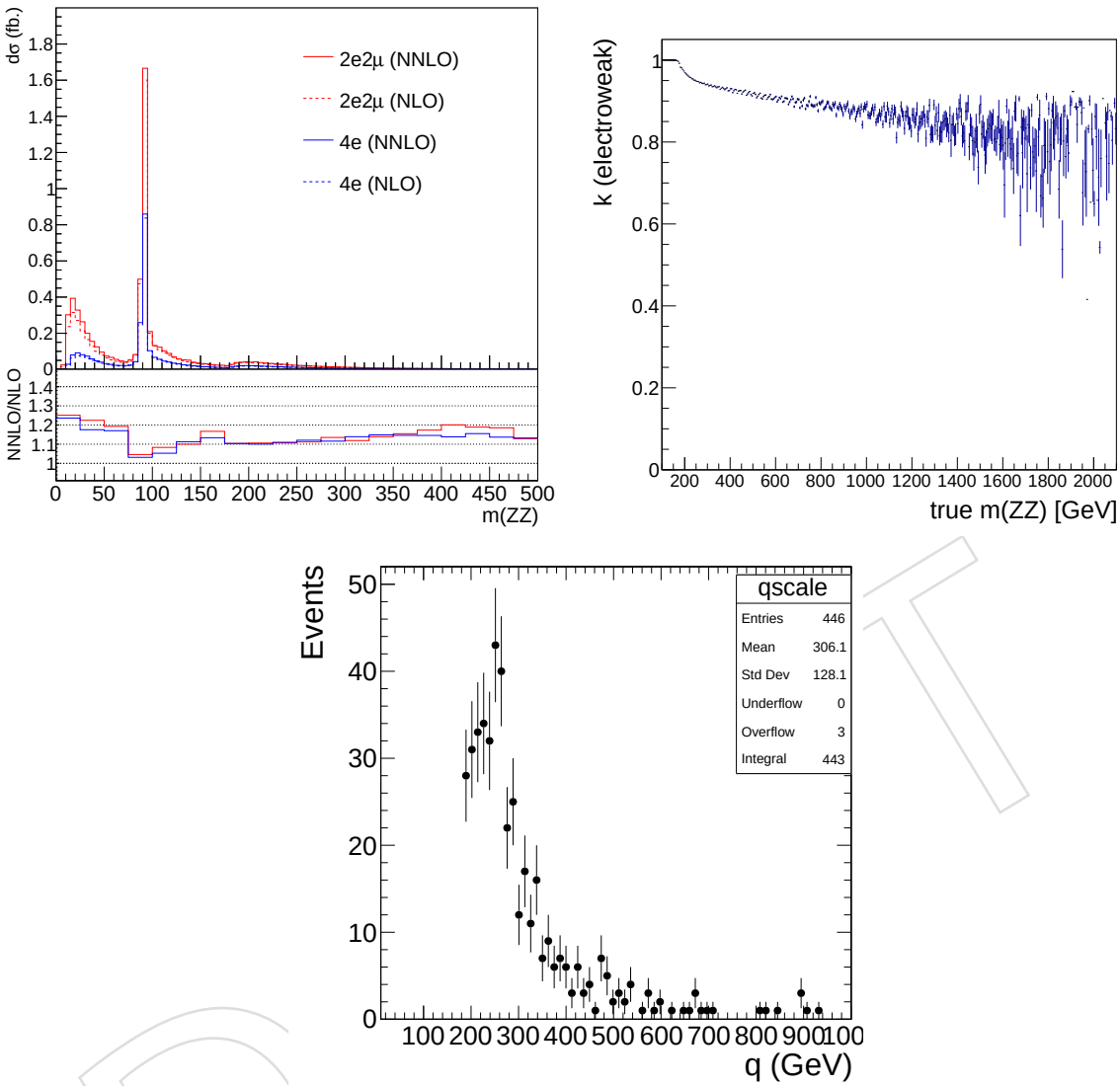


Figure 2: Left top: NNLO/NLO QCD k-factor for the qqZZ background as a function of $m(ZZ)$ for the 4ℓ and $2\ell 2\ell'$ final states. Right top: NLO/LO electroweak k-factor for the qqZZ background as a function of $m(ZZ)$. Bottom: QCD dynamic scale distribution for the ggZZ+1 jet MadGraph simulation.

- A NLO k-factor can be applied cleanly only to an inclusive sample at the LO, not to an exclusive jet-binned sample. Therefore we check that the cross-section of the MCFM sample matches the one from MadGraph in the fiducial phase-space of the analysis. This is shown in Table 14 in the following.
- The μ scale used in the MadGraph calculation must match the one in the calculation. The QCD scale in MadGraph is dynamically computed by the program and its distribution for events in the fiducial region is shown in Fig. 2. The QCD scales considered in the paper [17] are: $m_{ZZ}/2$, $m_{ZZ}/4$ and m_{ZZ}. Since the plot in Fig. 2 broadly resembles the m_{ZZ} distribution, we use the corresponding values: $\sigma_{\mathrm{LO}} = 3.53^{+0.74}_{-0.56}$ pb and $\sigma_{\mathrm{NLO}} = 4.59^{+0.32}_{-0.35}$ pb, from which we derive $k_{\mathrm{NLO/LO}} = 1.30^{+0.18}_{-0.11}$.
- We check the k-factor is constant in the main differential distributions, as shown in

the paper [16].

It has been shown [18, 19] that the soft collinear approximation is able to describe the ggZZ cross section and its interference term with the gluon-gluon Higgs production at NNLO. Since the estimated theory uncertainty from this approximation is 10%, which is similar to the NLO uncertainty, we do not use these results.

3 Object Reconstruction and Event Selection

The physics objects and the ZZ candidate selections used in this analysis are those of the $H \to 4\ell$ [3, 4] analysis, with minor changes. Here, we just report the main features, referring to that material for control plots/tables etc.

3.1 Electrons

3.1.1 Electron Reconstruction

Electron candidates are preselected using loose cuts on track-cluster matching observables, so as to preserve the highest possible efficiency while rejecting part of the QCD background. To be considered for the analysis, electrons are required to have a transverse momentum $p_T^e > 7$ GeV, a reconstructed $|\eta^e| < 2.5$, and to satisfy a loose primary vertex constraint defined as $d_{xy} < 0.5$ cm and $d_z < 1$ cm. Such electrons are called **loose electrons**.

The data-MC discrepancy is corrected using scale factors as is done for the electron selection with data efficiencies measured using the same tag-and-probe technique outlined later (see Section 3.1.5). These studies for reconstructions are carried out by the EGM POG and the results are only summarised here.

The electron reconstruction scale factors are applied as a function of the super cluster η and electron p_{T}.

3.1.2 Electron Identification

The electron selection is identical to that used in the $H \to ZZ \to 4\ell$ analysis [3, 4] and is based on a multivariate discriminant for all data taking periods.

Reconstructed electrons are now identified and isolated by means of an e**X**treme **G**radient **Boost**ing (XGBoost) optimized distributed gradient boosting library designed to be highly efficient, flexible and portable. It implements machine learning algorithms under the Gradient Boosting framework and exploits observables from the electromagnetic cluster, the matching between the cluster and the electron track, observables based exclusively on tracking measurements as well as particle flow isolation sums.

The relative isolation for electrons is defined as:

$$\mathrm{RelPFiso} = (\sum_{\mathrm{charged}} p_T + \sum_{\mathrm{neutral}}^{\mathrm{corr}} p_T)/p_T^{\mathrm{lepton}}. \quad (1)$$

where the corrected neutral component of isolation is then computed using the formula:

$$\sum_{\mathrm{neutral}}^{\mathrm{corr}} p_T = \max(\sum_{\mathrm{neutral}}^{\mathrm{uncorr}} p_T - \rho \times A_{\mathrm{eff}}, 0\,\mathrm{GeV}). \quad (2)$$

and the mean pile-up contribution to the isolation cone is obtained as :

$$PU = \rho \times A_{\mathrm{eff}} \quad (3)$$

where ρ is the mean energy density in the event and the effective area A_{eff} is defined as the ratio between the slope of the average isolation and that of ρ as a function of the number of vertices.

The full list of used features can be found in the Table 7.

The model is trained on 2016, 2017, and 2018 Drell-Yan with jets MC sample for both signal and background. The separate training for three periods guarantees optimal performance during the whole Run 2 data taking period.

Tables 8, 9 and 10 list the cuts values applied to the MVA output for 2016, 2017, 2018 training, respectively. For 2018, the corresponding signal and background efficiencies are given as examples.They are very similar for 2016 and 2017. For the analysis, loose electrons have to pass this MVA identification and isolation working point.

Observable type	Observable name
Cluster shape	RMS of the energy-crystal number spectrum along η and φ; $\sigma_{i\eta i\eta}$, $\sigma_{i\varphi i\varphi}$ Super cluster width along η and ϕ Ratio of the hadronic energy behind the electron supercluster to the supercluster energy, H/E Circularity $(E_{5\times5} - E_{5\times1})/E_{5\times5}$ Sum of the seed and adjacent crystal over the super cluster energy R_9 For endcap traing bins: energy fraction in pre-shower $E_{\mathrm{PS}}/E_{\mathrm{raw}}$
Track-cluster matching	Energy-momentum agreement E_{tot}/p_{in}, E_{ele}/p_{out}, $1/E_{tot} - 1/p_{in}$ Position matching $\Delta\eta_{in}$, $\Delta\varphi_{in}$, $\Delta\eta_{seed}$
tracking	Fractional momentum loss $f_{brem} = 1 - p_{out}/p_{in}$ Number of hits of the KF and GSF track N_{KF}, N_{GSF} $(\cdot)$ Reduced χ^2 of the KF and GSF track χ^2_{KF}, χ^2_{GSF} Number of expected but missing inner hits $(\cdot)$ Probability transform of conversion vertex fit χ^2 $(\cdot)$
isolation	Particle Flow photon isolation sum $(\cdot)$ Particle Flow charged hadrons isolation sum $(\cdot)$ Particle Flow neutral hadrons isolation sum $(\cdot)$
For PU-resilience	Mean energy density in the event: ρ $(\cdot)$

Table 7: Overview of input features to the identification classifier. Variables not used in the Run 2 MVA are marked with $(\cdot)$.

2016 Datasets			
minimum BDT score	$\lvert\eta\rvert < 0.8$	$0.8 < \lvert\eta\rvert < 1.479$	$\lvert\eta\rvert > 1.479$
$5 < p_T < 10$ GeV	0.9503	0.9461	0.9387
$p_T > 10$ GeV	0.3782	0.3587	-0.5745

Table 8: Minimum BDT score required for passing the electron identification, for 2016 samples.

2017 Datasets			
minimum BDT score	$\lvert\eta\rvert < 0.8$	$0.8 < \lvert\eta\rvert < 1.479$	$\lvert\eta\rvert > 1.479$
$5 < p_T < 10$ GeV	0.8521	0.8268	0.8694
$p_T > 10$ GeV	0.9825	0.9692	0.7935

Table 9: Minimum BDT score required for passing the electron identification, for 2017 samples.

	$\lvert\eta\rvert < 0.8$		
	Cut on BDT score	Signal eff.	Background eff.
$5 < p_T < 10$ GeV	0.8956	81.04%	4.4%
$p_T > 10$ GeV	0.0424	97.1%	2.9%
	$0.8 < \lvert\eta\rvert < 1.479$		
	Cut on BDT score	Signal eff.	Background eff.
$5 < p_T < 10$ GeV	0.9111	79.3%	4.6%
$p_T > 10$ GeV	0.0047	96.3%	3.6%
	$\lvert\eta\rvert > 1.479$		
	Cut on BDT score	Signal eff.	Background eff.
$5 < p_T < 10$ GeV	0.9401	72.97%	3.6%
$p_T > 10$ GeV	-0.6042	95.7%	6.7%

Table 10: Minimum MVA score required for passing the electron identification, together with the corresponding signal and background efficiencies, for 2018 samples.

3.1.3 Electron Impact Parameter Selection

In order to ensure that the leptons are consistent with a common primary vertex we require that they have an associated track with a small impact parameter with respect to the event primary vertex. We use the significance of the impact parameter to the event vertex, $|\mathrm{SIP}_{3D} = \frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}|$, where IP is the lepton impact parameter in three dimensions at the point of closest approach with respect to the primary interaction vertex, and σ_{IP} the associated uncertainty. Hereafter, a "primary lepton" is a lepton satisfying $|\mathrm{SIP}_{3D}| < 4$.

3.1.4 Electron Energy Calibrations

Electrons in data are corrected for features in ECAL energy scale in bins of p_{T} and $|\eta|$. Corrections are calculated on a $\mathrm{Z} \to \mathrm{ee}$ sample to align the dielectron mass spectrum in the data to that in the MC, and to minimize its width.

The $\mathrm{Z} \to \mathrm{ee}$ mass resolution in Monte Carlo is made to match data by applying a pseudorandom Gaussian smearing to electron energies, with Gaussian parameters varying in bins of p_{T} and $|\eta|$. This has the effect of convoluting the electron energy spectrum with a Gaussian.

The electron energy scale is measured in data by fitting a Crystal-ball function to the di-electron mass spectrum around the Z peak in the $Z \to \mathrm{ee}$ control region. Results of this procedure year per year can be found in [3, 20, 21].

3.1.5 Electron Efficiency Measurements

The Tag-and-Probe study was performed on the SingleElectron/EGamma dataset using the golden JSONs year per year.

Tag electrons need to satisfy the following quality requirements:

- trigger matched to single electron trigger (e.g HLT_Ele32_WPTight_Gsf_L1DoubleEG_v* for 2018 for instance)
- $p_T > 30$ GeV (tag), super cluster (SC) $\eta < 2.17$
- the tag and the probe need to have opposite charge.

For the bin between 7 and 20 GeV, additional criteria are required which help cleaning the background and makes the fits more reliable (and thus, the measurement more precise).

Probe electrons only need to be reconstructed as GsfElectron while the FSR recovery algorithm is not applied in efficiency measurement.

The nominal MC efficiencies are evaluated from the LO MadGraph Drell-Yan.

For the efficiency measurements a template fit is used. The m_{ee} signal shape of the passing and failing probes is taken from MC and convoluted with a Gaussian. The data is then fitted with the convoluted MC template and a CMSShape (an Error-function with a one-sided exponential tail). This change follows from the usage of the T&P tool developed by the EGM POG. For the low p_T bins, a gaussian is added to the signal model for the failing probes.

The electron selection efficiency is measured as a function of the probe electron p_T and its SC η, and separately for electrons falling in the ECAL gaps. Results of this procedure year per year can be found in [3, 20, 21].

The EGM recommendations on the evaluation of Tag-and-Probe uncertainties for efficiency measurements are followed. Specifically, we consider

- Variation of the signal shape from a MC shape to an analytic shape (Crystal Ball) fitted to the MC
- Variation of the background shape from a CMS-shape to a simple exponential in fits to data
- Using an NLO MC sample for the signal templates

The total uncertainty for the measurement of the scale factors is the quadratic sum of the statistical uncertainties returned from the fit and the aforementioned systematic uncertainties.

3.2 Muons

3.2.1 Muon Reconstruction and Identification

We define **loose muons** as the muons that satisfy $p_T > 5$, $|\eta| < 2.4$, $dxy < 0.5$ cm, $dz < 1$ cm, where dxy and dz are defined w.r.t. the PV and using the 'muonBestTrack'. Muons have to be reconstructed by either the Global Muon or Tracker Muon algorithm. Standalone Muon tracks that are only reconstructed in the muon system are rejected. Muons with `muonBestTrackType==2` (standalone) are discarded even if they are marked as global or tracker muons.

Loose muons with p_{T} below 200 GeV are considered identified muons if they also pass Muon BDT (see below). Loose muons with p_{T} above 200 GeV are considered identified muons if they pass the PF ID or the Tracker High-p_{T} ID, the definition of which is shown in Table 11. This relaxed definition is used to increase signal efficiency for the high-mass search. When a very heavy resonance decays to two Z bosons, both bosons will be very boosted. In the lab frame, the leptons coming from the decay of a highly boosted Z will be nearly collinear, and the PF ID loses efficiency for muons separated by approximately $\Delta R < 0.4$, which roughly corresponds to muons originating from Z bosons with $p_{\mathrm{T}} > 500$ GeV.

Table 11: The requirements for a muon to pass the Tracker High-p_{T} ID. Note that these are equivalent to the Muon POG High-p_{T} ID with the global track requirements removed.

Plain-text description	Technical description
Muon station matching	Muon is matched to segments in at least two muon stations **NB: this implies the muon is an arbitrated tracker muon.**
Good p_{T} measurement	$\frac{p_{\mathrm{T}}}{\sigma_{p_{\mathrm{T}}}} < 0.3$
Vertex compatibility $(x-y)$	$d_{xy} < 2$ mm
Vertex compatibility (z)	$d_z < 5$ mm
Pixel hits	At least one pixel hit
Tracker hits	Hits in at least six tracker layers

An additional "ghost-cleaning" step is performed to deal with situations when a single muon can be incorrectly reconstructed as two or more muons:

- Tracker Muons that are not Global Muons are required to be arbitrated.
- If two muons are sharing 50% or more of their segments then the muon with lower quality is removed.

3.2.2 Muon Identification and Isolation

The muon selection is identical to that used in the $H \to ZZ \to 4\ell$ analysis [3, 4]. The main sources of non-prompt muons are non-isolated muons coming from decays of heavy-flavour mesons and mis-reconstructed jets usually originating from light-flavour quarks. One of the main improvements brought in the analysis is development of new XGBoost multivariate discriminant for muon selection in all Run 2 data taking periods.

Reconstructed muons are now fully identified by means of an e**X**treme **G**radient **Boost**ing (XGBoost) gradient boosting library already used to identify electrons. This machine learning framework exploits observables based exclusively on tracking, information from the muon part of the detector as well as different components of the isolation.

Particle-Flow based isolation is used for the muons. The so-called $\Delta\beta$ correction is applied

in order to subtract the pileup contribution for the muons, whereby $\Delta\beta = \frac{1}{2}\sum_{\text{PU}}^{\text{charged had.}} p_{\text{T}}$ gives an estimate of the energy deposit of neutral particles (hadrons and photons) from pile-up vertices. The relative isolation for muons is then defined as:

$$\text{RelPFiso} = \frac{\sum^{\text{charged had.}} p_{\text{T}} + \max(\sum^{\text{neutral had.}} E_{\text{T}} + \sum^{\text{photon}} E_{\text{T}} - \Delta\beta, 0)}{p_{\text{T}}^{\text{lepton}}} \tag{4}$$

The full list of used features can be found in the Table 12.

The muons are preselected by requiring $p_T > 5$, $|\eta| < 2.4$, $dxy < 0.5$ cm, $dz < 1$ cm, where dxy and dz. Additionally, muons have to be reconstructed by either the Global Muon or Tracker Muon algorithm. The signal consists of prompt muons matched to truth muons while background is composed of unmatched and true but non-prompt muons. The MVA models are trained on 2016, 2017, and 2018 Drell-Yan with jets MC sample for both signal and background. The dedicated training for each of the Run 2 three data taking periods guarantees optimal performance.

Table 13 lists the cuts values applied to the Muon MVA output for 2016, 2017, 2018 trainings.

Observable type	Observable name
Kinematic variables	Pseudorapidity η
Global track quality variables	Global number of valid muon hits Normalized Chi2
Track quality variables	Number of valid hits Number of valid pixel hits SIP3D d_z d_{xy}
Isolation variables	Particle Flow photon isolation sum Particle Flow charged hadrons isolation sum Particle Flow neutral hadrons isolation sum
For PU-resilience	Mean energy density in the event: $\rho\ (\cdot)$

Table 12: Overview of input features passed to the identification classifier.

	Minimum BDT score		
	2016	2017	2018
$5 < p_T < 10$ GeV	0.8847	0.8836	0.9506
$p_T > 10$ GeV	-0.1939	-0.3831	-0.3629

Table 13: Minimum BDT score required for passing the muon identification, for 2016, 2017 and 2018 samples.

3.2.3 Muon Impact Parameter Selection

In addition to a cut to the Muon BDT, we apply an additional cut to the muon significance of impact parameter as for the electrons, as described in Sec. 3.1.3:

- $|\mathrm{SIP}_{3D} = \frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}| < 4$

3.2.4 Muon Energy Calibrations

Similar to electrons the muon momentum scale is measured in data by fitting a Crystal-ball function to the di-muon mass spectrum around the Z peak in the $Z \rightarrow \mu\mu$ control region. Results of this procedure year per year can be found in [3, 20, 21].

3.2.5 Muon Efficiency Measurements

Muon efficiencies are measured with the Tag and Probe (T&P) method performed on $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ events in bins of p_T and η. The Z sample is used to measure the muon reconstruction and identification efficiency at high p_T, and the efficiency of the isolation and impact parameter requirements at all p_T. The J/ψ sample is used to measure the reconstruction efficiency at low p_T, as it benefits from a better purity in that kinematic regime. In this case, events are collected using `HLT_Mu7p5_Track2_Jpsi_v*` when probing the reconstruction and identification efficiency in the muon system, and using the `HLT_Mu7p5_L2Mu2_Jpsi_v*` when probing the tracking efficiency.

Results for the muon reconstruction and identification efficiency for $p_T > 20\,\mathrm{GeV}$ have been derived by the Muon POG. The probe in this measurement are tracks reconstructed in the inner tracker, and the passing probes are those that are also reconstructed as a global or tracker muon and passing the Muon POG Loose muon identification. Results for low p_T muons were derived using J/ψ events, with the same definitions of probe and passing probes. The systematic uncertainties are estimated by varying the analytical signal and background shape models used to fit the dimuon invariant mass. The efficiency and scale factors used for low p_T muons are the ones derived using single muon dataset.

For the impact parameter requirements, the measurement is performed using Z events. Events are selected with `HLT_IsoMu27_v*` or `HLT_Mu50_v*` triggers. For this measurement, the probe is a muon passing the POG Loose identification criteria, and it is considered a passing probe if satisfies the SIP3D, dxy, dz cuts of this analysis. The efficiency to reconstruct a muon track in the inner detector is measured using as probes tracks reconstructed in the muon system alone. The efficiency and data to mc scale factors are measured from Z events as a function of η for $p_T > 10\,\mathrm{GeV}$ and $p_T < 10\,\mathrm{GeV}$.

Results of this procedure year per year can be found in [3, 20, 21].

3.3 Photons for FSR recovery

The FSR recovery algorithm was considerably simplified with respect to what was done in Run I, while maintaining similar performance. The algorithm is explained in [3] and has very little impact for this analysis.

3.4 Jets

3.4.1 Jet Identification

Jets are reconstructed by using the anti-k_T clustering algorithm out of particle flow candidates, with a distance parameter $R = 0.4$, after rejecting the charged hadrons that are associated to a pileup primary vertex.

Jet corrections are applied, following JetMET Physics Object Group recommendations. The corrections used are as follows:

- Jet energy scale corrections
  - 2016: Summer16_07Aug2017_V11_MC.db and Summer16_07Aug2017All_V11_DATA.db
  - 2017: Fall17_17Nov2017_V3_94X_MC.db and Fall17_17Nov2017_V32_94X_DATA.db
  - 2018: Autumn18_V8_MC.db and Autumn18_RunABCD_V8_DATA.db
- Jet energy resolution corrections
  - 2016: Summer16_25nsV1_MC.db
  - 2017: Fall17_V3_94X_MC.db
  - 2018: Autumn18_V1_MC.db

To reduce instrumental background, the tight working point jet ID suggested by the JetMET Physics Object Group is applied [22]. In addition, jets from Pile-Up are rejected using the PileUp jet ID criteria suggested by the JetMET POG [23]. It is to be noted that the PU JetID was only derived for 2016 conditions but is also applied to 2017 and 2018 samples.

In this analysis, the jets are required to be within $|\eta| < 4.7$ area and have a transverse momentum above 30 GeV. In addition, the jets are cleaned from any of the tight leptons (passing the SIP and isolation cut computed after FSR correction) and FSR photons by a separation criterion: $\Delta R(\text{jet,lepton/photon}) > 0.4$.

3.4.2 Jet Energy Corrections

The calorimeter response to particles is not linear and it is not straightforward to translate the measured jet energy to the true particle or parton energy, therefore we need Jet Energy Corrections. In this analysis, standard jet energy corrections are applied to the reconstructed jets, which consist of L1 Pileup, L2 Relative Jet Correction, L3 Absolute Jet Correction for both Monte Carlo samples and data, and also residual calibration for data [24].

3.4.3 Additional criteria on jets

The three data takins periods analyzed in this note suffered from issues during the data taking which impact the quality of the jet reconstruction. Some of these issues would need a complete re-reconstruction of the data to be fully fixed (the so-called "Ultra Legacy ReReco"), which is beyond the scope of the paper. In the mean time, following the guidance from the JetMET POG, we study the possibilty of adding some criteria on the jet to cope with these issues.

3.4.3.1 L1 pre-firing

In 2016 and 2017, the gradual timing shift of ECAL was not properly propagated to L1 trigger primitives (TP) resulting in a significant fraction of high eta TP being mistakenly associated to the previous bunch crossing. Since Level 1 rules forbid two consecutive bunch crossings to fire, an unpleasant consequence of this (in addition to not finding the TP in the bx 0) is that events can self veto if a significant amount of ECAL energy is found in the region of $2. < |\eta| < 3$. This effect is not described by the simulations [25]. A weight is thus calculating for each event, not to prefire, and apply to the simulation in 2016 and 2017 samples. The official tool is used for this purpose [25].

3.4.3.2 Removal of noisy jets

Increased jet multiplicity was reported for 2017 data, creating "horns" in the jet η distribution for $2.5 < |\eta_{jet}| < 3$. The issue was linked to an increase of the ECAL noise, PU and bunch-crossing dependent, thus getting worse as luminosity increases. The problem can only be fixed in the UL ReReco. For now, we checked the impact of rejecting jets with raw $p_T < 50$ GeV in $2.65 < |\eta| < 3.139$. As we see no significant impact in the data/MC agreement, we decided not to use these cuts.

3.5 Event Selection

The four-lepton candidates are built from selected leptons, where FSR photons are subtracted as described in Section 3.3. A lepton cross cleaning is applied by discarding electrons which are within $\Delta R < 0.05$ of selected muons.

The construction and selection of four-lepton candidates proceeds according to the following sequence:

1. **Z candidates** are defined as pairs of selected leptons of opposite charge and matching flavour (e^+e^-, $\mu^+\mu^-$) that satisfy $60 < m_{\ell\ell(\gamma)} < 120$ GeV/c^2, where the Z candidate mass includes the selected FSR photons if any.

2. **ZZ candidates** are defined as pairs of non-overlapping Z candidates. The Z candidate with reconstructed mass $m_{\ell\ell}$ closest to the nominal Z boson mass is denoted as Z_1, and the second one is denoted as Z_2. ZZ candidates are required to satisfy the following list of requirements:
   - **Ghost removal** : $\Delta R(\eta, \phi) > 0.02$ between each of the four leptons.
   - **lepton p_T**: Two of the four selected leptons should pass $p_{T,1} > 20$ GeV/c and $p_{T2j} > 10$ GeV/c.
   - **QCD suppression**: all four opposite-sign pairs that can be built with the four leptons (regardless of lepton flavor) must satisfy $m_{\ell\ell} > 4$ GeV/c^2. Here, selected FSR photons are not used in computing $m_{\ell\ell}$, since a QCD-induced low mass dilepton (eg. J/ψ) may have photons nearby (e.g. from π_0).
   - **Z mass**: $m_{Z1,Z2} > 60$ GeV/c^2 in order to comply with MC samples that do not describe the offshell ZZ* distributions.
   - **'smart cut'**: defining Z_a and Z_b as the mass-sorted alternative pairing Z candidates (Z_a being the one closest to the nominal Z boson mass), require NOT($|m_{Za} - m_Z| < |m_{Z1} - m_Z|$ AND $m_{Zb} < 12$). Selected FSR photons are included in m_Z's computations. This cut discards 4μ and $4e$ candidates where the alternative pairing looks like an on-shell Z + low-mass $\ell^+\ell^-$.
   - **four-lepton invariant mass**: > 180 GeV/c^2 in order to comply with MC samples that do not describe the offshell ZZ* distributions.

426 3. Events containing at least one selected ZZ candidate are kept.

427 In events where more than one viable ZZ candidate is selected, this analysis selects the one
428 with the largest scalar sum of the leptons of the Z_2 boson.

429 ### 3.5.1 ZZjj baseline selection

430 The ZZ event selection is supplemented by the dijet requirement for the VBS search. Specif-
431 ically, we require that the event feature at least two selected jets, and the leading plus sub-
432 leading jets are taken as the "tagging jets". The invariant mass of the tagging jets has to satisfy
433 $m_{jj} >$100 GeV, in order to suppress hadronic W/Z decays.

434 This **baseline selection region** is used to estimate significance of the EWK signal, in the mea-
435 surement of the EWK and total fiducial $4\ell jj$ cross-sections, and the aQGC search.

436 ### 3.5.2 VBS selections (and anti-selection)

437 Two VBS-enriched signal regions are defined for events that pass the ZZjj baseline selection. A
438 loose VBS-enriched signal region for which events also satisfy:

439 - $|\Delta\eta_{jj}| > 2.4$
440 - $m_{jj} > 400$ GeV.

441 and a tight VBS-enriched signal region for which events also satisfy:

442 - $|\Delta\eta_{jj}| > 5$
443 - $m_{jj} > 400$ GeV.

444 These signal-enriched regions are used in the measurement of the EWK and total fiducial $4\ell jj$
445 cross-sections, in addition to the ZZjj baseline selection.

446 Finally, a background control region is defined from events that satisfy the baseline ZZjj se-
447 lection but fail at least one of the criteria that defines the loose VBS signal-enriched selection
448 .

4 The MELA discriminant

The kinematic discriminants used in this study are computed using the MELA package [26–28], which uses MCFM matrix elements for the EWK signal and the dominant ZZjj background to describe process probabilities. The background includes gg or $q\bar{q} \to ZZ$ / $Z\gamma^*$ / $\gamma^*\gamma^*$ / $Z \to 4\ell$ processes.

Within the MELA framework, an analytic parameterization of the matrix elements is adopted and calculations are also tested against other implementations.

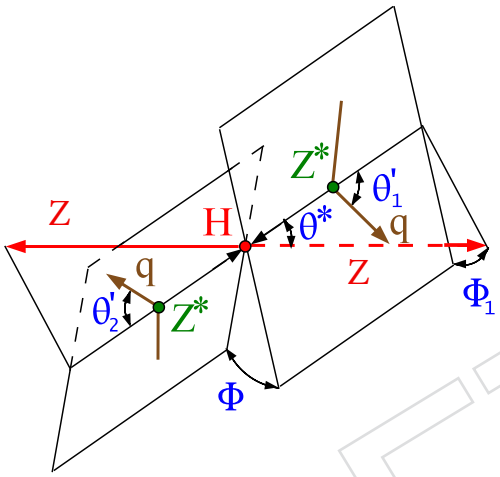


Figure 3: Illustrations of particle production and decay in VBS $qq' \to qq'ZZ \to qq'4\ell$. Angles and invariant masses fully characterize the orientation of the production and decay chain and are defined in the suitable rest frames.

The discriminant sensitive to the VBS signal topology with two associated jets is calculated as:

$$K_D = \left[1 + c(m_{4\ell}) \frac{\mathcal{P}_{\text{QCD-JJ}}(\vec{\Omega}^{4\ell+\text{JJ}}|m_{4\ell})}{\mathcal{P}_{\text{VBS+VVV}}(\vec{\Omega}^{4\ell+\text{JJ}}|m_{4\ell})}\right]^{-1}, \quad (5)$$

where $\mathcal{P}_{\text{VBS+VVV}}$ and $\mathcal{P}_{\text{QCD-JJ}}$ are probabilities obtained from the MCFM matrix elements for the EWK process and QCD-JJ, technically combination of gg/qg/qq parton collisions, in association with two jets within the MELA framework. $c(m_{4\ell})$ is a $m_{4\ell}$-dependent constant to calibrate the distribution in order to efficiently span the 0-1 interval and thus maximize the signal separation in binned templates.

This discriminant is efficient in separating VBS from either gg or $q\bar{q} \to 4\ell + 2$ jets background because jet correlations in these processes are distinct from the VBS process (see more details later).

5 MC Generator comparisons

5.1 Comparison for the VBS signal process

Using the above selection on reconstructed variables, we compare the prediction for the electroweak signal process obtained from the MadGraph and Phantom matrix-element generators.

Normalization and shapes of the VBS distributions for the two generators are in good agreement (see Fig. 4 for LHE level andFig. 5 for reconstruction level), with some exceptions that we explain below:

- Only at reconstruction level, a fairly large difference seems there at low m_{jj}, which also affects slightly other distributions. This is due to jet resolution effects, which populate the first bins of the distribution with events that have a true $m_{jj} < 100$ GeV. This is present in Phantom but not in MadGraph because m_{jj} has a lower cut of 100 GeV already at parton level. While this may be regarded as a problem, it is shown in the top left figures that this disagreement only affect the very low MELA region, where the search is not sensitive. This will be investigated further in the Results section.
- The $m_{\ell\ell}$ has different widths. MadGraph uses the leading-order result of the Z decay width ($\Gamma_Z = 2.4414$ GeV) due to internal consistency requirements, while Phantom uses the world best average as published by the Particle Data Group ($\Gamma_Z = 2.4952$ GeV), compatible with the effect found. This has no impact on the result, given the loose requirement $60 < m_Z < 120$ GeV.
- In the sensitive high-MELA region, an observed difference of 5-7% in the cross-secton is due to a combination of the different Z decay widths (explained above) and the different functional forms of the dynamic renormalization and factorization scales used in each generator. Aside from the small difference in the normalization, the shapes of the VBS distributions for the two generators are in excellent agreement in the high-MELA region, see Fig. 5.

5.2 Modelling of the loop-induced background process

Simulating the loop-induced $ZZ \to 4\ell$ production with two outgoing partons is a very time-consuming and CPU-expensive task. As a first attempt to model the loop-induced ZZ sample, the `MadGraph5_aMCatNLO` package was used to simulate the ZZjj final state where both Z bosons are at the pole mass and one extra jet is generated, but no tools exist to realistically decay the on-shell Z bosons and event generation requires about 20 minutes per event. This sample is produced privately and is compared to the results of the MCFM prediction.

The MCFM sample listed in Table 6 has been requested for this analysis, as the parton shower in the previously existing MCFM samples had been tuned with:

```
SpaceShower:pTmaxMatch = 1
```

which causes the shower evolution to stop at a p_T corresponding to the renormalization scale of the hard process, for the MCFM sample in question $\mu = m_{4l}/2$. This 'wimpy' shower causes gaps in the parton shower coverage. As a result of this configuration, the two-jet phase is poorly modelled.

Figure 6 shows ggZZjj as predicted by the regularly-showered MCFM sample and the matrix element prediction of the 1-jet sample state obtained from `MadGraph5_aMCatNLO`. Table 14 shows the ratio between the MCFM and MadGraph sample cross-sections. We conclude that:

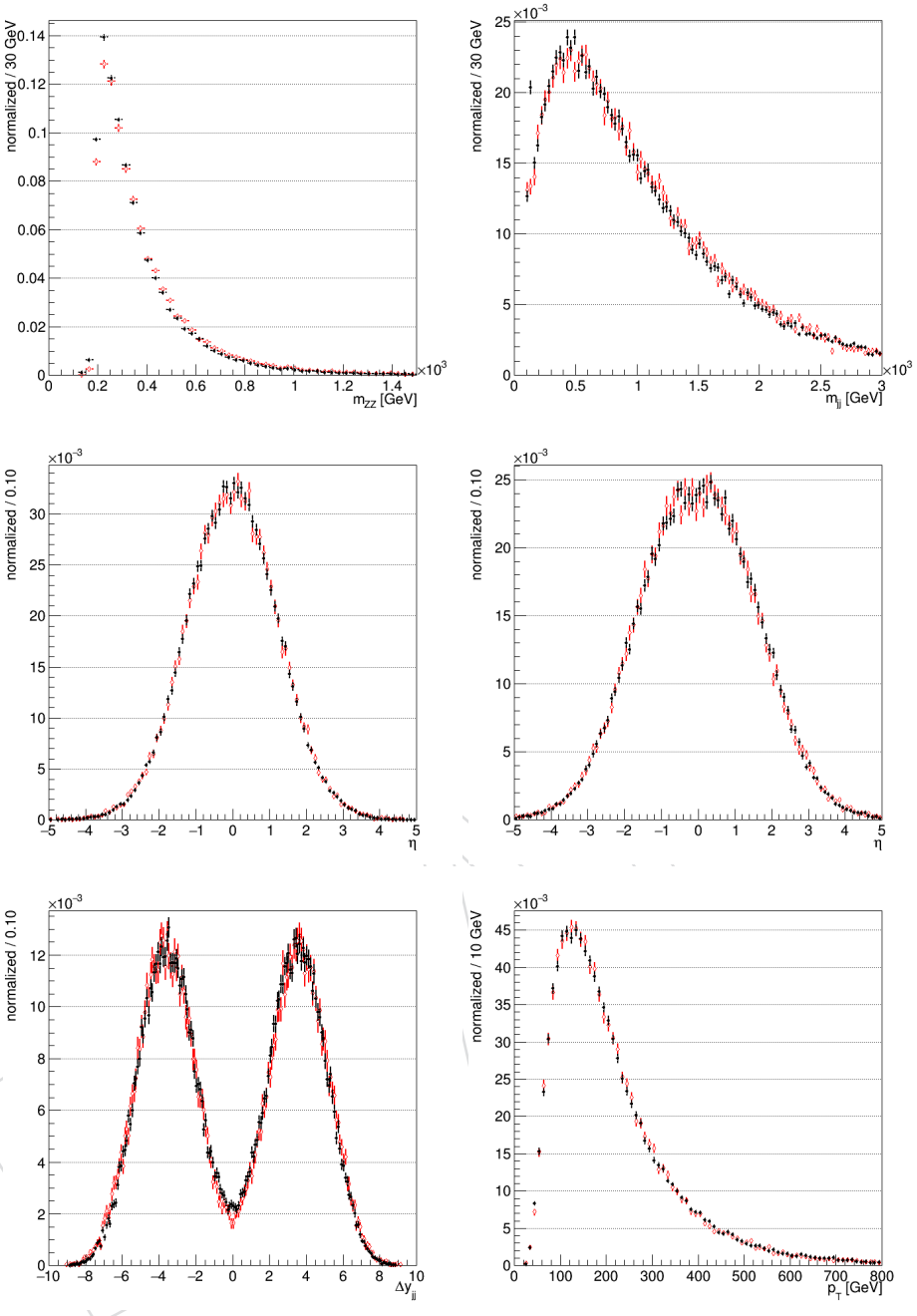


Figure 4: Comparison of the kinematics in the electroweak MadGraph (red) and Phantom MC (black) samples in the phase space defined by the ZZjj baseline selection at the LHE level. All distributions are normalized to unity. Top: $m_{\ell\ell}$ (left) and m_{jj} (right). Middle: Z_1 Zeppenfeld variable (left) and Z_2 Zeppenfeld variable (right). Low: $\Delta\eta_{jj}$ (right) and $p_{T,j}$ (right).

- The cross-sections in the baseline selection region agree at the 3% level. Of course the 4ℓ selection shows disagreement because the MadGraph sample lacks the 0-jet bin.
- The $m_{\ell\ell}$ has different widths. MadGraph uses zero because the decay is produced by PYTHIA at a lates stage, while MCFM uses the world best average. This has no impact on the result, given the loose requirement $60 < m_Z < 120$ GeV.

- There is a large disagreement in the jet-related distributions and consequently MELA. This has no explanation other than that MCFM, being inclusive in the number of jets at matrix-element, describes poorly the 2-jet phase space and will not be further used in the analysis. This is a noticeable improvement w.r.t. the 2016 analysis, because as shown in Table 14, corresponds to about 30% less contribution in the sensitive, high-MELA region.

Table 14: Ratios of ggZZ cross sections of the MCFM and MadGraph generators. The phase space is that of the generation, i.e. $m_{jj} > 100$ GeV and includes the branching ratios for the Z decays to electrons or muons.

4ℓ selection	baseline selection region	VBS region
1.53	1.03	1.31

A study was performed to produce a full matrix-element MadGraph 0/1/2-jet merged and matched to partons showers sample to better model this process. Details as well as comparison of main kinematic variables are discussed in Appendix B.

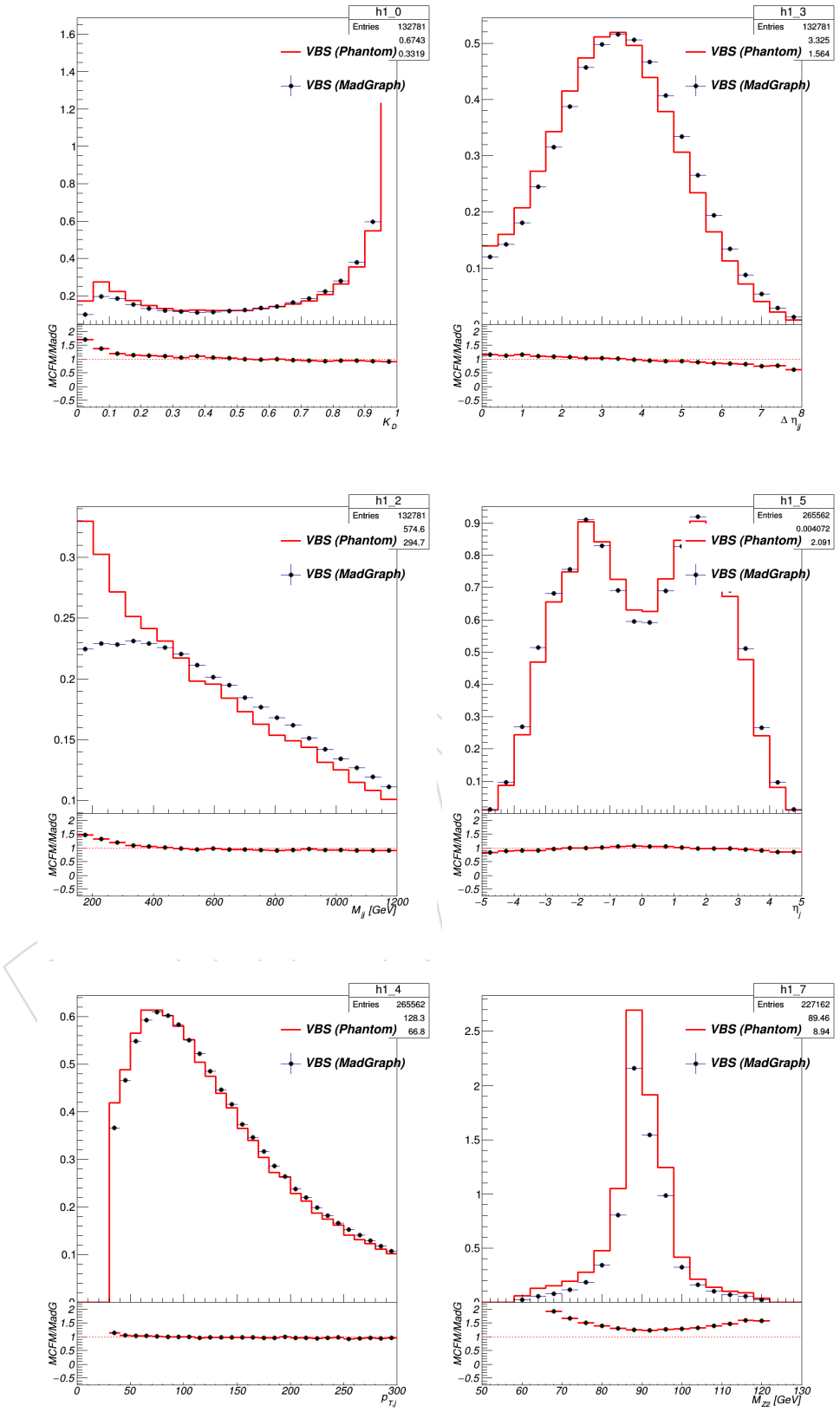


Figure 5: Comparison of the kinematics in the electroweak MadGraph (black points) and Phantom MC (red) samples in the phase space defined by the ZZjj baseline selection at the reconstruction level. All distributions are normalized to MC cross-sections. Top: MELA (left) and $\Delta\eta_{jj}$ (right). Middle: m_{jj} (left) and η_j (right). Low: $p_{T,j}$ (left) and $m_{\ell\ell}$ (right).

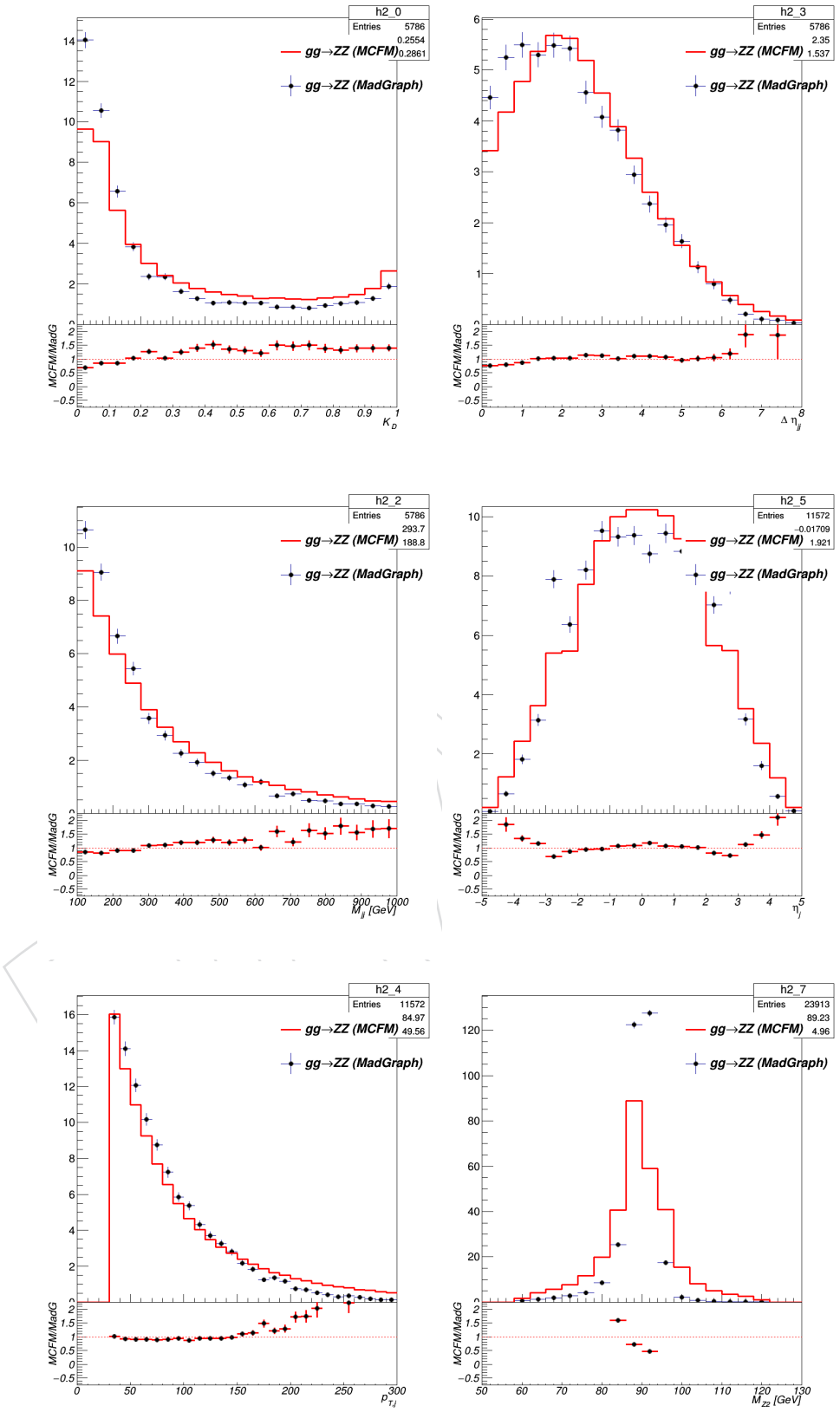


Figure 6: Comparison of the kinematics in the ggZZ MadGraph (black points) and MCFM MC (red) samples in the phase space defined by the ZZjj baseline selection at the reconstruction level. All distributions are normalized to MC cross-sections. Top: MELA (left) and $\Delta\eta_{jj}$ (right). Middle: m_{jj} (left) and η_j (right). Low: $p_{T,j}$ (left) and $m_{\ell\ell}$ (right).

6 Background estimation and Data/MC comparisons

6.1 Irreducible Backgrounds

Four-lepton events coming from ZZ production with 2 QCD vertices are estimated by separating their contribution in qqZZ and ggZZ, and applying higher-order corrections as explained in Sec. 2.

Multi-lepton processes with four or more leptons originating from non-Z decays, can contribute to the analysis if the leptons happen to satisfy the ZZ selection. The leading contribution arises from processes that feature one real on-shell Z boson (WWZ, $t\bar{t}Z$). These processes are sufficiently rare and contribute less than 3% in the final event selection. Their contributions are estimated from Monte-Carlo.

6.2 Reducible Background

The reducible background for the $ZZ \to 4\ell$ analysis, hereafter called $Z + X$, originates from processes which contain one or more non-prompt leptons in the four-lepton final state. The main sources of non-prompt leptons are non-isolated electrons and muons coming from decays of heavy-flavour mesons, mis-reconstructed jets (usually originating from light-flavour quarks) and electrons from γ conversions. In this discussion, we will consider a "fake lepton" any such event.

In this analysis, the rate of these background processes is estimated by measuring the f_e and f_μ ratios of fake electrons and fake muons which also pass the final selection criteria (defined in Section 3.5) over those which do pass the **loose** selection criteria (defined in Section 3.1.1 and 3.2.1). These ratios, hereafter referred to as fake rates, are applied in dedicated control samples in order to extract the expected background yield in the signal region. This is the same method used in [3] and will just be summarized here.

6.2.1 Reducible Background Estimate with Same-Sign Leptons

In order to measure the lepton fake ratios f_e and f_μ, we select samples of $Z(\ell\ell)$ + e and $Z(\ell\ell) + \mu$ events that are expected to be completely dominated by final states which include a Z boson and a fake lepton. These events are required to have two same flavour, opposite charge leptons with $p_T > 20/10$ GeV passing the tight selection criteria, thus forming the Z candidate. In addition, there is exactly one lepton passing the loose selection criteria as defined above. This lepton is used as the probe lepton for the fake ratio measurement.

The fake ratios are evaluated using the tight requirement $|M_{inv}(\ell_1, \ell_2) - M_Z| < 7$ GeV, to reduce the contribution from photon (asymmetric) conversions populating low masses, and $E_{\mathrm{T}}^{\mathrm{miss}} < 25$ GeV, separately for the 2016, 2017 and 2018 data. Measurements are in bins of the transverse momentum of the loose lepton and barrel and endcap region.

A control sample is obtained as a subset of the events that satisfy the first step of the selection, requiring an additional pair of loose leptons of same sign (to avoid signal contamination) and same flavour (SS-SF: $e^{\pm}e^{\pm}$, $\mu^{\pm}\mu^{\pm}$) . The SS-SF leptons are requested to pass the $\mathrm{SIP}_{3\mathrm{D}}$ cut, while no identification or isolation requirements are imposed. The reconstructed invariant mass of the SS-SF leptons has to satisfy 60 GeV $< m_{\ell\ell} <$ 120 GeV.

Starting from the control sample previously described, the final reducible background prediction in the signal region is given by the following expression:

$$N_{\mathrm{expect}}^{\mathrm{Z+X}} = \quad N^{\mathrm{DATA}} \times (\tfrac{\mathrm{OS}}{\mathrm{SS}})^{\mathrm{MC}} \times \quad f_1 \times f_2 \tag{6}$$

where:

- N^{DATA} is the number of events in the control region,
- $(\frac{\text{OS}}{\text{SS}})^{\text{MC}}$ is a correction factor between opposite sign and same sign control samples,
- f_1 and f_2 are the fake rates of each additional loose lepton, parameterised as a function of p_T and η.

This method only accounts for backgrounds with two fake leptons. Contributions with only one fake lepton mostly arise from rare WZjj events. With typical fake rates less or about equal to 1%, this contribution is negligible at low lepton p_{T} and becomes only a few % at high p_{T} (much less than the considered uncertainty from the comparison between the two methods).

The differences in rates between OS and SS samples are used to compute the correction factor in equation 6 for the final data-driven estimation. They are given for each year in table 15.

Channel	4e	2μ2e	4μ	2e2μ
2016	1.00 ± 0.01	1.00 ± 0.01	1.03 ± 0.03	1.04 ± 0.03
2017	1.01 ± 0.01	0.99 ± 0.01	1.04 ± 0.03	1.02 ± 0.03
2018	1.01 ± 0.01	1.00 ± 0.01	1.04 ± 0.02	1.04 ± 0.02

Table 15: The OS/SS ratios used in the SS method for each final state in all three years.

The event yields expected from Z+X in the signal region, are calculated for each final state. The statistical error is due to the event statistics in the control region, while the systematic one is mainly the uncertainty in the fake rates. The background is due to the systematic introduced when estimating the background composition. The total error is obtained with a quadrature sum for the statistical, background composition and correction systematics.

Table 17 shows the expected number of events in the signal regions from the reducible background processes at 13 TeV for all three years using the SS method.

6.2.2 Reducible Background Estimate with Opposite-Sign Leptons

In this other method, two control samples are obtained as subsets of four lepton events which pass the first step of the selection, requiring an additional pair of loose leptons of same flavour and opposite charge, that pass the SIP_{3D} cut. The first control sample is obtained by requiring that the two loose leptons which do not make the Z_1 candidate do not pass the final identification and isolation criteria (*2P+2F*) sample). The second control sample is obtained by requiring one of the four leptons not to pass the final identification and isolation criteria, while the other three do (*3P+1F* sample).

Combining the information from the two samples, the full expression for the prediction can be symbolically written as:

$$N_{\text{SR}}^{bkg} = (1 - \frac{N_{3P1F}^{ZZ}}{N_{3P1F}}) \sum_{j}^{N_{3P1F}} \frac{f_a^j}{1 - f_a^j} - \sum_{i}^{N_{2P2F}} \frac{f_3^i}{1 - f_3^i} \frac{f_4^i}{1 - f_4^i} \tag{7}$$

This second method is used to cross-check the nominal SS method. Differences beteween the two, ranging between 5 and 38% depending on the final state, are used as additional uncertainty on this estimate.

6.3 Data/MC comparisons for 2016, 2017, and 2018

Data are compared to signal and background estimates year by year, showing a good agreement in shape and normalization. Comparisons are shown in Fig. 7 for 2016 data/MC, in Fig. 8 for 2017 data/MC, and in Fig. 9 for 2018 data/MC. For 2017/2018, where the data are still unpublished, the region defined by MELA > 0.7 has been blinded in data.

Comparisons are also shown in Fig. 10 for summed data/MC of all years for illustrative purposes.

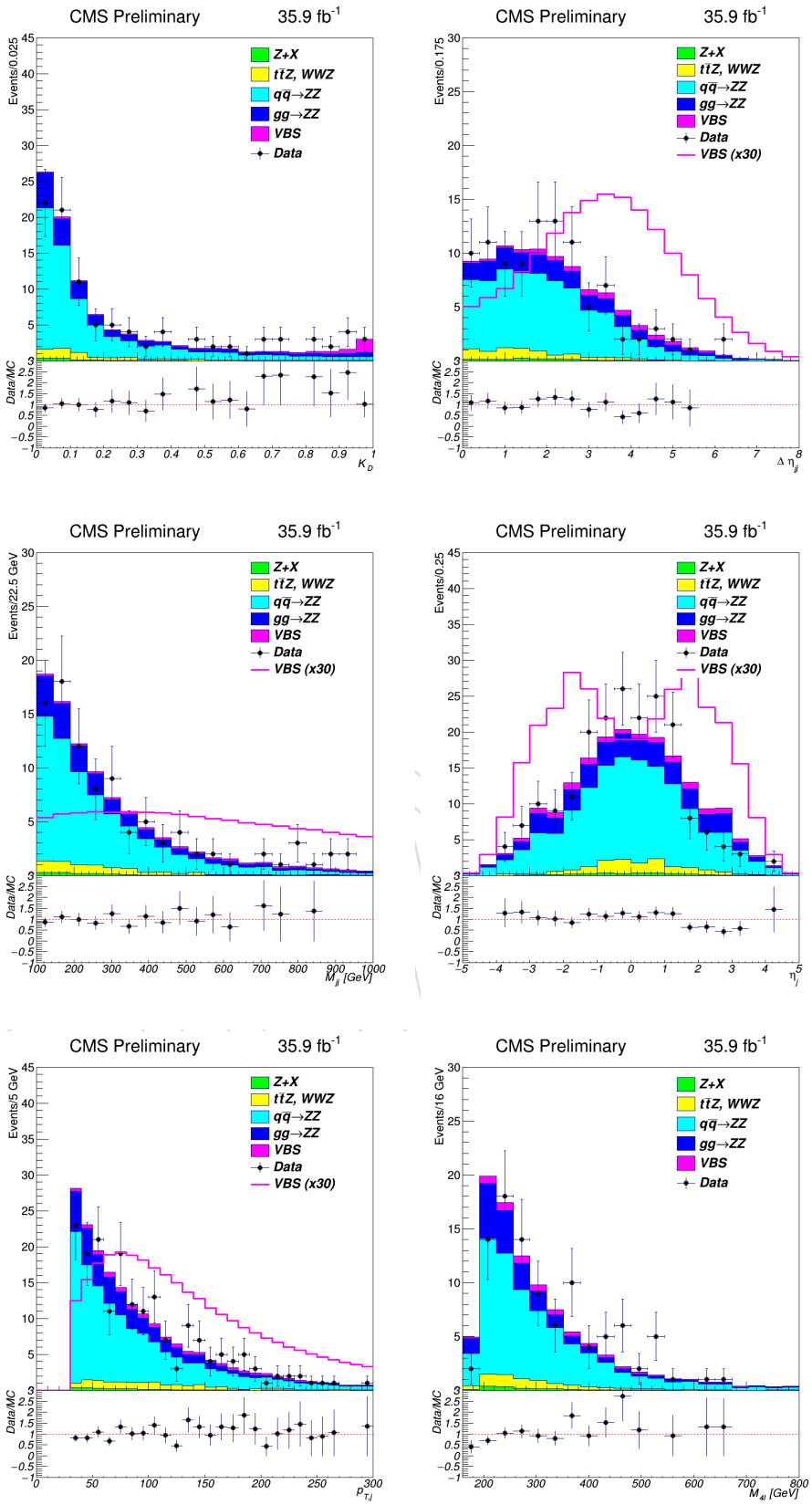


Figure 7: Comparison of data to background and signal estimations in 2016 samples. Top: MELA (left) and $\Delta\eta_{jj}$ (right). Middle: m_{jj} (left) and η_j (right). Low: $p_{T,j}$ (left) and $m_{4\ell}$ (right). Where relevant, the EWK signal is both added to the total distribution and superimposed multiplied by an illustrative scale factor of 30.

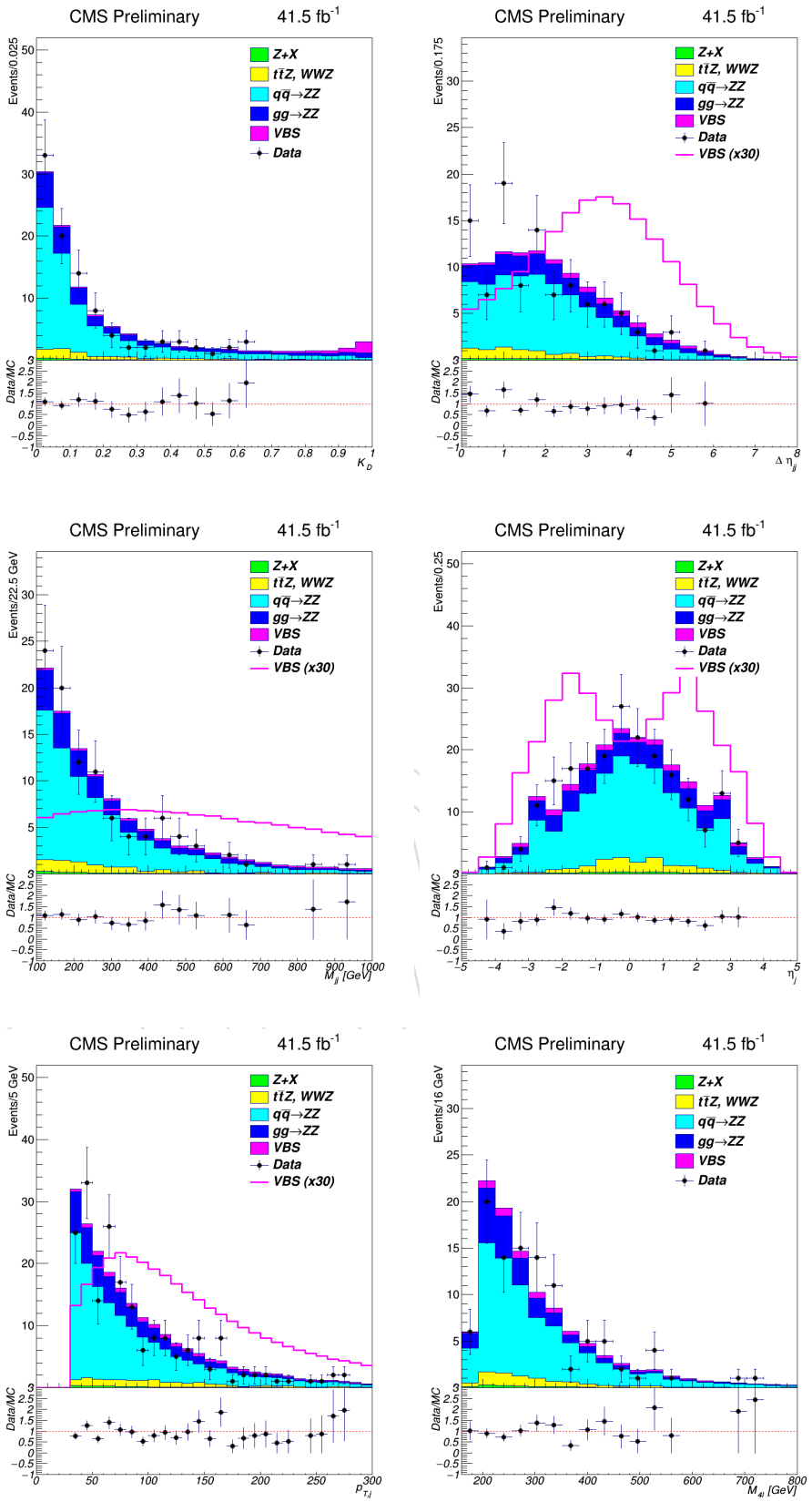


Figure 8: Comparison of data to background and signal estimations in 2017 samples. Top: MELA (left) and $\Delta\eta_{jj}$ (right). Middle: m_{jj} (left) and η_j (right). Low: $p_{T,j}$ (left) and $m_{4\ell}$ (right). Where relevant, the EWK signal is both added to the total distribution and superimposed multiplied by an illustrative scale factor of 30.

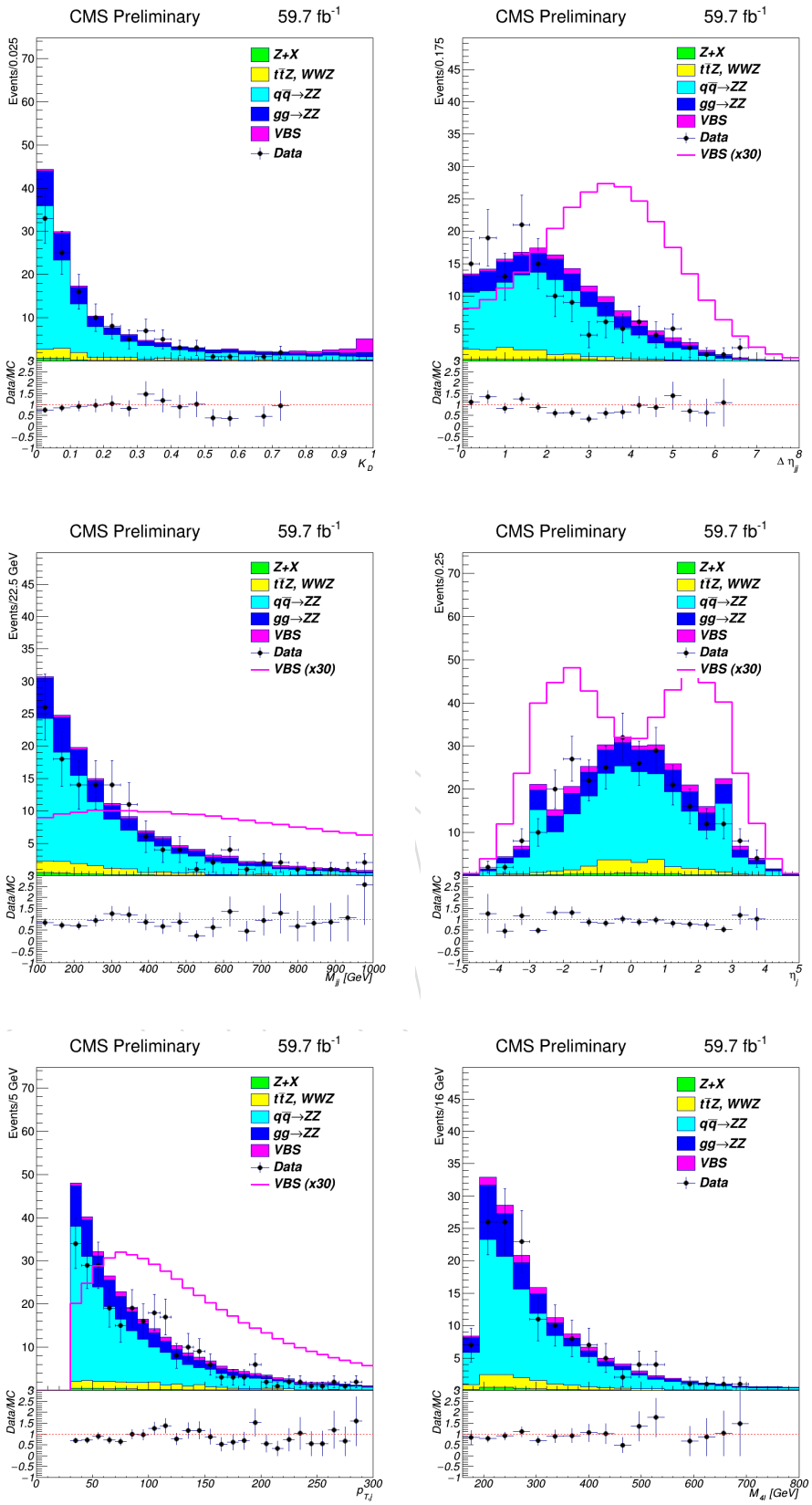


Figure 9: Comparison of data to background and signal estimations in 2018 samples. Top: MELA (left) and $\Delta\eta_{jj}$ (right). Middle: m_{jj} (left) and η_j (right). Low: $p_{T,j}$ (left) and $m_{4\ell}$ (right). Where relevant, the EWK signal is both added to the total distribution and superimposed multiplied by an illustrative scale factor of 30.

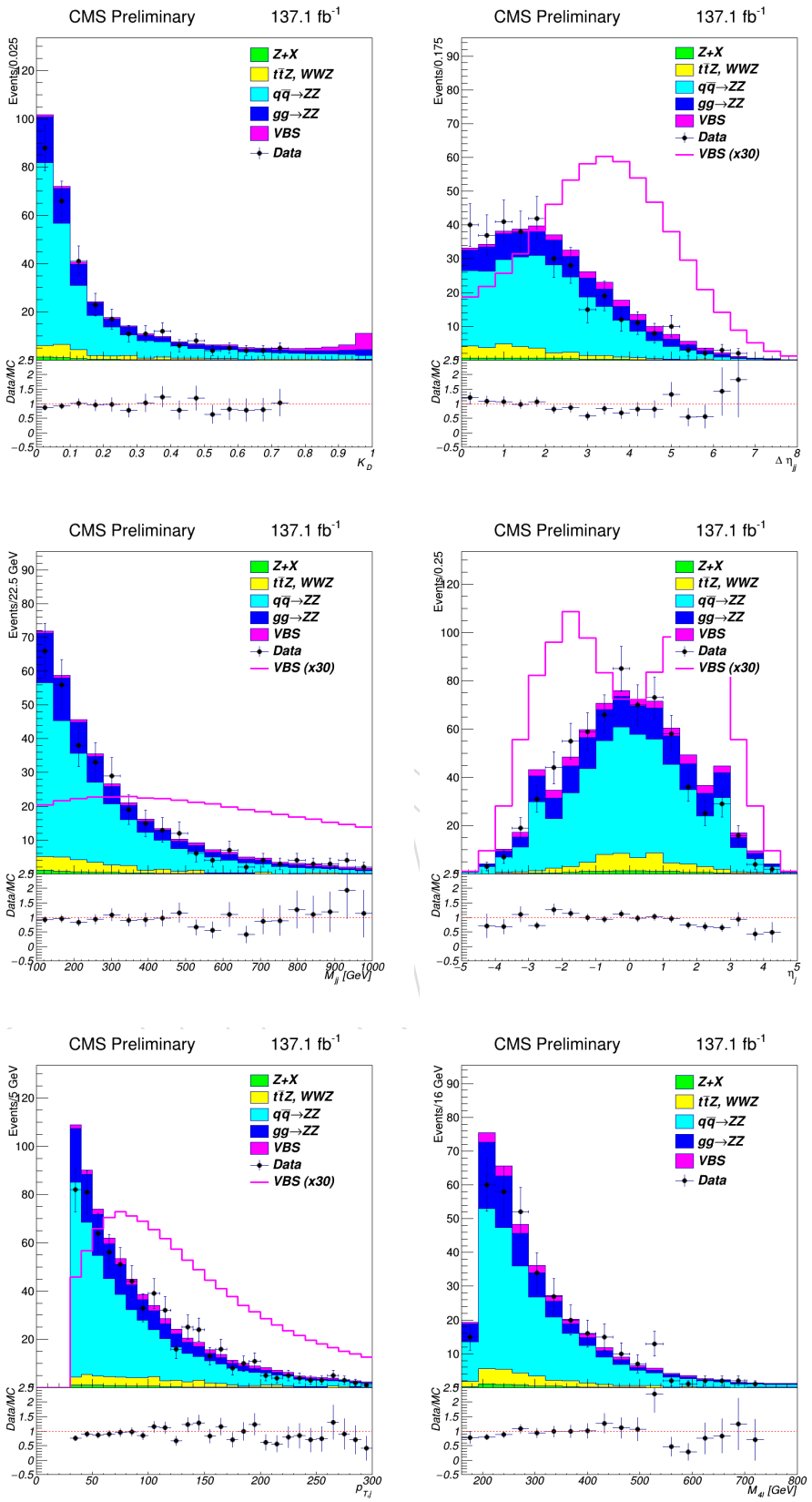


Figure 10: Comparison of data to background and signal estimations in samples from all years summed together. Top: MELA (left) and $\Delta\eta_{jj}$ (right). Middle: m_{jj} (left) and η_j (right). Low: $p_{T,j}$ (left) and $m_{4\ell}$ (right). Where relevant, the EWK signal is both added to the total distribution and superimposed multiplied by an illustrative scale factor of 30.

7 Systematic Uncertainties

The systematic uncertainties detailed below are taken into account in the statistical model via profiling of nuisance parameters according to a frequentist approach.

For the MELA-based signal extraction, both the yield variations and shape variations (if significant for a give uncertainty source) are considered and the resulting MELA output spectra are used in the template fit.

7.1 Theory Uncertainties

7.1.0.1 QCD scale uncertainties QCD scale uncertainties are estimated by simultaneously varying the renormalization and factorization scales, up and down by a factor of two with respect to the nominal values. Figure 11 shows the effect of the scale variations for the processes most relevant to the VBS search. As a pure electroweak process, the VBS signal exhibits a smaller uncertainty on the scale choice but there is a shape dependence, with highest-MELA events having a larger uncertainty. For this contribution, the MC result is used as systematic uncertainty. The same procedure is adopted for minor backgrounds. For dominant QCD processes, qqZZ and ggZZ, Figure 11 shows that scale uncertainties are MELA-independent. We therefore use MELA-independent uncertainties for qqZZ. For ggZZ, however, the uncertainty sizes relate to the QCD order used in MC (LO), which is corrected for higher orders in the analysis treatment. For these reasons, we use normalization-only uncertainties from our best estimate of the NLO-LO k-factor for qqZZ (Section 2.3).

7.1.0.2 PDF uncertainties Uncertainties related to the choice of the PDF and the strong coupling constant are evaluated following the prescriptions in [29]. Figure 12 shows the effect of the PDF variations for the processes most relevant to the VBS search. In principle, since PDFs are different in 2016 and 2017-18 (NNPDF3.0 NLO vs. NNPDF3.1 NNLO) these uncertainties could be year-dependent and therefore uncorrelated. We show in the top panel of Figure 12 that this is not the case, comparing uncertainties in 2016 and 2018 qqZZ samples. We do not use shape effects for these uncertainties, because the only small dependence (in VBS) is within the constant-fit errors. For correlation among processes, we assume as uncorrelated uncertainties those related to physics processes that originate from gg, qq or $q\bar{q}$ final states at the lowest order in QCD.

7.2 Experimental Uncertainties

7.2.0.1 Luminosity The uncertainty in the LHC integrated luminosity of the data sample is 2.3-2.5% [30]. Given that the correlated component among years is small and that the overal effect of systematic uncertainties in the measurements is also small, we take the uncertainty as uncorrelated among years.

7.2.0.2 Reducible background The uncertainty in the data-driven reducible background estimate is dominated by the statistical uncertainties in the control regions: it is final-state dependent and correspond to about 40% uncertainty in the yield.

7.2.0.3 MC sample size For the processes estimated from simulation the available statistics of the MC sample limits the precision of the modeling, and is therefore taken as a shape-dependent systematic uncertainty, uncorrelated over years.

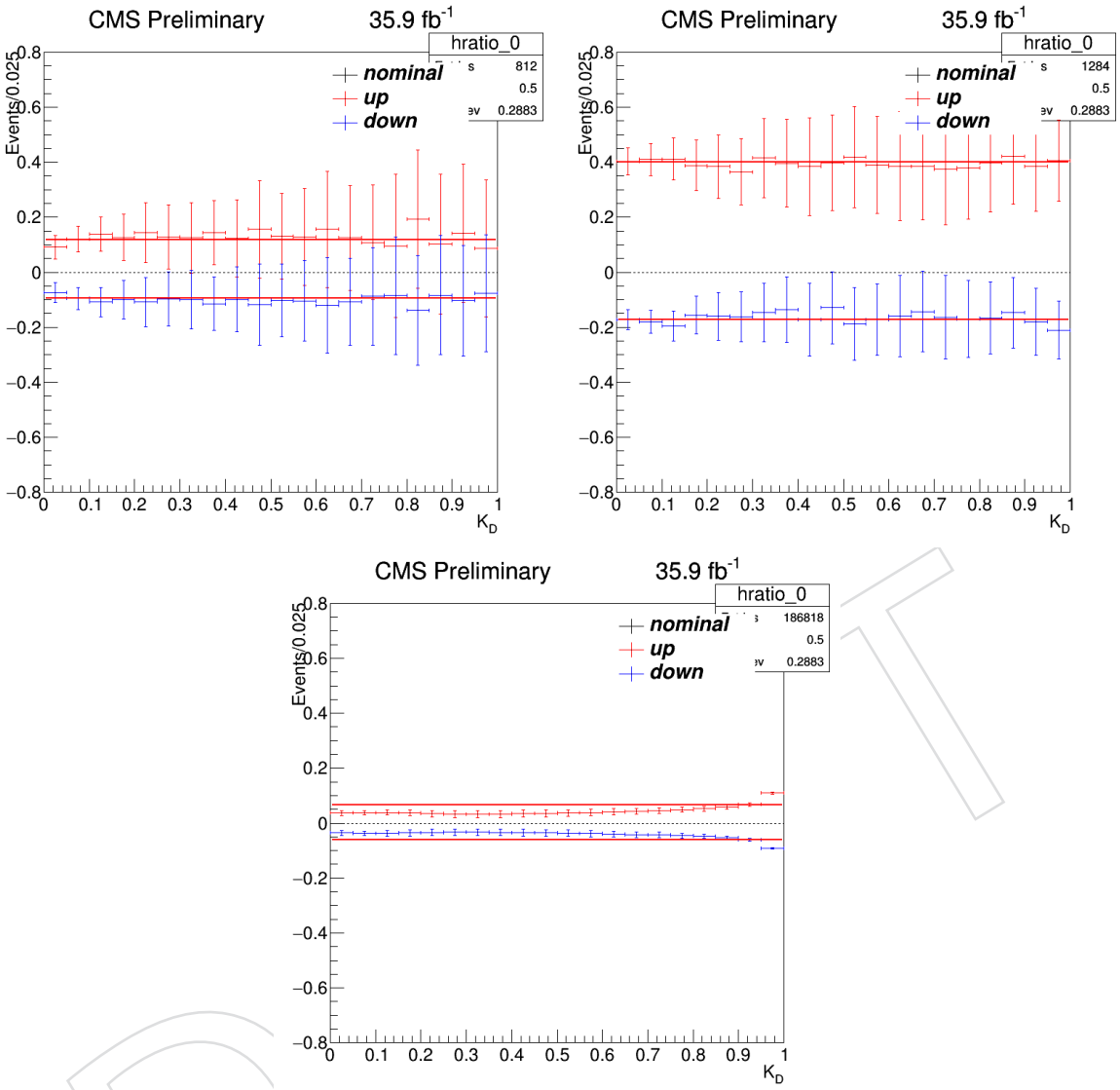


Figure 11: Systematic uncertainties due to the variation of the default factorizations and renormalization scales: non-loop-induced QCD background (left top), loop-induced QCD background (right top), and electroweak signal (bottom). All variations are fitted with a constant line.

7.2.0.4 Lepton trigger uncertainties Uncertainties arising from the trigger as well as lepton reconstruction and selection efficiencies range between 2.5% and 9%, depending on the final state [3].

7.2.0.5 JES/JER The JES and JER uncertainties are estimated in a similar way by varying the p_T of the tagging jets by their respective per-jet uncertainty obtained from the JETMET recipes [24]. The application of these factors affect yields (since events can migrate in and out the jet p_T thresholds) as well as shape, since MELA is recalculated for each event with the varied jet kinematics. The resulting templates include the shape and yield variations and are shown in Fig. 13 and 14 for the electroweak signal and and the dominant QCD backgrounds. For both, a small enhancement at low values of MELA is present for the "up" variation, indicating that new

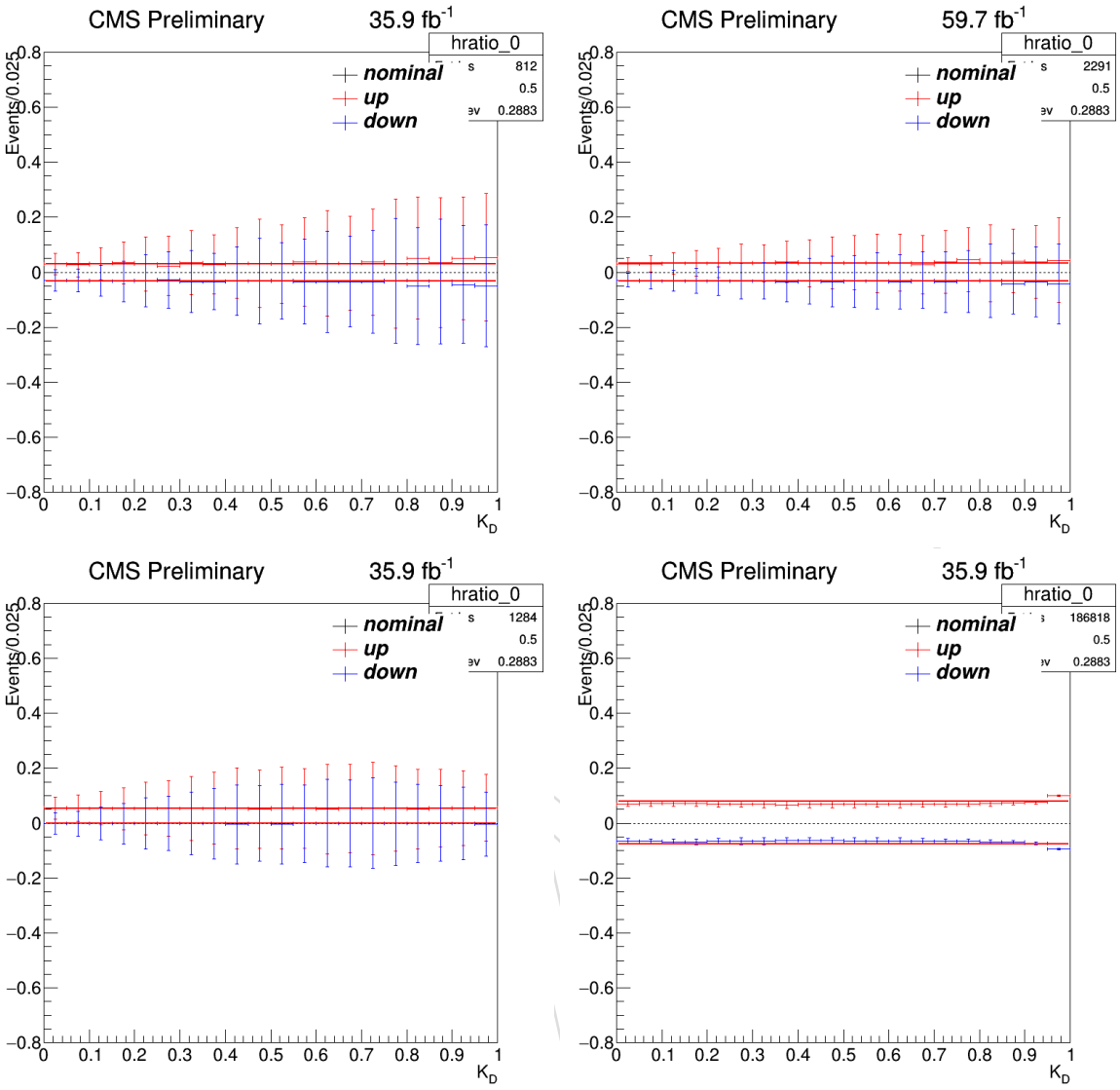


Figure 12: Systematic uncertainties due to the PDF and α_S variations: non-loop-induced QCD background in 2016 (left top), non-loop-induced QCD background in 2018 (right top), loop-induced QCD background (left bottom), electroweak signal (right bottom). All variations are fitted with a constant line.

events selected because more jets pass the p_T threshold are located at low VBS probabilities. For JER, the global effect is much smaller.

7.2.0.6 L1 prefiring

Weight variations for the L1 prefiring corrections are determined following the recipe in [25].

Uncertainty sizes are summarized in Table 16.

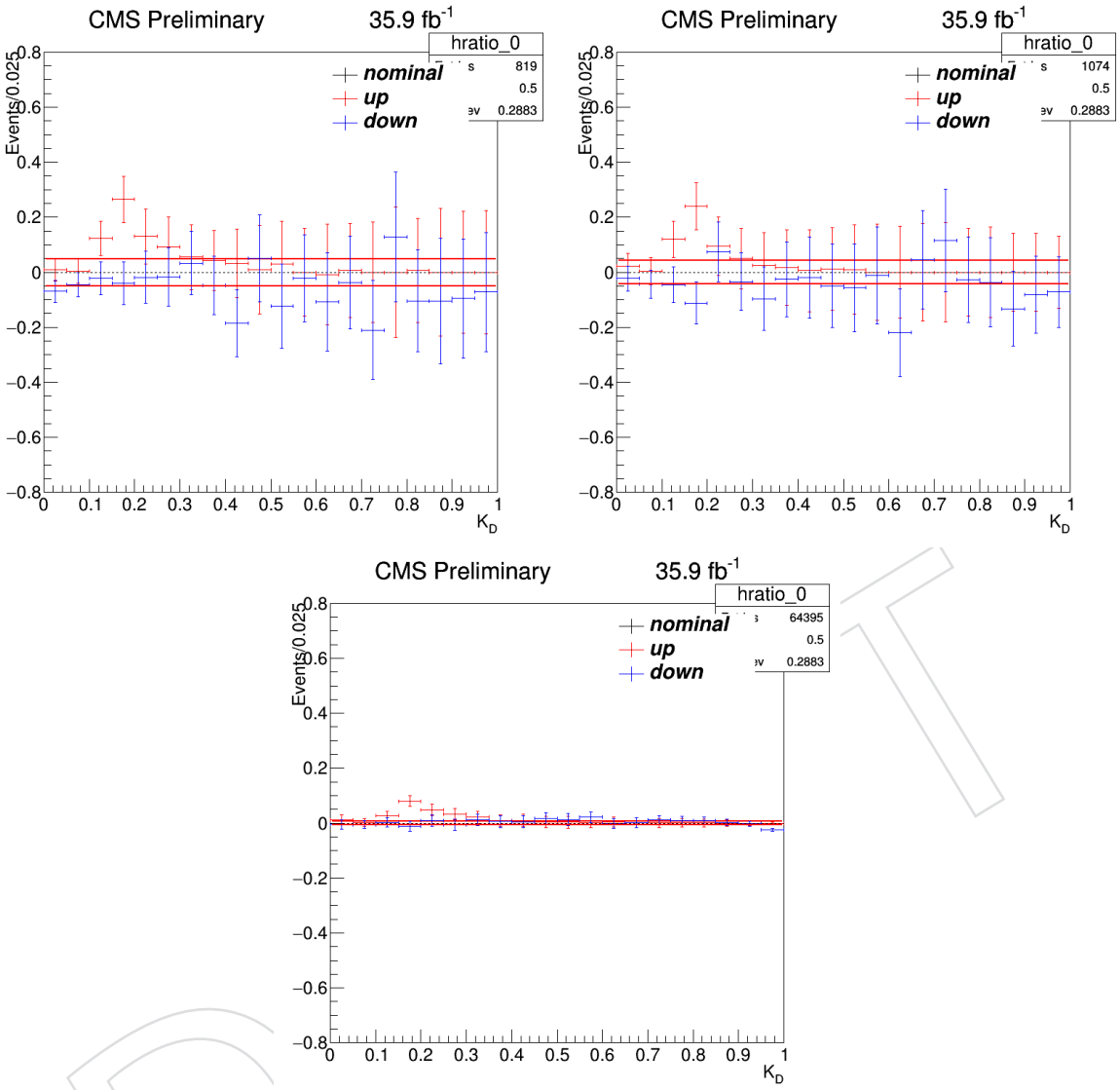


Figure 13: Systematic uncertainties due to the JES variations: non-loop-induced QCD background (top left) loop-induced QCD background (top right) and electroweak signal (bottom).

Table 16: Estimated systematic uncertainties on the signal yield. Minor backgrounds are not shown, the systematics being totally dominated by the MC sample size (19-24%).

Systematic source	qqZZ	ggZZ	VBS	Z+X	Shape	Years corr.?
Lepton trigger, reco, sel.	2.5-9%	2.5-9%	2.5-9%	-		x
Luminosity	2.3-2.5%	2.3-2.5%	2.3-2.5%	-		
Reducible background	-	-	-	33-45%		
QCD scales	10-12%	9-14%	6%	-	x	x
PDF + α_s	3.2%	5%	6.6%	-		x
JES	4.9-5.1%	2.4-2.6%	0.7%	-	x	
JER	2.2-2.4%	1.0-1.1%	0.2%	-		
MC samples	2.5-4.2%	3.2%	$\ll 1\%$	-	x	
L1 prefiring	0.6-1.0%	0.6%	1.8-3.0%	-		

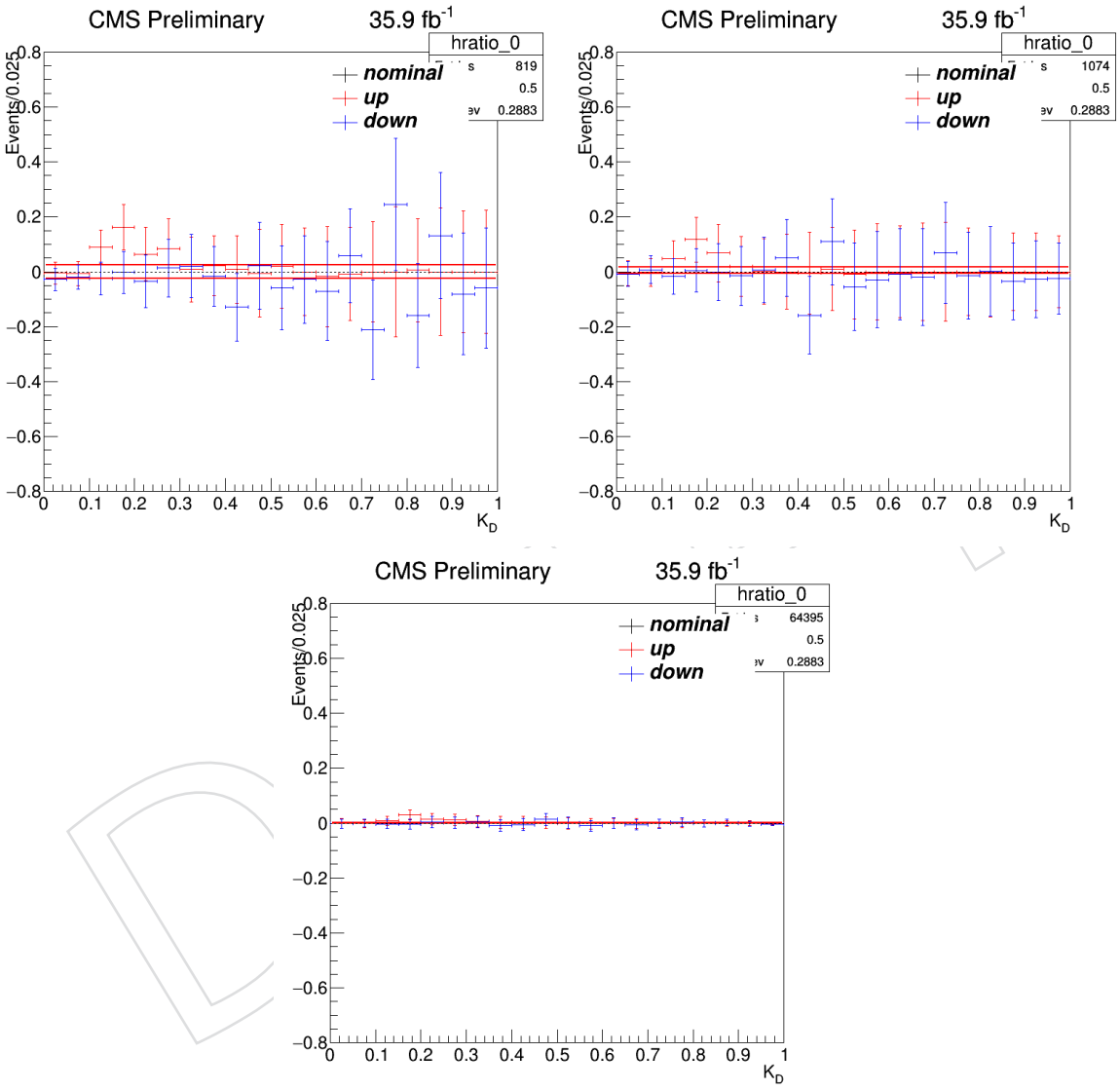


Figure 14: Systematic uncertainties due to the JER variations: non-loop-induced QCD background (top left) loop-induced QCD background (top right) and electroweak signal (bottom).

8 Significance of electroweak signal with the MELA and DNN methods

The signal and background yields for the ZZjj inclusive and VBS cut-based selection are summarized in Table 17.

Table 17: Signal and background yields for the ZZjj baseline and VBS selection.

Year	Signal (ZZjj EWK)	Z+X	$q\bar{q} \rightarrow$ ZZjj QCD	$gg \rightarrow$ ZZjj QCD	$t\bar{t}$ +WWZ	Data
			ZZjj baseline			
2016	5.6 ± 0.6	1.5 ± 0.6	61.4 ± 5.6	19.8 ± 2.7	6.3 ± 0.9	100
2017	6.1 ± 0.7	1.3 ± 0.5	67.9 ± 6.2	22.8 ± 3.1	8.1 ± 1.2	
2018	9.5 ± 1.1	2.5 ± 0.9	98.2 ± 9.0	32.9 ± 4.5	11.9 ± 1.7	
all	21.1 ± 2.3	5.3 ± 2.0	227.4 ± 20.8	75.5 ± 10.3	26.3 ± 3.8	
			VBS enriched			
2016	4.0 ± 0.4	0.2 ± 0.1	8.6 ± 0.8	4.6 ± 0.6	0.8 ± 0.1	19
2017	4.6 ± 0.5	0.3 ± 0.1	10.1 ± 0.9	5.4 ± 0.7	1.0 ± 0.1	
2018	6.3 ± 0.7	0.4 ± 0.2	16.5 ± 1.5	7.7 ± 1.1	1.7 ± 0.2	
all	14.9 ± 1.7	0.9 ± 0.3	35.2 ± 3.2	17.7 ± 2.4	3.4 ± 0.5	

The expected significances of the VBS for the full dataset with an integrated luminosity of $\mathcal{L} = 137.1\text{fb}^{-1}$ are calculated using the "combine" tool, where the test statistics is the profile log-likelihood in an asymptotic limit: each data-taking year and lepton final-state enters the combination as a different contribution.

The template analysis uses the MELA spectrum for all the contributions as determined in Section 6 and uses nuisance parameters defining all shape and normalization systematics as presented in Section 7. Preliminary systematic (no systematic) uncertainties yields an expected significance of 3.32 (3.51) standard deviations. For a cross-check of the past analysis, the same result with 2016 samples only is 1.76 (1.82), in good agreement with [1, 2] (considering analysis improvements, such as the new ggZZ description) and with a naive luminosity scaling.

9 Total and electroweak cross-sections in fiducial regions

The EWK and EWK+QCD cross-sections are estimated in fiducial regions, mimicking the selections at reconstruction level in order to reduce extrapolation to the minimum. The particle-level selections applied to define the fiducial regions in the baseline and VBS-enriched selections are detailed in Table 18.

The cross-section results with an integrated luminosity of $\mathcal{L} = 137.1\text{fb}^{-1}$ are calculated using the "combine" tool to perform a maximum-likelihood fit. The MELA spectrum is optimized to isolate the EWK signal, so the cross-section results for the EWK component use a MELA shape analysis, with all details identical to the previous section. Since the EWK+QCD determination is essentially background-free, we use for this measurement an event counting analysis which has less theoretical dependency.

Table 19 reports the SM cross-sections in the fiducial regions, the fitted value of the signal strength μ with its statistical and systematic uncertainty and the resulting measured cross-sections.

Table 18: Particle-level selections used to define the fiducial regions for EWK+QCD and EWK cross-sections.

Object	Selection
	ZZjj baseline
Leptons	$p_{\mathrm{T}}(\ell_1) > 20$ GeV
	$p_{\mathrm{T}}(\ell_2) > 10$ GeV
	$p_{\mathrm{T}}(\ell) > 5$ GeV
	$\lvert\eta(\ell)\rvert < 2.5$
	(γ with $\Delta R(\ell,\gamma) < 0.1$ added to ℓ 4-vector)
Z and ZZ	$60 < m(\ell\ell) < 120$ GeV
	$m(4\ell) > 180$ GeV
Jets	at least 2
	$p_{\mathrm{T}}(j) > 30$ GeV
	$\lvert\eta(j)\rvert < 4.7$
	$m_{jj} > 100$ GeV
	$\Delta R(\ell, j) > 0.4$ for each ℓ, j
	VBS-enriched
	All of the above +
Jets	$\Delta\eta(jj) > 2.4$
	$m_{jj} > 400$ GeV

Table 19: SM cross-sections in the fiducial regions, the fitted value of the signal strength μ with its total uncertainty (statistical only in parenthesis) and the resulting measured cross-sections.

	μ_{exp}	SM σ (fb)	μ_{obs}	Measured σ (fb)
	ZZjj baseline			
EWK	$1.00\ ^{+0.42}_{-0.36}\ (^{+0.40}_{-0.35})$	0.275 ± 0.021		
EWK+QCD	$1.00\ ^{+0.12}_{-0.11}\ (\pm 0.06)$	5.35 ± 0.21		
	VBS-enriched (loose)			
EWK	$1.00\ ^{+0.45}_{-0.38}\ (^{+0.41}_{-0.36})$	0.186 ± 0.015		
EWK+QCD	$1.00\ ^{+0.16}_{-0.15}\ (^{+0.13}_{-0.12})$	1.21 ± 0.05		
	VBS-enriched (tight)			
EWK	$1.00\ ^{+0.xx}_{-0.xx}\ (^{+0.xx}_{-0.xx})$			
EWK+QCD	$1.00\ ^{+0.xx}_{-0.xx}\ (^{+0.xx}_{-0.xx})$			

The impact of the various shape and normalization systematics can be seen in Figure 15.

10 Limits on anomalous quartic gauge couplings

The events in the baseline selection are used to constrain anomalous quartic gauge couplings in an effective field theory approach [31, 32]. The ZZjj channel is sensitive to the neutral tensor operators T8 and T9, as well as the tensor operators T0, T1, and T2, which increase the production cross section at large invariant masses of the ZZ system. Limits on the couplings $f_{\mathrm{T},i}/\Lambda^4$ are derived based on the invariant mass distribution of the 4 leptons, following previous analysis

of anomalous couplings in this channel [1].

The expected distributions for different values of the operator couplings are obtained using the reweighing feature of the MadGraph package. Dedicated MadGraph samples exploiting the SM_LT8_LT9 UFO model file is used in the aQGC analysis:

```
generate p p > z z j j QED=5 QCD=0 NP=1
```

The default coupling for the event generation is set to $f_{T8}/\Lambda^4 = 2$ TeV^{-4} in order to increase the statistics at large scattering energies. Alternative coupling strengths are then obtained by means of reweighting. The method uses event weights w_{new} to reweigh the nominal event sample to the alternative hypotheses of the coupling strength:

$$w_{new} = w_{old} \frac{|\mathcal{M}_{new}|^2}{|\mathcal{M}_{old}|^2},$$

where $\mathcal{M}_{old}$ is the nominal matrix element and $\mathcal{M}_{new}$ is the matrix element with the modified coupling strengths. Weights are generated on a 2D grid in f_{T8}/Λ^4 and f_{T9}/Λ^4 with coupling strengths equal to [0.25, 0.5, 1, 2, 4, 8, 16] TeV^{-4}, including the mixed and negative coupling strength configurations.

A semi-analytic description of the expected m_{ZZ} distribution as a function of the aQGC couplings is obtained by fitting quadratic functions to the ratio of the aQGC and Standard Model yield in each m_{ZZ} bin. Figures 16 and 17 show the expected yield ratio in m_{ZZ} bins for the parameter points and the result of the quadratic fit. As expected the quadratic function provides a good model for the yield ratio as a function of the coupling.

Figure 18 shows the expected m_{ZZ} distributions for the SM and for two aQGC scenarios as well as for the data, for the 2016 data set.

Confidence levels on the operator couplings are derived using the "combine" tool. The test statistics is the same log-likelihood ratio used for the EW signal significance, again with all systematic uncertainties profiled as nuisance parameters. The confidence limits are determined using Wilk's theorem and the assumption that the likelihood approaches a χ^2-distribution with one degree of freedom. The 95% confidence level is then determined by finding the coupling strength that yields a likelihood ratio of 3.84 for the 1D limits and 5.99 for the 2D limits.

Table 20 lists the individual confidence level (CL) obtained for the combined dataset, setting the other coupling to zero, as well as the unitarity limit . The same systematic uncertainties described in the EW signal significance determination are considered for the limit setting. The expected pre-fit yields are used for the backgrounds and the EW signal for the expected limits.

11 Summary

A search was made for the electroweak production of two jets in association with two Z vector bosons in the four-lepton final state in proton-proton collisions at 13 TeV. The data correspond to an integrated luminosity of 137 fb^{-1} collected with the CMS detector at the LHC.

The electroweak production of a pair of Z bosons in association with two jets is measured with an observed (expected) significance of XX.X (YY.Y) standard deviations. The fiducial cross section is measured to be $\sigma_{\text{fid}} = XXX^{+Y}_{-Z}(\text{stat})^{+Y}_{-Z}(\text{syst})$ fb, which is consistent with the standard model prediction.

Table 20: Observed and expected lower and upper 95% CL limits on the coupling of the quartic tensor operators T0, T1 and T2, as well as the neutral current operators T8 and T9. The unitarity limits are also listed. All couplings are in TeV^{-4}, the unitarity limits are in TeV.

Coupling	Exp. lower	Exp. upper	Obs. lower	Obs. upper	Unitarity limit
f_{T0}/Λ^4	-0.53	0.52	xx	xx	xx
f_{T1}/Λ^4	-0.71	0.71	xx	xx	xx
f_{T2}/Λ^4	-1.42	1.39	xx	xx	xx
f_{T8}/Λ^4	-0.99	0.99	xx	xx	xx
f_{T9}/Λ^4	-2.12	2.12	xx	xx	xx

Limits on anomalous quartic gauge couplings are set at 95% confidence level in terms of effective field theory operators, with units in TeV^{-4}:

$$-XXX < f_{T_0}/\Lambda^4 < YYY$$
$$-XXX < f_{T_1}/\Lambda^4 < YYY$$
$$-XXX < f_{T_2}/\Lambda^4 < YYY$$
$$-XXX < f_{T_8}/\Lambda^4 < YYY$$
$$-XXX < f_{T_9}/\Lambda^4 < YYY$$

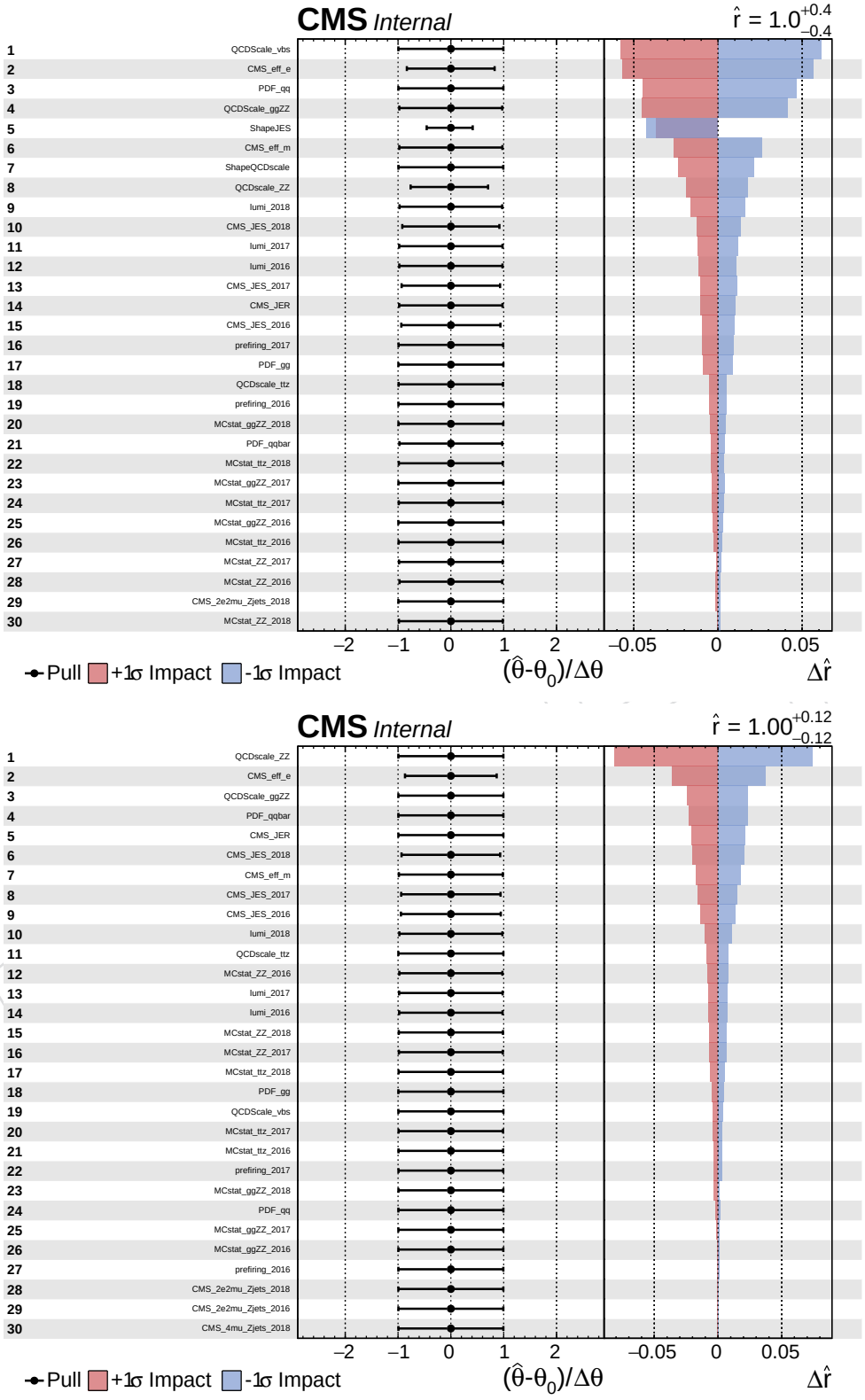


Figure 15: Impact of the various shape and normalization systematics for the EWK (top) and EWK+QCD cross-section fit (bottom) in the ZZjj baseline selection, as provided by the "combine" tool.

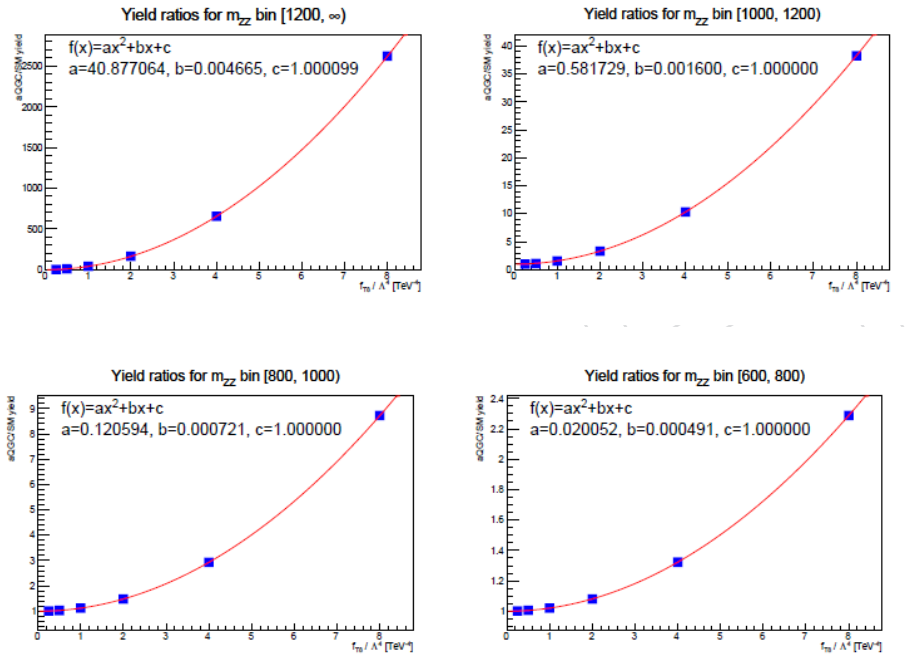


Figure 16: Yield ratios of the discrete operator couplings f_{T8}/Λ^4 obtained from the reweighing and the fitted quadratic interpolation for the most relevant mass bins used in the statistical analysis.

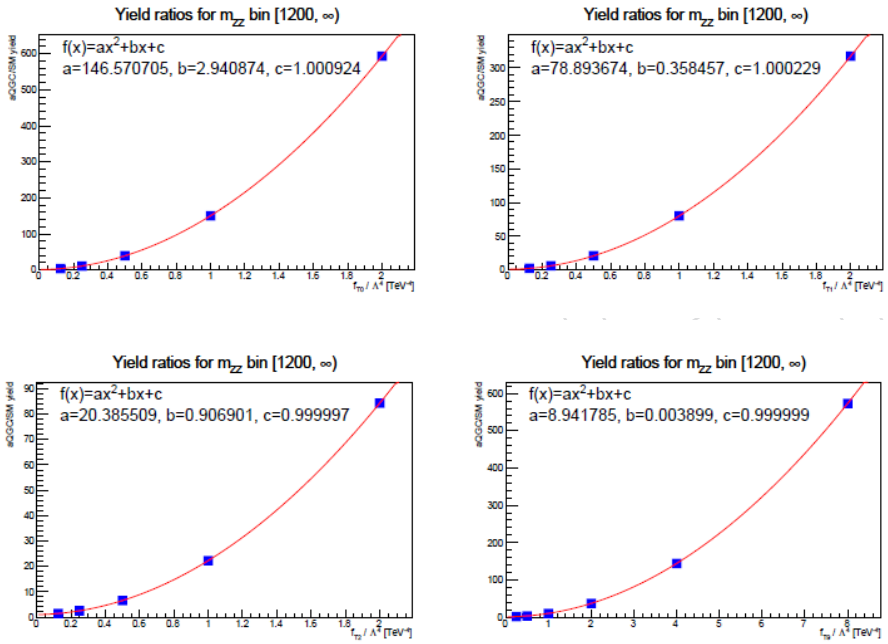


Figure 17: Yield ratios of the discrete operator couplings obtained from the reweighing and the fitted quadratic interpolation for each of the mass bins used in the statistical analysis. Shown is the last mZZ bin of the distribution for the f_{T0}/Λ^4 (top left), f_{T1}/Λ^4 (top right), f_{T2}/Λ^4 (bottom left), and f_{T9}/Λ^4 (bottom right) operators.

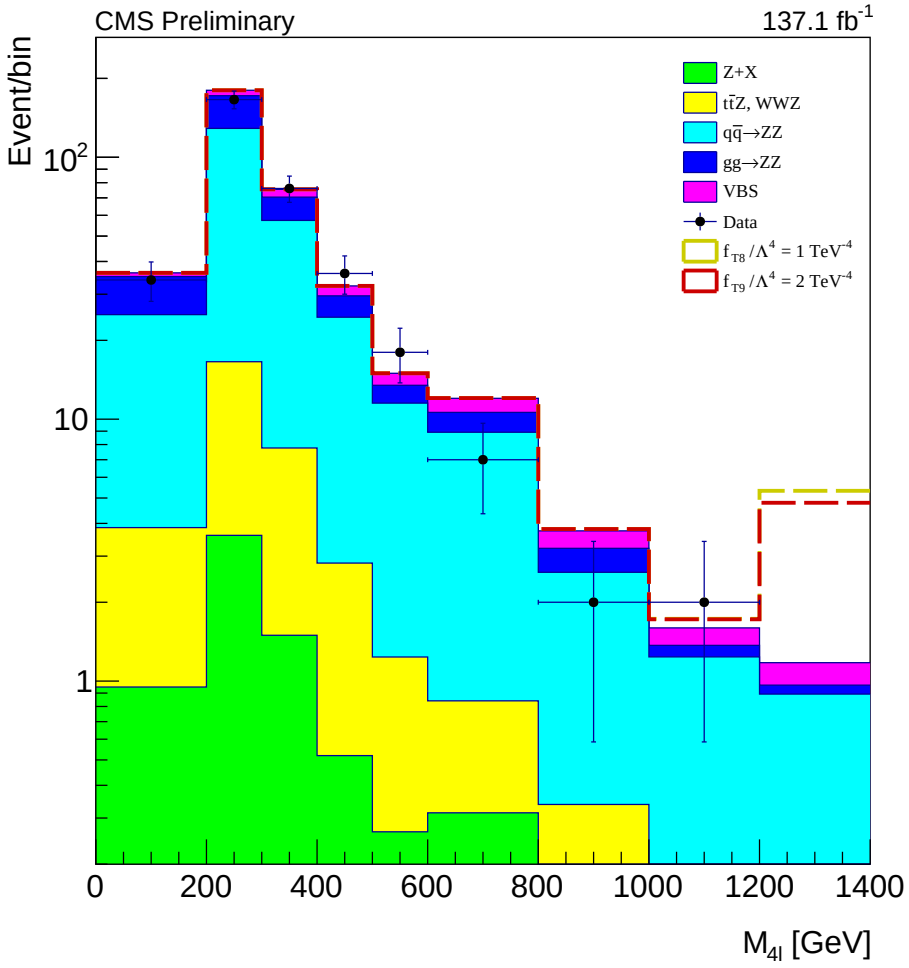


Figure 18: The m_{ZZ} distributions in the ZZjj selection for the full Run II dataset together with the SM prediction and two hypotheses for the aQGC coupling strength: $f_{T8}/\Lambda^4 = 1$ TeV (yellow dashed line) and $f_{T9}/\Lambda^4 = 2$ TeV (red dashed line).

A Alternative VBS signal extraction methods

A.1 Deep Neural Network signal extraction

We want to use DNN (Deep Neural Network) to distinguish signal and background events with a set of feature variables in each event. A deep neural network is a model consisting of multiple layers that is used to learn increasingly meaningful representations of data through successive layers (see [33]). Typically we feed training data into the neural network so that it can recognize patterns inside the data, and we monitor the training process with validation data. Then after verifying the trained model with testing data, we use it to make new predictions. To build and apply the neural network, we use Keras [34], a convenient high-level neural networks API written in Python, which in our case runs on top of Tensorflow [35], an open source machine learning platform.

With Keras, we build a simple neural network suitable for binary classification problem with 3 fully connected (Dense) layers. The first two layers have 32 hidden units and use relu activation, while the third layer outputs a single scalar value (predicted probability) with sigmoid activation. For optimizer we use RMSprop with learning rate 0.001, and for the loss function we use "binary_crossentropy". We preprocessed the input data samples by normalizing each feature variable to have mean equal to 0 and standard deviation equal to 1 across the training and validation data, and then repeated the same operation on the testing data (subtracting mean_training&val from each feature variable and then dividing it by std_training&val, instead of using mean_testing and std_testing for normalization). We trained for 50 epochs with batch size 512.

To prepare the samples, we mixed around 2.16 million ZZjj EWK, ggZZ, and $q\bar{q}$ZZ 2017 MC events in 4e,4m and 2e2m channel from previously produced ntuple files, each tagged as either signal or background event (1 or 0). Currently 9 feature variables are being used to distinguish signal and background: m_{jj}, $\Delta\eta_{jj}$, m_{4l}, $\Delta\phi(Z1, Z2)$, η_{j1}, η_{j2}, pt_{j1},pt_{j2}, n_{Jets}. Around 1.86 million of these events were put into the DNN built with Keras for training, with 50,000 additional events used for validation. Then around 250,000 remaining events were used for testing the trained model. The ratio of EWK: ggZZ: $q\bar{q}$ZZ test events is about 1:3:10. The test result has ROC auc 0.936 indicating reasonably good distinguishing power of the trained DNN model.

A.2 Boosted Decision Tree signal extraction

As a continuation of 2016 study, Boosted Decision Tree (BDT) method was employed to perform the signal extraction. A decision tree is a statistical method in which a set of features is used to split the input data based on those features. Various methods exist which utilize several decision trees instead of just one to maximize the predictive power and robustness of the decision tree.
In this study the adaptive boosting (AdaBoost) method was used. BDT classifier was trained to discriminate between the signal (EWK ZZjj) and the main QCD background (QCD ZZjj). Number of trees is set to 400 and the maximum depth to 3. Checks were made to make sure that the results are stable with respect to the change in hyperparameters.
On Figs. 20 - 24 we show all variables that were explored for training the BDT. Same plots are shown for 2017 and 2018 periods on Figs. 25 - 29 and 30 - 34 respectively.

In the 2016 study [1, 2], after studying variable importance and associated modelling uncertainties, 7 of them were finally selected. These are listed in the Table 21. The resulting BDT classifier distribution (BDT7) is shown on the Fig. 35 (left). The same figure shows that no overtraining was observed. The corresponding ROC curve is shown on the right plot.

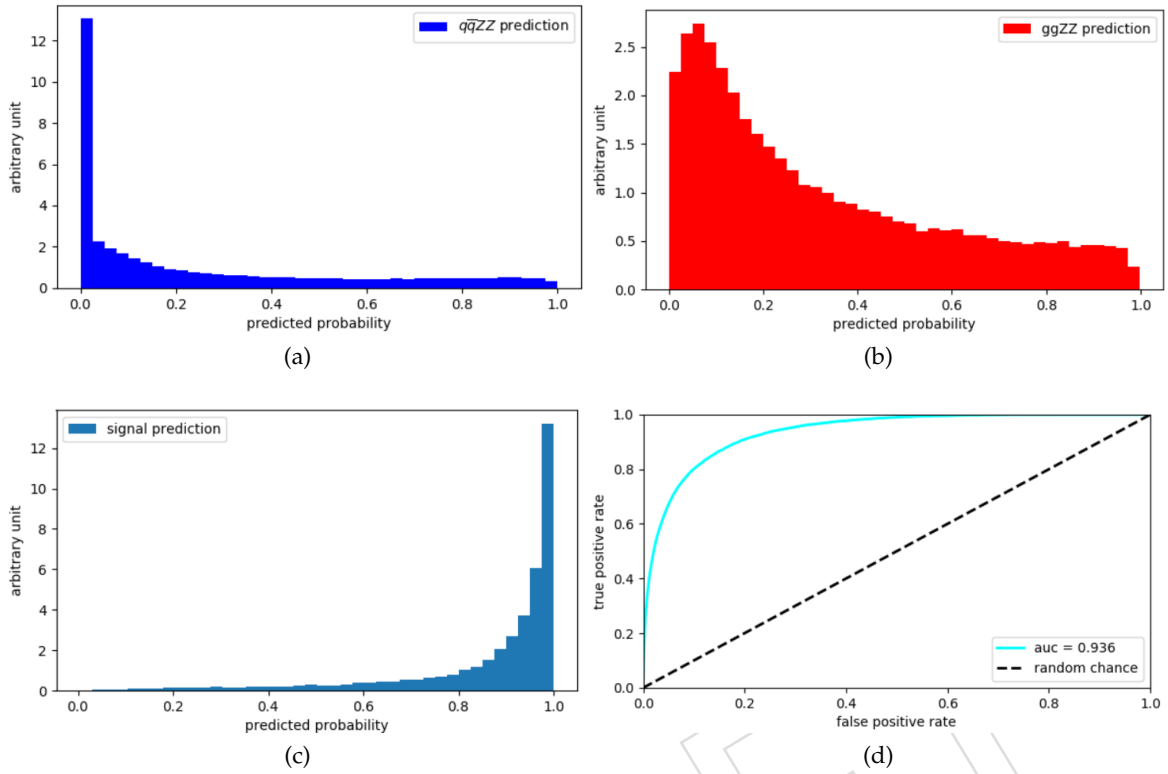


Figure 19: DNN shapes for qqZZ, ggZZ, EWK ZZ and ROC curve.

	Variable
1	m_{jj}
2	$\Delta\eta_{jj}$
3	m_{4l}
4	$\eta^*_{Z_1}$
5	$\eta^*_{Z_2}$
6	$R(p_T^{hard})$
7	$R(p_T^{jets})$

Table 21: List of final variables used in 2016 BDT [1, 2].

In order to assess the possible gain in the significance, 28 variables, including the ones used in BDT7 study, were used to train another BDT. A complete list of variables is shown in the Table 22. The resulting BDT classifier distribution (BDT28) with overtraining test is shown on the Fig. 36 (left) with the corresponding ROC curve (right). Again, no overtraining is observed. It should be noted that events with negative weight were included in the training of the BDT. It has been checked that removing the negative weights had no effect on the performance.

In this study, shape of the BDT classifier was used to derive the expected significance. BDT7 was trained for 2016, 2017 and 2018 periods and the resulting classifier distributions are shown on Fig. 37. Resulting significances for three periods and combined significance are shown in Table 23. Comparing these results with the ones obtained with MELA, we see the increase of 14% (12% without systematics) in expected significance.

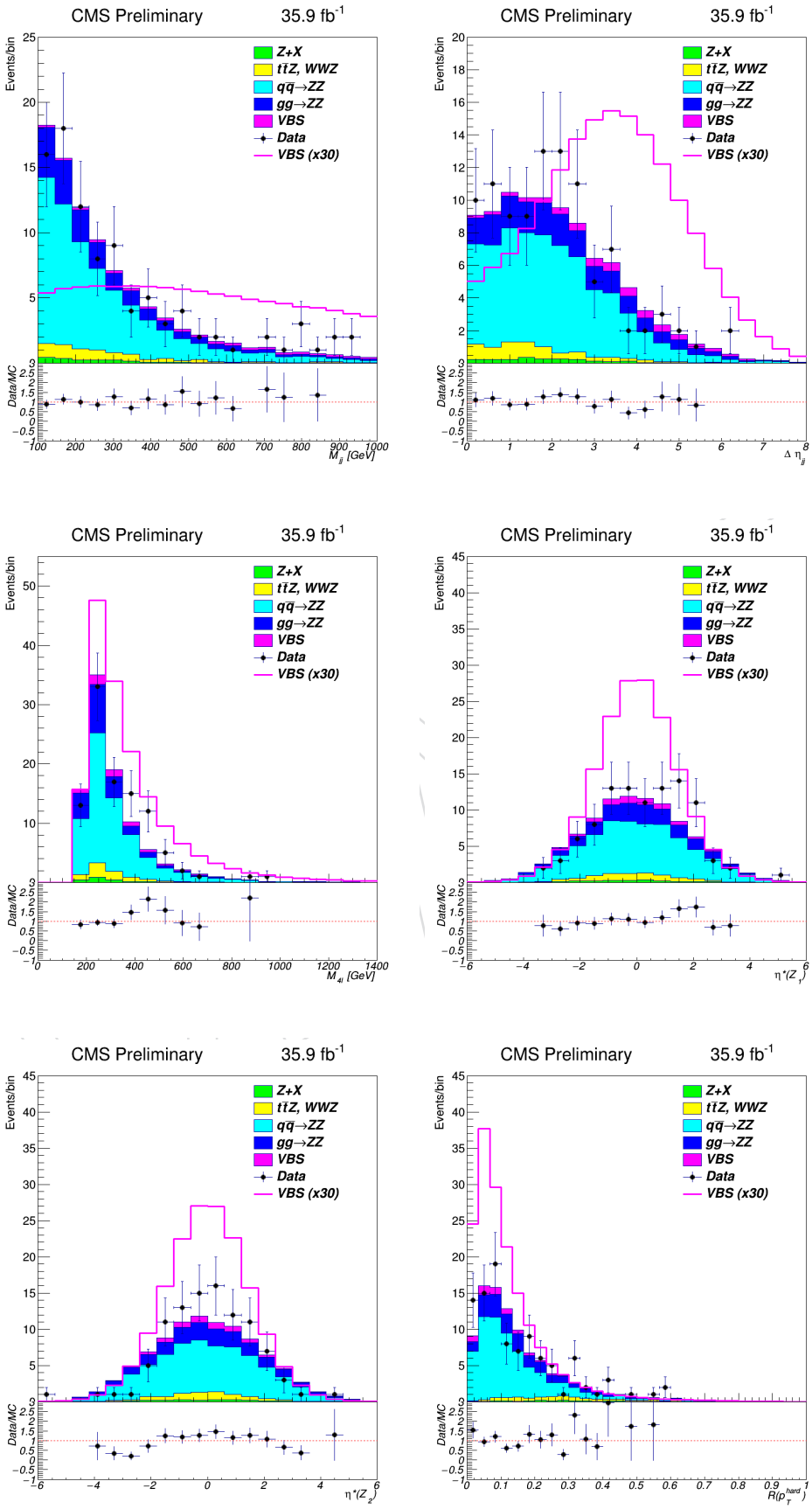


Figure 20: Variables considered for training the BDT classifier. Distributions for 2016 period are shown.

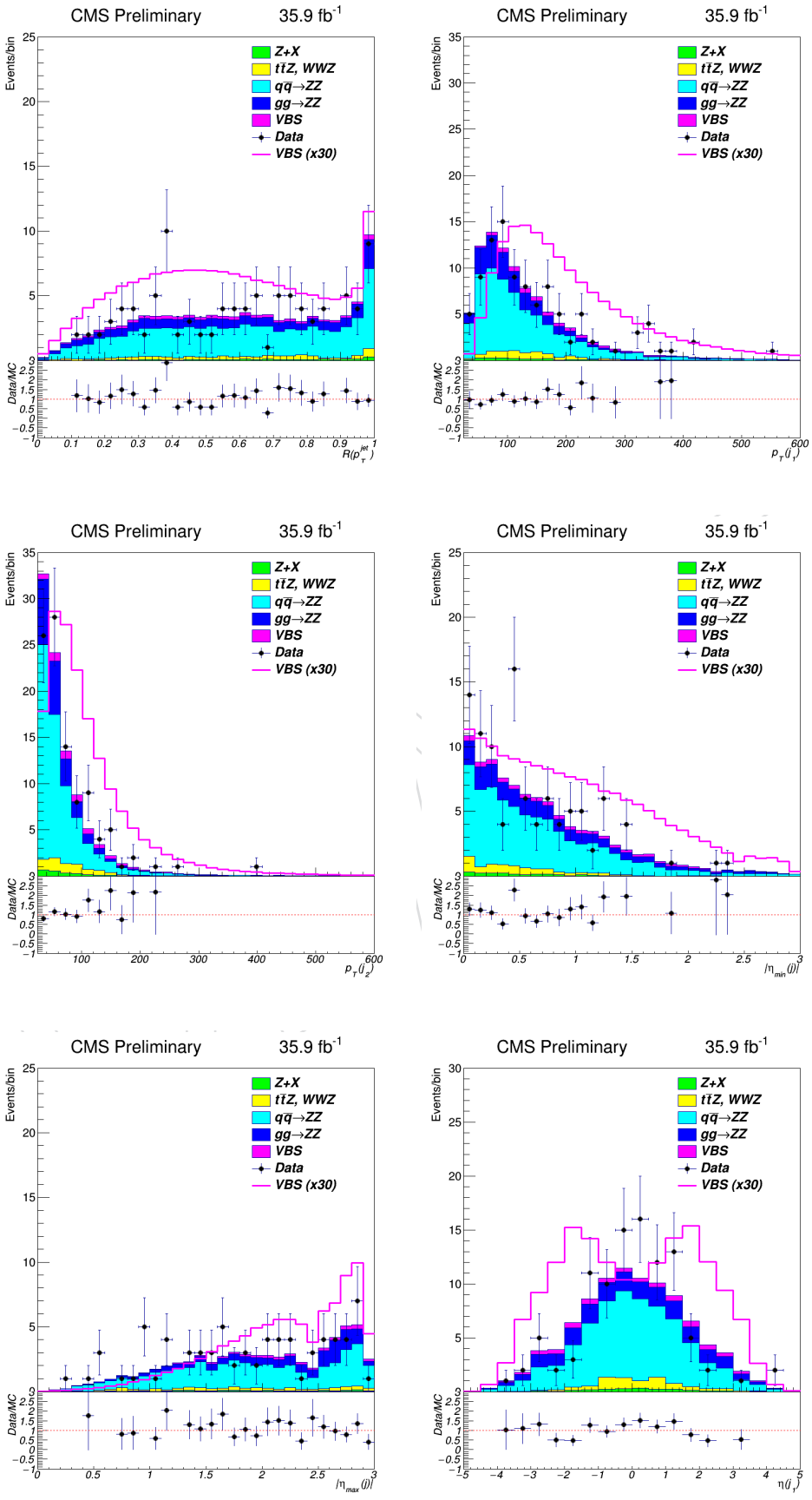


Figure 21: Variables considered for training the BDT classifier. Distributions for 2016 period are shown.

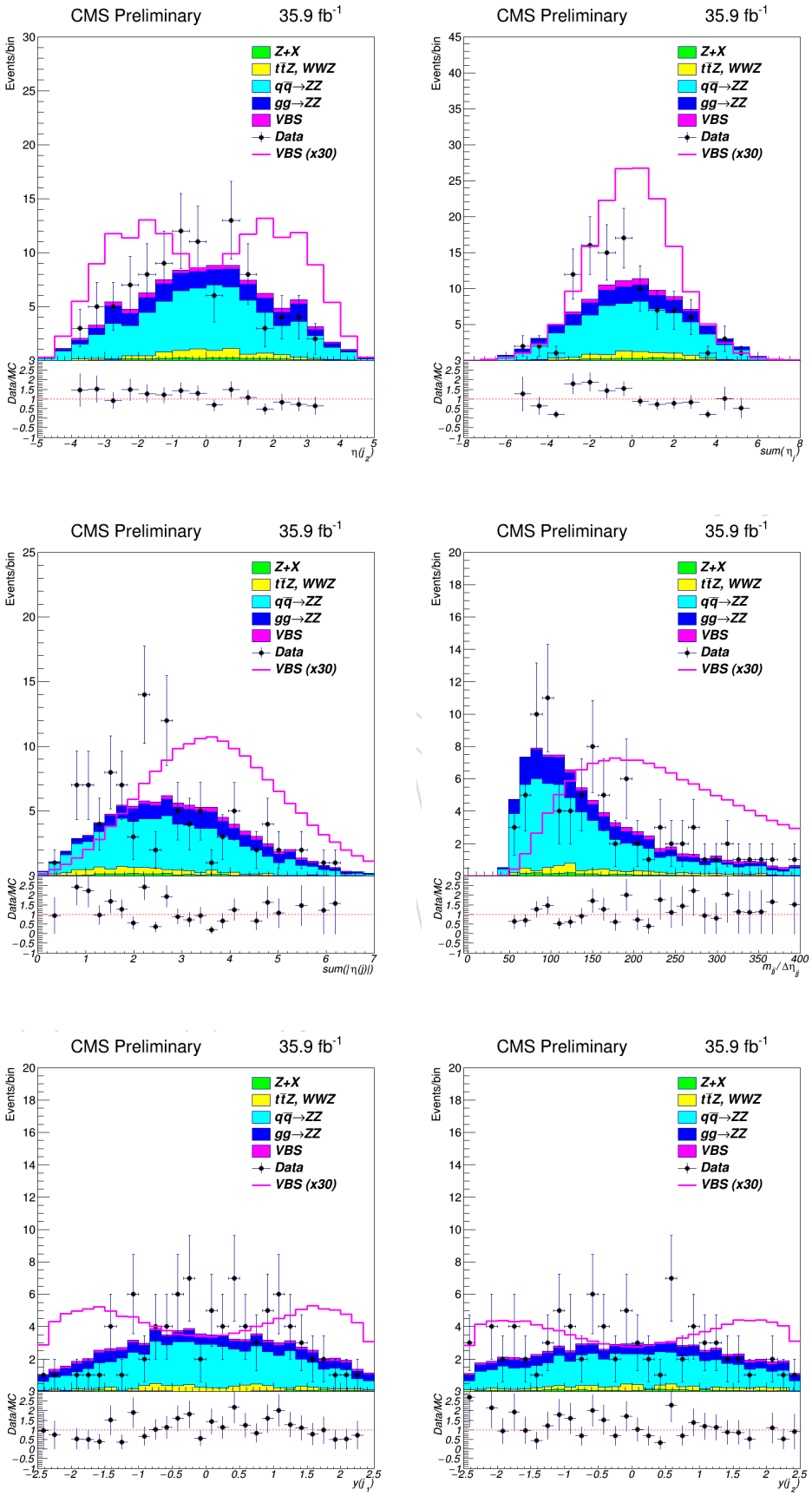


Figure 22: Variables considered for training the BDT classifier. Distributions for 2016 period are shown.

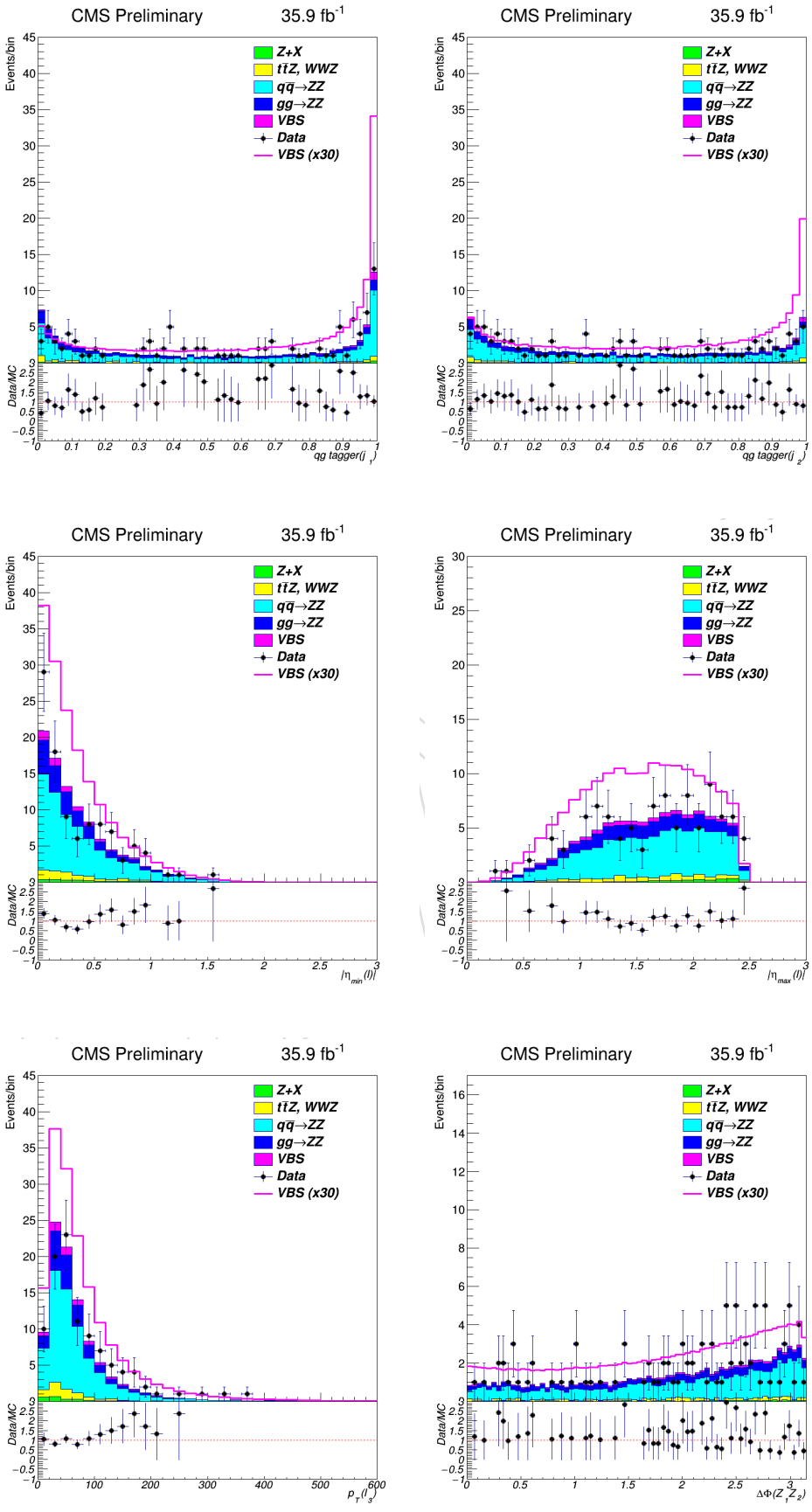


Figure 23: Variables considered for training the BDT classifier. Distributions for 2016 period are shown.

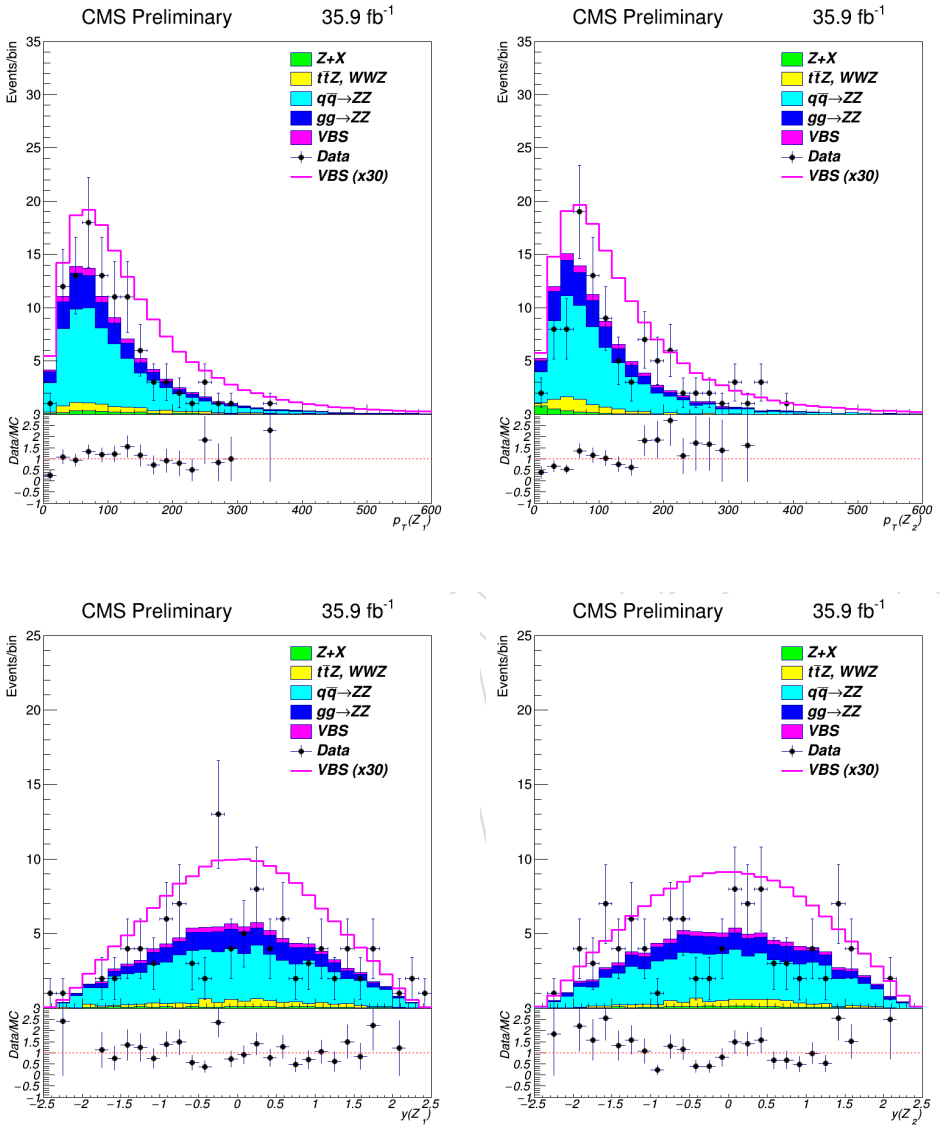


Figure 24: Variables considered for training the BDT classifier. Distributions for 2016 period are shown.

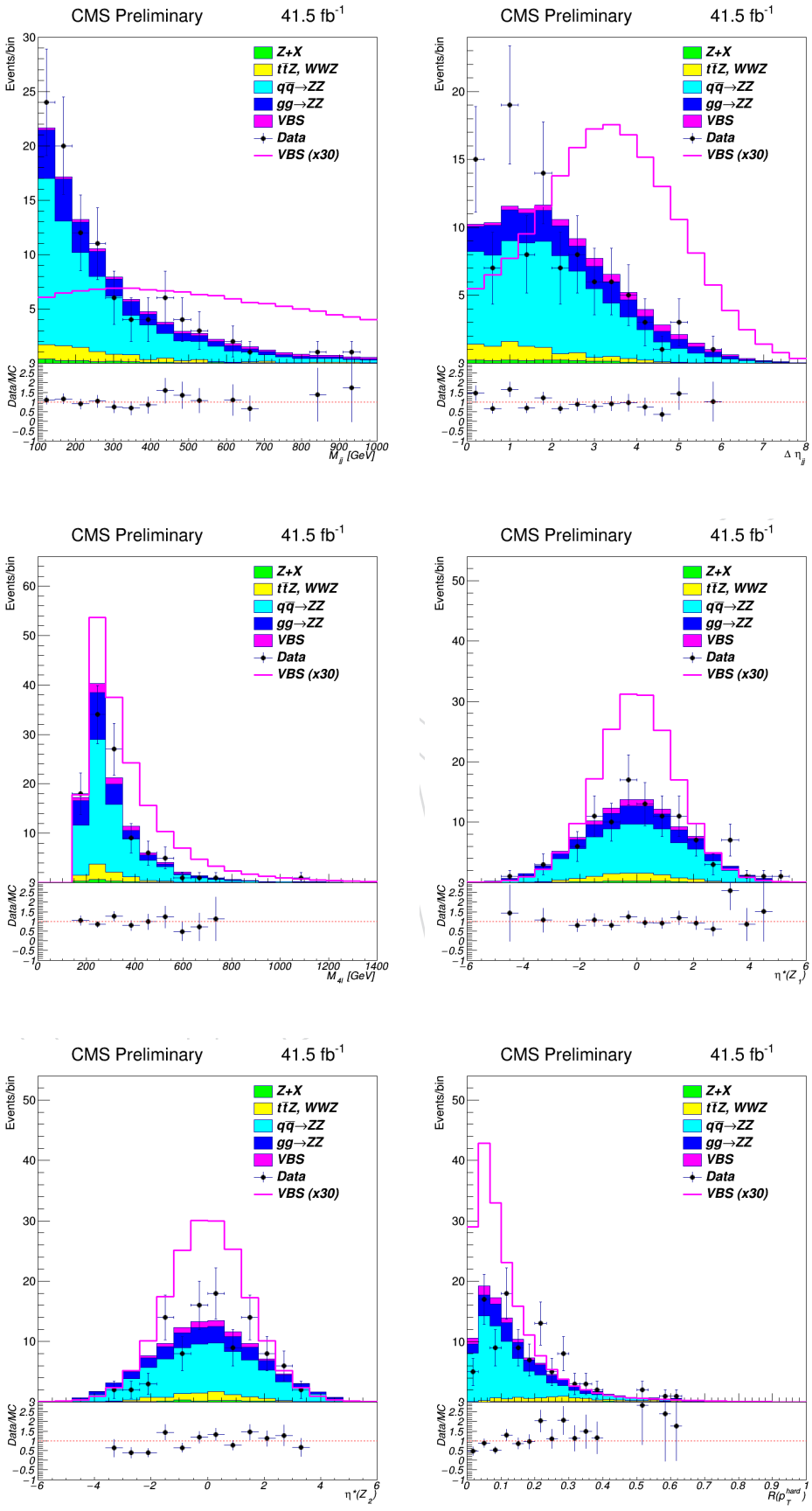


Figure 25: Variables considered for training the BDT classifier. Distributions for 2017 period are shown.

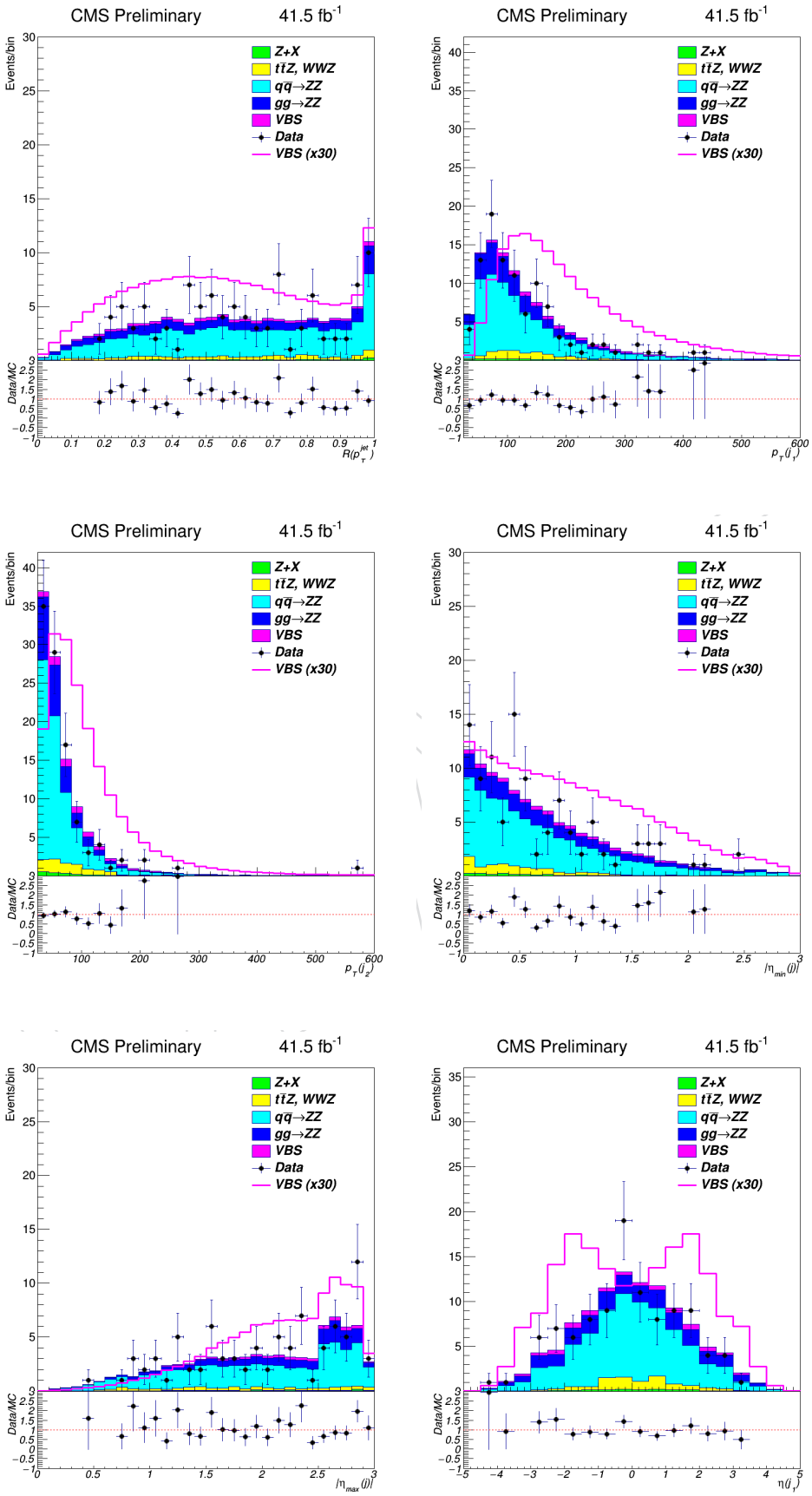


Figure 26: Variables considered for training the BDT classifier. Distributions for 2017 period are shown.

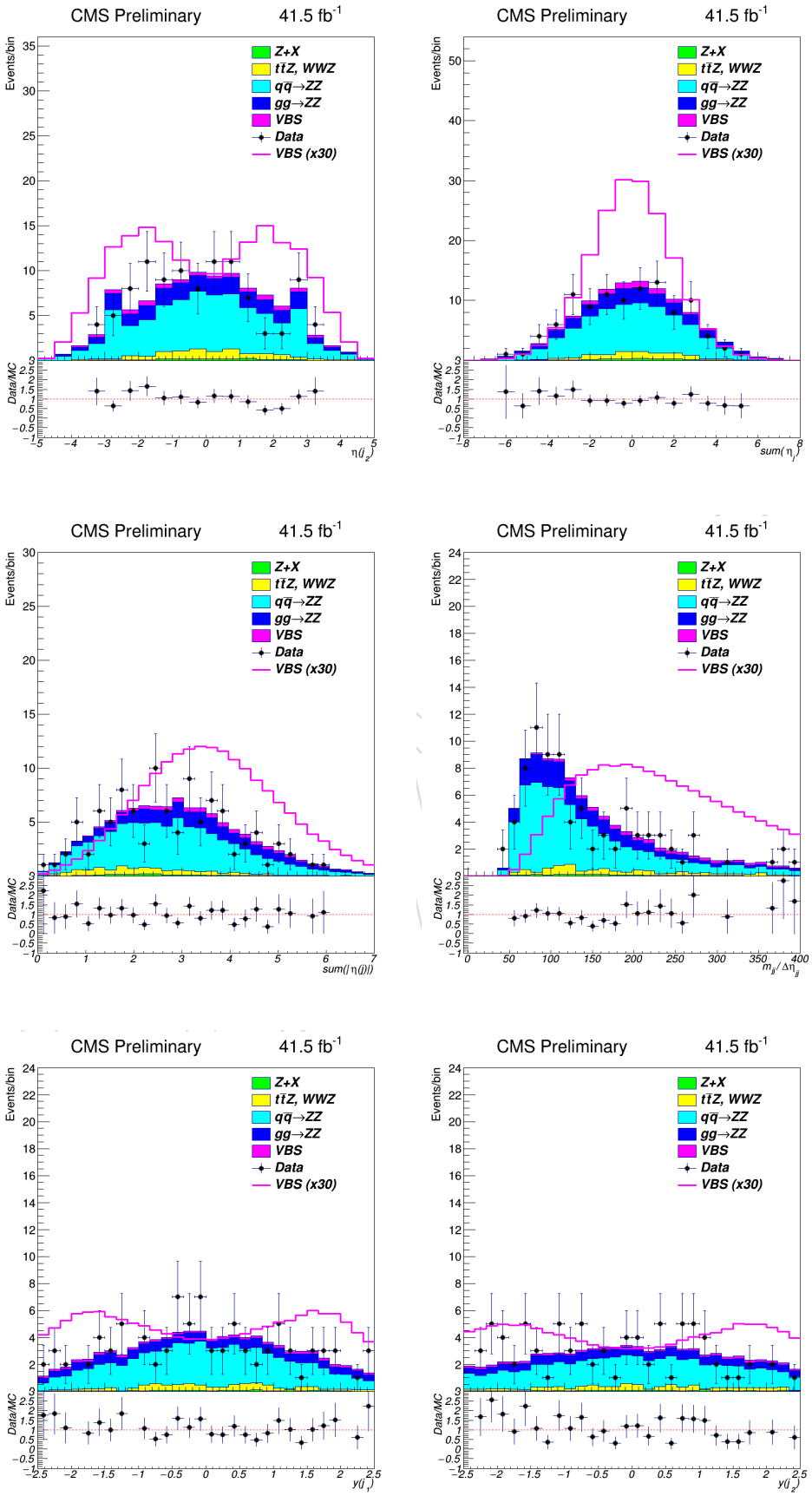


Figure 27: Variables considered for training the BDT classifier. Distributions for 2017 period are shown.

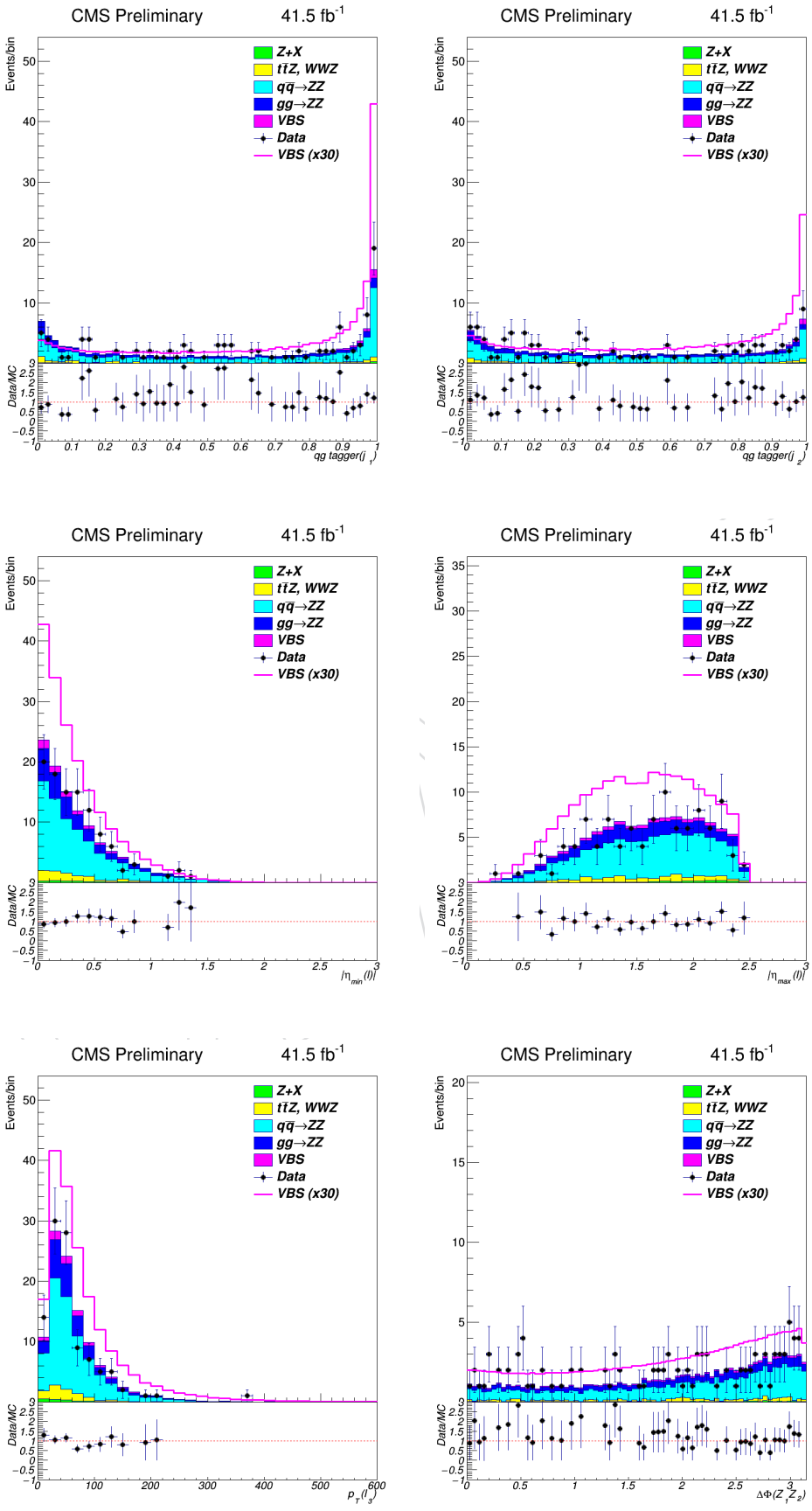


Figure 28: Variables considered for training the BDT classifier. Distributions for 2017 period are shown.

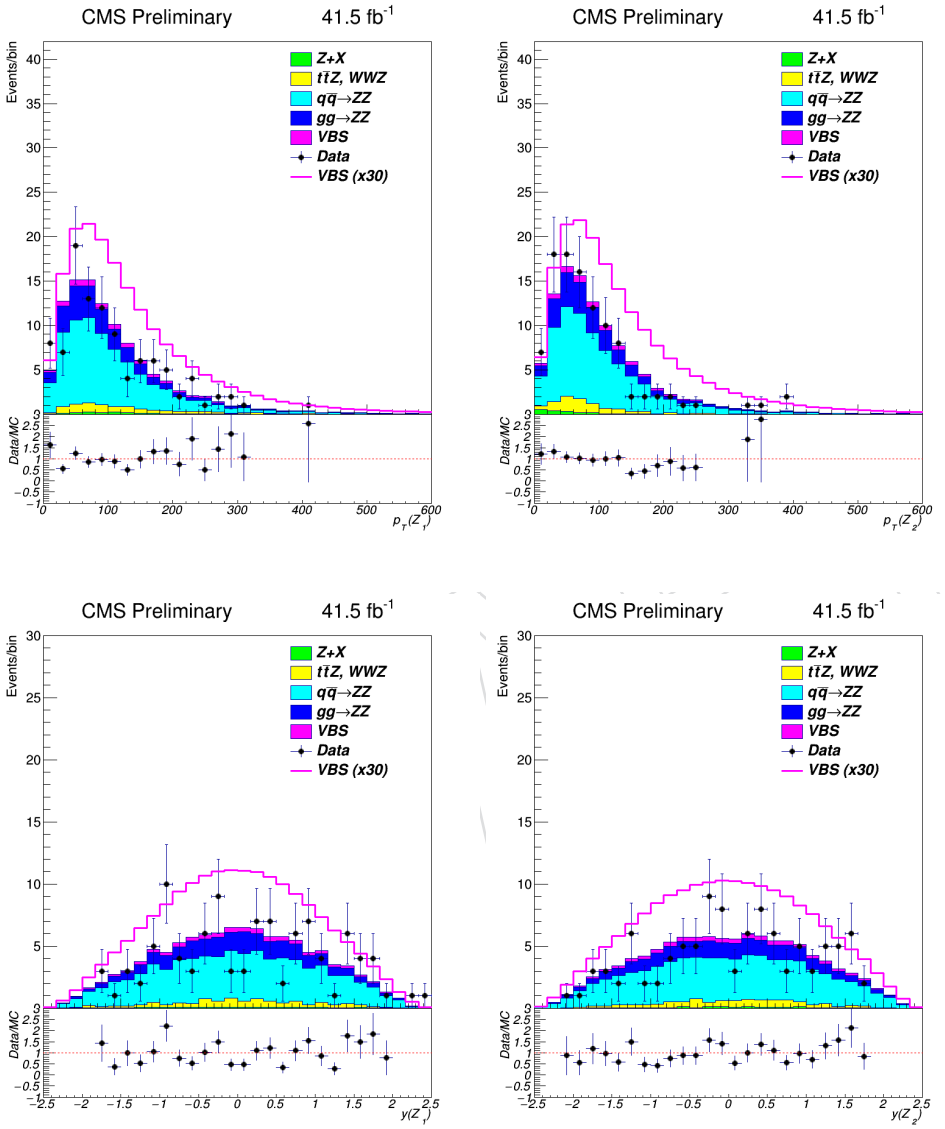


Figure 29: Variables considered for training the BDT classifier. Distributions for 2017 period are shown.

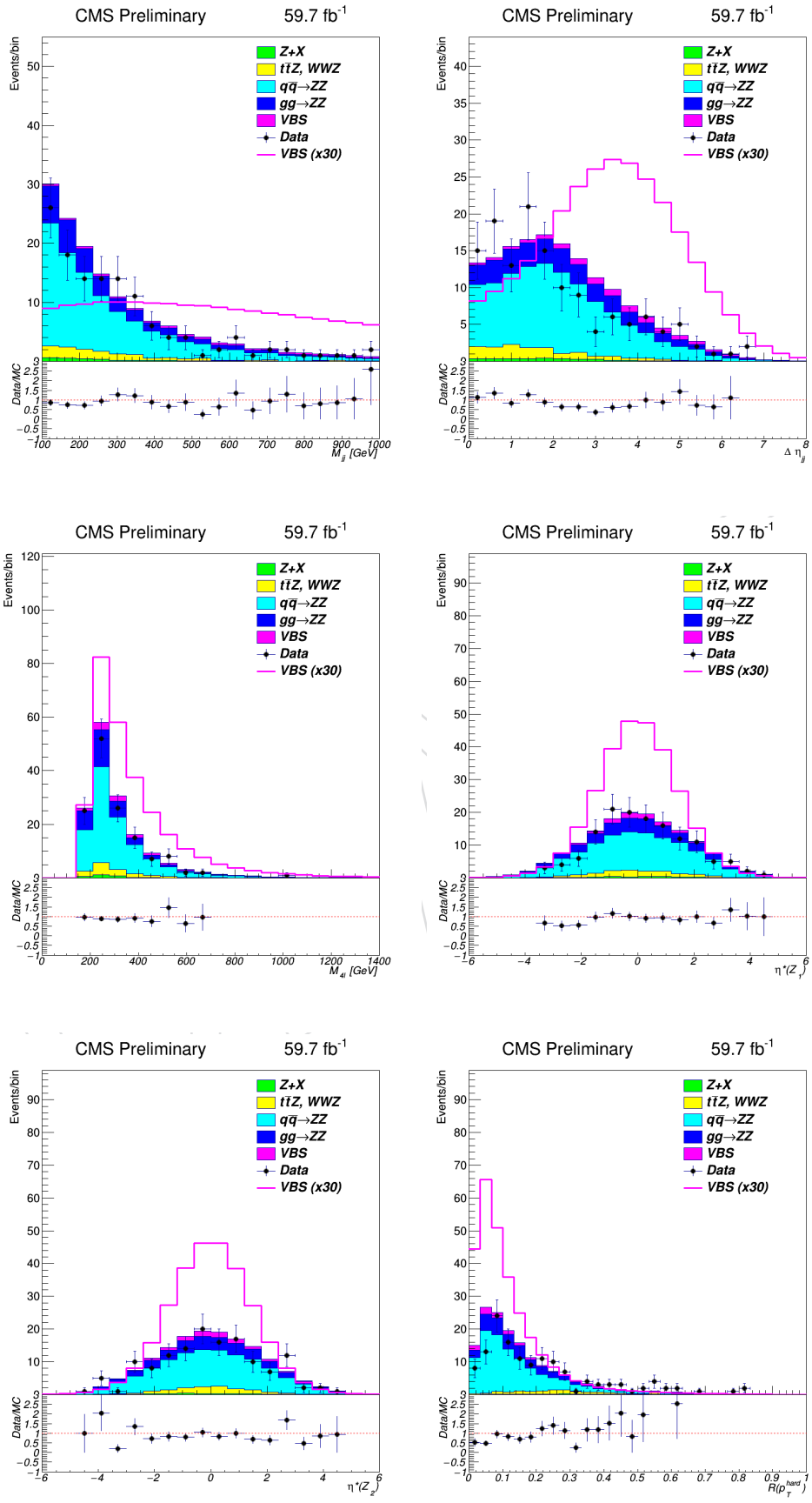


Figure 30: Variables considered for training the BDT classifier. Distributions for 2018 period are shown.

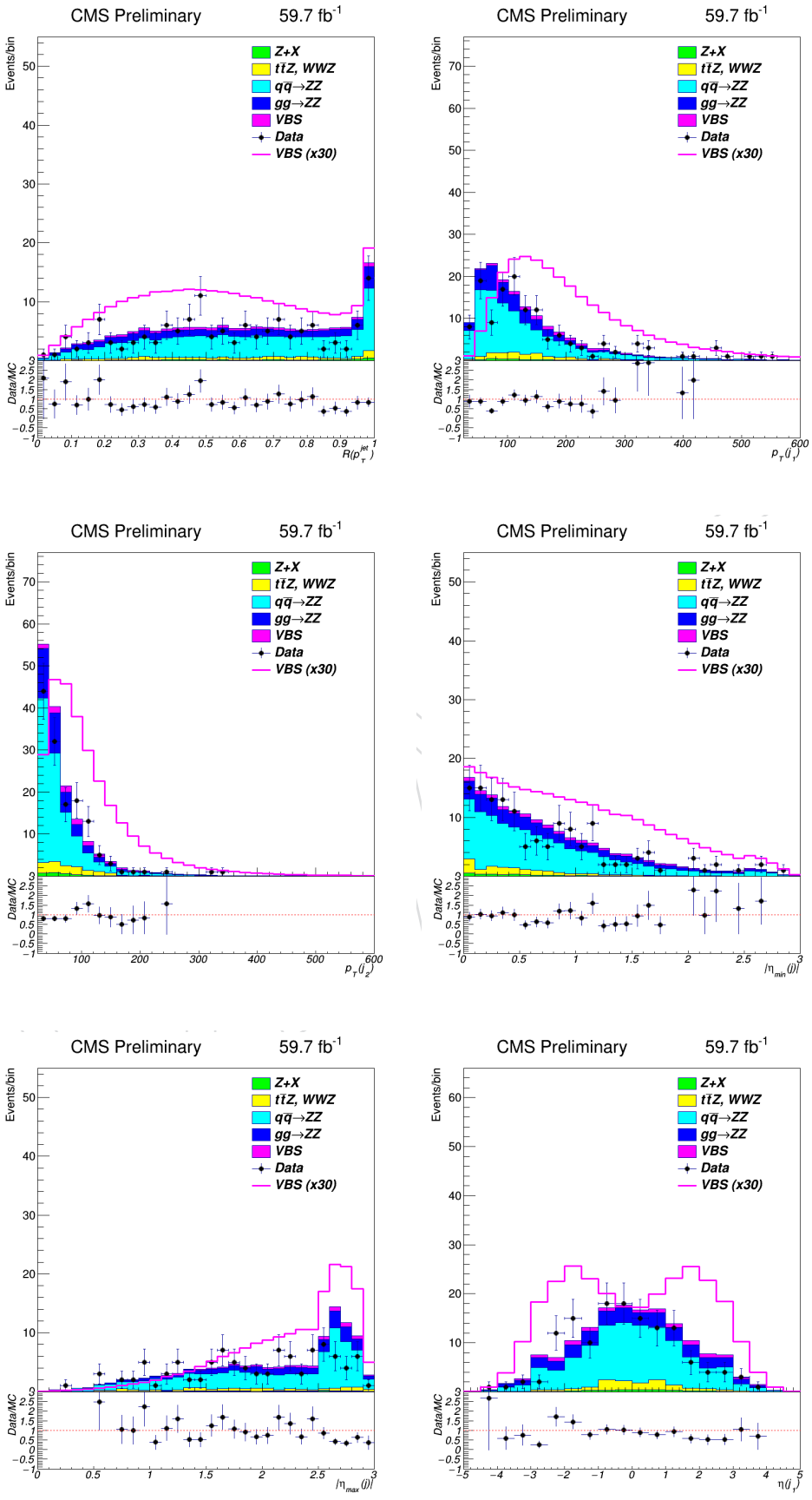


Figure 31: Variables considered for training the BDT classifier. Distributions for 2018 period are shown.

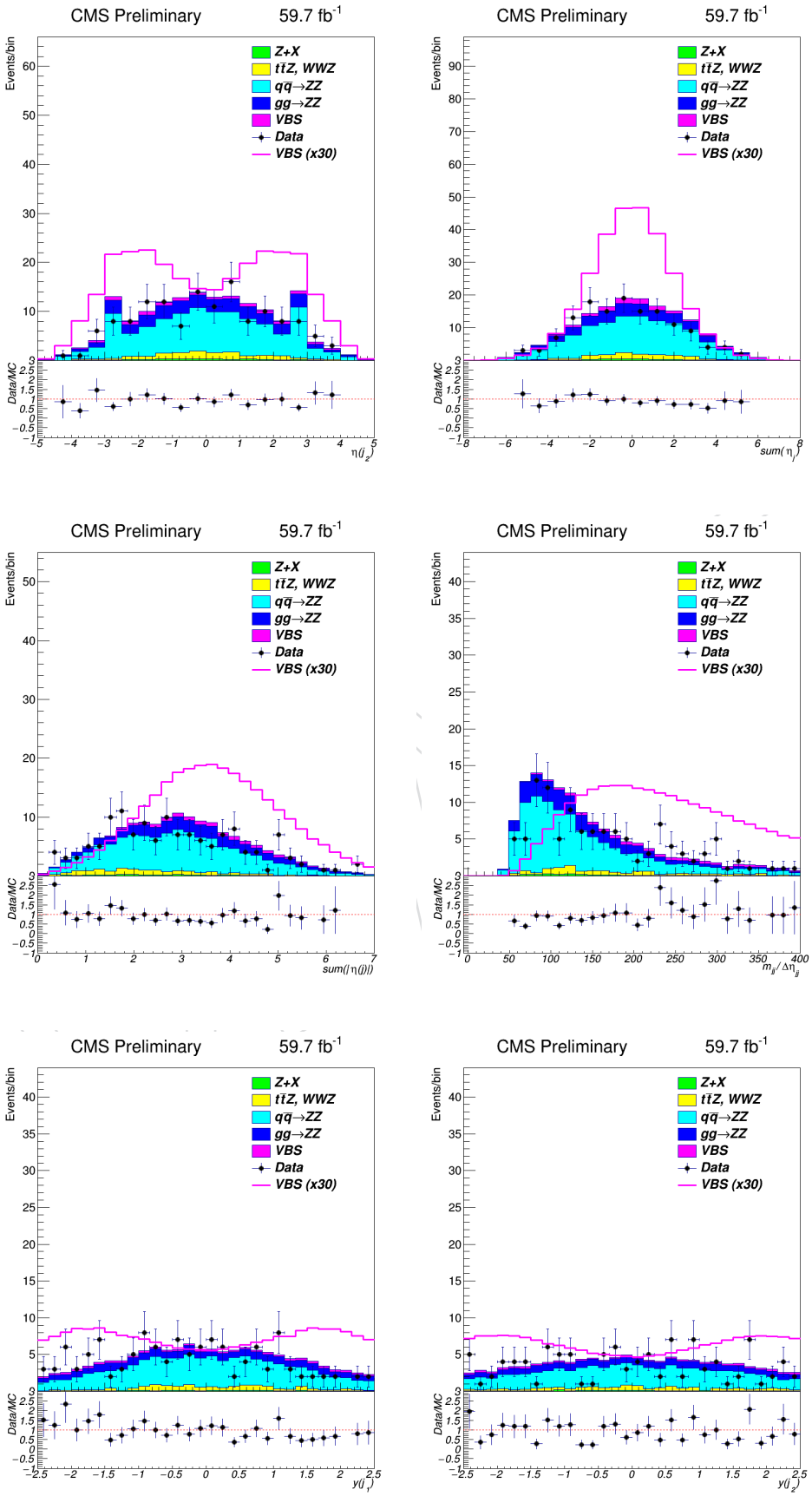


Figure 32: Variables considered for training the BDT classifier. Distributions for 2018 period are shown.

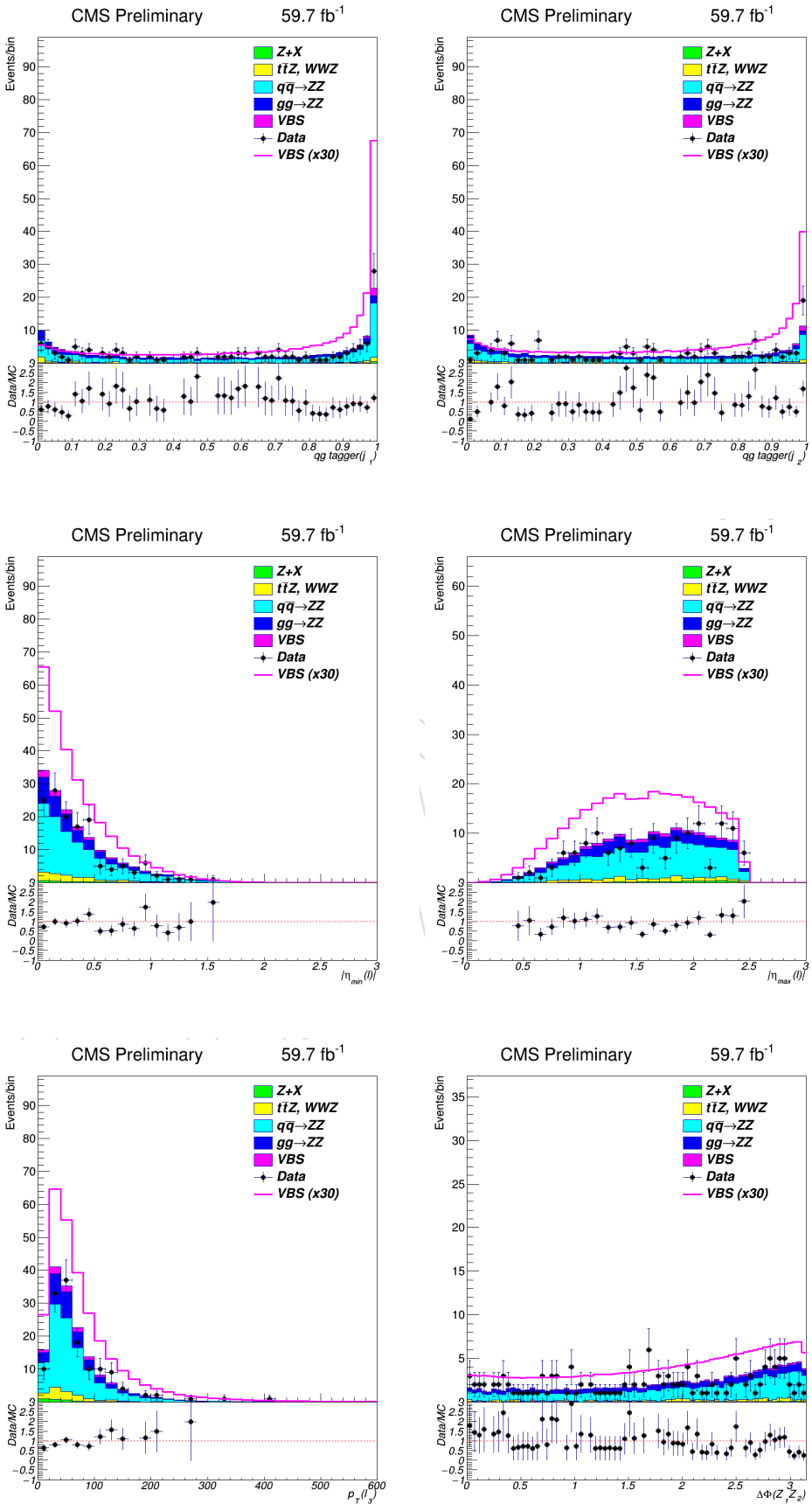


Figure 33: Variables considered for training the BDT classifier. Distributions for 2018 period are shown.

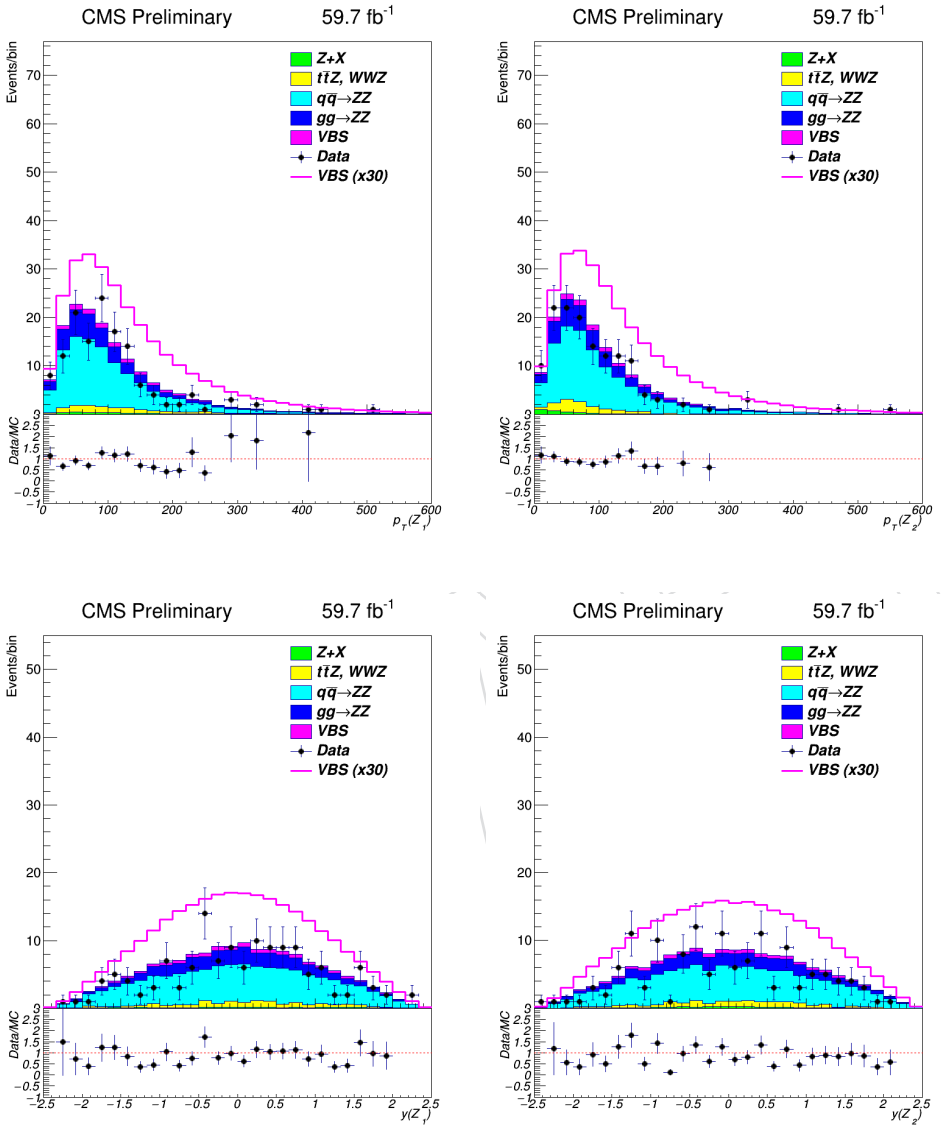


Figure 34: Variables considered for training the BDT classifier. Distributions for 2018 period are shown.

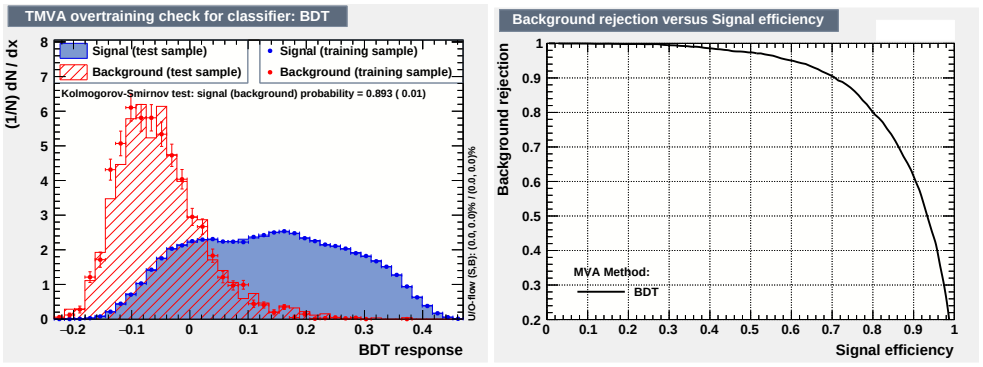


Figure 35: BDT output with overtraining test (left) and corresponding ROC curve (right) for BDT7. Plots are showing BDT results for 2016 period.

	Variable						
1	m_{jj}	8	$\Delta\phi(Z_1, Z_2)$	15	$\lvert\eta_{max}(lep)\rvert$	22	$\lvert\eta_{min}(l)\rvert$
2	$\Delta\eta_{jj}$	9	$p_T(j_1)$	16	$p_T(j_2)$	23	$p_T(Z_2)$
3	m_{4l}	10	$\lvert\eta(j_{min})\rvert$	17	$y(Z_2)$	24	$p_T(Z_1)$
4	$\eta^*_{Z_1}$	11	$y(j_1)$	18	$y(j_2)$	25	$p_T(l_3)$
5	$\eta^*_{Z_2}$	12	$qgtagger(j_2)$	19	$qgtagger(j_1)$	26	$\eta(j_1)$
6	$R(p_T^{hard})$	13	$y(Z_1)$	20	$\Sigma\lvert\eta(j)\rvert$	27	$\eta(j_2)$
7	$R(p_T^{jets})$	14	$\lvert\eta_{max}(j)\rvert$	21	$\Sigma\eta(j)$	28	$m_{jj}/\Delta\eta(jj)$

Table 22: List of variables used for training the BDT in this study. List contains original variables used in 2016 study and 21 variables added to assess potential significance gain.

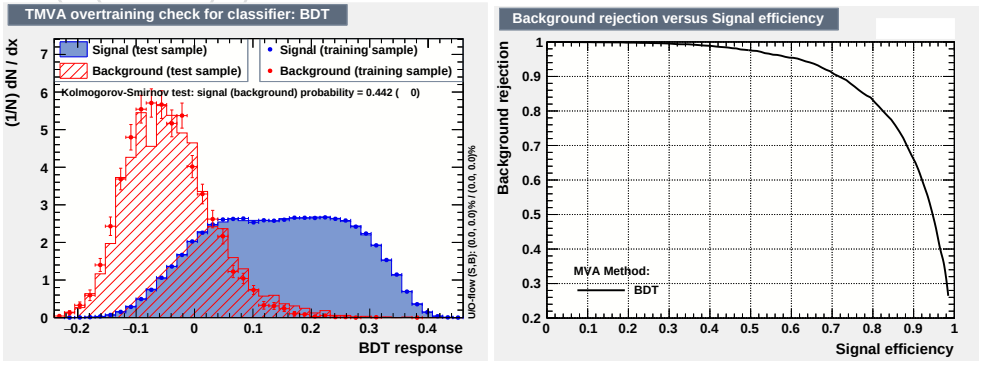


Figure 36: BDT output with overtraining test (left) and corresponding ROC curve (right) for BDT28. Plots are showing BDT results for 2016 period.

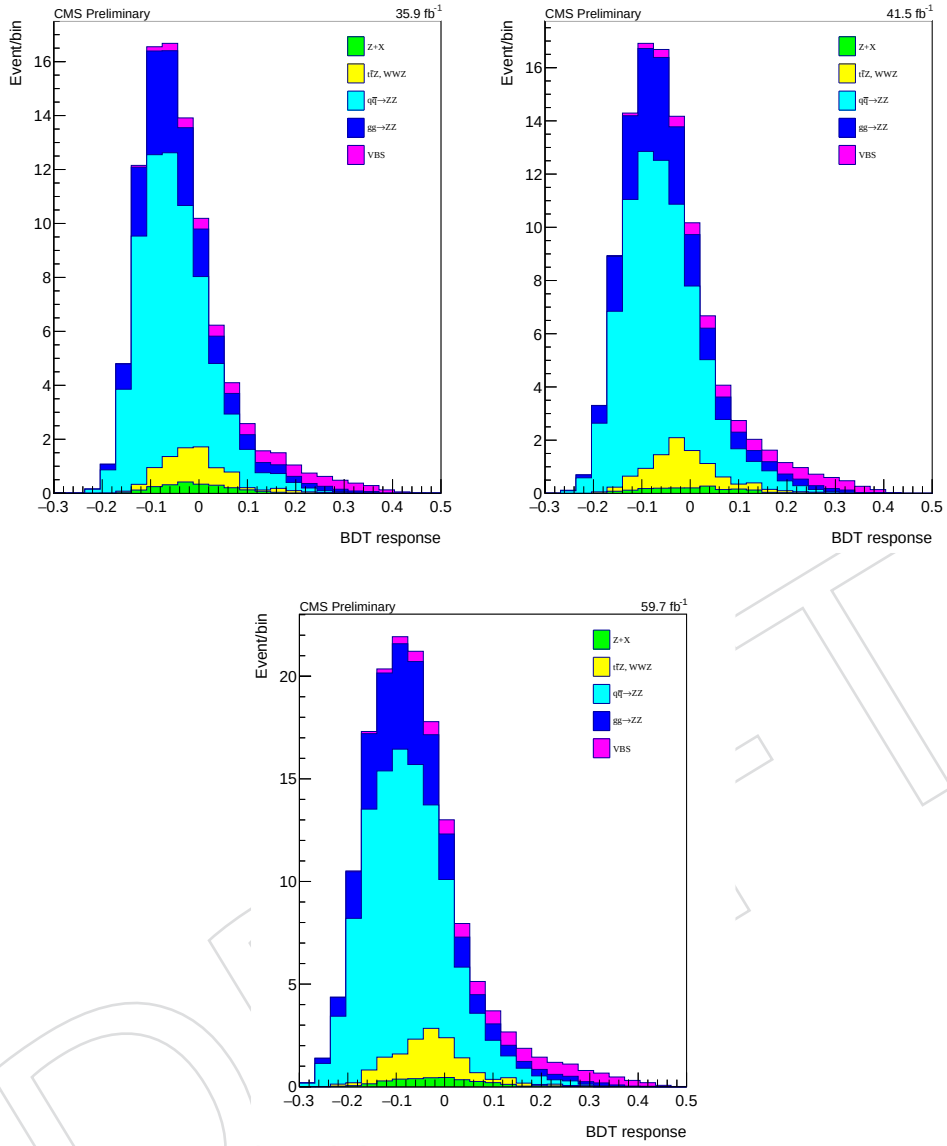


Figure 37: BDT classifier distributions are shown for 2016 (top left), 2017 (top right) and 2018 (bottom) periods.

Systematic uncertainties described in Section 7 were used with a difference that the three sub-channels ($4e/4\mu/2e2\mu$) were merged into one. It can be seen from the results with uncertainties from statistics only that the effect of systematic uncertainty is very small in this analysis.

BDT28 was trained to asses the significance and the significance gain compared to BDT7 for 2016 period giving 1.94 (stat. only) and 1.90 (with systematics) compared to 1.87 (stat. only) and 1.83 (with systematics) for BDT7. The gain with BDT28 was 3.7% (stat. only) and 3.8% (with systematics). Another check was made to see the effect of training the BDT on VBS selection and the resulting significance for 2016 period. VBS selection BDT7 resulted in significance of 1.86 (stat.only) and 1.83 (with systematics) resulting in change below 0.5% for both cases.

This study confirms that the 7 variables used in the 2016 analysis and listed in Table 21 capture essentialy all kinematics differences between signal and background.

year	stat. + sys.	stat. only	difference [%]
2016	1.83	1.87	2.6
2017	1.99	2.04	2.4
2018	2.71	2.78	2.5
combined	3.78	3.92	3.6

Table 23: Expected significance for different data taking periods and the combined results using BDT7 training.

B Alternative gluon loop-induced ZZ sample

B.1 Simulation details

An alternative ggZZ loop-induced sample produced by `MadGraph5_aMCatNLO` [6] is studied which simulates up to 2 jets by matrix-element and applied MLM matching scheme [9]. The process is generated at LO by requesting explicitly the loop-induced process:

```
generate g g > z z [noborn=QCD]
add process p p > z z j [noborn=QCD]
add process p p > z z j j [noborn=QCD]
```

Since the loop-induced mode of `MadGraph5_aMCatNLO` [6] is incapable to identify genuine loop-induced diagrams out of all one-loop diagrams, which also consists of the one-loop correction to the tree diagrams, a "diagram filter" is specially designed based on official suggestions [36], requiring that

- the loop must not contain any gluon line, such that vertex- and box-correction diagrams are discarded;
- the loop must attach to at least one Z, W boson or photon, to avoid diagrams concerning the gluon self-energy correction through quark lines, and diagrams mediated by a Higgs boson.

After applying the filter, only loop-induced diagrams can survive.

Besides, for 1- and 2-jet process, "pp" initial state rather than "gg" is used in order to involve the ISR process, where a quark as the initial state particle will first transform to a gluon through ISR, then the gluon takes part in the hard process. The use of "pp" initial state brings significantly more diagrams, including both loop-induced ones and loop corrections, but only the former will survive the diagram filter. Also, we note that "pp" is equivalent with "gg" for the 0-jet process, since it does not introduce extra loop-induced diagrams.

Furthermore, it is worth mentioning that the diagrams for 1/2-jet process not only consist of 0-jet ggZZ diagrams decorated with ISR, but also, include new types of diagrams with different structures, *e.g.* jets directly emitted from the loop. Fig. 38 shows some example diagrams for 1/2-jet process that cannot evolve from the basic 0-jet ggZZ diagrams. This brings possibilities to better model the dijet phase-space, rather than simply considering 0-jet ggZZ process in a matrix-element way, while modelling the two jets from parton shower.

However, due to the constraint of `MadSpin` generator which cannot decay particle in a loop-induced process from the matrix-element calculation, the decaying of Z bosons is still implemented in `PYTHIA`. Therefore the spin correlations of outcoming leptons are not simulated, and both Z bosons keep at its pole mass.

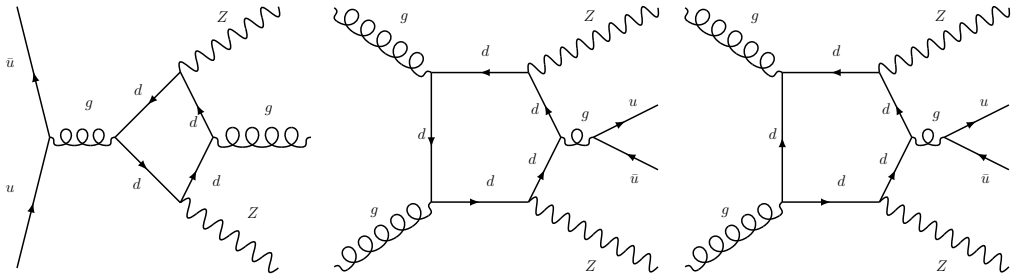


Figure 38: Example diagrams of loop induced ggZZ 1/2-jet process which have different structures from ggZZ 0-jet process.

Simulation of such sample is very expensive in time, which is at a rate of 8 min per raw LHE event, and, if considering an MLM matching rate of 8%, will reach a net rate of 100 min/event. Currently, the most time-consuming part, *i.e.* the private production of 12 M LHE events have been completed, and the subsequent generation steps were officially requested which aims to produce $\sim$ 1 M events in MiniAOD level. A medium-scale production was first completed for preliminary kinematics tests, containing $\sim$ 150 K events in MiniAOD.

B.2 MLM matching optimization

The MLM matching scheme [9] is applied to avoid phase-space overlapping of dijet modeled by matrix-element method and parton shower. In this analysis, we perform the first optimization study on the matching parameter for the loop-induced process. As in the MLM matching sample, the relatively harder jets are generated by matrix-element while softer jets produced by parton shower, the method introduces a cut-off scale q_{cut} to piece the emission jets from two approaches together. The q_{cut} value depends on the process, and conventionally is larger than 10 GeV. For the loop-induced process, emission partons consist of both the ordinary ISR partons and the ones emitted directly from the loop. It is investigated that emission partons from loop-induced events are generally softer. Therefore, a correspondingly small value of $q_{\text{cut}} = 5$ is studied. Fig. 39 shows the differential jet rate (DJR) plot for the first and second jet as $q_{\text{cut}} = 5$ is specified. The smoothness of DJR plots validates a good matching under such parameter.

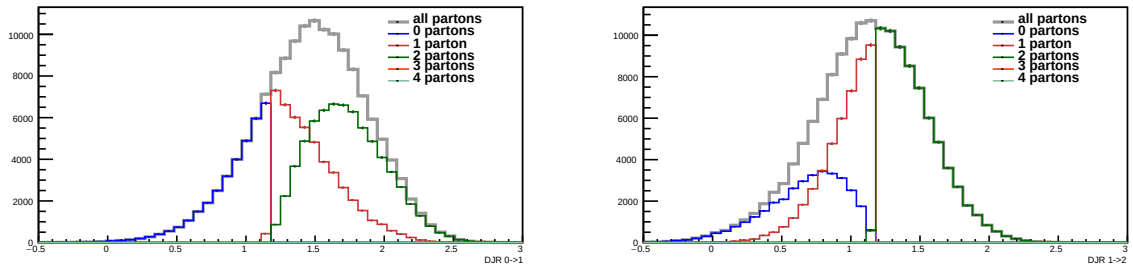


Figure 39: Differential jet rate (DJR) plot for the first and second jet in MadGraph 0/1/2-jet matching sample with $q_{\text{cut}} = 5$.

B.3 Kinematics comparison

To illustrate that jets produced by matrix-element is generally softer than that from parton shower in a loop-induced event, The p_{T,j_1}, p_{T,j_2} spectrum is compared among the new ggZZ

MadGraph 0/1/2-jet matching, the ggZZ MadGraph 0-jet, and the MCFM sample, as shown in Fig. 40. Note that both the latter two samples model the dijet phase-space fully by PYTHIA parton shower approach. The 0/1/2-jet matching sample, as previously mentioned, contains unique diagrams such as jet(s) directly emitted from the loop, which is beyond the scope of parton shower. Therefore we expect to see the most accurate modeling on dijet phase-space in the 0/1/2-jet matching sample.

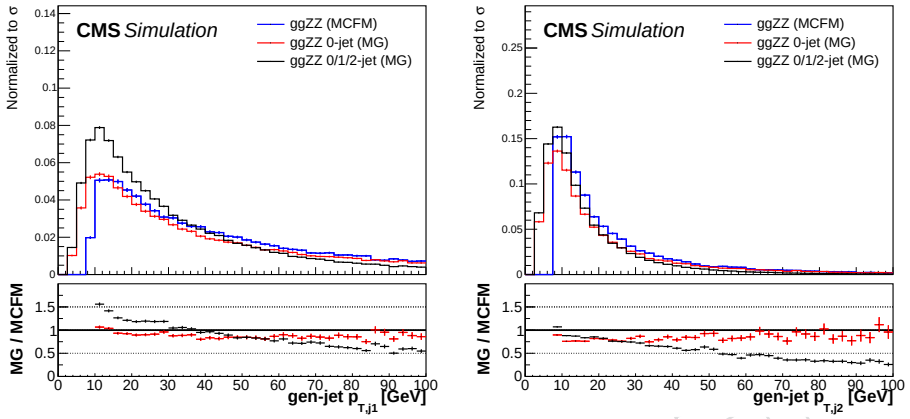


Figure 40: p_{T,j_1}, p_{T,j_2} spectrum comparison between the ggZZ MadGraph 0/1/2-jet matching, the ggZZ MadGraph 0-jet, and the MCFM sample, where the latter two cases model the dijet by parton shower. The softness of jet in the 0/1/2-jet matching sample illustrates that jets produced by matrix-element is generally softer than that from parton shower in a loop-induced event.

As a supplement to Fig. 6, the new ggZZ MadGraph 0/1/2-jet matching sample is compared with the MCFM and ggZZ 1-jet sample on several kinematic variables, as shown in Fig. 41.

Large discrepancy is spotted in the 0/1/2-jet matching sample. The reason can be concluded as:

- The softness of jets modeled in the 0/1/2-jet matching sample may cause lower baseline selection passing rate, hence smaller yields in each plots.
- The smallness of jet p_T may, in turn, result in larger $m_{4\ell}$ which stands for the recoiling part of emitted jets.

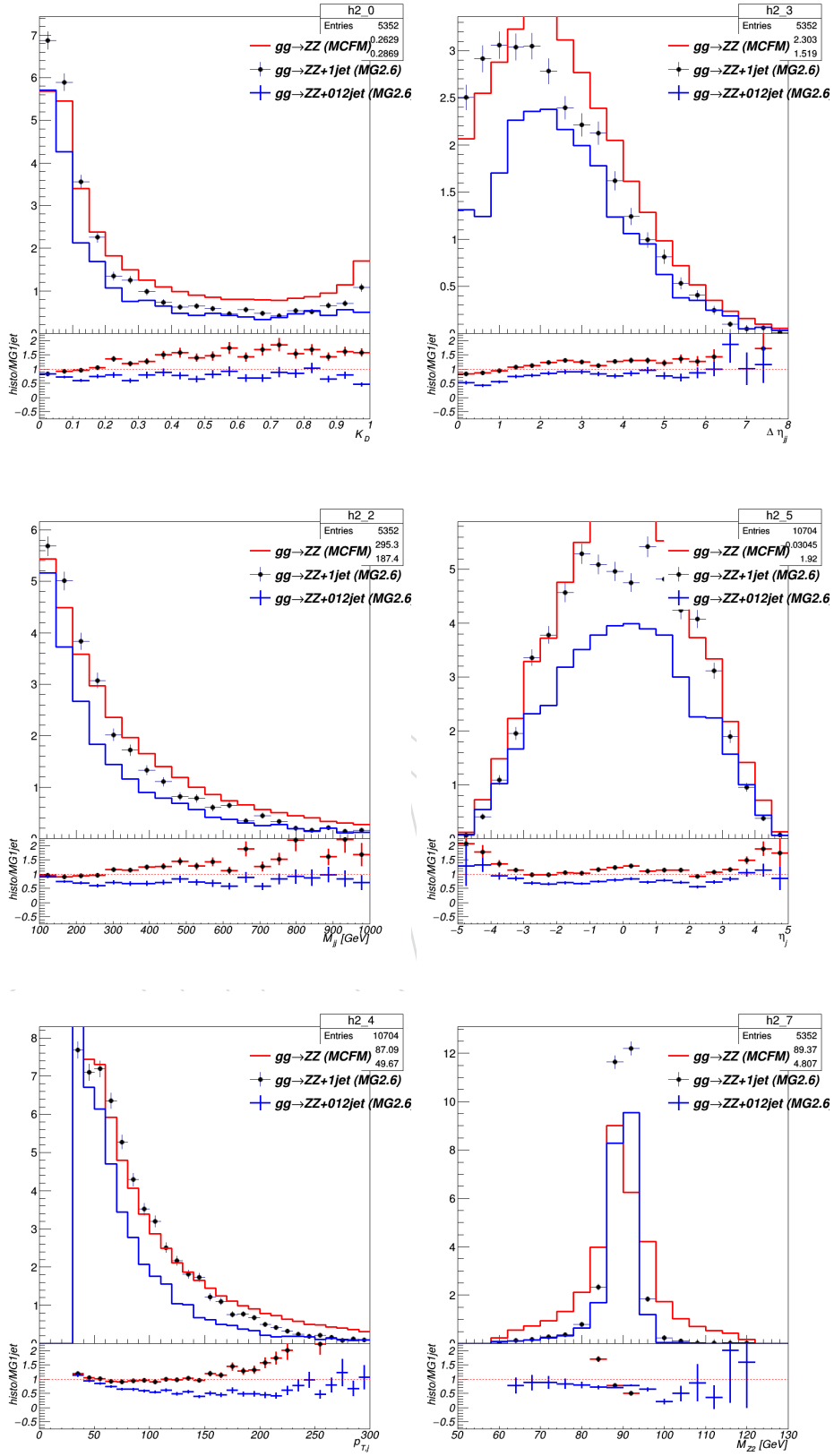


Figure 41: Comparison of the kinematics in the ggZZ MadGraph 0/1/2-jet matching (blue), ggZZ MadGraph 1-jet (black points), and MCFM MC (red) samples in the phase space defined by the ZZjj baseline selection at the reconstruction level. All distributions are normalized to MC cross-sections. Top: MELA (left) and $\Delta\eta_{jj}$ (right). Middle: m_{jj} (left) and η_j (right). Low: $p_{T,j}$ (left) and $m_{\ell\ell}$ (right).

References

[1] C. Charlot, P. Pigard et al., "Search for the electroweak production of two Z bosons in the 4ℓjj final state in pp collisions at $\sqrt{s} = 13$ TeV", CMS AN 2017/002.

[2] CMS Collaboration, "Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in protonproton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett.* **B774** (2017) 682–705, doi:10.1016/j.physletb.2017.10.020, arXiv:1708.02812.

[3] C. Ochando, T. Sculac, M. Xiao et al., "Measurements of properties of the Higgs boson in the four-lepton final state at $\sqrt{s} = 13$ TeV with full Run II data", CMS AN 2019/139.

[4] CMS Collaboration, "Measurements of properties of the higgs boson in the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ tev", CMS Physics Analysis Summary CMS-PAS-HIG-19-001, 2019.

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

[6] 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.

[7] A. Ballestrero et al., "PHANTOM: A Monte Carlo event generator for six parton final states at high energy colliders", *Comput. Phys. Commun.* **180** (2009) 401, doi:10.1016/j.cpc.2008.10.005, arXiv:0801.3359.

[8] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. B Proc. Suppl.* **205-206** (2010) 10, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.

[9] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *The European Physical Journal C* **53** (2009), no. 3, 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[10] 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.

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

[12] K. Arnold et al., "VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons", *Comput. Phys. Commun.* **180** (2009) 1661–1670, doi:10.1016/j.cpc.2009.03.006, arXiv:0811.4559.

[13] M. Chiesa, A. Denner, J.-N. Lang, and M. Pellen, "An event generator for same-sign W-boson scattering at the LHC including electroweak corrections", *Eur. Phys. J.* **C79** (2019), no. 9, 788, doi:10.1140/epjc/s10052-019-7290-6, arXiv:1906.01863.

[14] M. Grazzini, S. Kallweit, and D. Rathlev, "ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD", *Phys. Lett. B* **750** (2015) 407, doi:http://dx.doi.org/10.1016/j.physletb.2015.09.055, arXiv:1507.06257.

[15] S. Gieseke, T. Kasprzik, and J. H. Kuehn, "Vector-boson pair production and electroweak corrections in HERWIG++", *Eur. Phys. J.* **C74** (2014), no. 8, 2988, doi:10.1140/epjc/s10052-014-2988-y, arXiv:1401.3964.

[16] F. Caola, K. Melnikov, R. Rntsch, and L. Tancredi, "QCD corrections to ZZ production in gluon fusion at the LHC", *Phys. Rev.* **D92** (2015), no. 9, 094028, doi:10.1103/PhysRevD.92.094028, arXiv:1509.06734.

[17] F. Caola et al., "QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC", *JHEP* **07** (2016) 087, doi:10.1007/JHEP07(2016)087, arXiv:1605.04610.

[18] M. Bonvini et al., "Signal-background interference effects in $gg \to H \to WW$ beyond leading order", *Phys. Rev. D* **88** (2013) 034032, doi:10.1103/PhysRevD.88.034032, arXiv:1304.3053.

[19] C. S. Li, H. T. Li, D. Y. Shao, and J. Wang, "Soft gluon resummation in the signal-background interference process of gg(h) ZZ", *JHEP* **08** (2015) 065, doi:10.1007/JHEP08(2015)065, arXiv:1504.02388.

[20] R. Salerno and M. Ahmad, "Measurements of properties of the Higgs boson in the four-lepton final state at $\sqrt{s}$=13 TeV", cms an, 2016.

[21] T. Sculac and D. Sperka, "Measurements of properties of the Higgs boson in the four-lepton final state at $\sqrt{s}$=13 TeV with 2017 data", cms an, 2017.

[22] JetMETPOG, "Jet ID for 2018", *CMS Twiki* (2018) doi:https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2018.

[23] JetMETPOG, "Jet PU ID for 2017", *CMS Twiki* (2017) doi:https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJetID#Information_for_13_TeV_data_anal.

[24] JetMETPOG, "JEC MC for 2018", *CMS Twiki* (2018) doi:https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC#2018_Data.

[25] JetMETPOG, "Reweighting recipe to emulate Level 1 ECAL prefiring", *CMS Twiki* (2018) doi:https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1ECALPrefiringWeightRecipe.

[26] Y. Gao et al., "Spin determination of single-produced resonances at hadron colliders", *Phys. Rev. D* **81** (2010) 075022, doi:10.1103/PhysRevD.81.075022, arXiv:1001.3396. [Erratum: doi:10.1103/PhysRevD.81.079905].

[27] S. Bolognesi et al., "On the spin and parity of a single-produced resonance at the LHC", *Phys. Rev. D* **86** (2012) 095031, doi:10.1103/PhysRevD.86.095031, arXiv:1208.4018.

[28] I. Anderson et al., "Constraining anomalous HVV interactions at proton and lepton colliders", *Phys. Rev. D* **89** (2014) 035007, doi:10.1103/PhysRevD.89.035007, arXiv:1309.4819.

[29] GENPOG, "PDF weights in MC Simulations", *CMS Twiki* (2018) doi:https://twiki.cern.ch/twiki/bin/view/CMS/HowToPDF.

[30] CMS Collaboration, "CMS luminosity measurement for the 2016/2017/2018 data-taking period at $\sqrt{s} = 13$ TeV", technical report.

[31] A. S. Belyaev et al., "Strongly interacting vector bosons at the CERN LHC: Quartic anomalous couplings", *Phys. Rev.* **D59** (1999) 015022, doi:10.1103/PhysRevD.59.015022, arXiv:hep-ph/9805229.

[32] O. J. P. Éboli, M. C. Gonzalez-Garcia, and J. K. Mizukoshi, "$pp \to jj\, e^{\pm}\mu^{\pm}\nu\nu$ and $jj\, e^{\pm}\mu^{\mp}\nu\nu$ at $\mathcal{O}(\alpha_{\text{em}}^6)$ and $\mathcal{O}(\alpha_{\text{em}}^4\alpha_{\text{s}}^2)$ for the study of the quartic electroweak gauge boson vertex at CERN LHC", *Phys. Rev. D* **74** (2006) 073005, doi:10.1103/PhysRevD.74.073005, arXiv:hep-ph/0606118.

[33] F. Chollet, "Deep Learning with Python". Manning Publications, 2018.

[34] F. Chollet et al., "Keras". https://keras.io, 2015.

[35] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems", 2015. Software available from tensorflow.org. https://www.tensorflow.org/.

[36] V. Hirschi and P. Pigard, "Discussions on 'gluon-loop induced ZZ+2jets'", *MadGraph5_aMC@NLO Launchpad* (2016) doi:https://answers.launchpad.net/mg5amcnlo/+question/402723.