

CMS Draft Analysis Note

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

2022/07/08

Archive Hash: cf3b440

Archive Date: 2022/07/08

Search for exotic decays of the Higgs boson to a pair of new light bosons with two muons and two b jets in the final states at $\sqrt{s} = 13$ TeV

E. Khazaie¹, M. Zeinali², H. Bakhshiansohi³, and A. Jafari³

¹ Isfahan University of Technology (IUT)

² Isfahan University of Technology (IUT) and Sharif University of Technology (SUT)

³ IUT and DESY

Abstract

We report the results of a search for exotic decays of a Higgs boson with $m_H = 125$ GeV to a pair of new light bosons, a_1 , where one of the light bosons decays to a pair of muons and the other one decays to a pair of b quarks. A data sample corresponding to an integrated luminosity of 138 fb^{-1} recorded with the CMS detector in years 2016–2018 is exploited in $\mu^+\mu^-b\bar{b}$ final states where no statistically significant excess is observed with respect to the standard model backgrounds for different m_{a_1} hypotheses above 15 GeV and below $m_H/2$. Upper limits are set on $\sigma_{\text{ggF}} \times \text{Br}(H \rightarrow \mu^+\mu^-b\bar{b})$ and on the branching ratio itself with former ranging between xxx to yyy fb, depending on the m_{a_1} values. The limit on the branching ratio is $\text{Br}(H \rightarrow \mu^+\mu^-b\bar{b}) < (0.17 - 3.3) \times 10^{-4}$ for the given mass range.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	E. Khazaie, M. Zeinali, H. Bakhshiansohi, A. Jafari
PDFTitle:	Search for exotic decays of the Higgs boson to a pair of new light bosons with two muons and two b jets in the final states at sqrt=13 TeV
PDFSubject:	CMS
PDFKeywords:	CMS, Exotic Higgs decays

Please also verify that the abstract does not use any user defined symbols

1 Introduction

The Higgs-like boson discovered by ATLAS and CMS experiments [1, 2] at the CERN LHC [3], has been under detailed studies over the past years. The mass of the boson, referred to as “the Higgs” or “H” hereafter, is confirmed to be around 125 GeV at per-mil level [4, 5]. At this point, the majority of the Higgs boson productions mechanisms as well as decay modes are discovered with $t\bar{t}H$ production [6, 7], Higgs to $b\bar{b}$ [8, 9], and Higgs to $\mu^+\mu^-$ [10, 11] decays being the latest. Measurements of the Higgs boson couplings [11–13] confirm, within uncertainties, the predictions of the standard model (SM) of particle physics. Yet, the possibility of the exotic Higgs decays to new lighter bosons is not entirely excluded.

The current data allows for Higgs decays to states beyond the standard model (BSM) with a rate of $\mathcal{O}(20\%–50\%)$. There are several experimental and theoretical studies that fit the couplings of the Higgs to SM and constrain the Higgs branching ratio (Br) into a so far undetected or an invisible final state [11–17]. Assuming the SM production for the Higgs boson, a $\text{Br}(H \rightarrow \text{BSM})$ of 20% is allowed at 95% CL. This value increases to 30% if some new physics modifies the loop-induced couplings to gluons and photons. More conservative approaches allow for even larger rates [18, 19].

Such decays are well motivated in the context of the next-to-minimal supersymmetric standard model, NMSSM, and extensions to Two Higgs Doublet models (2HDM) where the existence of a scalar singlet is foreseen [20]. The 2HDM, and hence 2HDM+S, are categorized into four types depending on the interaction of SM fermions with the Higgs doublet structure [21]. All SM particles couple to the first Higgs doublet, Φ_1 , in type I models. In type II, up-type quarks couple to Φ_1 while leptons and down-type quarks couple to the second Higgs doublet, Φ_2 . Type II models have NMSSM as a special case where the additional singlet is introduced to resolve the so-called μ -problem of the MSSM superpotential. Quarks couple to Φ_1 and leptons couple to Φ_2 in type III and finally in type IV, leptons and up-type quarks couple to Φ_1 , while down-type quarks couple to Φ_2 . After the electroweak symmetry breaking, the Two Higgs Doublet models allow for a pair of charged Higgs bosons, H^\pm , a neutral pseudoscalar, A , and two neutral scalar mass eigenstates, H and h . One can assume that the model is in a decoupling limit such that the lighter scalar eigenstate, H , is the discovered boson with $m_H = 125$ GeV. Furthermore, to allow for a wider variety of exotic Higgs decay phenomenology, a complex scalar singlet, $S_R + iS_I$ is added to the model which has no direct Yukawa couplings. Hence, it is expected to decay to SM fermions by virtue of mixing with the Higgs sector. This mixing is small enough to preserve the SM-like nature of the H boson. The Higgs boson can decay to fermions through a_1 (s_1) where a_1 (s_1) is a (pesudo)scalar mass eigenstate mostly composed of S_R (S_I).

The possibility of the Higgs boson decaying into pseudoscalar a_1 is studied in this note for the final state of $H \rightarrow a_1 a_1 \rightarrow b\bar{b}\mu\mu$, as illustrated in Fig. 1. The gluon fusion production mechanism is considered as the dominant Higgs production process with the next-to-leading-order production rate of $\sigma_{ggH} \simeq 48.58 \pm 1.89$ pb [22]. The contribution of the vector boson fusion (VBF) Higgs boson production process is also taken into account with the next-to-leading-order production cross section $\sigma_{qqH} = 3.93 \pm 0.08$ pb [22]. As a benchmark, a branching fraction 10% is assigned to $H \rightarrow a_1 a_1$ where assuming $\tan\beta = 2$, one can obtain $2 \times \text{Br}(a_1 \rightarrow b\bar{b})\text{Br}(a_1 \rightarrow \mu^+\mu^-) = 1.7 \times 10^{-3}$ for $m_{a_1} = 30$ GeV in type-III 2HDM+S [20]. For the set of parameters and the mass range under discussion, no strong dependence to m_{a_1} is expected for $\text{Br}(a_1 \rightarrow f\bar{f})$, with f being muon or b quark [20]. The signal cross section times branching ratios is therefore approximated to be about 9 fb for all m_{a_1} values considered in this analysis.

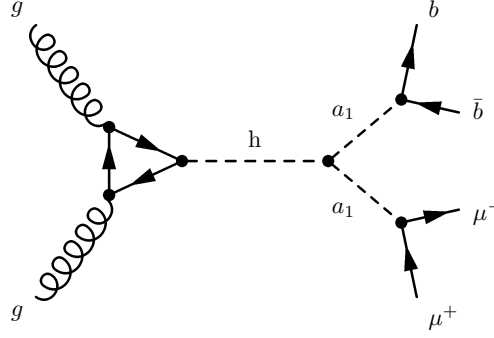


Figure 1: Feynman diagram for the signal topology.

The search for the exotic a_1 particle is performed with $15 \leq m_{a_1} \leq 62.5$ GeV. The lower bound is set in order to ignore intermediate states with quarkonia, whose interpretation is complicated by non-perturbative QCD effects. Moreover, the sensitivity of the search largely decreases towards $m_{a_1} \approx 15$ GeV and lower. The upper bound is imposed by the Higgs mass. The signal selection is optimized for a $H \rightarrow a_1 a_1 \rightarrow \mu \mu b \bar{b}$ signal. The analysis is performed using the data collected with the CMS detector during Run-II LHC, corresponding to an integrated luminosity of 138 fb^{-1} .

Similar searches are performed at the LHC, at different center-of-mass energies and data samples. The latest analysis by ATLAS [23] has placed a strong bound of $\text{Br}(H \rightarrow b \bar{b} \mu^+ \mu^-) < (0.22 - 4) \times 10^{-4}$ using the entire LHC Run-II data. This was an update on a former search with the partial data sample [24]. The CMS search at $\sqrt{s} = 13 \text{ TeV}$ [25] is based on 36 fb^{-1} of data and results in an upper limit of $(1 - 7) \times 10^{-4}$ for the Higgs boson branching ratio in $b \bar{b} \mu \mu$ final state. At 8 TeV, the experiment has provided an upper bound of about 2.5 times the expectation for the same branching ratio [21]. The analysis presented in this note is an update of Ref. [25] including a more in-depth study of the signal to achieve more stringent limits than what is offered by pure addition of the full Run-II LHC data. The note is organized as follows: Section 2 is a brief description of data and simulated samples used in the analysis. In Section 3, physics objects are defined and the preselection requirements are presented. Studies to optimize the identification requirements of physics object are shown in Section 4 while final event selection and categorization are detailed in Section 5. Sections 6 and 7 contain details on signal modeling and background determination studies. Systematic uncertainties are described in Section 8 and the results are presented in Section 9. A flowchart of the analysis is presented in Fig. 2 to visualize the steps.

2 Data samples and simulation

This analysis is performed using the data from the LHC proton-proton collisions at 13 TeV center-of-mass energy. The data sample, corresponding to an integrated luminosity of 138 fb^{-1} for a mixture of double-muon and single muon triggers, was collected with the CMS detector during Run II of the LHC. The data correspond to DoubleMuon and SingleMuon primary datasets and are used for the signal enriched region as well as a region to validate Drell-Yan background. In addition, MuonEG primary dataset, triggered by the presences of an electron-muon pair is used to construct a region for $t \bar{t}$ validation in the $e \mu$ final state. All data have been (re)reconstructed using our knowledge about the physics object reconstructions, energy corrections and additional pp interactions within the same bunch crossing (pileup) in campaigns 17Jul2018, 31Mar2018 and 17Sep2018 for 2016, 2017 and 2018 respectively. Version seven

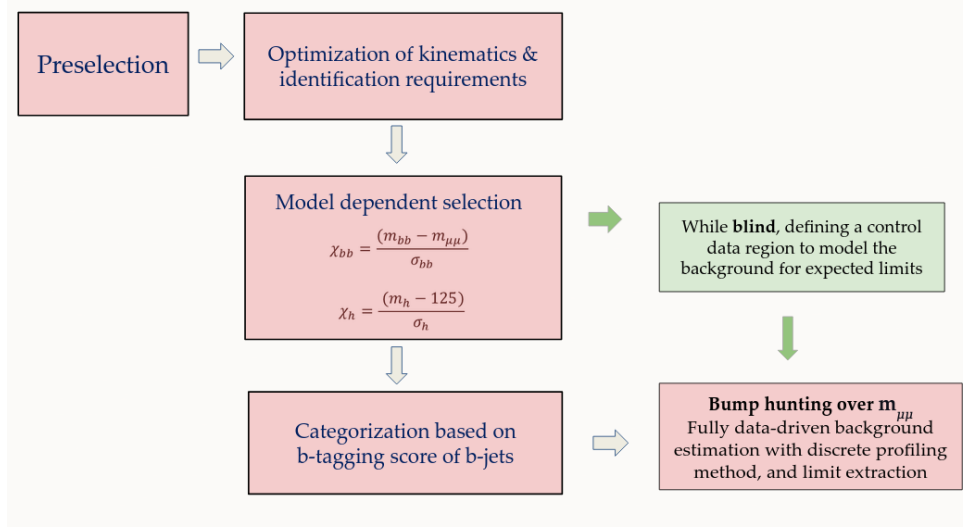


Figure 2: Analysis steps in a nutshell. The green box corresponds to the CR defined in Section 7 and it is there only for the blind analysis, i.e., will be removed after unblinding.

of the so-called NanoAOD event format, produced in 02Apr2020 campaign, is used for both data and simulated samples.

The NMSSMHET model [20] is used to generate signal samples with MADGRAPH5_AMCATNLO [26] at leading order (LO). The ggF signal samples are produced for $m_{a_1} = 15$ GeV to $m_{a_1} = 60$ GeV with 5 GeV steps. The VBF signal samples are generated at $m_{a_1} = 20$ GeV, $m_{a_1} = 40$ GeV and $m_{a_1} = 60$ GeV. Signal samples are produced privately using the exact Monte Carlo (MC) matrix element generator setting as that of Ref. [25]. Latest CMS recommendations and guidelines are used for showering, tune, pile up profile, etc. Samples are validated with respect to the central production of Ref. [25]. Figure 3 shows some validation distributions. Background samples in this analysis are only used for optimization as well as for validation in those steps of selection that yield enough statistics. The final contribution of backgrounds to the selected sample is directly extracted from data with no reference to simulation. The Drell-Yan process, $Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$, is modeled with the same event generator at either leading or next-to-leading order (NLO). The matching scheme for the LO samples follows the MLM [27] prescription while the FxFx merging scheme [28] is used at NLO. Based on the $m_{\ell\ell}$ threshold at generator level, two Drell-Yan samples are exploited for this analysis, one with $m_{\ell\ell} > 50$ GeV and the other with $10 < m_{\ell\ell} < 50$ GeV. The high-mass Drell-Yan samples are produced at NLO in all years, exclusive in number of additional partons (up to two). In low mass, NLO samples are only available in 2016. Furthermore in 2016, an LO sample exclusive in number of additional partons (up to four) exists. The inclusive LO samples are used in all years for low-mass Drell-Yan to harmonize the usage in all years. A k-factor of 1.5 (NLO-to-LO) is extracted from 2016 samples and applied in all years. To account for the limited size of the 2016 samples, an uncertainty of 30% is considered on the k-factor. Fig. 4 shows NLO-to-LO ratios where both inclusive and jet-binned samples at LO are studied.

The top quark samples, $t\bar{t}$ and single-top, are produced with POWHEG 2.0 [29–32] at NLO. Backgrounds from diboson (WW, WZ, ZZ) production are generated with PYTHIA 8.212 [33] at LO. For all samples, PYTHIA with tune CUETP8M1 [34] is used to model the showering. The simulated minimum bias interactions are added to the simulated events to model the effect of pileup. Simulated events are then re-weighted to reproduce the pileup distribution in data. The full CMS detector simulation based on GEANT4 [35] is implemented for all MC generated

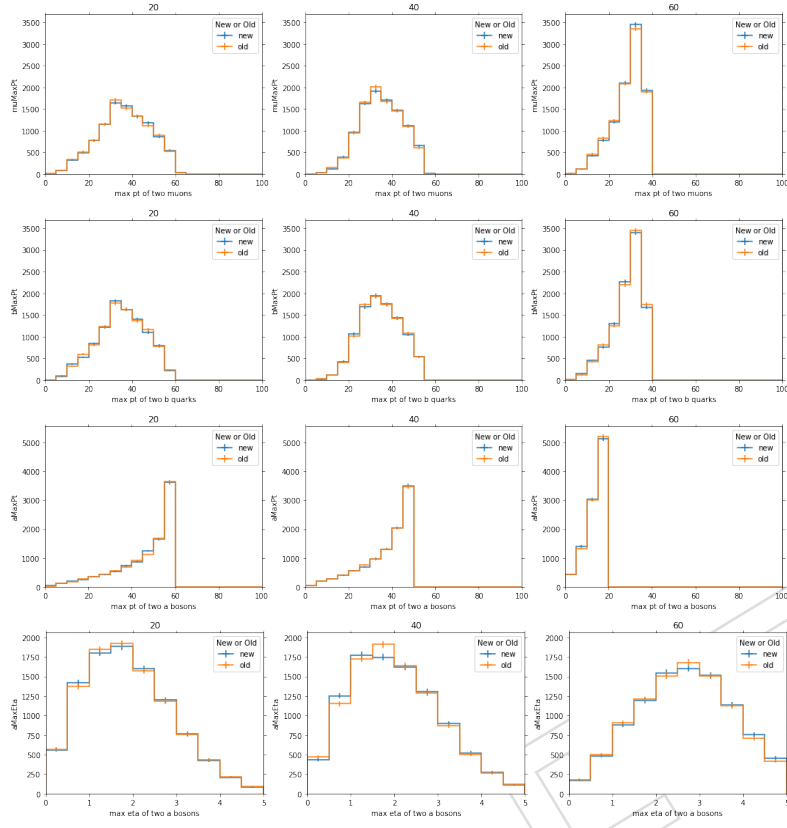


Figure 3: Comparison between official (old) and private (new) signal productions for $m_{a_1} \in \{20, 40, 60\}$ GeV: p_T of leading muon (top), p_T of leading b quark (second row), p_T of the leading a_1 boson (third row) and $|\eta|$ of the more forward a_1 boson.

112 event samples.

113 Table 1 contains the list of signal samples. In Tables 2, 3, and 4 a summary of simulated back-
 114 ground samples in different years together with their theoretical cross sections are shown.

Table 1: Privately produced signal samples in all years.

Simulated process	Sample name	Cross section (pb)
RunIISummer16, RunIIFall17, RunIIAutumn18 – NanoAODv7-Nano02Apr2020		
NMSSM $\mu\mu b\bar{b}$ ggF signal, $m_{a_1} \in \{15, 20, \dots, 60\}$ GeV	SUSYGluGluToHToAA_AToMuMu_AToBB_M- m_{a_1} - TuneCP5_13TeV_madgraph_pythia8 private	$\sigma \times BR \approx 0.007$ [20]
NMSSM $\mu\mu b\bar{b}$ VBF signal, $m_{a_1} \in \{20, 40, 60\}$ GeV	SUSYVBFToHToAA_AToMuMu_AToBB_M- m_{a_1} - TuneCP5_13TeV_madgraph_pythia8 private	$\sigma \times BR \approx 0.0004$ [20]

Table 2: Summary of simulated samples in 2016. The NLO low-mass Drell-Yan sample is used to extract the k-factor. The LO jet-binned sample of low-mass Drell-Yan is used for cross checks.

Simulated process	Sample name	Cross section (pb)
RunIISummer16NanoAODv7-PUMoriond17_Nano02Apr2020_102X_mckRun2_asymptotic		
$Z/\gamma^* + \text{jets}, Z/\gamma^* \rightarrow l\bar{l}$ $m_{l\bar{l}} > 50 \text{ GeV}$	DYToLL_0j_13TeV-amcatnloFXFX-pythia8 (ext1, backup) DYToLL_1j_13TeV-amcatnloFXFX-pythia8 (ext1, backup) DYToLL_2j_13TeV-amcatnloFXFX-pythia8 (ext1)	4755.0 885.6 340.3
$Z/\gamma^* + \text{jets}, Z/\gamma^* \rightarrow l\bar{l}$ $10 < m_{l\bar{l}} < 50 \text{ GeV}$	DYjetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DY1(2)jetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DY3(4)jetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYjetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	19536.0 875.8 (464.5) 114.1 (43.9) 22500.0
$t\bar{t}$ single top (tW) single top (t-channel) WW WZ ZZ	TTTo2L2Nu_TuneCP5_PSwights_13TeV-powheg-pythia8 ST_tW_(anti)top_5f_inclusiveDecays_TuneCP5_PSwights_13TeV-powheg-pythia8 ST_t-channel_(anti)top_4f_inclusiveDecays_TuneCP5_PSwights_13TeV-powheg-pythia8 WW_TuneCUETP8M1_13TeV-pythia8 WZ_TuneCUETP8M1_13TeV-pythia8 ZZ_TuneCUETP8M1_13TeV-pythia8	831.80 35.8 136.0 (80.9) 118.7 47.13 16.52

Table 3: Summary of simulated samples in 2017.

Simulated process	Sample name	Cross section (pb)
RunIIFall17NanoAODv7-PU2017_12Apr2018_Nano02Apr2020_102X_mc2017_realistic		
$Z/\gamma^* + \text{jets}, Z/\gamma^* \rightarrow ll$ $m_{ll} > 50 \text{ GeV}$	DYToLL_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8 DYToLL_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8 DYToLL_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	4755.0 885.6 340.3
$Z/\gamma^* + \text{jets}, Z/\gamma^* \rightarrow ll$ $10 < m_{ll} < 50 \text{ GeV}$ $t\bar{t}$ single top (tW) single top (t-channel) WW WZ ZZ	DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8 (ext1, pmx) TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8 ST_tW_(anti)top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8 ST_t-channel_(anti)top_4f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8 WW_TuneCP5_13TeV-pythia8 WZ_TuneCP5_13TeV-pythia8 ZZ_TuneCP5_13TeV-pythia8	19536.0 831.80 35.8 136.0 (80.9) 118.7 47.13 16.52

Table 4: Summary of simulated samples in 2018.

Simulated process	Sample name	Cross section (pb)
RunII Autumn18 NanoAOD v7-Nano02 Apr2020_102X_upgrade2018_realistic		
$Z/\gamma^* + \text{jets}, Z/\gamma^* \rightarrow ll$ $m_{ll} > 50 \text{ GeV}$	DYToLL_01_TuneCP5_13TeV-amcatnloFXFX-pythia8 DYToLL_11_TuneCP5_13TeV-amcatnloFXFX-pythia8 DYToLL_21_TuneCP5_13TeV-amcatnloFXFX-pythia8	4755.0 885.6 340.3
$Z/\gamma^* + \text{jets}, Z/\gamma^* \rightarrow ll$ $10 < m_{ll} < 50 \text{ GeV}$ $t\bar{t}$ single top (tW) single top (t-channel) WW WZ ZZ	DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8 (ext1) TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8 ST_tW_(anti)top_5f_inclusiveDecays_TuneCP5_P5weights_13TeV-powheg-pythia8 (ext1) ST_t-channel_(anti)top_4f_inclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8 WW_TuneCP5_P5weights_13TeV-pythia8 WZ_TuneCP5_P5weights_13TeV-pythia8 ZZ_TuneCP5_13TeV-pythia8	19536.0 831.80 35.8 136.0 (80.9) 118.7 47.13 16.52

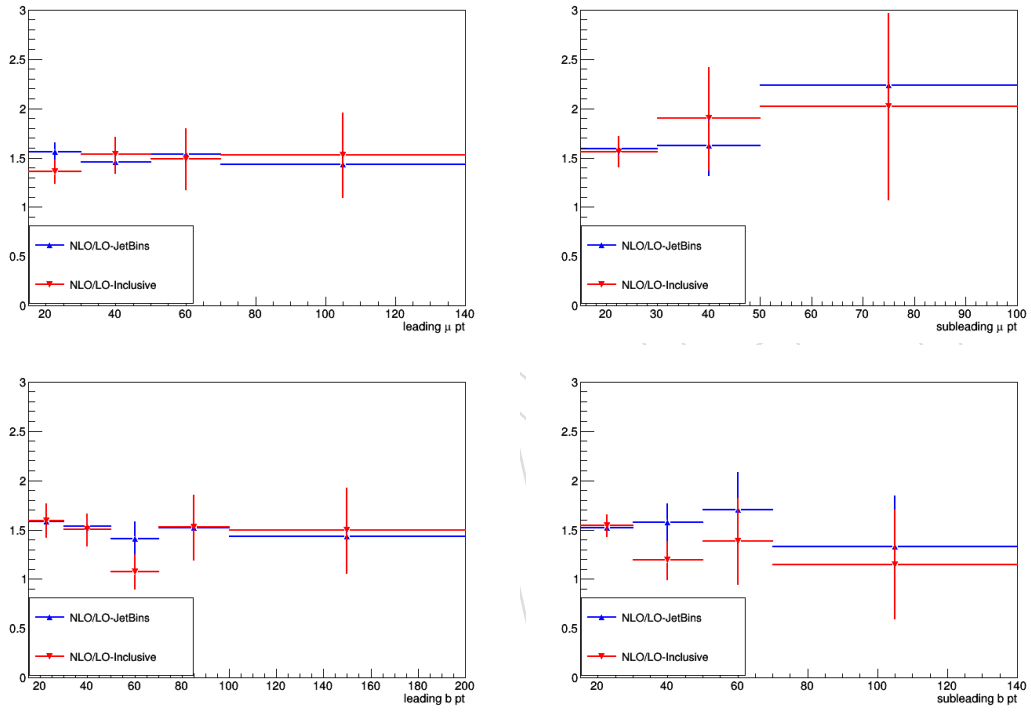


Figure 4: The NLO-to-LO k-factors in Drell-Yan low mass for the leading muon p_T (top left), subleading muon p_T (top right), leading b-jet p_T (bottom left) and subleading b-jet p_T (bottom right).

Table 5: Summary of HLT triggers for different primary datasets used in the signal and Drell-Yan validation regions for different years.

year	primary dataset	HLT path
2016	Double Muon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL
		HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
		HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL
		HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ
	Single Muon	HLT_Mu30_TkMu11
		HLT_IsoMu24
		HLT_IsoTkMu24
2017	Double Muon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8
		HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
		HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_Mass3p8
		HLT_Mu37_TkMu27
	Single Muon	HLT_IsoMu24
		HLT_IsoMu27
		HLT_Mu50
2018	Double Muon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
		HLT_Mu37_TkMu27
	Single Muon	HLT_IsoMu24
		HLT_IsoMu27
		HLT_Mu50

3 Physics object definition and preselection

This section is devoted to details about object reconstruction and preselection. The final event selection is equivalent to what is summarized in Table 9. This selection will result in a sample with few data and simulated events where the statistics for simulation is not enough to draw any conclusion about data-MC agreement. While this is not an issue since the analysis is fully based on data to estimate the background and derive the limits, we show distributions at few steps earlier, with enough statistics, to ensure that we have control on backgrounds and the optimization studies are reliable enough. Considering the dominant backgrounds to our search are $t\bar{t}$ and Drell-Yan, validation regions enriched with either of the two are also defined to bring more confidence into the optimization procedure in signal region (Section 3.1).

Events are filtered using a high-level trigger (HLT) requirement based on the presence of two muons with $p_T > 17$ and 8 GeV as listed in Table 5. In order to gain a bit in efficiency, single muon triggers detailed in the same table are also used with a logical OR. No correction is applied for difference in trigger efficiencies between data and MC since the data-to-simulation correction factors for this set of triggers are found to be close to unity (99%) [36].

For offline selection, events must contain at least one primary vertex, considered as the vertex of the hard interaction. At least four tracks must be associated to the selected primary vertex. The longitudinal and radial distances of the vertex from the center of the detector must be smaller than 24 cm and 2 cm, respectively. For events with more than one selected primary vertex, the one with the largest Σp_T^2 of the associated tracks is chosen for the analysis. Extra selection criteria are applied to leptons and jets reconstructed using the CMS particle flow algorithm [37].

The selection requires at least two isolated muons originating from the selected primary vertex

with $|\eta| < 2.4$. The leading muon p_T should exceed a threshold of 17 GeV whereas the sub-leading one is accepted with $p_T > 15$ GeV. Muons are identified with the loose criteria defined by the CMS “Muon POG”. The isolation variable I_{rel} is calculated by summing the transverse energy deposited by other particles in a cone of size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ around the muon, divided by the muon p_T . This quantity is required to be less than 0.15 (0.25) for a tight (loose) selection. In case more muons with those criteria appear in the event, the two with highest p_T are considered. Following the Muon POG recommendations, correction factors for reconstruction and selection efficiencies are applied on simulated events to be comparable with the data.

Jets are reconstructed by clustering the charged and neutral particles using an anti- k_T algorithm [38, 39] with a distance parameter of 0.4. The reconstructed jet energy is corrected for effects from the detector response as a function of the jet p_T and η . Furthermore, contamination from additional interactions (pileup), underlying events, and electronic noise are subtracted [40, 41]. To achieve a better agreement between data and simulation, an extra η -dependent smearing is performed on the jet energy of the simulated events [40, 41]. Events are required to have at least two jets with $|\eta| < 2.4$ and $p_T > 15$, where both jets must be separated from the selected leptons ($\Delta R > 0.4$). The DeepJet and DeepCSV algorithms are tried to identify b jets where in each case, the two selected jets in the event are required to pass the loose working point (see Section 4 where the decision is made to use DeepJet). Corrections are applied on simulation to bring the full b-tag discriminator shape in agreement with data. The use of shape-based corrections is motivated by the later categorization of events that relies on the b-tag discriminator value.

The signal events do not have any genuine neutrino and are not expected to produce any imbalance in the event transverse momentum. The missing transverse energy, \cancel{E}_T , is defined as the modulus of \vec{p}_T^{miss} , which is the negative vector p_T sum of all reconstructed particles. The jet energy calibration therefore introduces corrections to the \cancel{E}_T measurement. The optimized threshold for \cancel{E}_T is found to be 60 GeV.

The search for a new scalar, and therefore the final limits, are restricted to $m_{a_1} \in [15, 62.5]$. However for the selection, optimization and eventual background modeling, a slightly wider range is used to provide the full coverage at the boundaries. This means events with $m_{\mu\mu}$ out of $[15, 70]$ GeV are discarded. Figure 5 shows the distributions of the p_T of the leading and subleading muons and that of leading and subleading jets where b-jets are identified with the DeepCSV algorithm. Similar distributions are provided in Fig. 6 using DeepJet b-tagging algorithm is employed.

3.1 Validation regions

The main purpose of looking into validation regions is to make sure that optimization studies are robust enough. Other than that, no information of these regions are used in the analysis since background determination is purely based on data. Studies in $t\bar{t}$ and Drell-Yan validation regions do not point to any significant mismodeling. As long as the shapes are described well, possible normalization effects have no impact on the optimization results.

The $t\bar{t}$ validation region is defined using similar requirements as those of the signal region. The main difference is that one muon is replaced by an electron at online and offline selections. Events in this region should pass triggers listed in Table 6. Electrons with $p_T > 20$ GeV and $|\eta| < 2.4$ where $\Delta R(e, \mu) > 0.4$ are selected. Electrons within the calorimeter gap are discarded. The working point with 80% efficiency of Fall17v2 identification is used. It is based on a multivariate training and is recommended by the EGamma POG. There is no \cancel{E}_T requirement

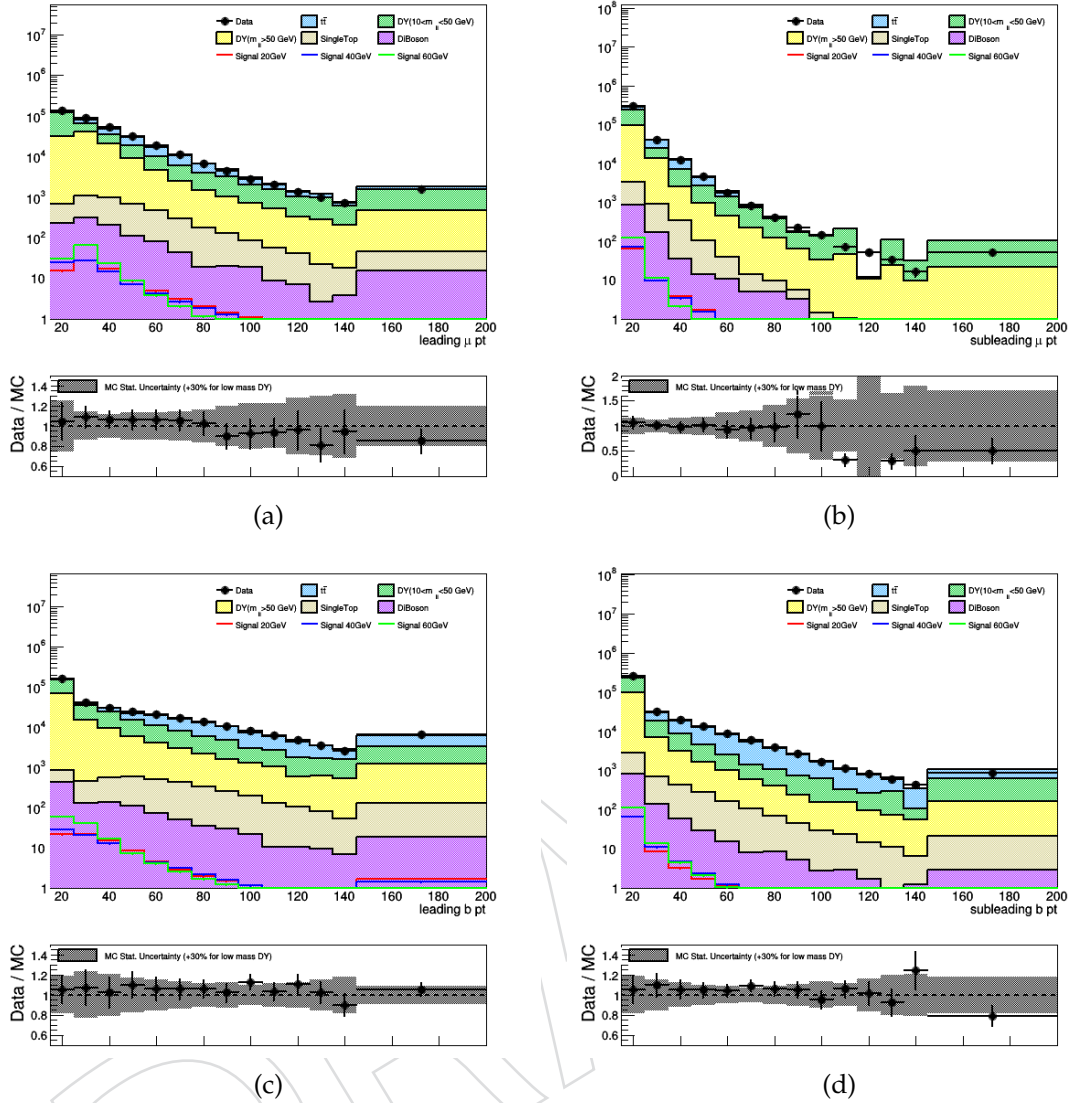


Figure 5: The distribution of leading and subleading muon p_T (a,b) and leading and subleading jet p_T (c,d) after asking for two muons and two b jets with the loose working point of the DeepCSV algorithm. Uncertainty band represents the limited size of MC samples together with 30% uncertainty on low mass DY cross section. Simulated samples are normalized to 138 fb^{-1} with their theoretical cross sections.

in this region. In terms of normalization, simulation is found to underestimate the data by about 10% while shapes are described well. The $t\bar{t}$ normalization is increased by a factor of 1.1 in Fig. 7 where distributions for muon p_T , $m_{e\mu}$, and leading and subleading b-jets in this region are shown.

To construct the the Drell-Yan region, the selection requirements of signal region are preserved except the boundaries on $m_{\mu\mu}$ which are changed to be compatible with the Z-peak, i.e., $75 < m_{\mu\mu} < 105 \text{ GeV}$. Figure 8 shows similar distributions as those of Fig. 6 in this region. No major disagreement is observed.

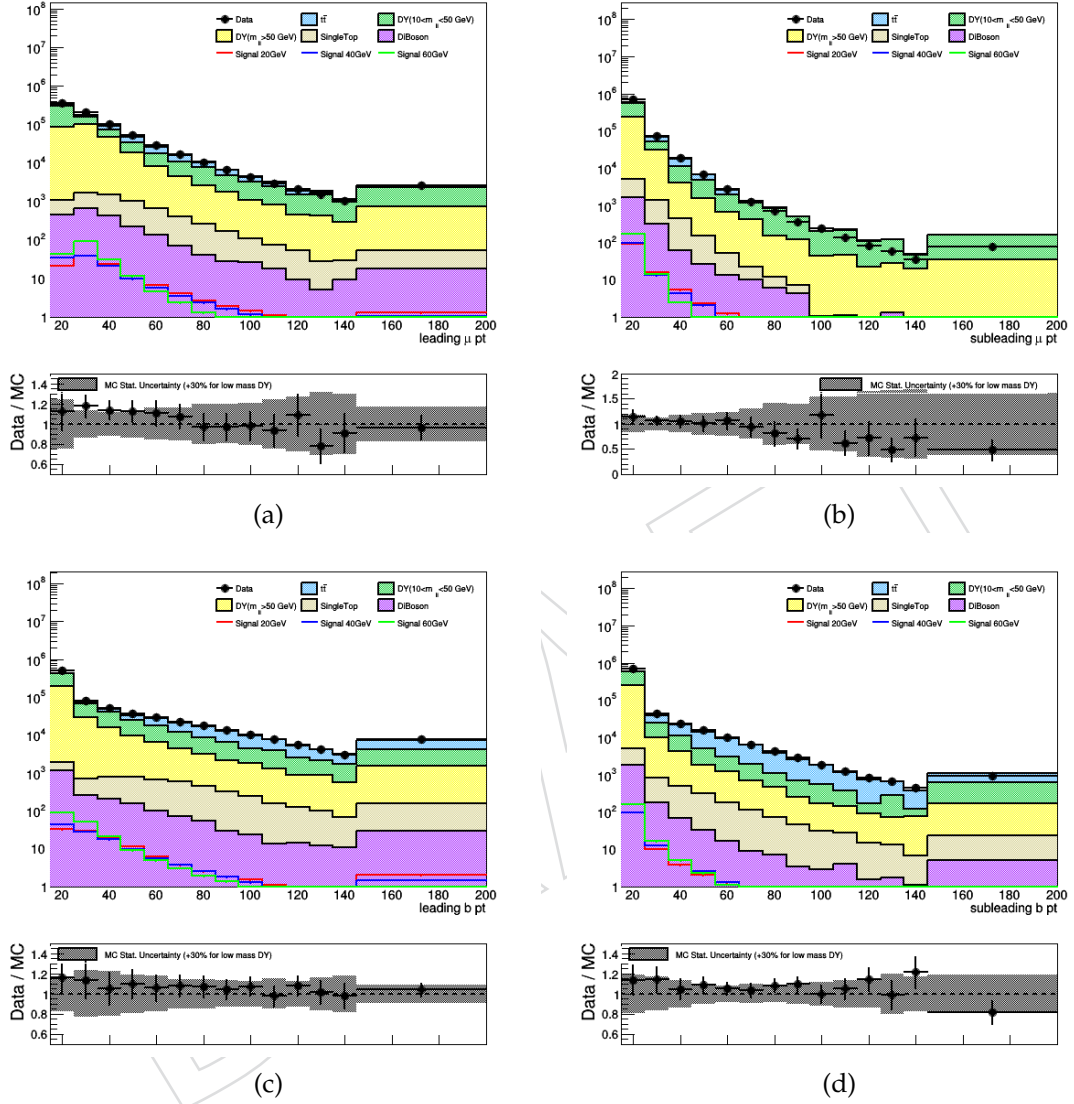


Figure 6: The distribution of leading and subleading muon p_T (a,b) and leading and subleading jet p_T (c,d) after asking for two muons and two b jets with the loose working point of the DeepJet algorithm. Uncertainty band represents the limited size of MC samples together with 30% uncertainty on low mass DY cross section. Simulated samples are normalized to 138 fb^{-1} with their theoretical cross sections.

Table 6: List of high level triggers used to select events in $t\bar{t}$ validation region.

year	primary dataset	HLT path
2016	Muon EG	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL
	Single Muon	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu33_Ele33_CaloIdL_GsfTrkIdVL HLT_IsoMu24 HLT_IsoTkMu24 HLT_Mu50
2017	Muon EG	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Single Muon	HLT_IsoMu24 HLT_IsoMu27 HLT_Mu50
2018	Muon EG	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Single Muon	HLT_Mu27_Ele37_CaloIdL_MW HLT_Mu37_Ele27_CaloIdL_MW HLT_Ele32_WPTight_Gsf_L1DoubleEG HLT_IsoMu24 HLT_IsoMu27 HLT_Mu50

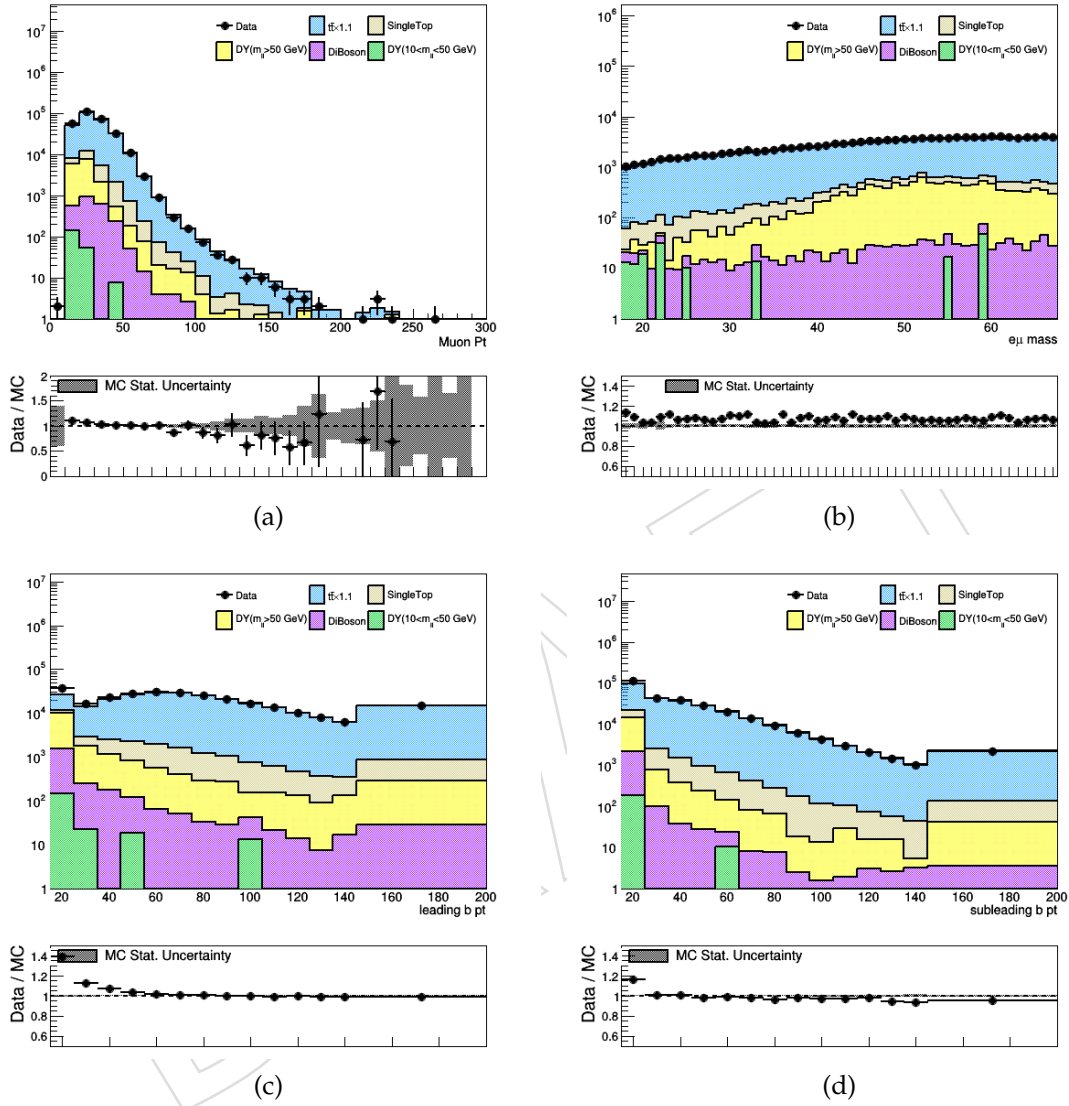


Figure 7: The distribution of the muon p_T , the invariant mass of electron and muon (a,b) and leading and subleading jet p_T (c,d) in $t\bar{t}$ validation region after asking for one muons and one electron and two b jets with the loose working point of the DeepJet algorithm. Uncertainty band represents the limited size of MC samples. Except $t\bar{t}$, simulated samples are normalized to 138 fb^{-1} with their theoretical cross sections.

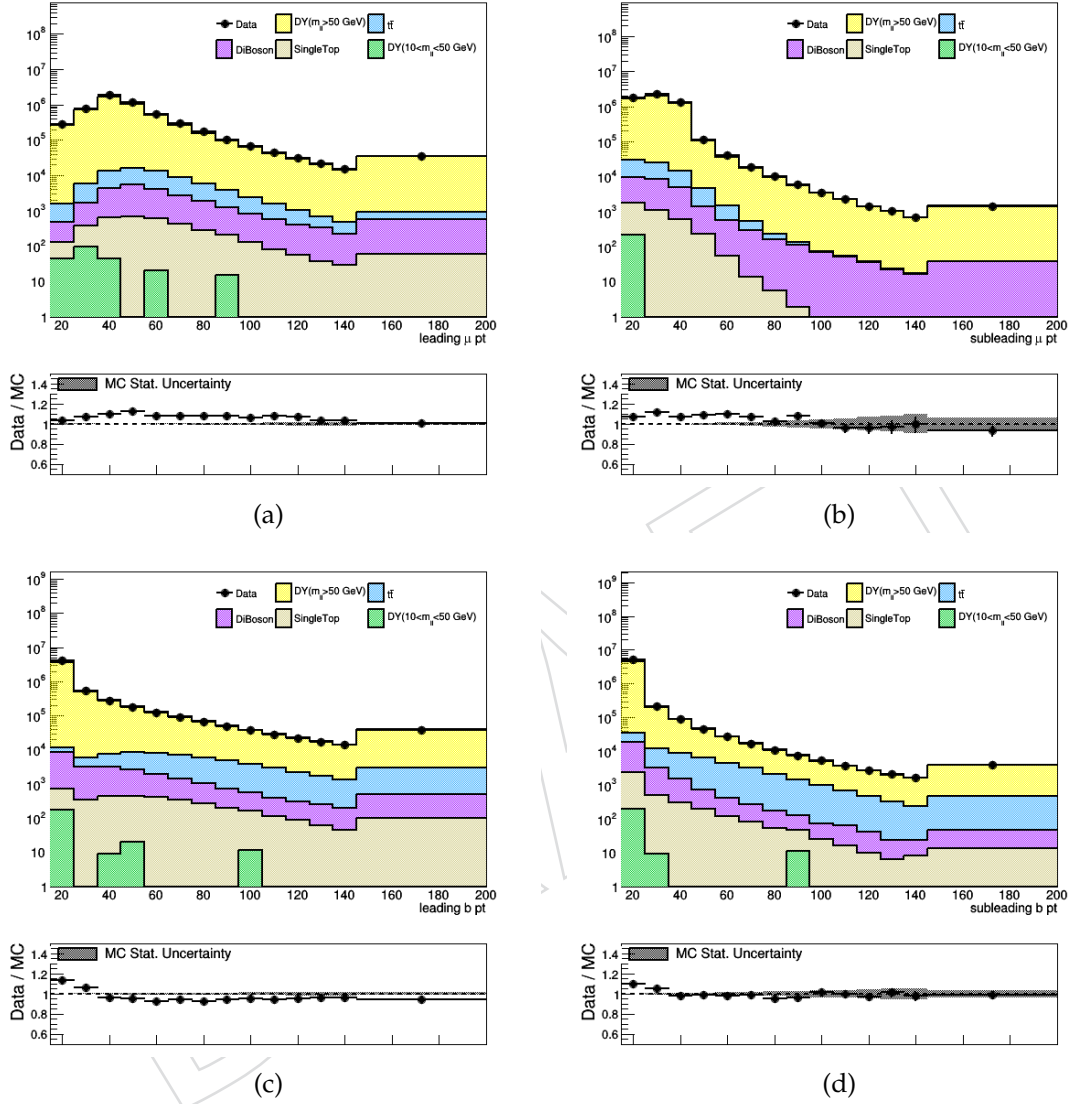


Figure 8: The distribution of leading and subleading muon p_T (a,b) and leading and subleading jet p_T (c,d) in Drell-Yan validation region after asking for one muons and one electron and two b jets with the loose working point of the DeepJet algorithm. Uncertainty band represents the limited size of MC samples. Simulated samples are normalized to 138 fb^{-1} with their theoretical cross sections.

4 Object-level optimization

Distributions of simulated preselected events agree with the data using loose muon identification and isolation as well as loose and b-tagging requirements. Preselected events in simulation are used to find optimal thresholds of those criteria. The figure of merit is a significance defined as $S_{m_{a_1}} / \sqrt{B}$ where $S_{m_{a_1}}$ is the expected signal yield for a given m_{a_1} hypothesis and B is the background yield (mainly $t\bar{t}$ and Drell-Yan) in a 5 GeV interval around m_{a_1} . As an example, Fig. 9 shows significance as a function of m_{a_1} for different muon isolation scenarios. The identification requirement of muon meets the tight criteria. Optimization is performed at every year, leading to keep the muons loosely isolated.

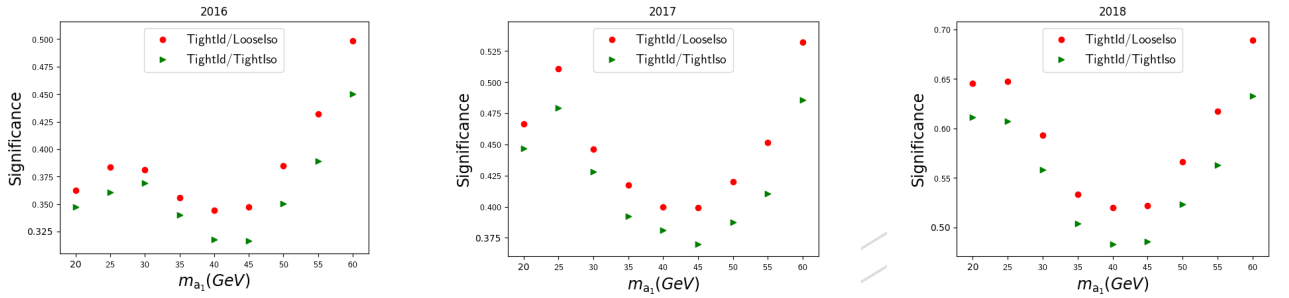


Figure 9: Significance as a function of m_{a_1} for different choices of muon isolation in years 2016 (left), 2017 (middle) and 2018 (right).

In Fig. 10 (upper row) the difference in significance is shown between DeepJet and DeepCSV algorithms for different years where DeepJet outperforms the other. In lower row, various combinations of working points for the two b-jets are compared. Starting from events with two “loose” b-jets we study if tightening the b-tag criteria on the jets could increase the sensitivity. In each combination, the more stringent criterion is applied on the jet with the higher b-tag score. There is a competition between medium-loose (ML) tight-loose (TL). The TL combination seems to work slightly better.

Table 7 shows the number of data events together with the expected yield for different backgrounds. The yields in data and simulation are in agreement within 5%. For different m_{a_1} values, expected number of signal events are also provided. It can be seen that for a generic signal cross section, the signal-to-background ratio is low in the preselected sample and there is room for improvement.

Using the optimized preselection criteria, the p_T distributions of (sub)leading muons and b-jets are shown in Fig. 11. In Fig 12, the p_T and the mass ($m_{\mu\mu}$) of the dimuon system, the p_T and the mass (m_{bb}) of the di-b-jet system, as well as the p_T and the mass of the Higgs candidate ($m_{\mu\mu bb}$) are presented. The agreement between data and simulation is overall satisfactory.

In summary, the optimization at object level leads to tight identification and loose isolation for muons together with tight-loose b-jet identification for the two jets using DeepJet b-tagging algorithm. Optimization continues at event level in the next section where in the end, a summary is presented in Table 9.

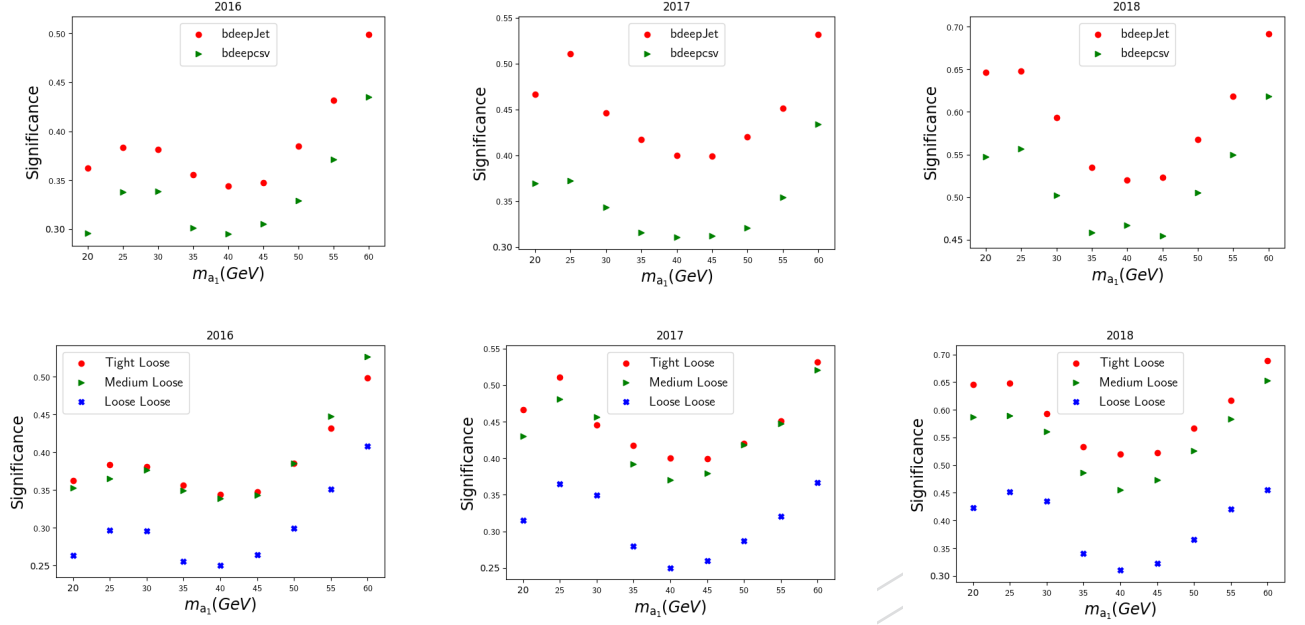


Figure 10: Significance as a function of m_{a_1} for different choices of b-algorithms (top) and Deep-Jet working point combinations (bottom) in years 2016 (left), 2017 (middle) and 2018 (right).

Table 7: Event yields for data and simulated processes using optimized object selection. The expected number of simulated events is normalized to the integrated luminosity of 138 fb^{-1} .

(*) : $\sigma_{\text{sig.}} \times \text{Br} \approx 7 \text{ fb}$

Process		Yield
$t\bar{t}$		56701 ± 56
Single top		2892 ± 25
Drell-Yan ($10 < m_{\ell\ell} < 50$)		25491 ± 7327
Drell-Yan ($50 < m_{\ell\ell}$)		18523 ± 283
Diboson		267 ± 11
Total expected background		103874 ± 7332
Data		109821
Signal for ggF(*)		
$m_{a_1} = 20 \text{ GeV}$ 75.2	$m_{a_1} = 25 \text{ GeV}$ 70.8	$m_{a_1} = 30 \text{ GeV}$ 78.0
$m_{a_1} = 35 \text{ GeV}$ 80.6	$m_{a_1} = 40 \text{ GeV}$ 69.2	$m_{a_1} = 45 \text{ GeV}$ 71.6
$m_{a_1} = 50 \text{ GeV}$ 77.5	$m_{a_1} = 55 \text{ GeV}$ 82.17	$m_{a_1} = 60 \text{ GeV}$ 90.9
Signal for VBF		
$m_{a_1} = 20 \text{ GeV}$ 7.2	$m_{a_1} = 40 \text{ GeV}$ 7.2	$m_{a_1} = 60 \text{ GeV}$ 7.5

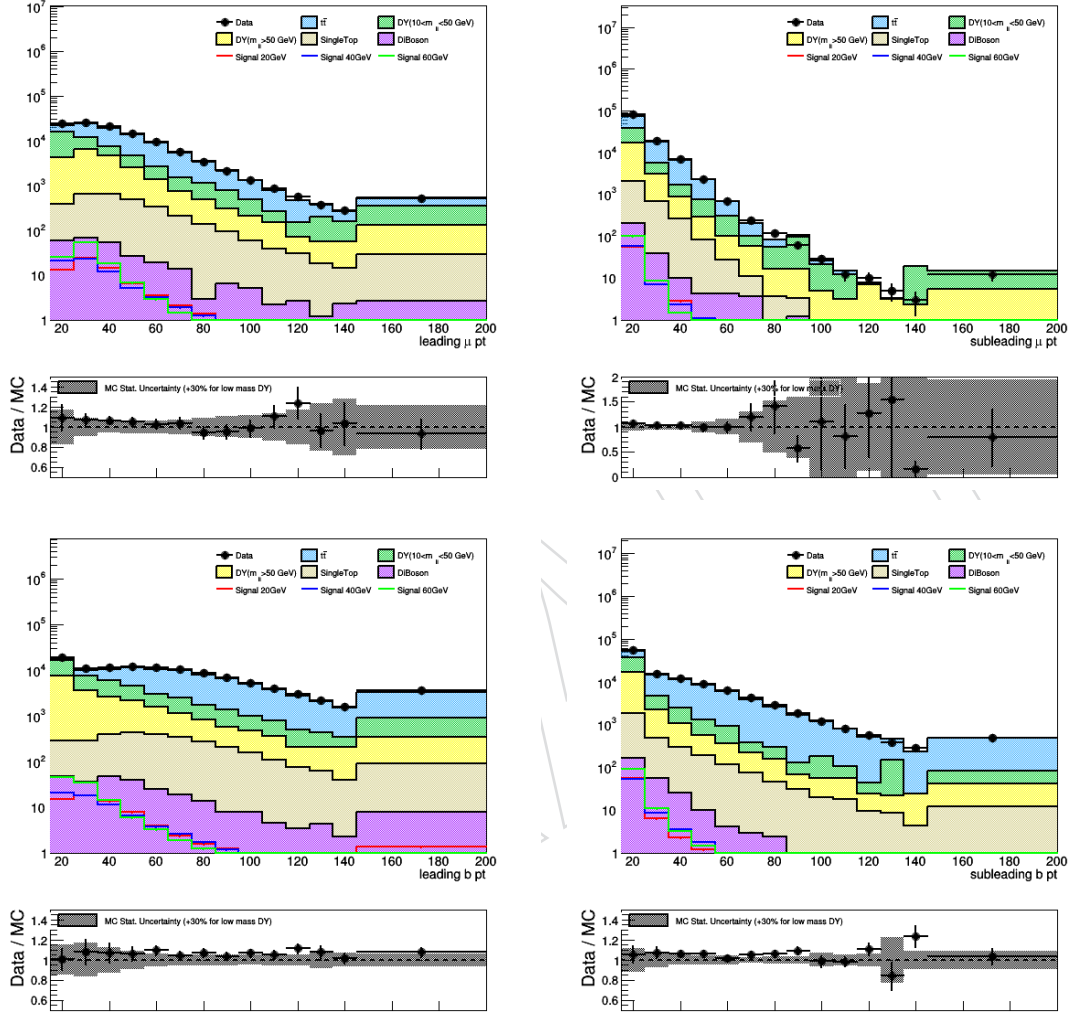


Figure 11: The p_T distributions of the leading (left) and subleading (right) muon (top) and b-jets (bottom) after using the optimized preselection requirements. Simulated samples are normalized to 138 fb^{-1} . The band in the lower panel of each plot represents the MC statistical uncertainty.

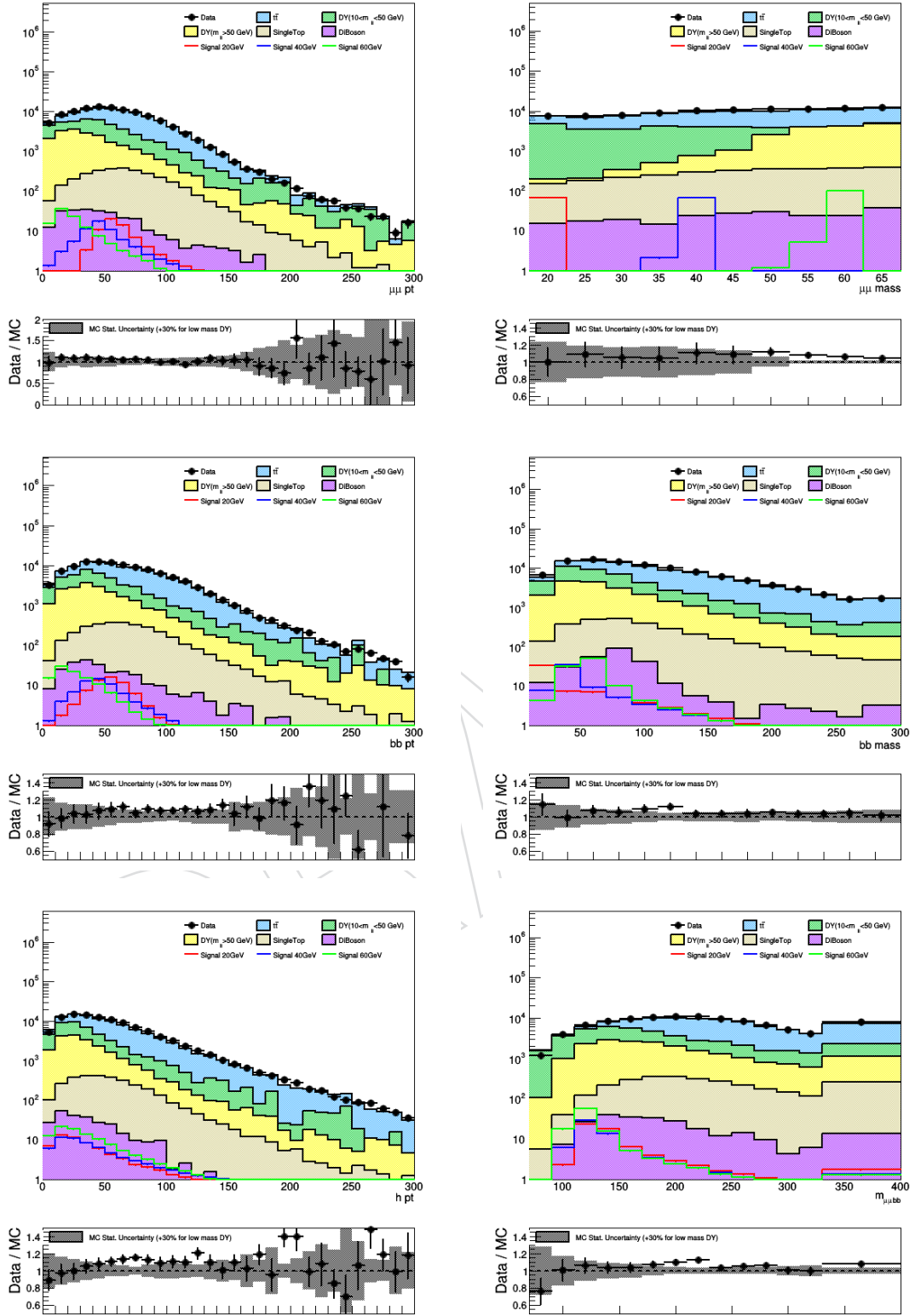


Figure 12: The p_T (left) and the invariant mass (right) distributions of the dimuon system (top), di-b-jets system (middle) and the Higgs candidate (bottom) after using the optimized preselection requirements. Simulated samples are normalized to 138 fb^{-1} . The band in the lower panel of each plot represents the MC statistical uncertainty.

5 Final selection and categorization

In order to reject more backgrounds and improve the sensitivity of the search, a χ_{tot}^2 quantity is defined as,

$$\chi_{\text{tot}}^2 \equiv \chi_{\text{bb}}^2 + \chi_{\text{H}}^2 \quad (1)$$

$$\chi_{\text{bb}} \equiv \frac{(m_{\text{bb}} - m_{\mu\mu})}{\sigma_{\text{bb}}} \quad (2)$$

$$\chi_{\text{H}} \equiv \frac{(m_{\mu\mu\text{bb}} - 125)}{\sigma_{\text{H}}} \quad (3)$$

where m_{bb} and $m_{\mu\mu}$ are the masses of di-b-jet and dimuon systems, respectively. The resolution of di-b-jet system is denoted by σ_{bb} , and σ_{H} is the resolution of the $\mu\mu\text{bb}$ system. Resolutions are derived from Gaussian fits to m_{bb} and $m_{\mu\mu\text{bb}}$ where objects are matched to the MC truth. Variable χ_{tot} is calculated on an event-by-event, encoding some crucial features of the signal event which are not necessary true for backgrounds. It is expected to be small in signal events since one expects $m_{\text{a}_1}(\text{bb}) = m_{\text{a}_1}(\mu\mu)$ and $m_{\mu\mu\text{bb}} = m_{\text{H}}$. For the analysis of 2016 data [25], the study of resolution terms showed that σ_{bb} increases with m_{a_1} while σ_{H} remains almost unchanged. This feature is reconfirmed using new simulations for different years, as shown in Fig. 13. In order to place a single cut on χ_{tot} over the entire m_{a_1} range, σ_{bb} is modeled with m_{a_1} and the dependence is included in Eq. 1.

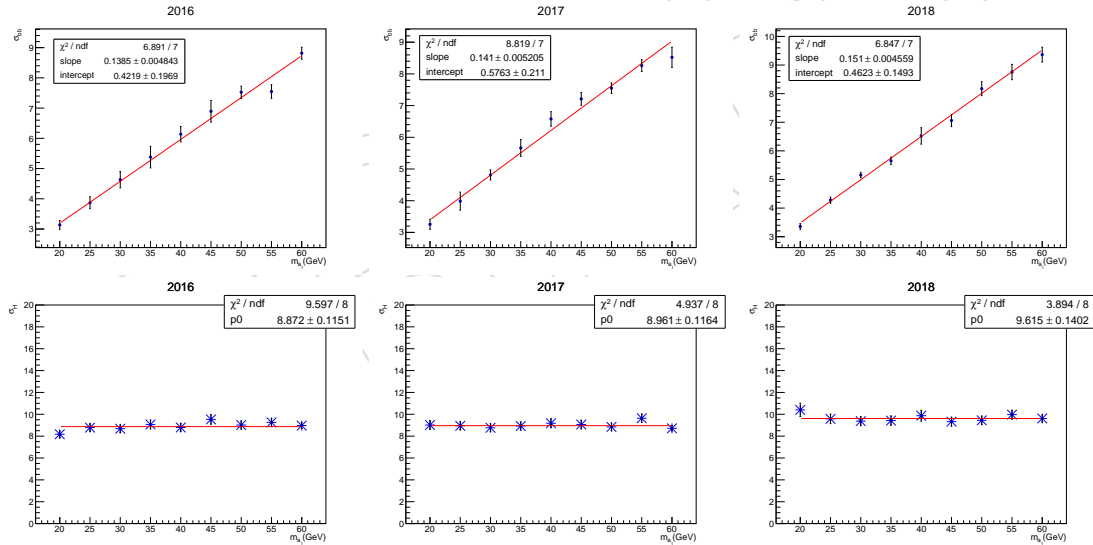


Figure 13: Dependence of σ_{bb} (top) and σ_{H} (bottom) on m_{a_1} in signal samples of 2016 (left), 2017 (middle) and 2018 (right).

As presented in Fig. 14, the χ_{bb} and χ_{H} are correlated. Also, it looks like the mean value of these quantities is not exactly placed at zero. While the bidimensional distribution of χ_{bb} versus χ_{H} looks elliptical in signal samples, placing a circular cut, i.e. $\chi_{\text{tot}}^2 < X$, is sub-optimal. Also, such circular cut assumes signal being centered at zero which is not exactly the case. These two effects were ignored in the previous analysis [25]. Here however a detailed study is pursued to improve these shortcomings.

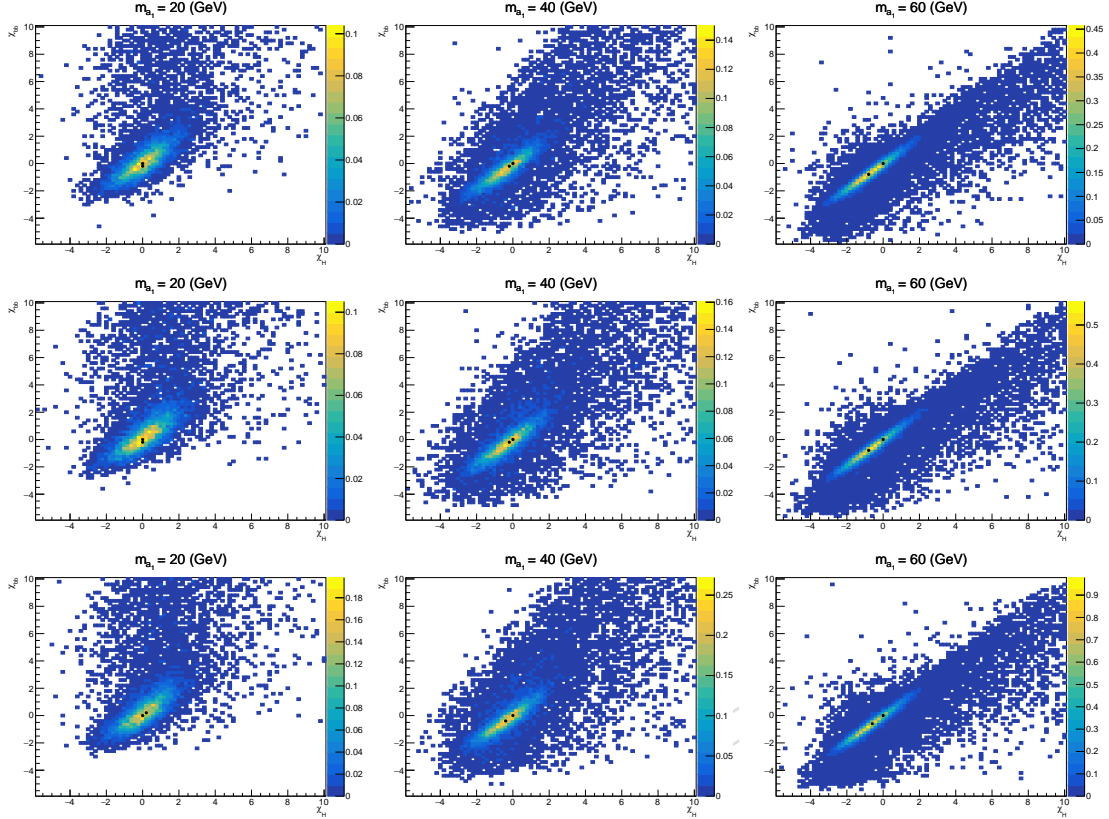


Figure 14: Bidimensional distributions of χ_{bb} versus χ_H in signal samples with different m_{a_1} values of 2016 (top), 2017 (middle) and 2018 (bottom).

5.1 Offset correction

From Fig. 14 it is clear that the offset in the mean value changes with m_{a_1} . In order to correct this effect, the difference of m_{bb} and $m_{\mu\mu bb}$ with their nominal values (m_{a_1} and 125 respectively) are modeled by a line as function of m_{a_1} . Line parameters for each year are obtained by fitting on the simulated mass points. The offset values can be found in Fig. 15.

Shifted χ values are defined as

$$\chi_{bb,s} \equiv \chi_{bb} - (a \times m_{\mu\mu} + b) \quad (4)$$

$$\chi_{H,s} \equiv \frac{(m_{\mu\mu bb} - 125)}{\sigma_H} - (c \times m_{\mu\mu} + d) \quad (5)$$

$$(6)$$

where a, b, c and d are the results of the linear fit. The distribution of the $\chi_{bb,s}$ versus $\chi_{H,s}$ are shown in Fig. 16. It can be seen from the plots that shifted values are more centered around zero.

5.2 Decorrelation of χ_{bb} and χ_H

In order to decorrelate $\chi_{bb,s}$ and $\chi_{H,s}$ the principal component analysis (PCA) method is used. The idea behind this method is to use the inverse of the covariance matrix of two variables to define two new uncorrelated variables. The covariance matrix is a 2×2 symmetric matrix, so it

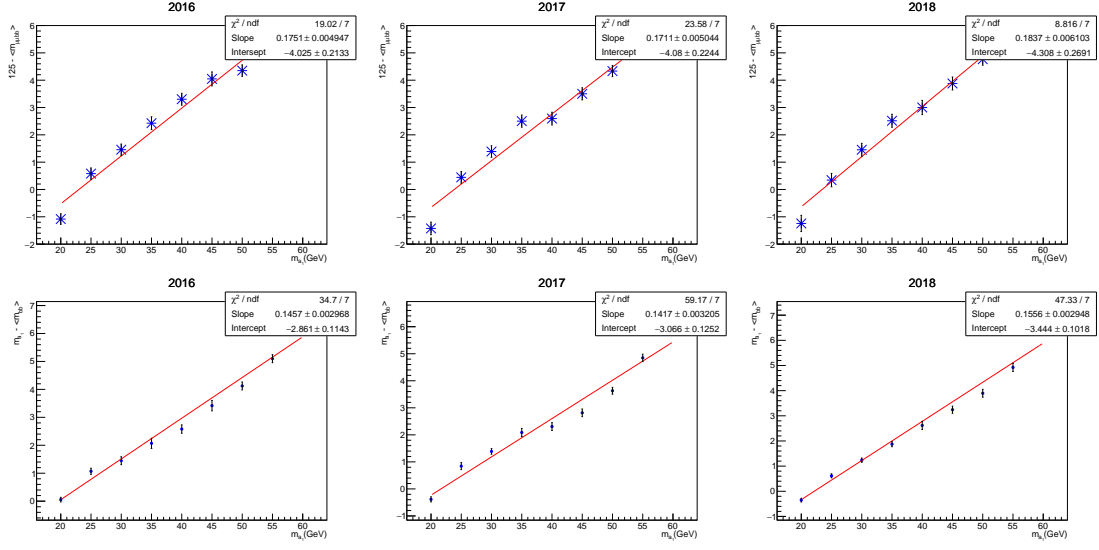


Figure 15: Dependence of offset in χ_H (top) and χ_{bb} (bottom) on m_{a_1} in signal samples of 2016 (left), 2017 (middle) and 2018 (right).

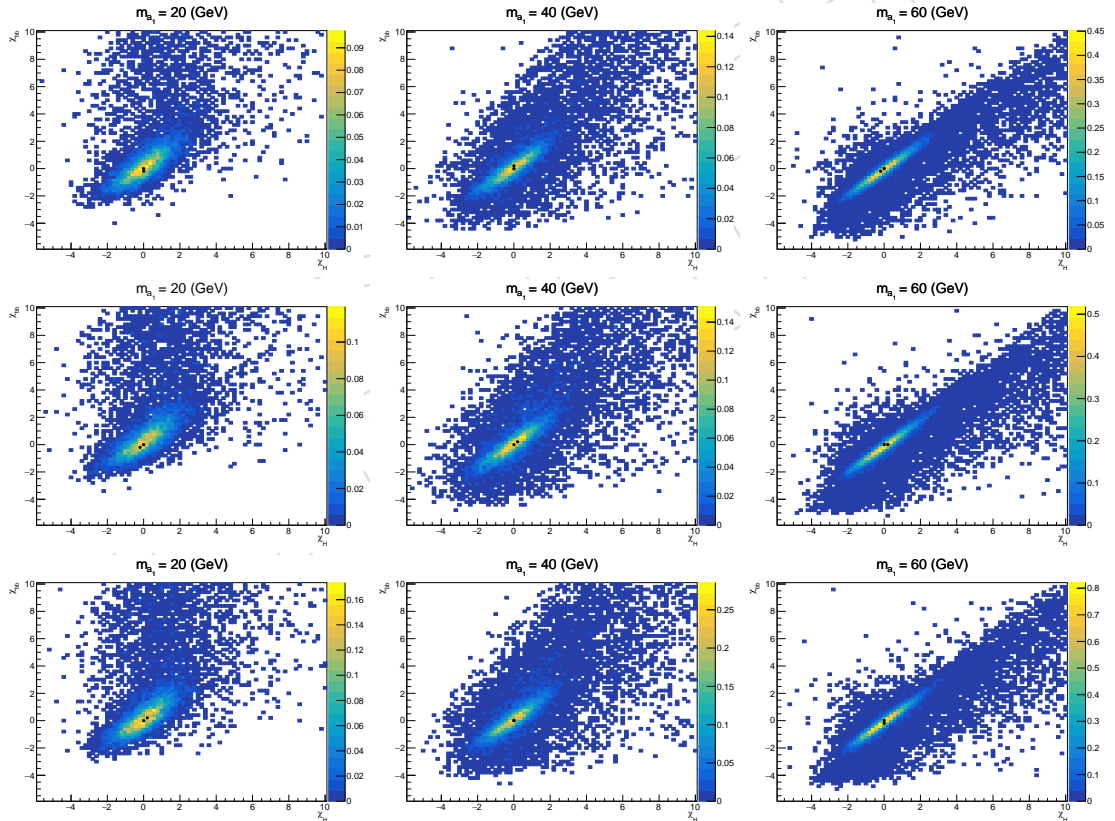


Figure 16: Bidimensional distributions of $\chi_{bb,s}$ versus $\chi_{H,s}$ in signal samples with different m_{a_1} values of 2016 (top), 2017 (middle) and 2018 (bottom).

252 can be parametrized by its eigen values (λ_1, λ_2) and the components of one of its eigen vectors
 253 ($\begin{pmatrix} a \\ b \end{pmatrix}$). It can be shown that the two following variables are decorrelated:

$$\chi_{H,D} \equiv \frac{a}{\sqrt{\lambda_1}} (\chi_{H,s} + \frac{b}{a} \chi_{bb,s}) \quad (7)$$

$$\chi_{bb,D} \equiv \frac{a}{\sqrt{\lambda_2}} (-\frac{b}{a} \chi_{H,s} + \chi_{bb,s}) \quad (8)$$

$$\chi_D^2 \equiv \chi_{H,D}^2 + \chi_{bb,D}^2 \quad (9)$$

So this transformation can be parametrized using three parameters: $\frac{a}{\sqrt{\lambda_1}}$, $\frac{a}{\sqrt{\lambda_2}}$ and $\frac{b}{a}$ where we call it $\tan(\alpha)$. Polynomial functions are used to model the dependence of these variables to m_{a_1} . The fitted functions are shown in Fig 17 for all the years. To perform the transformation, the transformation matrix is calculated for each at $m_{\mu\mu}$. Correlation of $\chi_{H,D}$ and $\chi_{bb,D}$ for $m_{a_1} = 20, 40, 60\text{GeV}$ is shown in Fig 18.

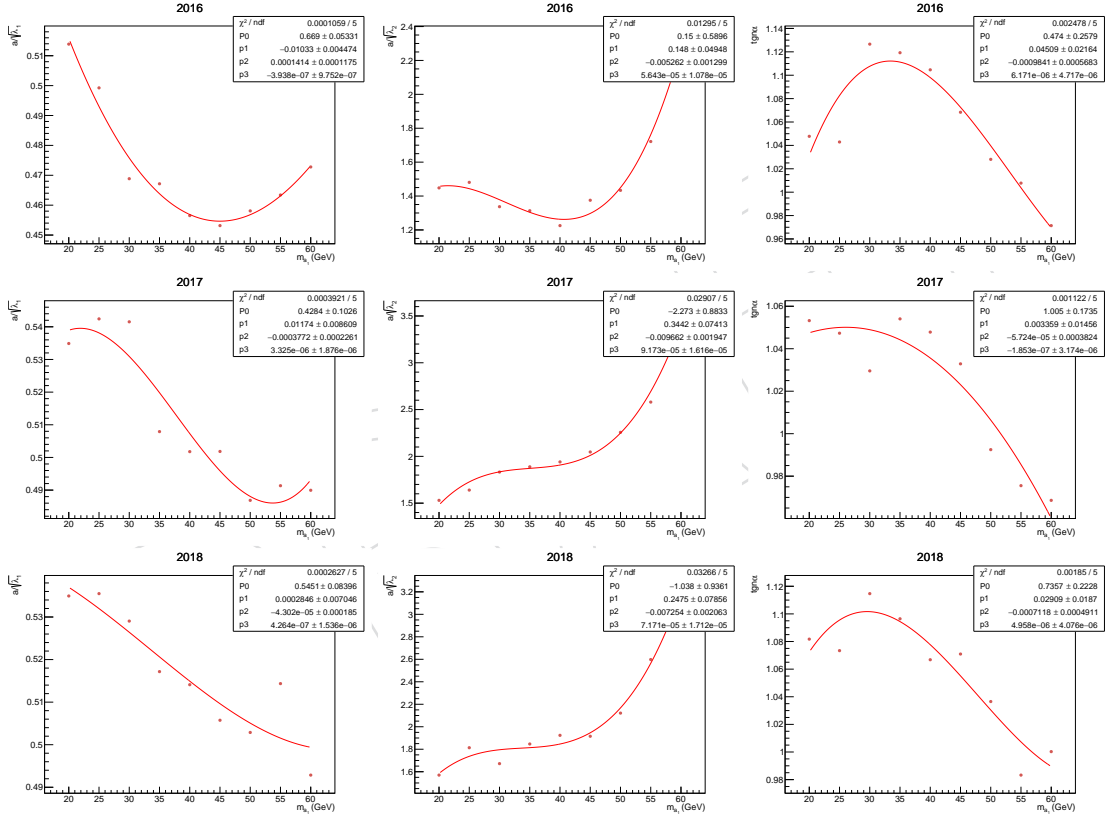


Figure 17: Dependence of correlation matrix parameters on m_{a_1} , $\frac{a}{\sqrt{\lambda_1}}$ (left), $\frac{a}{\sqrt{\lambda_2}}$ (middle) and $\tan(\alpha)$ (right) in signal samples of 2016 (top), 2017 (middle) and 2018 (bottom).

5.3 Optimized χ_D

The distribution of the χ_D^2 for signal and backgrounds and a zoomed version of it are shown in Fig 19. Data in these plots are blinded for $\chi_D^2 < 3$. A similar procedure to that of Section 4, the optimal cut is found to be $\chi_D^2 < 1.5$ (Fig. 20). Table 8 summarizes the expected yields after this requirement. A clear improvement is visible in comparison with Table 7. In the final extraction of limits, the signal yield for mass points without simulation is obtained using Spline. Figure 21 shows the interpolation of yields for the g g H and VBF signals.

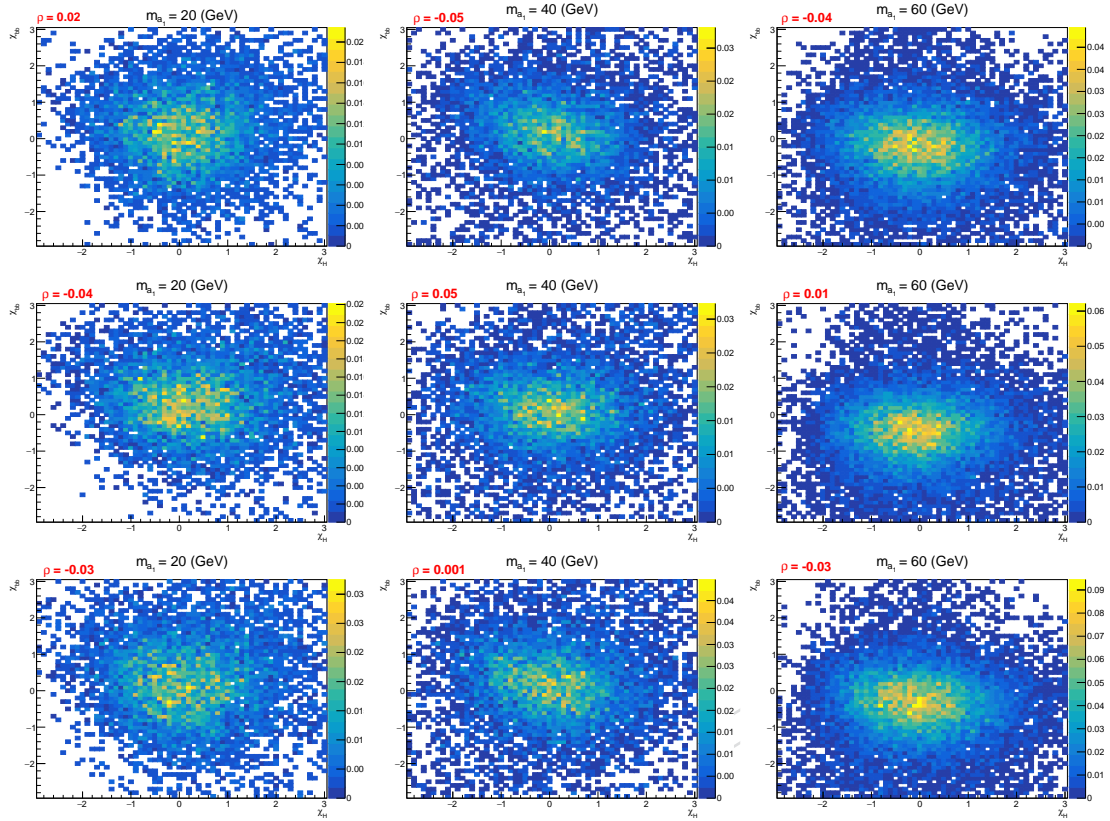


Figure 18: Bidimensional distributions of $\chi_{bb,D}$ versus $\chi_{H,D}$ in signal samples with different m_{a1} values of 2016 (top), 2017 (middle) and 2018 (bottom).

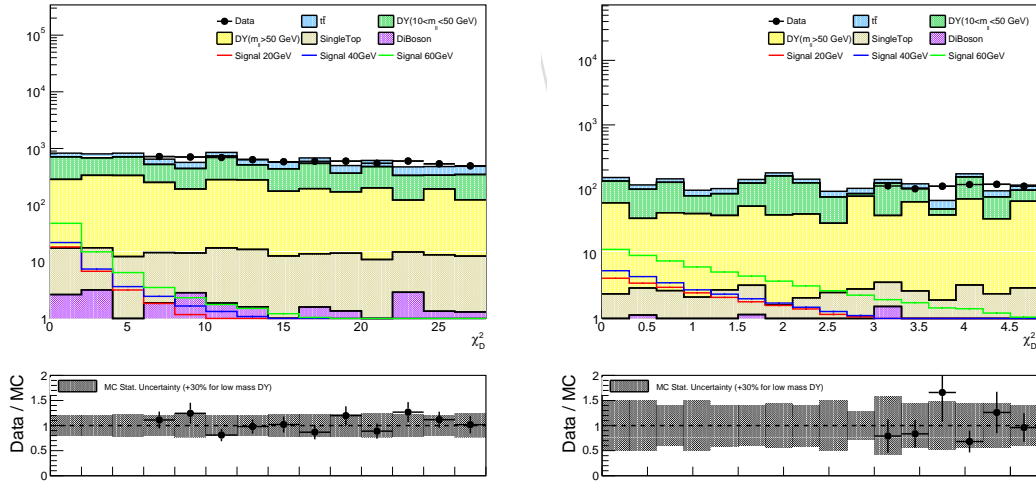


Figure 19: Distribution of χ_D^2 for the full Run-II dataset in a wide range (left) and zoomed view (right).

5.4 Categorization

Finally a categorization of events is performed. First, events with at least one jet below 20 GeV are placed in a single category called Low p_T . The b-tagging SFs at 20 GeV were applied to these

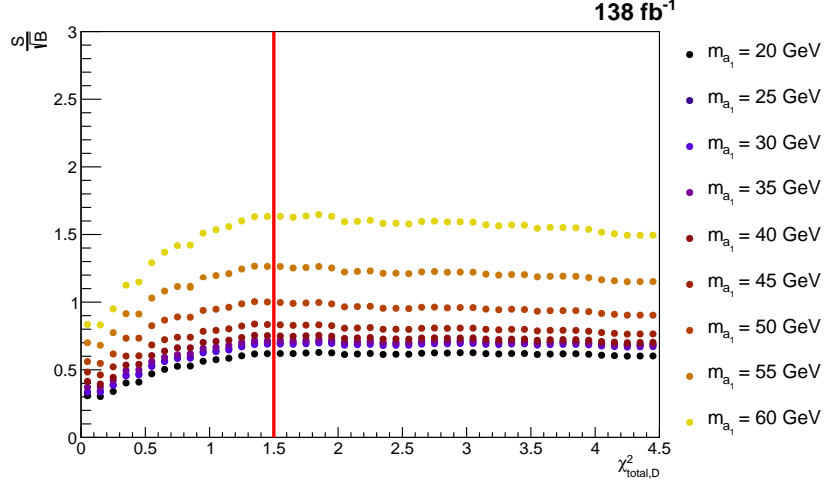


Figure 20: Significance as a function of the cut on the χ^2_D value for all m_{a_1} masses. The results are based on the full Run-II simulation.

Table 8: Event yields for simulated processes after applying $\chi^2_D < 1.5$ cut. The expected number of simulated events is normalized to the integrated luminosity of 138 fb^{-1} . The type-III parametrization of 2HDM+S with $\tan \beta = 2$ is used to evaluate $\text{Br}(a_1 \rightarrow ff)$.

Process		Yield
Top		86.3 ± 2.25
Drell-Yan ($10 < m_{\ell\ell} < 50$)		289.6 ± 89.5
Drell-Yan ($50 < m_{\ell\ell}$)		200.2 ± 31.9
Diboson		1.5 ± 0.9
Single Top		11.4 ± 1.6
Total expected background		589.05 ± 95.09
Data		641
Signal for ggH ($\mu\mu b\bar{b}$)		
$m_{a_1} = 20 \text{ GeV}$	$m_{a_1} = 40 \text{ GeV}$	$m_{a_1} = 60 \text{ GeV}$
15.4 ± 0.2	18.7 ± 0.2	40.5 ± 0.3

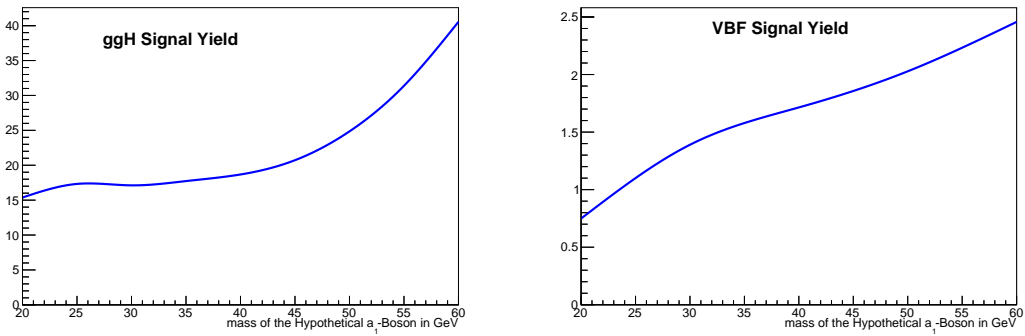


Figure 21: Interpolation of the expected signal yield for the g g H (left) and VBF (right) Higgs production mechanisms.

events and their uncertainties were doubled. These events don't categorize based on their B-tag value. This approach was presented at the BTV-WG meeting, and the group was satisfied with the application of SFs and uncertainties in our analysis. See further details in Appendix A. Then, events containing two additional jets events with $|\eta| < 4.7$ and $p_T > 30$ GeV are separated from the rest if the invariant mass of the two jet meets the requirement of $m_{jj} > 250$ GeV. Rest of the events are categorised based on the b-tag discriminator value of the loose b-jet. Three exclusive categories are defined where the loose b-jet is not medium (TL), is medium but not tight (TM), and is tight (TT). Such categorization relies on the fact that events with real and misidentified b-quark jets distribute differently in those categories. As shown in Fig.22, the majority of background events ($\approx 70\%$) fall into the Low p_T category, and for signal(ggH) about 40% of signal events end up in the Low p_T category.

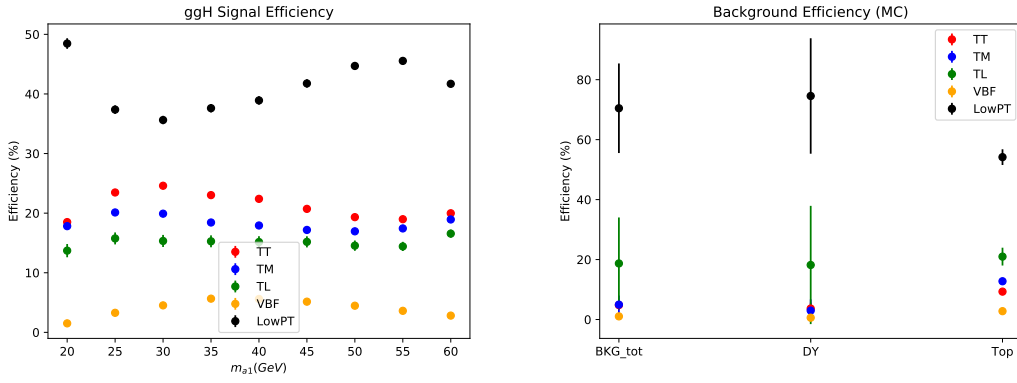


Figure 22: Fraction of ggH (left) and background events (right) in each category. For signal samples, fractions are provided for different m_{a_1} values.

We calculated the expected limit in order to assess how sensitivity increases after categorization while blinded to the data in signal regions. We used the signal and background models described respectively in Sections 6 and Appendix B. Figure 23 illustrates the expected limits for individual categories and for their combination, and it is apparent that categorization increased sensitivity. Limits include systematic uncertainties from Section 8.

Table 9 summarizes the final selection and categories of the current analysis, and the yield of expected background and signal samples in each category are summarized in Table 10.

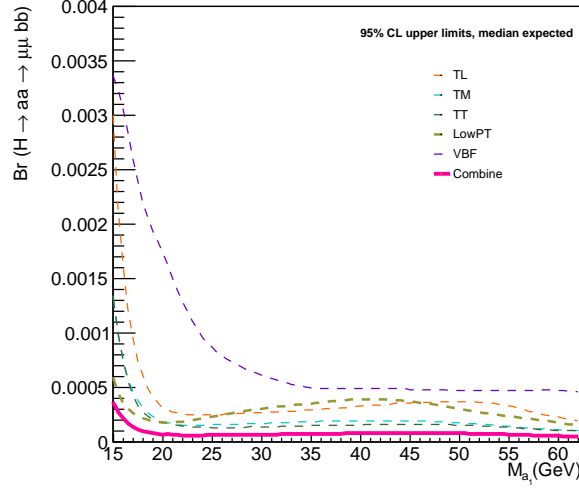


Figure 23: Upper limits on $\text{Br}(H \rightarrow a_1 a_1 \rightarrow \mu^+ \mu^- b \bar{b})$ at 95% CL for individual categories and their combination as a function of m_{a_1} .

Table 9: Summary of object and event selection requirements together with categorization, following optimization studies.

Physics objects and event kinematic requirements	
Leading and subleading muon p_T (GeV)	> 17 and 15
Muons identification and isolation	tight-loose
Leading and subleading jet p_T (GeV)	> 15
b-tagging	tight-loose
\cancel{E}_T (GeV)	< 60
χ_D^2	< 1.5
Categories for selected events	
Low p_T , VBF, two add. jets TT, TM, TL	Events with at least one jet below 20 GeV $p_T > 30$ GeV, $ \eta < 4.7$ and $m_{jj} > 250$ GeV !VBF and T(T, M!T, L!M)

Table 10: The yield of expected background and signal samples in each category. Since the individual numbers were rounded, the total is slightly different from the sum of the individual numbers.

Category	Signal for $ggH(\mu\mu bb)$			Expected background	Data
	$m_{a_1} = 20 \text{ GeV}$	$m_{a_1} = 40 \text{ GeV}$	$m_{a_1} = 60 \text{ GeV}$		
TL	2.1 ± 0.06	2.8 ± 0.07	6.7 ± 0.11	109 ± 30	121
TM	2.7 ± 0.06	3.3 ± 0.07	7.7 ± 0.11	27 ± 15	40
TT	2.8 ± 0.06	4.2 ± 0.08	8.1 ± 0.1	28 ± 10.8	26
VBF	0.2 ± 0.02	1.0 ± 0.04	1.1 ± 0.4	5 ± 2	8
Low p_T	7.4 ± 0.11	7.3 ± 0.11	17 ± 0.17	421 ± 88	446
Total	15.4 ± 0.2	18.7 ± 0.2	40.5 ± 0.3	589.05 ± 95.09	641

6 Signal modeling

The procedure of signal modeling has already been validated in Run I and 2016 Run II publications [21, 25]. The shape of signal is derived from simulation and is a combination of a Voig profile, V , and a Crystal ball profile, CB ,

$$S(m_{\mu\mu}|f, p_V, p_{cb}) \equiv f \cdot V(m_{\mu\mu}|p_V) + (1 - f) \cdot CB(m_{\mu\mu}|p_{cb}), \quad (10)$$

with f being the relative contribution of the Voig profile, p_V being the set of parameters for the Voig and p_{cb} being the set of parameters for the Crystal ball profile functions. Figures 24, 25, and 26 show the signal distributions together with fitted functions where a reasonable signal description is obtained by the model for simulation in all three years.

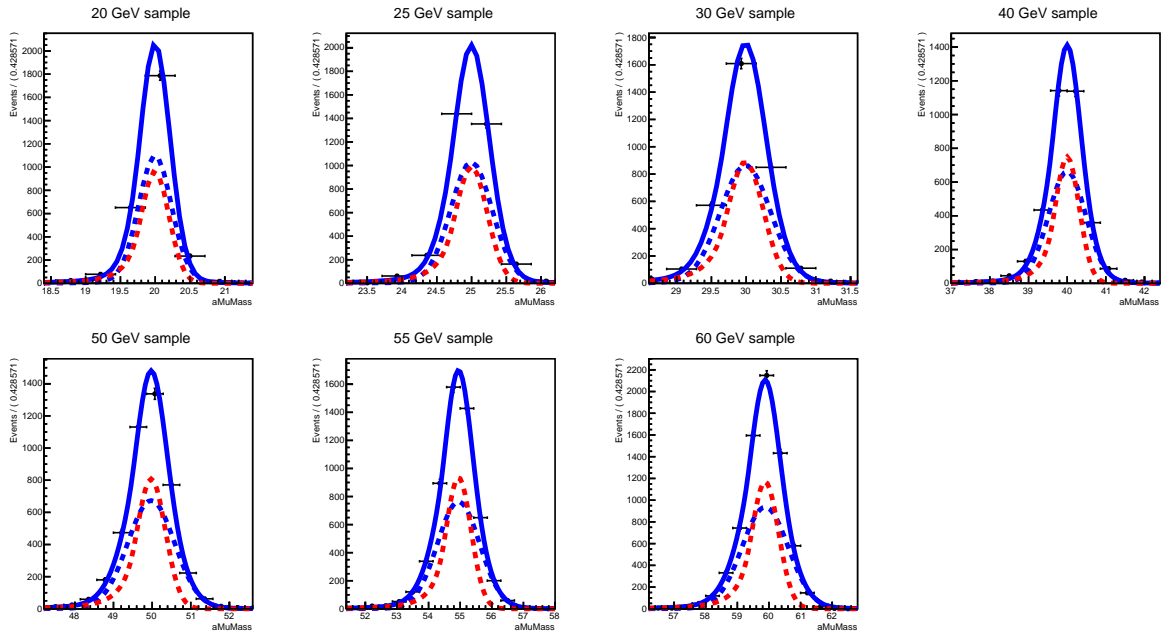


Figure 24: Results of the signal model $S(m_{\mu\mu}|f, p_V, p_{cb})$, fitted independently to $m_{\mu\mu}$ distributions at different mass points in 2016 simulated samples. Solid line is the constructed model. Dotted red line is the crystal-ball and dotted blue line is the Voig component.

In the signal model, $S(m_{\mu\mu}|f, p_V, p_{cb})$, there are correlations amongst f , p_V and p_{cb} and they could in principal change with m_{a_1} .

We account for these correlations in a similar way as we did in the previous publications [21, 25] while we preserve also the mass dependencies of relevant parameters. The only difference is that the same exercise is repeated for each year, considering possible differences in reconstructions. There are eventually three free parameters in the signal model: one is a perturbation to the mass of the hypothetical signal, $\tilde{m}_{a_1} = m_{a_1} + \epsilon_{m_{a_1}}$, and the other two, a_1 and b_1 , are used to linearly model the resolutions, σ and σ_{cb} , with mass. As an example, the linear behavior of the resolution of the Voigt profile, σ , is shown in Fig. 27 for the three years.

Table 11 illustrates the parameters for each signal component and how they modeled with the mass. With this parametrization, it is possible to interpolate the signal shape into regions for which the simulated samples are not available. Details on this study can be found in Section 5 of Ref. [42]. Values for the fixed parameters are summarized in Table 12.

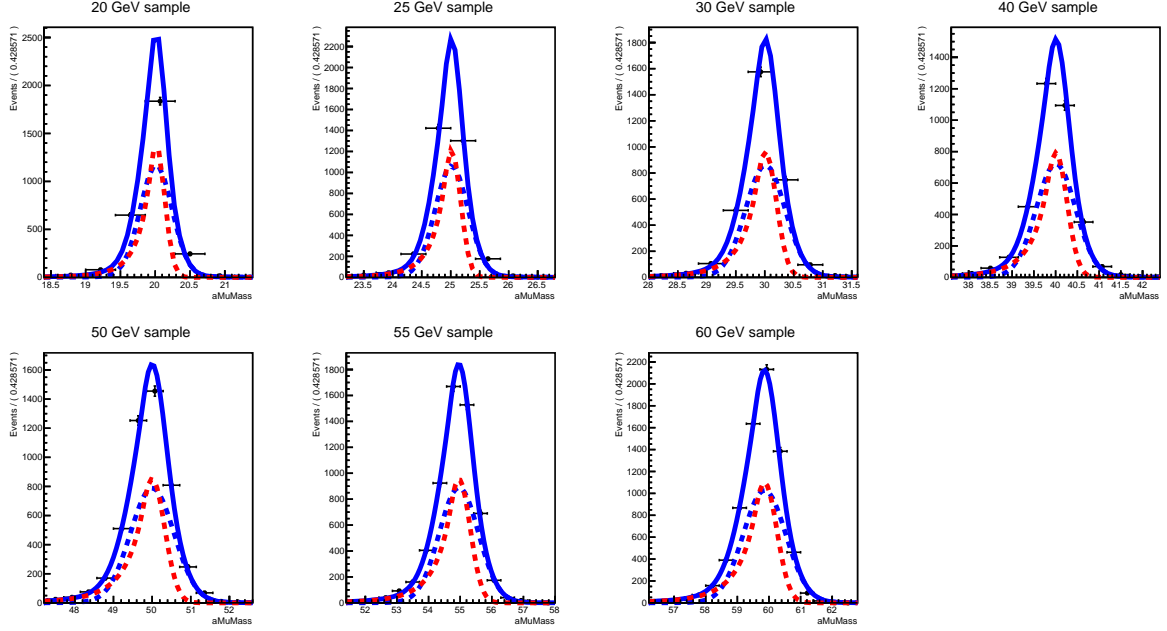


Figure 25: Results of the signal model $S(m_{\mu\mu}|f, p_V, p_{cb})$, fitted independently to $m_{\mu\mu}$ distributions at different mass points in 2017 simulated samples. Solid line is the constructed model. Dotted red line is the crystal-ball and dotted blue line is the Voigtion component.

Table 11: The parameters for the signal model and the way they are modeled with the mass: \hat{f} , $\hat{\gamma}$, $\hat{\alpha}$ and \hat{n} are fixed where σ and σ_{cb} are modeled linearly. Apart from the signal mass perturbative term, $\tilde{m}_{a_1} = m_{a_1} + \epsilon_{m_{a_1}}$, the model has two free parameters: a_1 and b_1 .

$S(m_{\mu\mu} f, p_V, p_{cb}) _{f=\hat{f}}$	$\hat{f} \cdot V(m_{\mu\mu} \tilde{m}_{a_1}, \sigma, \gamma)$	$\sigma = \hat{\sigma} + a_1 \cdot m_{\mu\mu}$	a_1 is free
		$\gamma = \hat{\gamma}$	fixed
	$(1 - \hat{f}) \cdot CB(m_{\mu\mu} \tilde{m}_{a_1}, n, \sigma_{cb}, \alpha)$	$\sigma_{cb} = \hat{\sigma}_{cb} + b_1 \cdot m_{\mu\mu}$	b_1 is free
		$\alpha = \hat{\alpha}$	fixed
		$n = \hat{n}$	fixed

6.1 Parameter correlations and signal validation studies

As already introduced, a mass-shifting model is utilized to describe the signal distributions. The signal model must be interpolated through the entire mass range. Obviously mean value changes linearly with m_{a_1} , i.e. $\tilde{m}_{a_1} = m_{a_1} + \epsilon_{m_{a_1}}$. The $\epsilon_{m_{a_1}}$ parameter is common and will be floating in the limit extraction, and is denoted by mS(standing for a mean shift) in the correlation plots discussed below.

Before taking $\epsilon_{m_{a_1}}$ as a common variable, there is no significant correlation between parameters as understood from the left plot in Fig. 28. The right plot of the same figure focuses on a typical

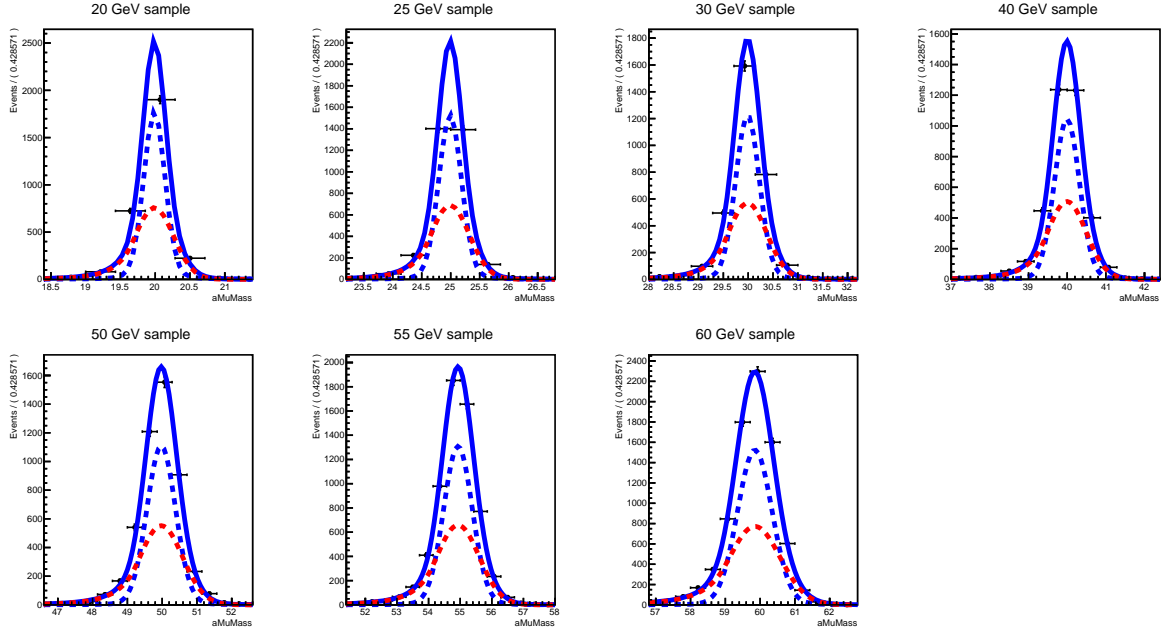


Figure 26: Results of the signal model $S(m_{\mu\mu}|f, p_V, p_{cb})$, fitted independently to $m_{\mu\mu}$ distributions at different mass points in 2018 simulated samples. Solid line is the constructed model. Dotted red line is the crystal-ball and dotted blue line is the Voigtion component.

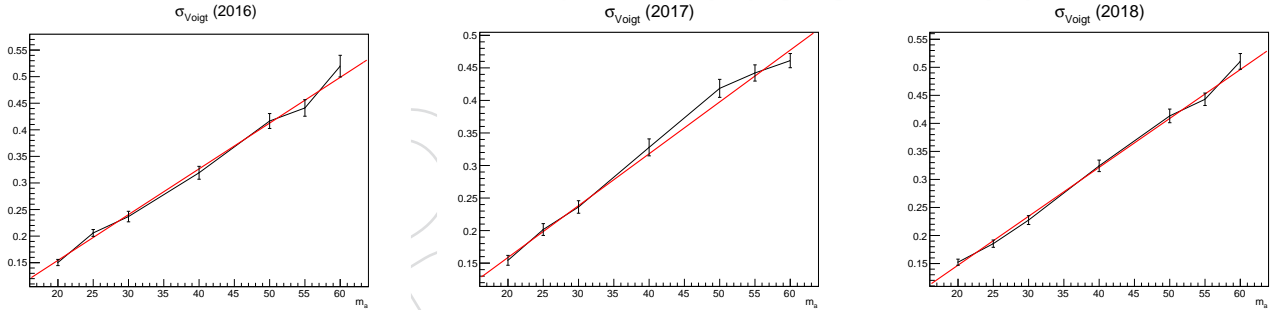


Figure 27: An example showing the linear behaviour of Voigt function resolution in years 2016 (left), 2017 (middle) and 2018 (right).

316 mass point.

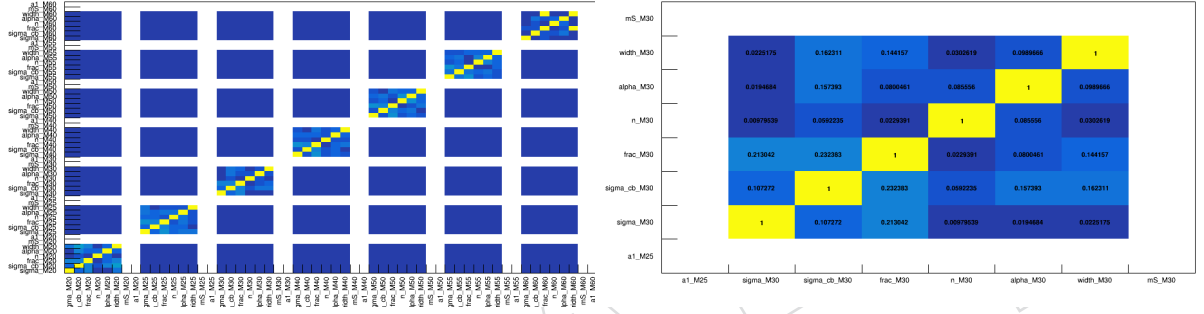
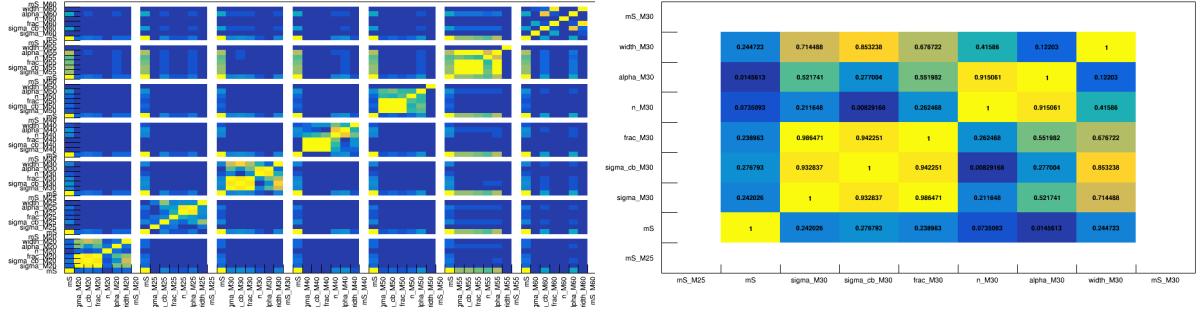
317 Once $\epsilon_{m_{a_1}}$ parameter is considered to be common, correlations introduced between different
 318 masses as observed from the left plot in Fig. 29. Also parameters of a single mass get correlated
 319 as shown in the right plot of the same figure.

320 An iterative procedure is followed, using simultaneous fit to all masses to extract signal model
 321 parameters. Having fixed some parameters of the model, as summarized in Table 11, the re-
 322 maining parameters are now showing minor correlations, as illustrated in Fig. 30.

323 Validation of the procedure to model the signal shape is needed to first ensure the parametriza-
 324 tions found by the fit works fine, and second to verify whether simultaneous fit can properly
 325 describe the intermediate masses. MC samples (e.g. 35 and 45 GeV) are excluded from simul-
 326 taneous fit and parameter extraction. We then apply the model on excluded signal samples to

Table 12: The fixed parameters of the model per year.

	\hat{f}	$\hat{\gamma}$	$\hat{\alpha}$	\hat{n}	$\hat{\sigma}$	$\hat{\sigma}_{cb}$
2016	0.56	0.041	1.24	3.66	-0.018	0.089
2017	0.56	0.028	0.71	6.3	-0.001	-0.0128
2018	0.56	0.042	1.21	6.3	-0.027	0.049

Figure 28: Correlation between parameters of the fit method before taking $\epsilon_{m_{a_1}}$ as a common variable, for different masses (left) and for a typical mass (right) using 2016 MC sample.Figure 29: Correlation between parameters of the fit method once $\epsilon_{m_{a_1}}$ is considered common, for different masses (left) and for a typical mass (right) using 2018 MC sample.

327 compare the results with individual fit.

328 Figures 31, 32 and 33 show the validation plots as well as pull distributions for the method
 329 applied on 2016, 2017 and 2018 signal MC samples respectively. In each figure, the model
 330 from simultaneous fit (orange distribution) is compared to the individual fit of the model (blue
 331 distribution) where all 7 parameters float. No major difference in the performance from pull
 332 distributions is observed.

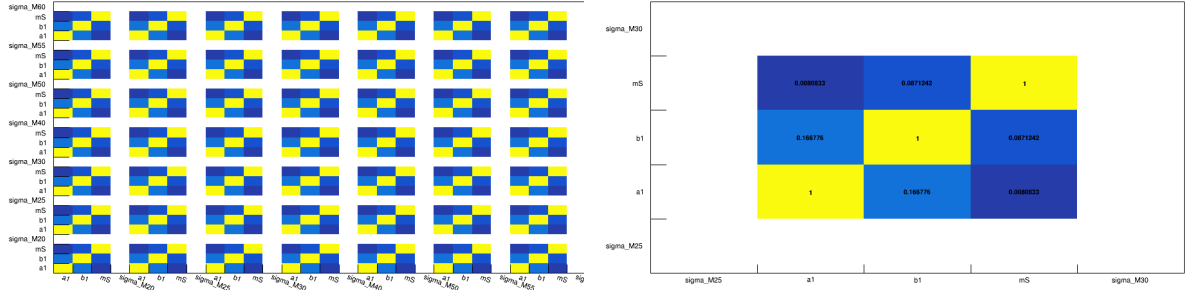


Figure 30: Correlation between parameters of the fit method after fixing floating variables, for different masses (left) and for a typical mass (right) using 2017 MC sample.

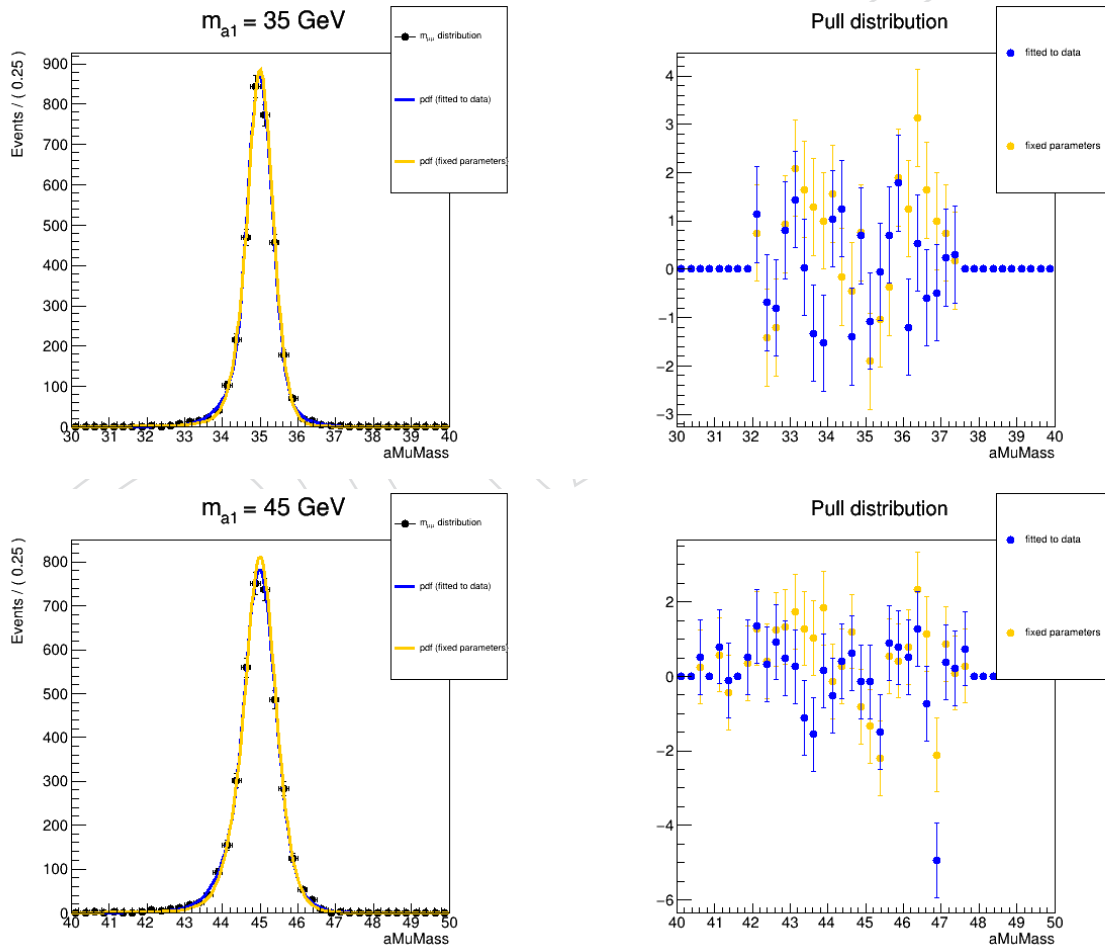


Figure 31: Validation plot (left) and pull distribution (right) for $m_{a_1} = 35 \text{ GeV}$ (top) and for $m_{a_1} = 45 \text{ GeV}$ (bottom) for 2016 sample.

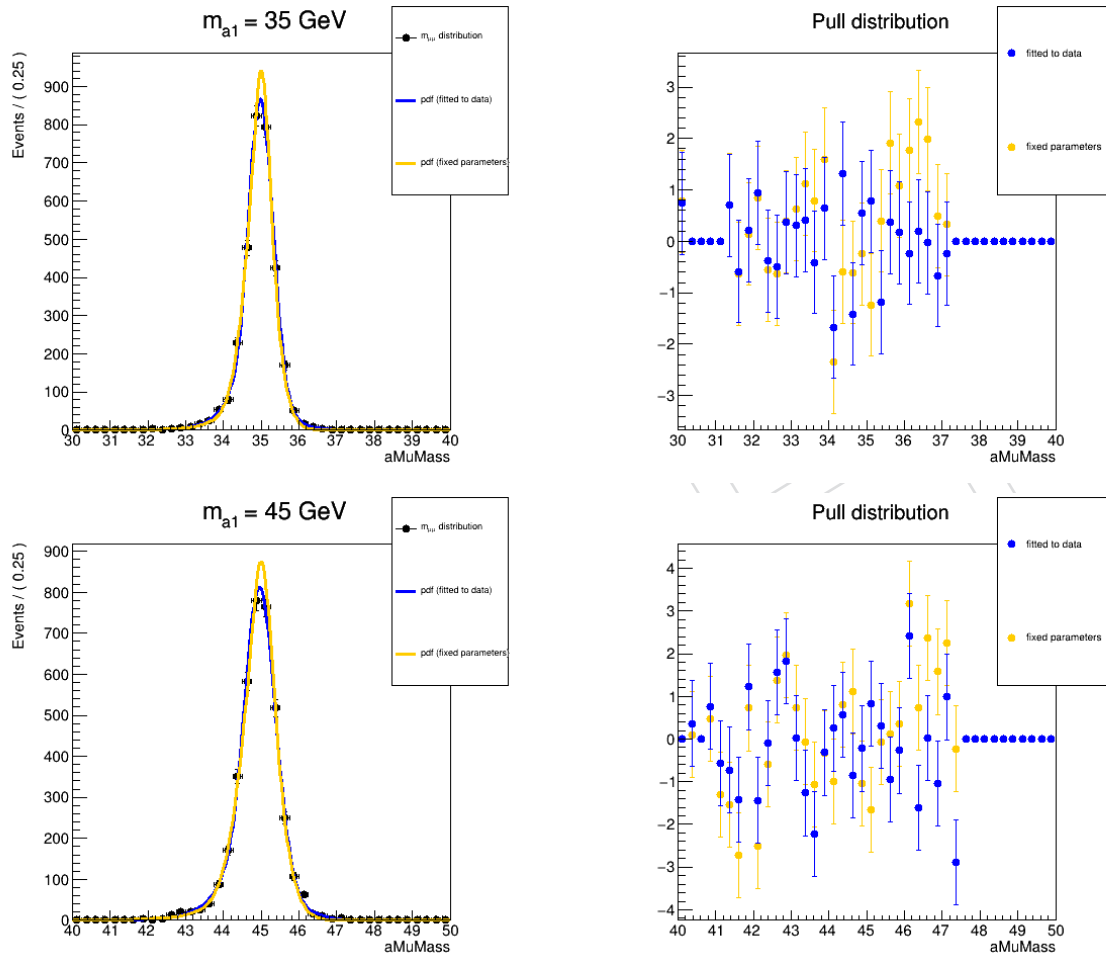


Figure 32: Validation plot (left) and pull distribution (right) for $m_{a1} = 35$ GeV (top) and for $m_{a1} = 45$ GeV (bottom) for 2017 sample.

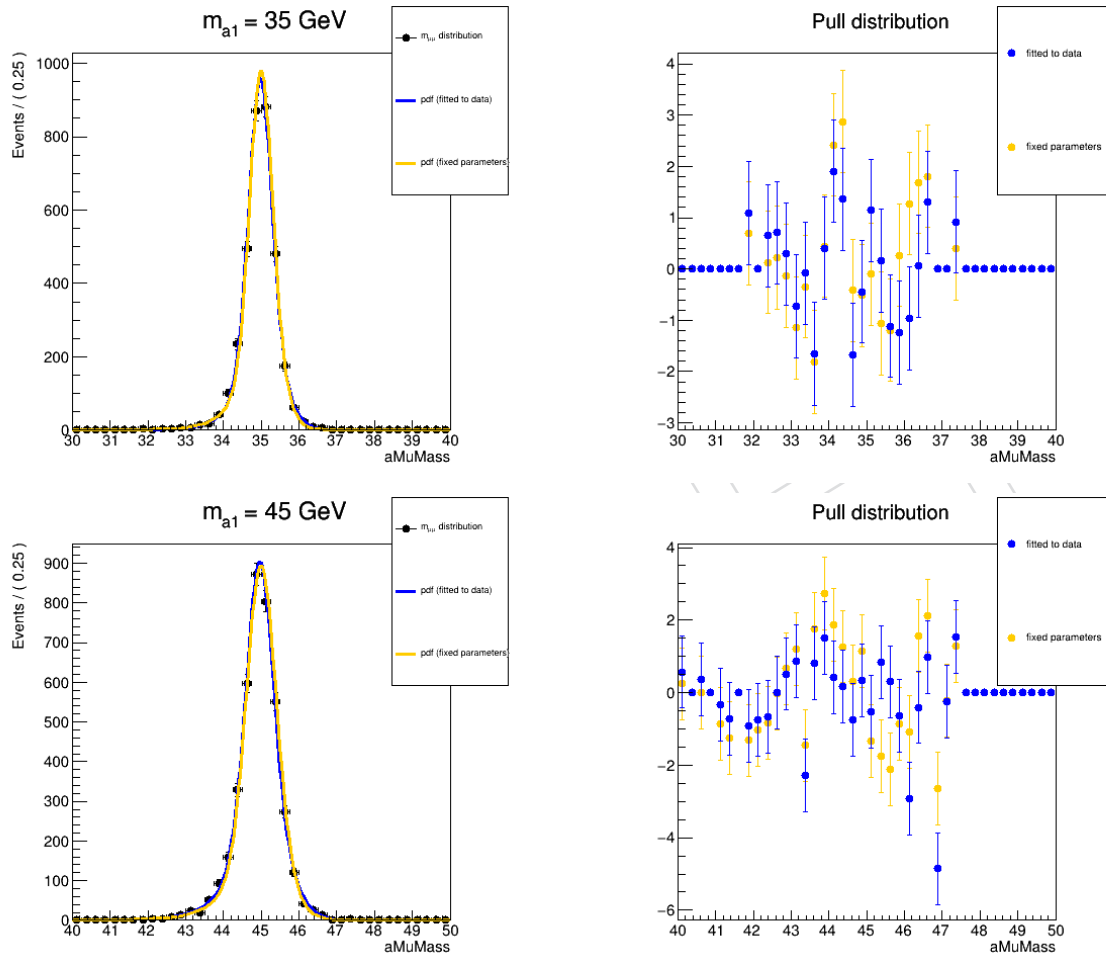


Figure 33: Validation plot (left) and pull distribution (right) for $m_{a_1} = 35$ GeV (top) and for $m_{a_1} = 45$ GeV (bottom) for 2018 sample.

7 Background determination in signal region

Similar to previous analyses [21, 25], the background is evaluated through a fit to the $m_{\mu\mu}$ distribution in data, without any reference to simulation. The uncertainty associated with the choice of the background model is treated in a similar way to other uncertainties for which there are nuisance parameters in the fit. The likelihood function for the signal-plus-background fit has the form of

$$\mathcal{L}(\text{data}|s(p, m_{\mu\mu}) + b(m_{\mu\mu})), \quad (11)$$

where $s(p, m_{\mu\mu})$ is the parametric signal shape with a set of parameters indicated by p , and $b(m_{\mu\mu})$ is the background model. The shape for the background is modeled with a set of analytical functions, using the discrete profiling method [43–45]. In this approach, the background shape is treated as a discrete nuisance parameter in the fit where its phase space contains multiple models, each including their own parameters. It is noteworthy to say that while the analysis was blind, a control region was defined and a similar method for background estimation was employed using the data in that side-band. This was mainly to estimate the sensitivity of the analysis prior to unblinding (see Appendix B).

7.1 Input background models

To provide the input background models to the discrete profiling method, the background data are modeled with different parametrization of polynomials (`RooPolynomial`, `RooBernstein`, `RooChebychev`) as well as $1/P_n(x)$ functions where $P_n(x) \equiv x + \sum_{i=2}^n \alpha_i x^i$. The degree of polynomials in each category are determined by F-test to ensure the sufficiency of number of parameters and the fact that the data are not over-fit. In each category of models, we start with the function containing the lowest number of parameters, and we consider it as justified based on the closure of the fit to the data, i.e., having a $\chi^2/\text{ndf} \sim 1$. We have also cross checked different categories versus each other to make sure that no pdf is repeated in the final pool of background models.

Tables 13, 14, 15, 16, and 17 summarize functions in different categories together with the statistical F-test results corresponding to each parameterization.

Figure 34 illustrates the summary of background models found suitable to describe the data.

The input background functions are tried in the minimization of the negative logarithm of the likelihood with a penalty term added to account for the number of free parameters in the background model. The likelihood ratio for the penalized likelihood function can be written as

$$-2 \ln \frac{\tilde{\mathcal{L}}(\text{data}|\mu, \hat{\theta}_\mu, \hat{b}_\mu)}{\tilde{\mathcal{L}}(\text{data}|\hat{\mu}, \hat{\theta}, \hat{b})}, \quad (12)$$

where μ is the measured quantity. The numerator is the maximum penalized likelihood for a given μ , at the best-fit values of nuisance parameters, $\hat{\theta}_\mu$ and of the background function, \hat{b}_μ . The denominator is the global maximum for $\tilde{\mathcal{L}}$, achieved at $\mu = \hat{\mu}$, $\theta = \hat{\theta}$ and $b = \hat{b}$. A confidence interval on μ is obtained with the background function maximizing $\tilde{\mathcal{L}}$ for any value of μ . This interval is always wider than those evaluated with the fixed functional form from the global best-fit, $b = \hat{b}$ [43].

Table 13: Trial functions to model background in the TL category. The following orders are not selected because of the failure of the F-test.

The TL category			
Model	χ^2/ndf	F-test probability (> 0.05)	Decision
Inv.Poly IV	1.07	–	✓
Inv.Poly V	0.98	0.008	×
Polynomial I	0.91	–	✓
Polynomial II	0.67	0.37	✓
Polynomial III	0.27	0.03	×

Table 14: Trial functions to model background in the TM category. The following orders are not selected because of the failure of the F-test.

The TM category			
Model	χ^2/ndf	F-test probability (> 0.05)	Decision
Polynomial 0	1.16	–	✓
Polynomial I	1.0	0.26	✓
Polynomial II	0.86	0.15	✓
Polynomial III	0.63	0.02	×
Inv.Poly I	1.05	–	✓
Inv.Poly II	0.64	0.004	×

Table 15: Trial functions to model background in the TT category.

†: The pdf is not selected as it was repeated in another category. The following orders are not selected either for the same reason or because of the failure of the F-test.

The TT category			
Model	χ^2/ndf	F-test probability (> 0.05)	Decision
Polynomial 0	0.98	–	✓
Polynomial I [†]	0.77	0.08	×
Polynomial II	0.71	0.5	✓
Polynomial III	1.19	0.03	×
Inv. Poly I	0.84	–	✓
Inv. Poly II	0.6	0.05	✓
Inv. Poly III	0.58	0.0	×
Bernstein I [†]	0.93	–	×
Bernstein II	0.73	0.07	✓
Bernstein III	0.89	0.02	×

Table 16: Trial functions to model background in the VBF category.

The VBF category			
Model	χ^2/ndf	F-test probability (> 0.05)	Decision
Polynomial I	0.65	–	✓
Polynomial II	0.42	0.04	×

Table 17: Trial functions to model background in the Low p_T category.

†: The pdf is not selected because the value of pdf becomes negative at some mass intervals.

The Low p_T category			
Model	χ^2/ndf	F-test probability (> 0.05)	Decision
Inv.Poly III	1.11	–	✓
Inv.Poly IV	1.05	0.23	✓
Inv.Poly V [†]	1	0.09	×
Inv.Poly VI	0.98	0.0	×

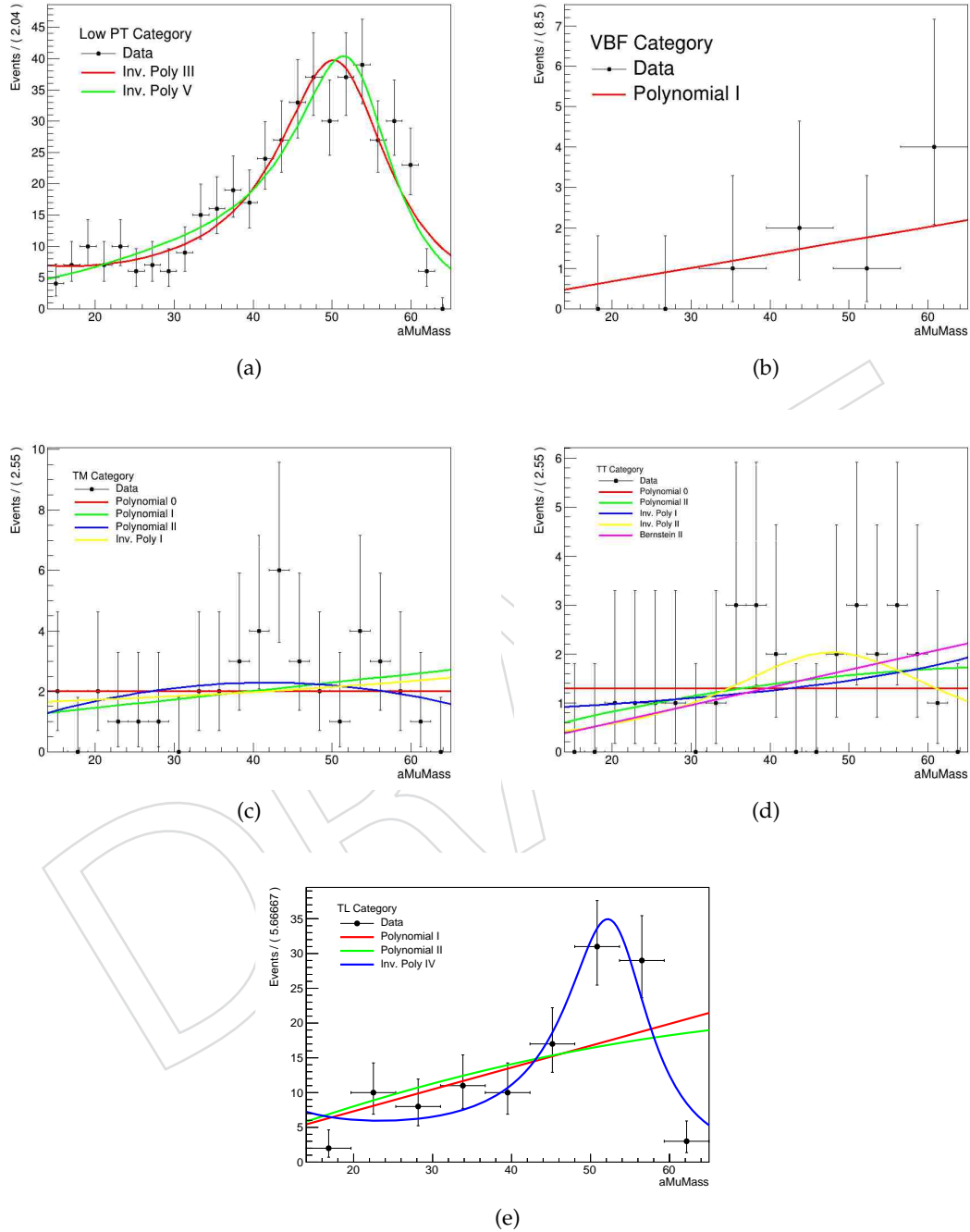


Figure 34: Suitable pdf's describing the background from signal region in data in the Low p_T (a), VBF (b), TM (c), TT (d), and TL (e) categories.

8 Systematic uncertainties

The statistical interpretation of the analysis takes into account several sources of systematic uncertainties related to the accuracy in the signal modelling and uncertainties in the signal acceptance. Separate simulations for signal are employed for the data taken in the three years in order to match the varying detector performance and data taking conditions. Therefore the level of the systematic correlations between variations from the three periods is also specified. The imprecise knowledge of the background contributions is taken into account by the discrete profiling method described in Section 7.

Uncertainties on Higgs production cross section: to evaluate the upper limit on $\text{Br}(\text{H} \rightarrow a_1 a_1 \rightarrow \mu^+ \mu^- b \bar{b})$, the Higgs boson production cross section is set to the SM prediction where an uncertainty of 3.6% is considered for the sum of the ggF and VBF production cross sections, accounting for QCD scale, PDF, and α_s uncertainties [46].

Experimental uncertainties: the uncertainty on the integrated luminosity [47–49] has a component that is uncorrelated across the years. It amounts to 1.0%, 2.0% and 1.5% respectively for 2016, 2017 and 2018 periods. Another component is correlated across all three years: 0.6%, 0.9%, and 2.0% respectively for years 2016, 2017 and 2018. Additionally, the luminosity measurements in 2017 and 2018 are accompanied by an additional uncertainty, 0.6% and 0.2% respectively, that is correlated between the two years. The uncertainty in the number of pileup interactions per event is estimated by varying the total inelastic pp cross section by $\pm 4.6\%$ [50], fully correlated across the years. An uncertainty of 2% is assigned to the trigger efficiency. The data-to-simulation correction factors for the muon reconstruction and selection efficiencies are estimated using a “tag-and-probe” method [51] in Drell–Yan data and simulated samples. These uncertainties include the pileup dependence of the correction factors and are considered correlated across the years since common procedural uncertainties are the dominant source. For the jet energy scale (JES), the variations are made according to the η - and p_T -dependent uncertainties and propagated to the p_T^{miss} of the event. These uncertainties include several sources and the correlation among the years is evaluated separately for each source. An additional uncertainty, arising from unclustered energies in the event, is assessed for p_T^{miss} , considered uncorrelated across the years. For the jet energy resolution, the smearing corrections are varied within their uncertainties [40], uncorrelated across the years. Systematic uncertainty sources that affect the data-to-simulation corrections of the b tagging discriminator distribution are JES, the contaminations from light flavor (LF) jets in the b-jet sample, the contaminations from heavy flavor (HF) jets in the light-flavor jet sample, and the statistical fluctuations in data and MC [52]. The JES variations b-tagging are performed consistently with uncertainties on jets and follow the same correlation pattern across the years. The statistical components of the b-tagging Uncertainties are considered uncorrelated while the rest are assumed to be correlated between different periods. During the 2016 and 2017 data taking periods, a gradual shift in the timing of the inputs of the ECAL Level-1 trigger in the forward endcap region ($|\eta| > 2.4$) led to a specific inefficiency. A correction for this effect was determined using an unbiased data sample and is found to be relevant in events with high- p_T jets with $2.4 < |\eta| < 3.0$. This predominantly affect the VBF category of this analysis which has very low statistics. Although the effect on the final result is not expected to be large, a systematic variation of 20% is considered for this correction.

Modeling uncertainties: uncertainties in this category are considered fully correlated across the years. This is because the same model assumptions and procedure are employed in deriving variations. Uncertainties in the choice of the normalization, μ_R , and factorization, μ_F , scales are estimated by doubling and halving μ_R and μ_F simultaneously in the signal sample. For the

parton-shower simulation, uncertainties are separately assessed for initial- and final-state radiations, by varying the respective scales up and down by factors of two. Finally, uncertainties arising from the limited understanding of the PDFs [53] are taken into account.

Based on the studies on 2016 data [25] uncertainties have a negligible effect on the shape of the signal. Their effects on the yield are taken into account by introducing nuisance parameters with log-normal distributions into the fit.

Figure 35 shows the background only and signal-plus-background impact of different sources of uncertainties using Asimove data for signal hypotheses with $m_{a_1} = 35$ GeV.

In Fig. 36, the observed impact is presented for two signal hypotheses with $m_{a_1} = 20$ and 35 GeV. For the two masses, it can be seen that the effects of different sources of uncertainty are different. While blind to the data in the signal region we calculated the impact of different sources of uncertainties while blind to the data in the signal region (Appendix C). Overall, systematic uncertainties do not play a significant role.

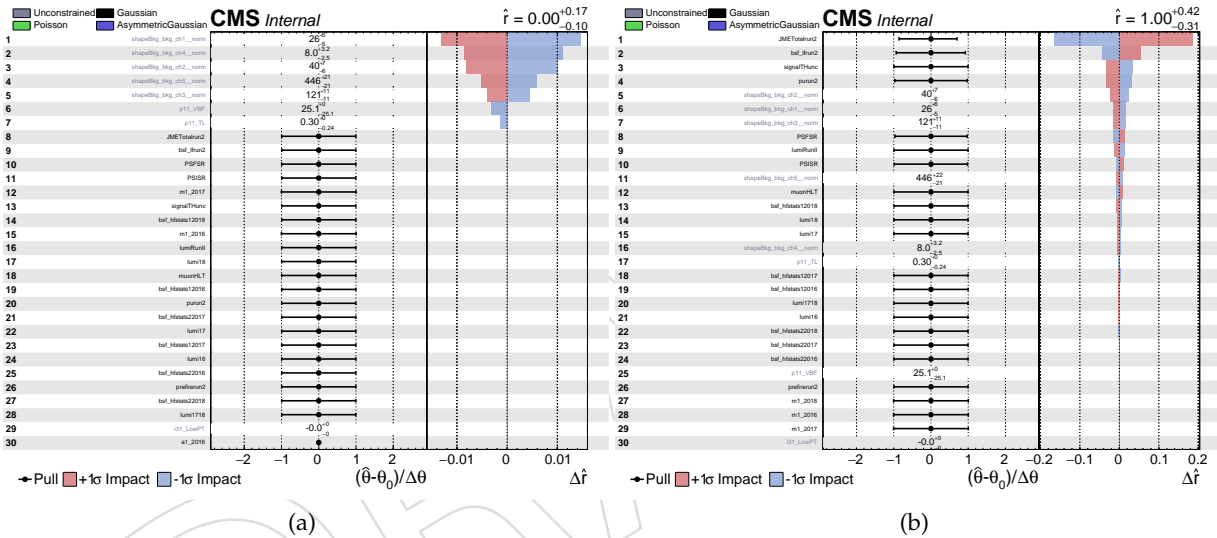


Figure 35: The impact of systematic uncertainties for a signal with $m_{a_1} = 35$ GeV using Asimov data for background only (left) and signal-plus-background (right) hypotheses.

9 Results

Using the asymptotic CL_s method, the observed upper limits are set, at 95% confidence level (CL), on the signal strength where signal strength, r , is defined as the measured cross section in the $\mu\mu b\bar{b}$ final state with respect to the theory expectation from type-III 2HDM+S at $\tan\beta = 2$. Figure 37 shows the observed limits, at 95% CL, on $\text{Br}(H \rightarrow a_1 a_1 \rightarrow \mu^+ \mu^- b \bar{b})$.

Current limits are more stringent than those of the CMS analysis using 2016 data [25]. The main improvement comes from the modification of the χ^2_{tot} definition, and the new categorization improved the results for low m_{a_1} values as well. In comparison with ATLAS [23], expected results are compatible within one standard deviation almost over the entire mass range. Fig. 38 shows a slight improvement is however observed from the current analysis. At 95% CL, the observed upper limits on $\text{Br}(H \rightarrow a_1 a_1 \rightarrow \mu^+ \mu^- b \bar{b})$ are $(0.17 - 3.3) \times 10^{-4}$ for the mass range 15 to 62.5 GeV, whilst the expected limits are $(0.35 - 2.6) \times 10^{-4}$. The improvement over the

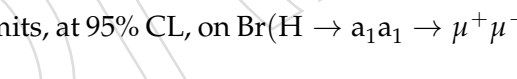


Figure 39 shows the best-fit background pdf as extracted from the background only fit to data and Fig. 40 shows the post-fit background pdf including background model uncertainties as extracted from the background only fit to data in each category.

Figure 41 shows the post-fit signal-plus-background pdf and Fig. 42 shows the post-fit signal-plus-background pdf including background model uncertainties as extracted from the fit to data for $m_{a_1} = 35$ GeV in each category.

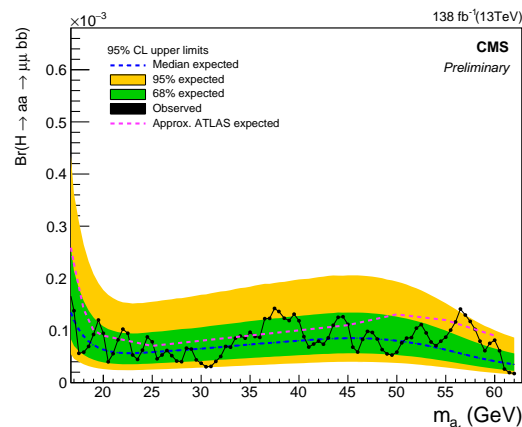


Figure 38: Observed limits, at 95% CL, on $\text{Br}(H \rightarrow a_1 a_1 \rightarrow \mu^+ \mu^- b \bar{b})$ at 13 TeV using full Run-II data. ATLAS results are shown for comparison.

doi:<http://dx.doi.org/10.1016/j.physletb.2012.08.020>.

- [2] CMS Collaboration, “Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc”, *Physics Letters B* **716** (2012)

doi:<http://dx.doi.org/10.1016/j.physletb.2012.08.021>.

- [3] L. Evans and P. Bryant, “Lhc machine”, *JINST* **3 S08001** (2008).

- [4] ATLAS Collaboration, “Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector at the LHC”,.

- [5] CMS Collaboration, “A measurement of the Higgs boson mass in the diphoton decay channel”, *Phys. Lett. B* **805** (2020) doi:[10.1016/j.physletb.2020.135425](https://doi.org/10.1016/j.physletb.2020.135425), arXiv:2002.06398.

- [6] CMS Collaboration Collaboration, “Observation of $t\bar{t}H$ production”, *Phys. Rev. Lett.* **120** (Jun, 2018) doi:[10.1103/PhysRevLett.120.231801](https://doi.org/10.1103/PhysRevLett.120.231801).

- [7] ATLAS Collaboration, “Observation of higgs boson production in association with a top quark pair at the lhc with the atlas detector”, *Phys. Lett. B* **784** (2018)

doi:<https://doi.org/10.1016/j.physletb.2018.07.035>.

- [8] ATLAS Collaboration, “Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector”, *Phys. Lett. B* **786** (2018) doi:[10.1016/j.physletb.2018.09.013](https://doi.org/10.1016/j.physletb.2018.09.013), arXiv:1808.08238.

- [9] CMS Collaboration, “Observation of Higgs boson decay to bottom quarks”, *Phys. Rev. Lett.* **121** (2018) doi:[10.1103/PhysRevLett.121.121801](https://doi.org/10.1103/PhysRevLett.121.121801), arXiv:1808.08242.

- [10] ATLAS Collaboration, “A search for the dimuon decay of the standard model higgs boson with the atlas detector”, *Physics Letters B* **812** (2021)

doi:<https://doi.org/10.1016/j.physletb.2020.135980>.

- [11] CMS Collaboration, “Evidence for Higgs boson decay to a pair of muons”, *JHEP* **01** (2021) doi:[10.1007/JHEP01\(2021\)148](https://doi.org/10.1007/JHEP01(2021)148), arXiv:2009.04363.

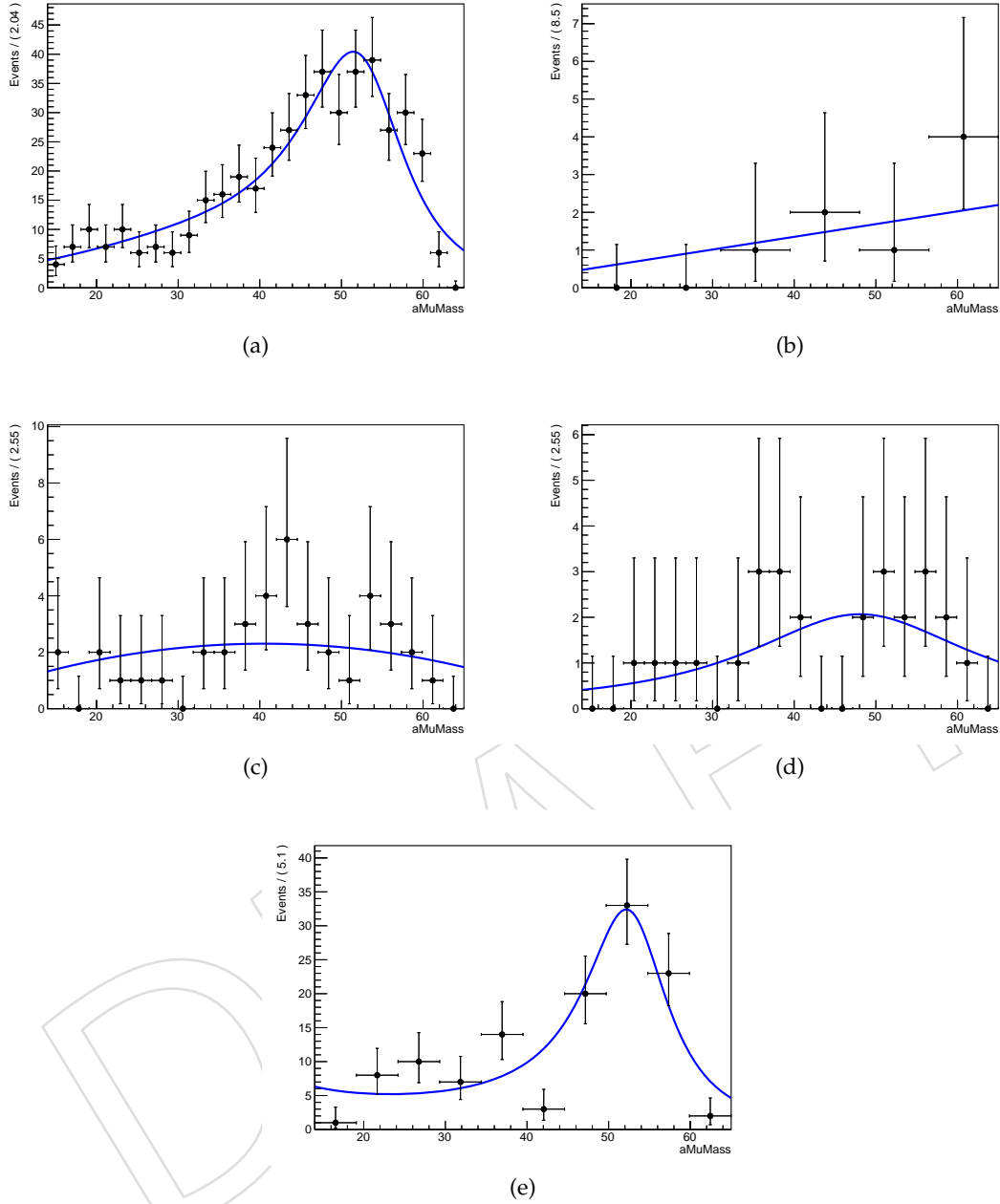


Figure 39: The post-fit signal-plus-background pdf for $m_{a_1} = 35$ GeV in the Low p_T (a), VBF (b), TM (c), TT (d), and TL (e) categories.

[12] ATLAS Collaboration, “Combined measurements of higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the atlas experiment”, *Phys. Rev. D* **101** (Jan, 2020)
doi:10.1103/PhysRevD.101.012002.

[13] CMS Collaboration, “Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **79** (2019), no. 5,
doi:10.1140/epjc/s10052-019-6909-y, arXiv:1809.10733.

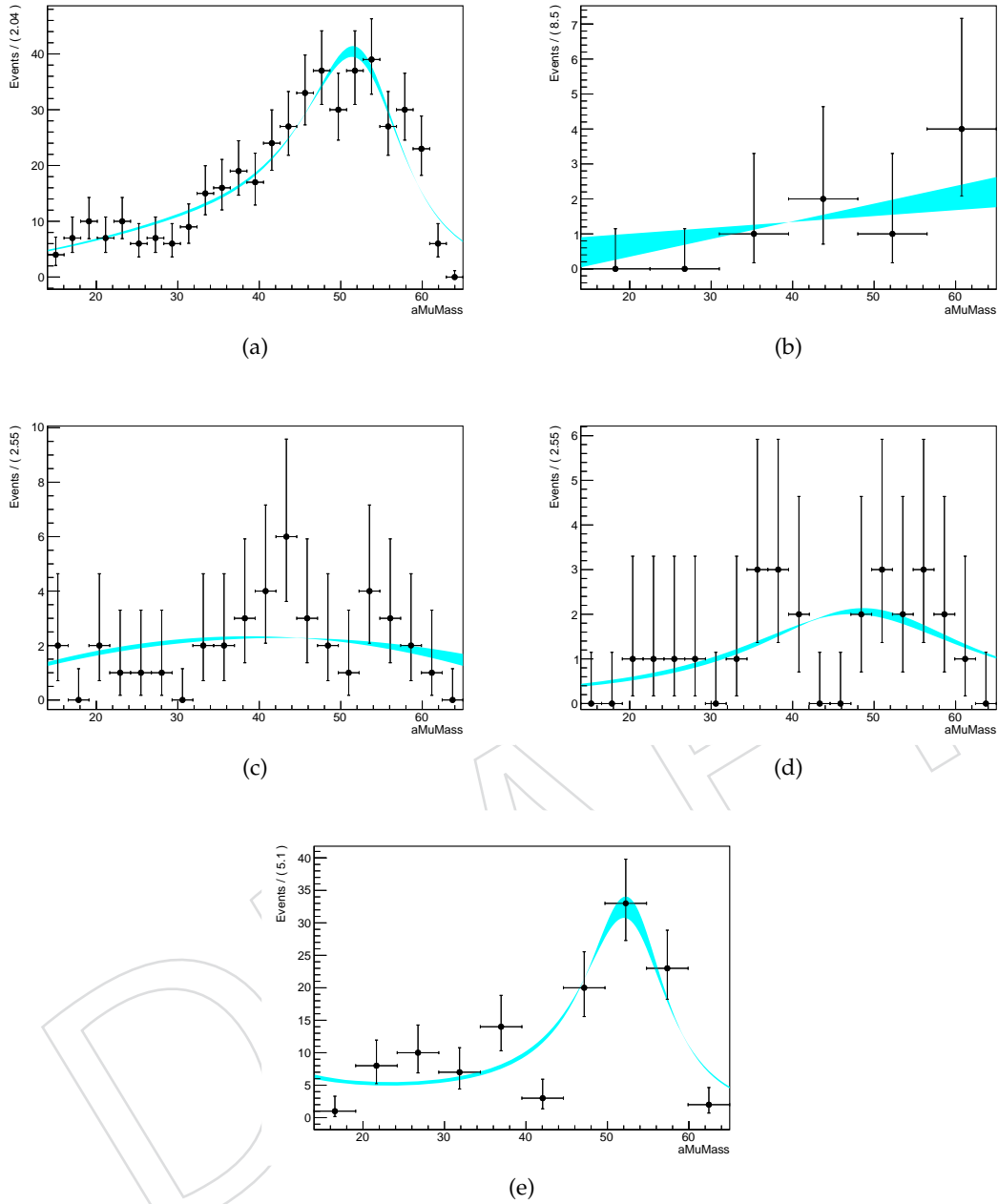


Figure 40: The post-fit signal-plus-background pdf including background model uncertainties as extracted from the fit to data for $m_{a_1} = 35$ GeV in the Low p_T (a), VBF (b), TM (c), TT (d), and TL (e) categories.

[14] G. Belanger et al., “Status of invisible Higgs decays”, *Phys.Lett.* **B723** (2013)

doi:10.1016/j.physletb.2013.05.024, arXiv:1302.5694.

[15] P. P. Giardino et al., “The universal Higgs fit”, *JHEP* **1405** (2014)

doi:10.1007/JHEP05(2014)046, arXiv:1303.3570.

[16] J. Ellis and T. You, “Updated Global Analysis of Higgs Couplings”, *JHEP* **1306** (2013)

doi:10.1007/JHEP06(2013)103, arXiv:1303.3879.

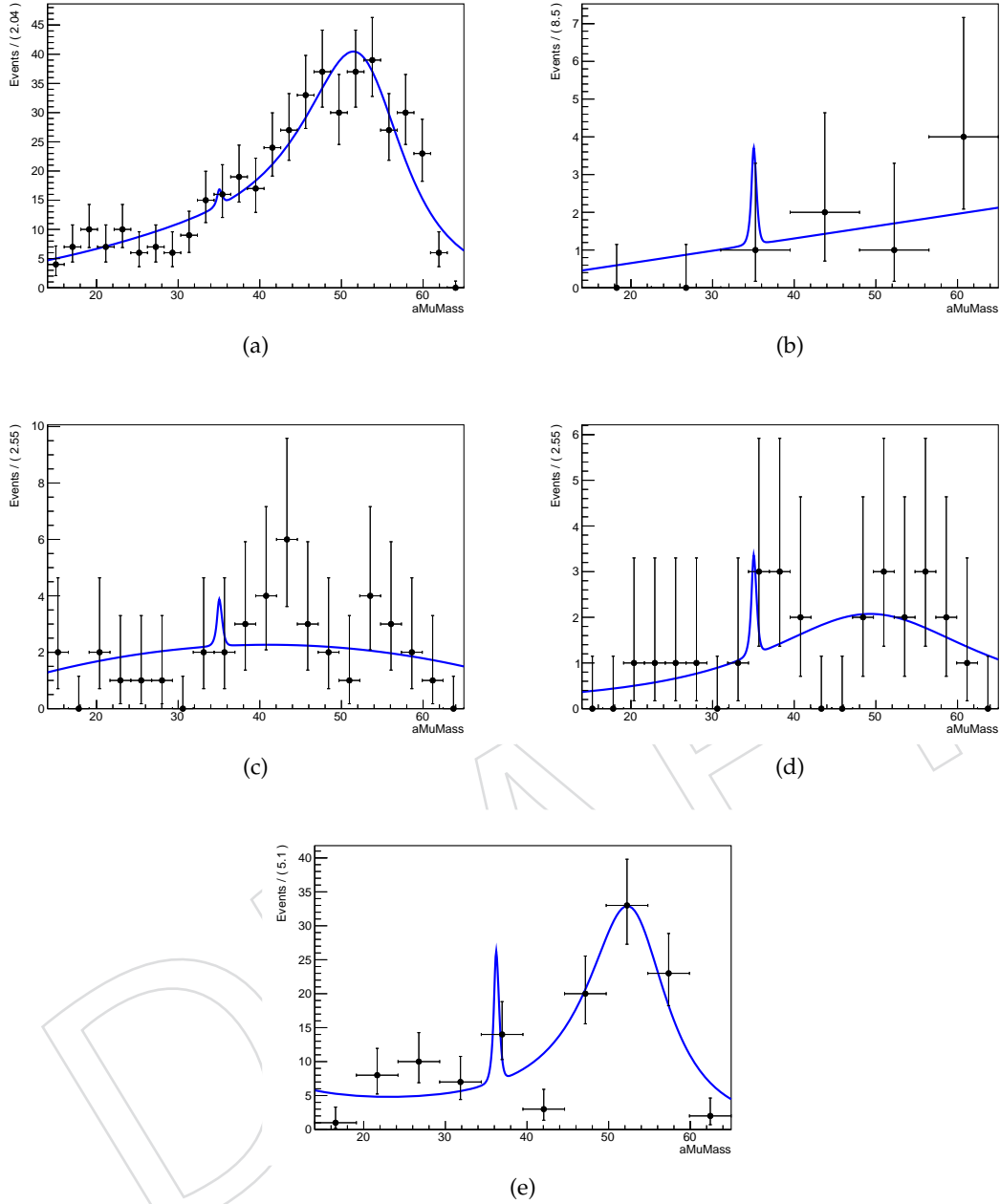


Figure 41: The post-fit signal-plus-background pdf for $m_{a_1} = 35$ GeV in the Low p_T (a), VBF (b), TM (c), TT (d), and TL (e) categories.

[17] K. Cheung, J. S. Lee, and P.-Y. Tseng, “Higgs Precision (Higgcision) Era begins”, *JHEP* **1305** (2013) doi:10.1007/JHEP05(2013)134, arXiv:1302.3794.

[18] A. Djouadi and G. Moreau, “The couplings of the Higgs boson and its CP properties from fits of the signal strengths and their ratios at the 7+8 TeV LHC”, *Eur.Phys.J.* **C73** (2013) doi:10.1140/epjc/s10052-013-2512-9, arXiv:1303.6591.

[19] B. A. Dobrescu and J. D. Lykken, “Coupling spans of the Higgs-like boson”, *JHEP* **1302** (2013) doi:10.1007/JHEP02(2013)073, arXiv:1210.3342.

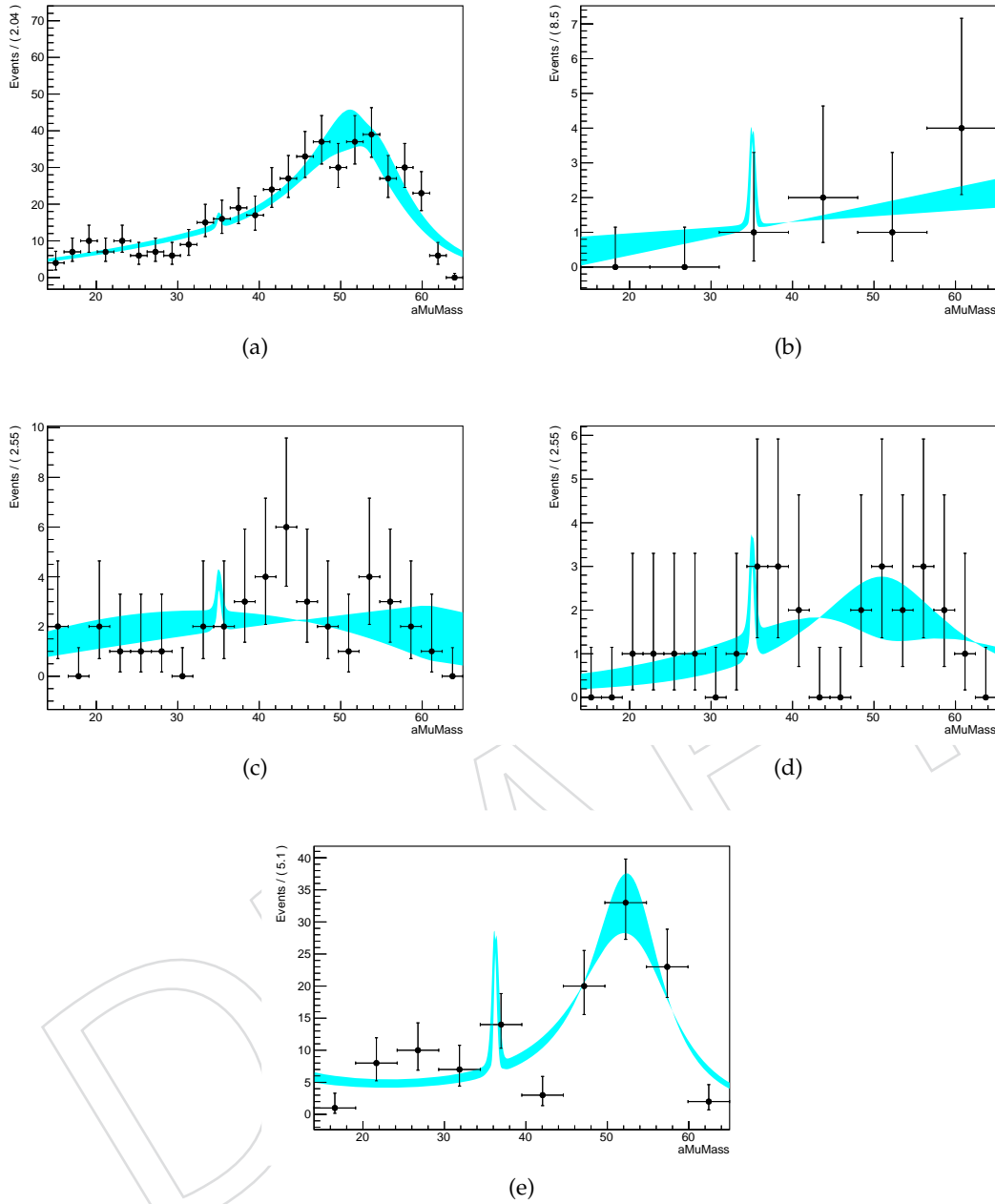


Figure 42: The post-fit signal-plus-background pdf including background model uncertainties as extracted from the fit to data for $m_{a_1} = 35$ GeV in the Low p_T (a), VBF (b), TM (c), TT (d), and TL (e) categories.

[20] D. Curtin et al., “Exotic decays of the 125 gev higgs boson”, *Phys. Rev. D* **90** (Oct, 2014) doi:10.1103/PhysRevD.90.075004.

[21] CMS Collaboration, “Search for light bosons in decays of the 125 GeV Higgs boson in proton-proton collisions at $\sqrt{s} = 8$ TeV”, arXiv:1701.02032.

[22] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, “ SM Higgs production cross sections at $\sqrt{s} = 13$ TeV (update in CERN Report4 2016) ”, technical report, CERN.

- [23] ATLAS Collaboration, “Search for Higgs boson decays into two spin-0 particles in the $b\bar{b}\mu\mu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV”,
- [24] ATLAS Collaboration, “Search for Higgs boson decays into a pair of light bosons in the $b\bar{b}\mu\mu$ final state in pp collision at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Lett. B* **790** (2019) 1, doi:10.1016/j.physletb.2018.10.073, arXiv:1807.00539.
- [25] CMS Collaboration, “Search for an exotic decay of the higgs boson to a pair of light pseudoscalars in the final state with two muons and two b quarks in pp collisions at 13 tev”, *Phys. Lett. B* **795** (2019) doi:https://doi.org/10.1016/j.physletb.2019.06.021.
- [26] 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* **1407** (2014) doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [27] M. L. Mangano et al., “Matching matrix elements and shower evolution for top-pair production in hadronic collisions”, *JHEP* **2007** (2007) doi:10.1088/1126-6708/2007/01/013.
- [28] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [29] E. Re, “Single-top Wt -channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J.* **C71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [30] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [31] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG: s- and t-channel contributions”, *JHEP* **09** (2009) 111, doi:10.1007/JHEP02(2010)011, 10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP02,011(2010)].
- [32] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [33] T. Sjöstrand, S. Mrenna, and P. Skands, “A brief introduction to PYTHIA 8.1”, *Comput.Phys.Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [34] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune”, *Eur. Phys. J.* **C74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.
- [35] GEANT4 Collaboration, “GEANT4: A simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [36] S. Chenarani, R. Goldouzian, G. Negro, and J. R. Thieman, “Measurement of the trigger efficiencies for a dilepton selection for a $t\bar{t}$ analysis with the full run 2 dataset”, CMS AN-2019/140, (2019).

- [37] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *arXiv:1706.04965*.
- [38] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_T jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, *arXiv:0802.1189*.
- [39] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, *arXiv:1111.6097*.
- [40] CMS Collaboration, “Jet algorithms performance in 13 TeV data”, Technical Report CMS-PAS-JME-16-003, CERN, Geneva, 2017.
- [41] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS”, *JINST* **6** (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, *arXiv:1107.4277*.
- [42] A. Jafari and A. Giammanco, “Search for exotic decays of the discovered higgs boson to a pair of new light bosons with two muon and two b jets in final states”, CMS AN-2014/266, (2014).
- [43] P. Dauncey, M. Kenzie, N. Wardle, and G. Davies, “Handling uncertainties in background shapes: the discrete profiling method”, *arXiv:1408.6865*.
- [44] CMS Collaboration, “Observation of the diphoton decay of the Higgs boson and measurement of its properties”, *Eur. Phys. J. C* **74** (2014), no. 10, 3076, doi:10.1140/epjc/s10052-014-3076-z, *arXiv:1407.0558*.
- [45] ATLAS, CMS Collaboration, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”, *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, *arXiv:1503.07589*.
- [46] D. de Florian et al., “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector”, CERN Report CERN-2017-002-M, 2016. doi:10.23731/CYRM-2017-002, *arXiv:1610.07922*.
- [47] CMS Collaboration, “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS”, (2021). *arXiv:2104.01927*. Submitted to *Eur. Phys. J. C*.
- [48] CMS Collaboration, “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.
- [49] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.
- [50] CMS Collaboration, “Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV”, *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, *arXiv:1802.02613*.
- [51] CMS Collaboration, “Measurements of inclusive w and z cross sections in pp collisions at $\sqrt{s} = 7$ tev”, *JHEP* **01** (2011) doi:10.1007/JHEP01(2011)080, *arXiv:1012.2466*.
- [52] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, *arXiv:1712.07158*.
- [53] J. Butterworth et al., “PDF4LHC recommendations for LHC Run II”, *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, *arXiv:1510.03865*.

A Study of b-tagging below 20 GeV

We are using b-tagging in the p_T range between 5 and 20 GeV with scale factors at 20 GeV and doubled uncertainties. In our analysis, no major discrepancy observed below 20 GeV. Fig 43 shows the sub-leading jet p_T distribution and it can be seen that signal for all masses has a peak below 20 GeV, Background shows a similar trend. It should be emphasized that the uncertainty bands are only MC statistics.

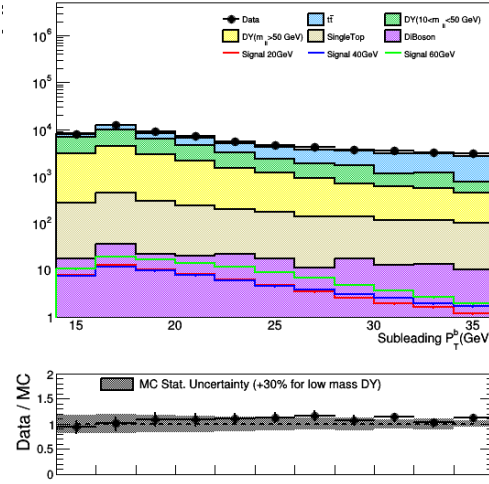


Figure 43: The sub-leading jet p_T distribution. There is no major discrepancy below 20 GeV. Simulated samples are normalized to 138 fb^{-1} . The band in the lower panel of each plot represents the MC statistical uncertainty.

The main concern is, deep jet b-tagging SF's are not available below 20 GeV. Doubling systematic uncertainties may not cover the possible issues. Therefore it motivates the study of events where leading b-jet has p_T below 20 GeV. For this study we use loose-loose b-working point. The questions are:
Is there any major mis-modeling of b discriminator in our phase space?
Are current b-tag systematics sufficient?

Events with leading b-jet $p_T < 20 \text{ GeV}$ are selected and then samples divided into prompt and non-prompt b-jet (light/gluon jets). From Fig 44, it's clear that the statistics from leading b-jet is less than sub-leading b-jet, so leading b-jets are mostly above 20 GeV and not many leading jets are with p_T between 5 and 20 GeV. It can be seen that there is no major mis-modeling within the uncertainties of data and MC statistics.

In Fig 45 all different source of systematic uncertainties of bSF's are shown where $p_T < 20 \text{ GeV}$ and uncertainties are doubled. The total of them is shown with the green line. Obviously, these systematic uncertainties cover the residuals. It looks safe for $p_T < 20 \text{ GeV}$ to apply bSF's at 20 GeV and doubled the uncertainties. It should be mentioned that we should use the shape scale factors since events are later categorized based on the b-tag score. Therefore we would need to make sure that shape-based uncertainties cover any possible mis-modeling. BTV is satisfied with this process.

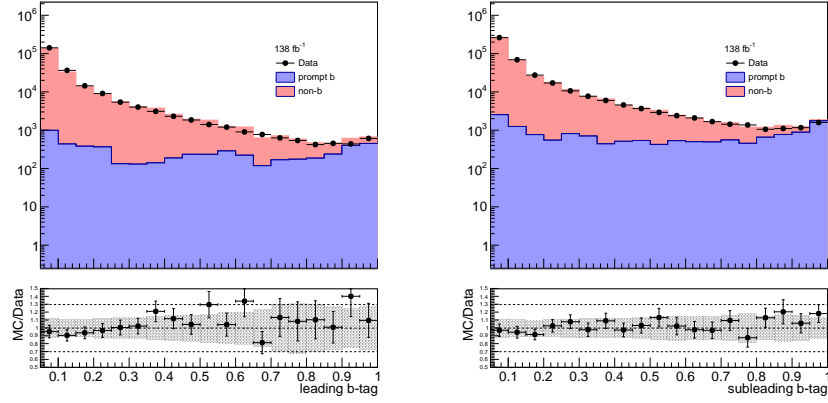


Figure 44: The distribution of b-tag discriminator for leading b-jet (left) and sub-leading b-jet(right).

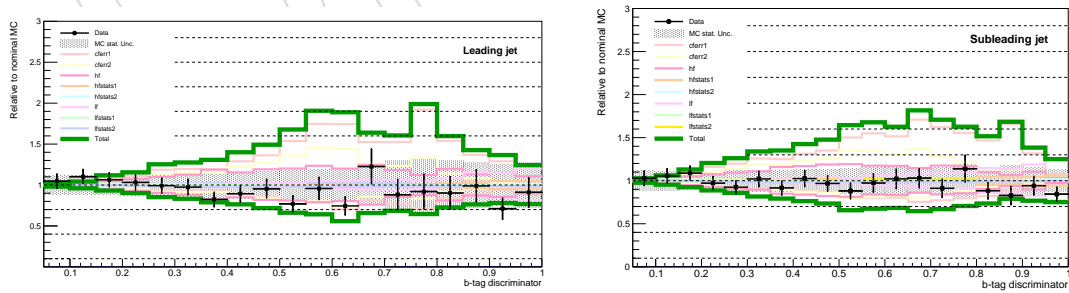


Figure 45: Relative to nominal MC versus b-tag discriminator for leading b-jet (left) and sub-leading b-jet (right).

B Background model before unblinding

B.1 Control data region for the blind analysis

We identify a control data region that is similar to the expected background in signal region (SR) in order to examine the background determination method and estimate the sensitivity of the analysis while blind to the data in signal region (SR). By *similar* we mean both the yield and the $m_{\mu\mu}$ distribution should be close enough to the background in SR. The χ_D variable is found to have small correlation with $m_{\mu\mu}$, meaning that for backgrounds the shape of the dimuon mass remains stable in different intervals of χ_D . Figure 46 shows the scattered plot of χ_D versus $m_{\mu\mu}$, including the correlations.

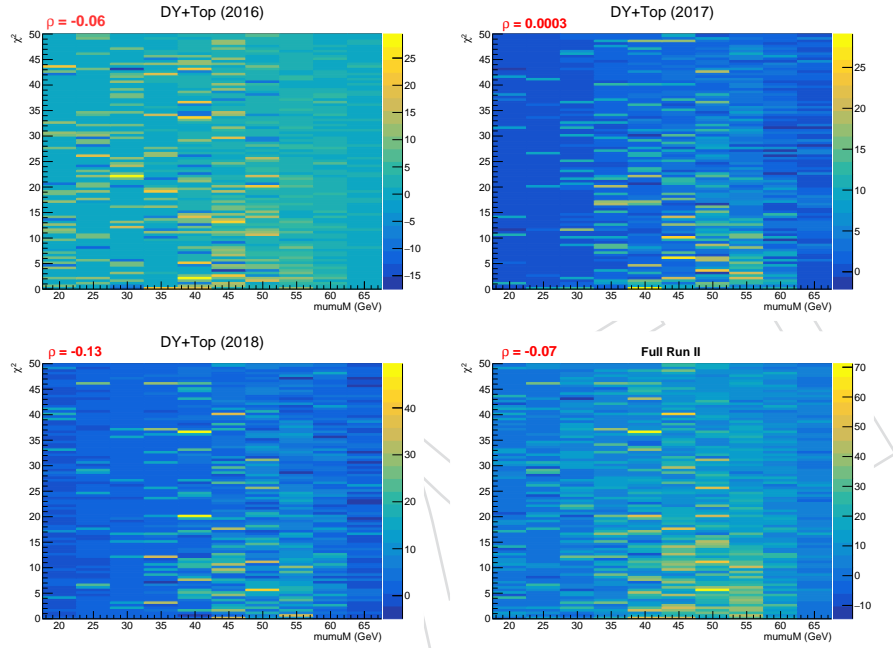


Figure 46: distribution of χ_D vs $m_{\mu\mu}$ for simulated backgrounds in 2016 (top left) 2017 (top right) 2018 (bottom left) and full Run II (bottom right).

It was shown earlier that the signal contamination is negligible for $\chi_D > 10$. A detailed study of simulation suggests a sideband of $12.5 < \chi_D < 14.5$ for 2017 together with $13.5 < \chi_D < 15$ in 2016 and 2018. These data regions are used to determine input models for backgrounds and the signal extraction fit while being blind to data in SR. The information from sideband is not used in the analysis after unblinding. Figure 47 shows the ratio of simulation in SR over proposed control data region.

In Figure 48 we compared data in control region with MC in SR for each category, it can be seen that CR data have similar behavior even in categorization within statistics.

B.2 Input background models

As long as the analysis is blind, the data in the control region are used for the data-driven background modeling. Using the procedure described in Section 7, different types of functional forms are used on data in CR. The F-test is used to determine the collection of pdfs for each family.

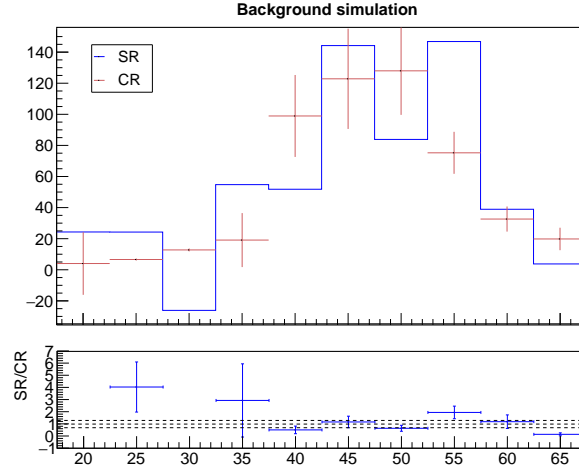


Figure 47: Comparing the shape of background in SR and control data region.

Figure 49 illustrates the summary of background models found suitable to describe the data in control region.

The input background functions are tried in the minimization of the negative logarithm of the likelihood with a penalty term added to account for the number of free parameters in the background model. The likelihood ratio for the penalized likelihood function can be written as

$$-2 \ln \frac{\tilde{\mathcal{L}}(\text{data}|\mu, \hat{\theta}_\mu, \hat{b}_\mu)}{\tilde{\mathcal{L}}(\text{data}|\hat{\mu}, \hat{\theta}, \hat{b})}, \quad (13)$$

where μ is the measured quantity. The numerator is the maximum penalized likelihood for a given μ , at the best-fit values of nuisance parameters, $\hat{\theta}_\mu$ and of the background function, \hat{b}_μ . The denominator is the global maximum for $\tilde{\mathcal{L}}$, achieved at $\mu = \hat{\mu}$, $\theta = \hat{\theta}$ and $b = \hat{b}$. A confidence interval on μ is obtained with the background function maximizing $\tilde{\mathcal{L}}$ for any value of μ . This interval is always wider than those evaluated with the fixed functional form from the global best-fit, $b = \hat{b}$ [43].

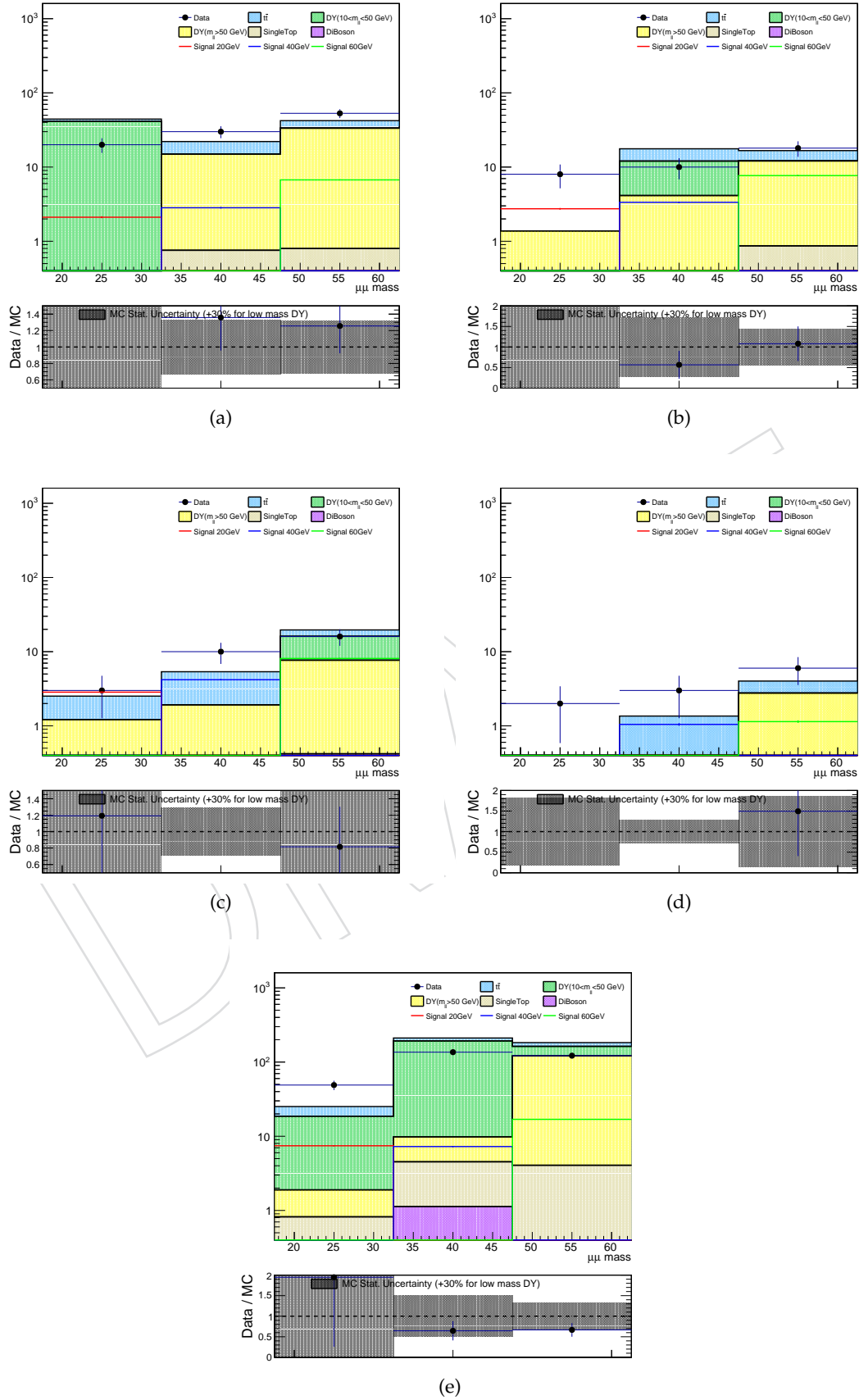


Figure 48: CR Data compared with SR MC in TL category (a) TM category (b) TT category (c) and VBF category (d) and Low p_T category (e).

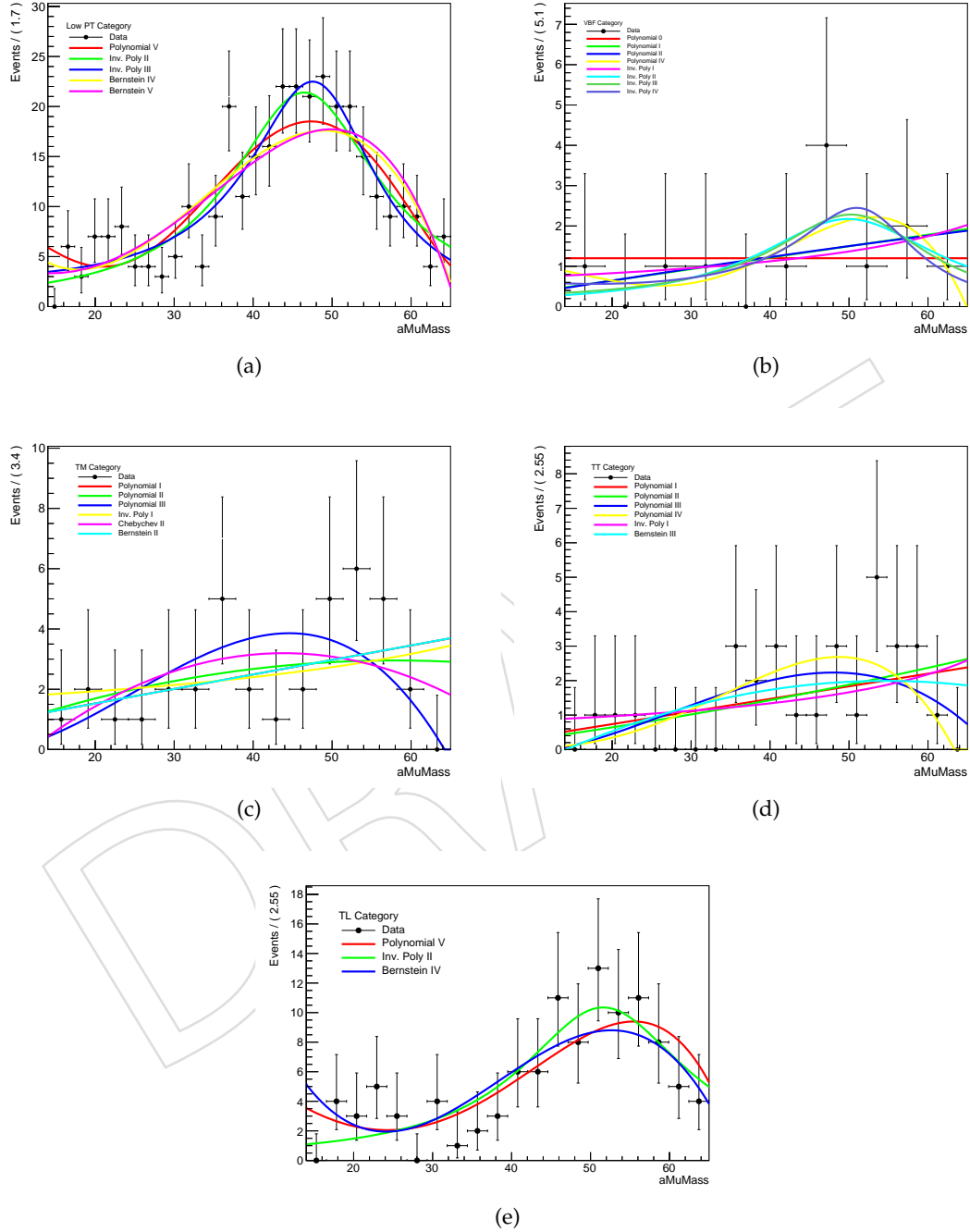


Figure 49: Suitable pdf's describing the background from control region in data in the Low p_T (a), VBF (b), TM (c), TT (d), and TL (e) categories.

C Systematic uncertainties before unblinding

The information here is provided before unblinding and therefore is based on the data in the control region. Figures 50 and 51 show the expected impact of different sources of uncertainties are shown for two signal hypotheses with $m_{a_1} = 19$ and 35 GeV. Overall, systematic uncertainties do not play a significant role. The leading systematic uncertainties are those associated with the modeling of the background which also encode part of the statistical limitation of the analysis. Other uncertainties are negligible in comparison. A zoomed view of their contribution is provided.

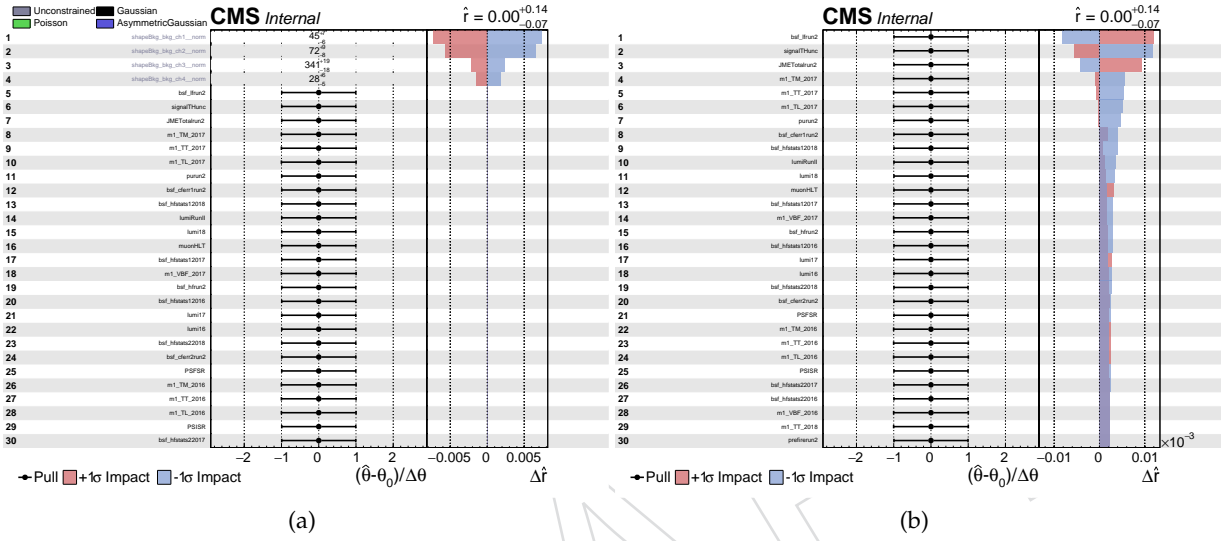


Figure 50: Expected impact of systematic uncertainties for a signal with $m_{a_1} = 19$ GeV including (left) all and (right) negligible contributions.

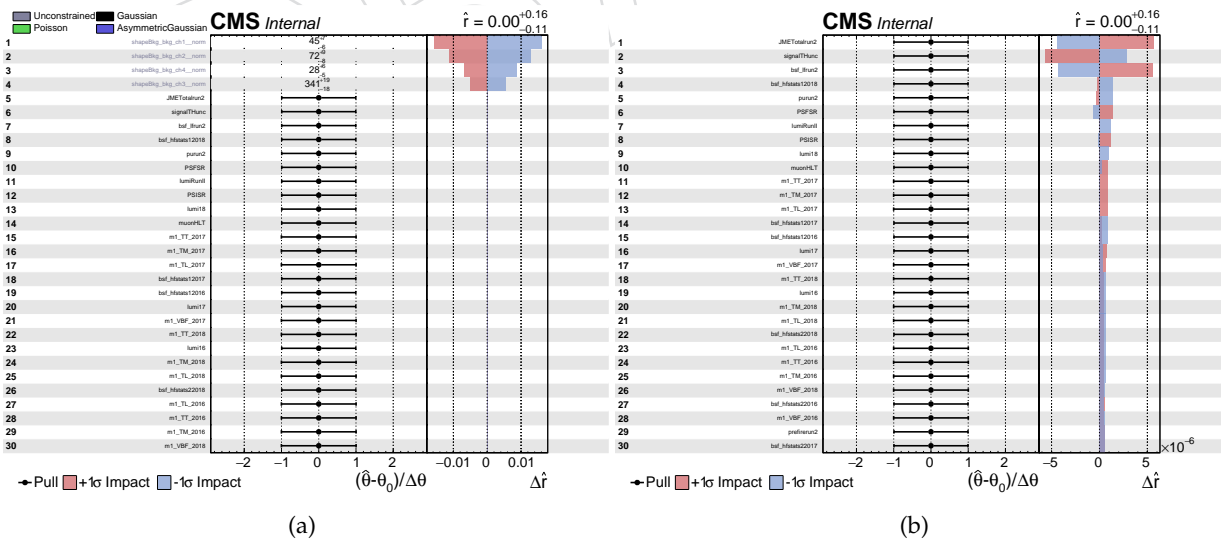


Figure 51: Expected impact of systematic uncertainties for a signal with $m_{a_1} = 35$ GeV including (left) all and (right) negligible contributions.

D Expected results before unblinding

The results here are for the blind analysis and are therefore based on the data in the control region. In comparison with ATLAS [23], expected results are compatible within one standard deviation almost over the entire mass range. A slight improvement is however observed from the current analysis. As demonstrated in Fig. 53, current limits are more stringent than those of

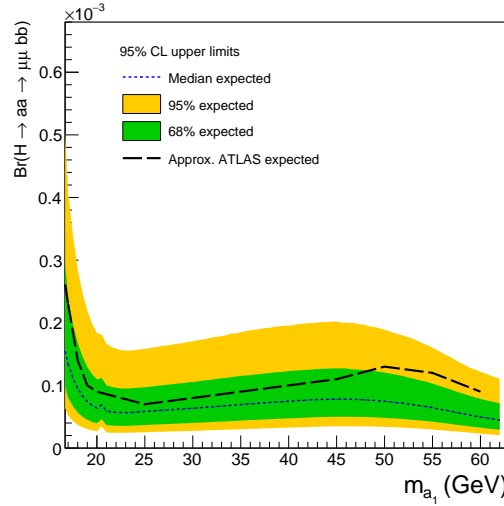


Figure 52: Expected limits, at 95% CL, on $\text{Br}(H \rightarrow a_1 a_1 \rightarrow \mu^+ \mu^- b \bar{b})$ at 13 TeV using full Run-II data. ATLAS results are shown for comparison.

the CMS analysis using 2016 data [25]. This goes beyond the increase of the luminosity, thanks to optimized usage of signal properties over the mass range.

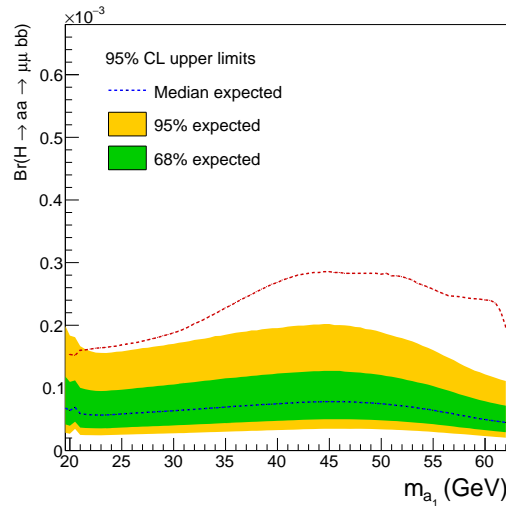


Figure 53: Comparison of the power of the current analysis with that of 2016 data in terms of expected limits, at 95% CL, on $\text{Br}(H \rightarrow a_1 a_1 \rightarrow \mu^+ \mu^- b \bar{b})$.