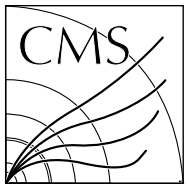
Available on CMS information server

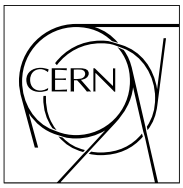
CMS AN -2022/115



The Compact Muon Solenoid Experiment

Analysis Note

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



16 August 2022 (v12, 17 May 2023)

Combined search for $\tau \to 3\mu$ decay using full Run II data

Caterina Aruta, Vladimir Cherepanov, Luca Guzzi, Bhargav Joshi, Sandra Malvezzi, Federica Maria Simone, Rosamaria Venditti, Jian Wang

Abstract

This note presents a search for Lepton Flavour Violating $\tau \to 3\mu$ decays, using τ leptons coming from the $W \to \tau\nu_\tau$ process and from D and B meson decays using data collected by the CMS detector during the full Run II. No evidence of signal is found and an upper limit on the $\tau \to 3\mu$ branching ratio is set to 2.9×10^{-8} (expected 2.4×10^{-8}) at 90% confidence level.

Contents

1 Introduction 2
2 Results of $\tau \to 3\mu$ search in the HF channel with 2017 and 2018 CMS data 3
2.1 Statistical analysis 3
2.2 Signal and background modelling 5
2.3 Summary of the systematic uncertainties 6
3 Results of $\tau \to 3\mu$ search in the W channel with 2017 and 2018 CMS data 8
3.1 Statistical analysis 8
3.2 Signal and background modeling 9
3.3 Summary of the systematic uncertainties 9
4 Combination of HF and W analyses on 2017 and 2018 datasets 11
5 Full Run II combination 25
6 Conclusion 27
7 Acknowledgments 27

1 Introduction

This Analysis Note documents the combined statistical interpretation of the results of the $W \to \tau\nu$, $\tau \to 3\mu$, and $D/B \to \tau X$, $\tau \to 3\mu$ (sometimes referred to as heavy flavour - HF) analyses, described in detail in AN-2020/102 [1] and AN-2020/127 [2] respectively. The analyses are conducted on data collected by CMS during the LHC proton-proton collisions of 2017 and 2018 at the centre of mass energy of 13 TeV. The combination includes also the published result of the 2016 analysis [3] (notes: [4] [5] [6]). The CMS' software package combine [7] has been used for all statistical interpretations presented in this document.

2 Results of $\tau \to 3\mu$ search in the HF channel with 2017 and 2018 CMS data

This section briefly describes the search for the $\tau \to 3\mu$ process in the HF channel. A complete description of this analysis can be found in the dedicated note [1]. D and B decays are the most abundant source of τ leptons at the LHC. On the other hand, the produced τ have low momenta and are significantly boosted in the forward direction.

The analysis targets three muons forming a good common vertex in the kinematic phase space of the signal, the 3μ vertex being displaced with respect to the primary vertex, with small pointing angle.

2.1 Statistical analysis

The analysis is conducted on data collected by CMS during the LHC proton-proton collisions of 2017 and 2018 at the centre of mass energy of 13 TeV, corresponding to an integrated luminosity of $37.9\text{fb}^{-1} + 59.7\text{fb}^{-1}$. Data has been collected with a dedicated high-level trigger using a three-muon selection path:

- 2017: HLT_DoubleMu3_Trk_Tau3mu_v*
- 2018: HLT_DoubleMu3_TkMu_DsTau3Mu_v*

Events are categorised based on the relative resolution on the 3μ invariant mass (ρ) and split in three exclusive categories:

- Category A: $\rho < 0.0070$
- Category B: $0.0070 \leq \rho \leq 0.0105$
- Category C: $\rho > 0.0105$

The categories A, B, C correlate to the pseudorapidity of the most forward final-state muon, roughly corresponding to the central, overlap and forward regions of the tracker detector.

Data and MC events firing those triggers undergo a set of offline pre-selection requirements, which mainly reproduce the HLT requirements to identify the 3μ candidates in terms of kinematics, vertexing and displacement.

The final-state muons are then required to fulfill some reconstruction and identification algorithms. At least two muons reconstructed with the Global Muon and Particle Flow algorithms are required in each event. Events are then splitted in two exclusive categories based on the reconstruction of the remaining muon:

- third muon reconstructed as a Global and Particle Flow: **Three Global Muons** category (3Glb),
- third muon reconstructed as a Tracker Muon and Particle Flow but not as a Global Muon: **Two Global Plus One Tracker Muon** category (2Glb+1Trk).

Events are further skimmed by a BDT discriminator, trained using real data falling outside the signal region (background) and simulated $\tau \to 3\mu$ signal samples (signal). The mass resolution categories (A, B, C) and the muon ID categories (3Glb, 2Glb+1Trk) have dedicated BDT trainings to exploit the CMS reconstruction capability and the signal kinematics thus maximising the sensitivity in each sub-category.

Based on the BDT score, the most backgroud-like events are discarded and the remaining events are splitted into three (two) sub-categories for 3Glb (2Glb+1Trk), which are labelled with

numbers (”1”, ”2”, ”3”). The boundaries between the categories are determined by computing the combined signal significance, while varying the three thresholds simultaneously over the full BDT score range. Since the algorithm only aims at maximising the combined AMS, the sub-categories are not sorted by significance.

Events passing the BDT selections are used to compute the strength of the $\tau \to 3\mu$ process with an unbinned maximum likelihood fit to the three-muon invariant mass distribution, or the upper limit to the process using the CLs approach.

Normalising the simulated signal yields to data would require the knowledge of the absolute production cross section of heavy-flavoured mesons. Moreover, it would involve a precise determination of the reconstruction and trigger efficiencies. Therefore, the $\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ channel is used to normalise the simulated signal yields to data with reduced systematic uncertainties.

The expected $\tau \to 3\mu$ signal event yield from $\mathrm{D_s^+} \to \tau^+\nu$ decays can be written as:

$$N_{\mathrm{sig(D)}} = \mathcal{L}\ \sigma(pp \to D_s)\ \mathcal{B}(\mathrm{D_s^+} \to \tau^+\nu)\ \mathcal{B}(\tau \to 3\mu)\ \mathcal{A}_{3\mu(D)}\ \epsilon^{3\mu}_{\mathrm{reco}} \epsilon^{2\mu}_{\mathrm{trig(sig)}} \tag{1}$$

while the number of $\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ events is given by:

$$N = \mathcal{L}\ \sigma(pp \to D_s)\ \mathcal{B}(D_s \to \phi\pi \to \mu\mu\pi)\ \mathcal{A}_{2\mu\pi}\ \epsilon^{2\mu\pi}_{\mathrm{reco}}\ \epsilon^{2\mu}_{\mathrm{trig}(\mu\mu\pi)} \tag{2}$$

where $\mathcal{L}$ is the integrated luminosity, $\sigma(pp \to D_s)$ is the D_s production cross section, $\mathcal{B}(\mathrm{D_s^+} \to \tau^+\nu)$, $\mathcal{B}(\tau \to 3\mu)$, $\mathcal{B}(\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+)$ are the decay branching fractions, $\mathcal{A}$ is the event selection acceptance, ϵ_{trig} and ϵ_{reco} are instrumental trigger and reconstruction efficiencies.

By measuring $\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ decay rate N, the expected $\tau \to 3\mu$ signal yield N_{sig} from $\mathrm{D_s^+} \to \tau^+\nu$ decays can be estimated with significantly reduced systematic uncertainties:

$$N_{\mathrm{sig(D)}} = N \cdot \frac{\mathcal{B}(\mathrm{D_s^+} \to \tau^+\nu)}{\mathcal{B}(\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+)} \frac{\mathcal{A}_{3\mu(D)}}{\mathcal{A}_{2\mu\pi}} \frac{\epsilon^{3\mu}_{\mathrm{reco}}}{\epsilon^{2\mu\pi}_{\mathrm{reco}}} \frac{\epsilon^{2\mu}_{\mathrm{trig(sig)}}}{\epsilon^{2\mu}_{\mathrm{trig}(\mu\mu\pi)}} \mathcal{B}(\tau \to 3\mu). \tag{3}$$

In fact, the trigger efficiencies mostly cancel out for events collected exactly with the same set of triggers, while in the ratio of the reconstruction efficiencies the reconstruction of one muon with respect to the track has to be taken into account.

Direct $B \to \tau + ...$ decays are the second largest source of τ-leptons. Their contribution to the signal can be written as follows:

$$N_{\mathrm{sig(B)}} = \mathcal{L}\ \sigma(pp \to B)\ \mathcal{B}(B \to \tau X)\ \mathcal{B}(\tau \to 3\mu)\ \mathcal{A}_{3\mu(B)}\ \epsilon^{3\mu}_{\mathrm{reco}}\ \epsilon^{2\mu}_{\mathrm{trig(sig)}}. \tag{4}$$

The expected $\tau \to 3\mu$ signal yield $N_{\mathrm{sig(B)}}$ from B decays cannot be directly normalized by $\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+$. Nevertheless, the consistency between the $B^0, B^\pm$ and the D_s meson production rates can be tested as follows.

The D_s production is a mixture of prompt D_s and $B \to D_s$. These two components can be discriminated in terms of decay length which can be used to measure the ratio between the prompt and $B \to D_s$ yields in data and MC. The $B \to D_s$ yield fraction (f) with respect to the total D_s yield can be defined as:

$$f = \frac{\sigma(pp \to B)\mathcal{B}(B \to D_s X)}{\sigma(pp \to D_s)} \tag{5}$$

By combining Equations 4, 2 and 5, the predicted contribution from the direct B-decays to the signal is found to be:

$$N_{\mathrm{sig(B)}} = N \cdot f \cdot \frac{\mathcal{B}(B \to \tau X)}{\mathcal{B}(D_s \to \phi\pi \to \mu\mu\pi)\mathcal{B}(B \to D_s X)} \frac{\mathcal{A}_{3\mu(B)}}{\mathcal{A}_{2\mu\pi}} \frac{\epsilon^{3\mu}_{\mathrm{reco}}}{\epsilon^{2\mu\pi}_{\mathrm{reco}}} \frac{\epsilon^{2\mu}_{\mathrm{trig,sig}}}{\epsilon^{2\mu}_{\mathrm{trig}(\mu\mu\pi)}} \mathcal{B}(\tau \to 3\mu), \quad (6)$$

where N is again the $\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ decay rate. The ratio (f) between the two production rates can be measured in both data and MC again using the $\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ channel. Therefore, if the MC prediction is in agreement with the measured ratio in data, the data-MC normalisation is also validated for B decays and $N_{\mathrm{sig(B)}}$ can be indirectly normalised to the $\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ channel.

In the normalisation of the signal MC the branching fraction of the $\tau \to 3\mu$ process is arbitrarily assumed to be 10^{-7}.

The simulated sample is corrected using per-event scale factors computed comparing data and MC in the control channel. The corrections cover for the online and offline muon reconstruction and identification efficiencies, the trigger efficiencies, the pile-up profile and the shape of per-muon MVA variables.

2.2 Signal and background modelling

The signal $m(3\mu)$ shapes are modelled in each subcategory using the signal MC events, while data events from the two sideband region are fitted to access the background model, where the tau mass region is blinded based on the mass resolution category as described in Table 1.

Category	Width	Sideband Mass Range	
		Low	High
A	48 MeV	[1.62, 1.753]	[1.801, 2.0]
B	76 MeV	[1.62, 1.739]	[1.815, 2.0]
C	100 MeV	[1.62, 1.727]	[1.827, 2.0]

Table 1

To choice of the p.d.f. to model the background and to assess a systematic uncertainty we use the discrete profiling method, where the choice of background function is handled as a discrete nuisance parameter.

A "multi-pdf" is built using exponential, powerlaw and polynomial functions up to the 6th order. For each family of p.d.fs, an F-test is used to select the maximum order. A penalty of magnitude equal to the order is added to the N+1-order negative-log-likelihood, so that the difference between the order N and N+1 is reduced. This disfavours the order N+1, that is higher orders.

The best-fit pdf is then chosen based on the g.o.f. The multi-pdf and the index of the best-fit function are passed to the workspaces. This procedure is repeated independently for each event subcategory.

Tables 2 and 3 show which pdf has been used in the different subcategories for 3Glb and 2Glb-Trk respectively.

The signal model is obtained by fitting the MC. For all signal parameters, the fitted value is set to constant and used to build the signal model passed to Combine. Therefore, the signal parameters are fixed in the ML fit.

	3Glb																
	2017									2018							
	A1	A2	A3	B1	B2	B3	C1	C2	C3	A1	A2	A3	B1	B2	B3	C1	C2	C3
expo	X			X		X										X		
powerlaw		X					X			X			X					
poly1			X		X			X	X		X	X		X	X			X
poly2																	X	
poly3..6																		

Table 2

	2Glb+1Trk											
	2017						2018					
	A1	A2	B1	B2	C1	C2	A1	A2	B1	B2	C1	C2
expo	X	X				X		X			X	
powerlaw							X					
poly1			X	X	X				X	X		X
poly2												
poly3..6												

Table 3

The chosen pdf is a gaussian and a crystal ball (CB) function to account for shape asymmetry due to radiative energy losses.

The signal shapes are fit independently in each mass-resolution category (A, B, C). Across the subcategories (e.g. A1, A2, A3), a simultaneous fit is performed where the CB tail parameters (n, α) and the central mass value (m_0) are shared, while the sigma parameters and the gaussian/CB relative contribution are left unconstrained across the subcategories. From the fit, the parameters of the shapes are extracted. The signal in each subcategory is refit using those fixed parameters.

2.3 Summary of the systematic uncertainties

This analysis is driven by statistical uncertainties, incorporated in the statistical analysis assuming a Poisson distribution for data. Nevertheless, several sources of systematic uncertainties can be found, mostly related to the normalization of the signal samples, directly affecting the signal yields.

Table 4 summarizes all sources of systematic uncertainties associated with modelling the signal event yields and trimuon mass distribution shapes.

These uncertainties are treated as nuisance parameters and profiled during the limit computation. The nuisance paramters related to the signal normalization are log-normal constrained, with a constraint strength equal to the associated systematic uncertainty.

The branching fractions uncertainties are correlated across the analysis categories and years. On the other hand, the corrections computed using the $D_s \to \phi(\mu\mu)\pi$ control channel are studied independently for 2017 and 2018 samples. The related uncertanties are therefore uncorrelated among the years. Moreover, the muon-ID specific corrections (muon identification and reconstruction efficiencies, MVA ID shape) are independent in the two main analysis categories (3Glb and 2Glb+1Trk). Finally, the systematics covering for BDT cut efficiencies in data and MC are taken as fully correlated across the different eras and subcategories.

Table 4: Sources of systematic uncertainties affecting signal modeling and their impact on the expected signal event yield and trimuon mass distribution shape.

Source of uncertainty	Yield (2017)	Yield (2018)	Shape
Uncertainty on D_s normalization [6.2% (2017), 4% (2018)]	6.2%	4.0%	
Uncertainty on W normalization [100%]	3%	5%	
Relative uncertainty in $\mathcal{B}(D_s \to \tau\nu)$ [4%]	3%	3%	
Relative uncertainty in $\mathcal{B}(D_s \to \phi\pi \to \mu\mu\pi)$ [8%]	8%	8%	
Relative uncertainty in $\mathcal{B}(B \to D_s X)$ [16%]	5%	5%	
Relative uncertainty in $\mathcal{B}(B \to \tau X)$ [11%]	3%	3%	
Uncertainty on measuring f (B/D ratio) [7%]	2%	2%	
Uncertainty on scaling D_s to include D^+ [4%]	3%	3%	
Uncertainty on scaling B^0 and B^+ to include B_s [12%]	4%	4%	
Uncertainty on n. events triggered by TripleMu [5% (2017), 2% (2018)]	5%	2%	
Uncertainty on 2018 HLT efficiency [5%]	–	5%	
Uncertainty on the ratio of acceptances $\mathcal{A}_{\text{sig}}/\mathcal{A}_{2\mu\pi}$ [1%]	1%	1%	
Muon reco. efficiency [*] [1.5% (2017), 1.6% (2018)]	1.5%	1.6%	
Pion reco. efficiency [2.2% (2017), 2.1% (2018)]	2.2%	2.1%	
Tracker muon reco. efficiency [**] [4% (2017), 8% (2018)]	4%	8%	
BDT cut efficiency [depends on subcategory A1-B1-C1/A2-B2-C2/A3-B3-C3]	20%/5%/0%	15%/7%/0%	
Global MVA muon ID [*] [3% (2017), 6% (2018)]	3%	6%	
Tracker MVA muon ID [**] [4% (2017), 4% (2018)]	4%	4%	
Muon momentum scale uncertainty [0.09%]	–	–	yes
Muon momentum resolution uncertainty [2%]	–	–	yes

[*] Applied to the "Three Global" category
[**] Applied to the "Two Global and 1 Tracker" category

3 Results of $\tau \to 3\mu$ search in the W channel with 2017 and 2018 CMS data

This section briefly describes the search for the $\tau \to 3\mu$ process in the $W \to \tau\nu$ τ production channel. A complete description of this analysis can be found in the dedicated note [2]. Although less abundant with respect to the HF channel, the W channel offers some characteristics which allow for a better mitigation of the background contamination and make the W channel a key element in the study of the $\tau \to 3\mu$ process; mainly the large missing transverse momentum generated by the W decay, the presence of an isolated and collimated three-muon final state, and the presence of a displaced secondary vertex.

3.1 Statistical analysis

The analysis is conducted on data collected by CMS during the LHC proton-proton collisions of 2017 and 2018 at the centre of mass energy of 13 TeV. Data has been collected with a dedicated high-level trigger using a three-muon selection path and are further skimmed by a BDT discriminator. The BDT is trained using real data falling outside the signal region (background) and simulated $\tau \to 3\mu$ signal samples (signal), exploiting the topological and kinematic properties of the $W \to \tau\nu$, $\tau \to 3\mu$ decay, as well as the muon track quality information. The electroweak background $W \to 3\mu\nu_\mu$ is rejected introducing a cut on the displacement significance between the secondary vertex and the beamspos in the transverse plane $l/s > 2$.
Data passing the BDT selection are used to compute the strength of the $\tau \to 3\mu$ process with an unbinned maximum likelihood fit to the three-muon invariant mass distribution, or the upper limit to the process using the CLs approach.
The expected number of signal events is predicted using the simulated $\tau \to 3\mu$ signal sample, normalized using Eq. 7.

$$N = \mathcal{L} \cdot \sigma(\mathrm{pp} \to W(\to \mu\nu_\mu) + X) \cdot \frac{BR(\mathrm{W} \to \tau\nu_\tau)}{BR(\mathrm{W} \to \mu\nu_\mu)} \cdot r \cdot BR(\tau \to 3\mu) \tag{7}$$

where r is the signal strength (the parameter of interest of this analysis) and the branching fraction of the $\tau \to 3\mu$ process is arbitrarily assumed to be 10^{-7}.
Particular care has been taken into ensuring that the simulated signal samples correectly reflects the real detector condition during the data taking; the simulated sample is corrected using per-event scale factors affecting the online and offline muon reconstruction and identification efficiencies, the trigger efficiencies, the pile-up profile and the generated W spectrum.
The analysis is performed using three independent event categories per year, based on the relative three-muon invariant mass resolution, as shown in Fig. 1. For each category, a dedicated working point of the BDT is computed against the expected upper limit.

All the selection criteria described above were determined in a blinded way, without looking at data events with invariant mass in the range [1.74, 1.82] GeV. Once the data were unblinded, it was observed that a significant fraction of events had the trimuon vertex not displaced from the primary vertex. These events are consistent with the MC simulation prediction of a $W^+ \to \mu^+\mu^-\mu^+\nu_\mu$ sample (where a W boson decays to a muon and a neutrino, and final state radiation off the muon leads to another two muons). Therefore a further requirement is added, without introducing bias, that the distance in the transverse plane between the interation point and the trimuon vertex should be larger than twice of its uncertainty.

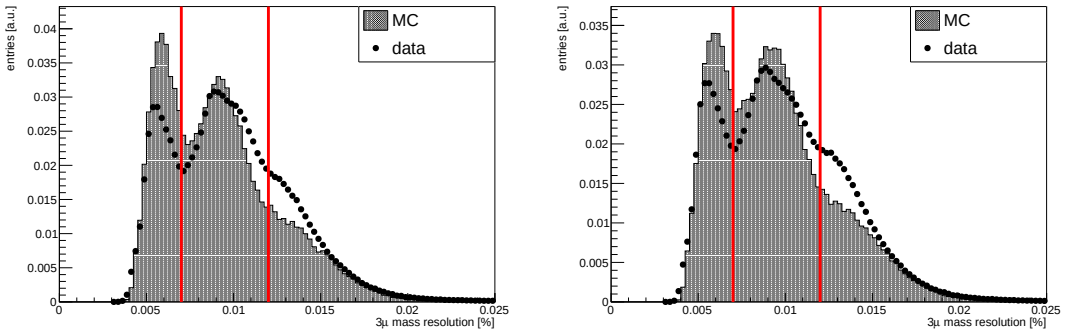


Figure 1: Relative mass resolution of the three-muon candidates on background (points) and signal (bars) events for 2017 (left) and 2018 (right)

3.2 Signal and background modeling

The signal is modelized using a gaussian function. The mean of the gaussian is a fixed parameter; its value is initialized to the τ mass value obtained from the simulated signal sample. The width of the gaussian is a floating parameter; its value is initialized by the simulated sample and it is gaussian-constrained during the fit (the strength of the constraint depends on the category and it is equal to the discrepancy observed on the data-MC comparison of $D_s^+ \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ events). Both the signal mean and width are independent among the analysis categories.

The background is a flat polinomial (constant). The background normalization is an unconstrained parameter initialized to the epected number of background events, obtained by the fit to the signal side-bands, independent among the analysis categories. To count for the uncertainties of the background estimation in categories with a low sideband yield, an hybrid-bayesian method is used during the limit computation. The nuisance parameter values are randomized during the toy generation; in particular, background normalizations are sampled from a flat distribution defined in a range equal to the 99% binomial confidence interval around the pre-fit value.

3.3 Summary of the systematic uncertainties

The systematic uncertainties of the analysis include uncertainties associated to the signal sample normalization (the integrated luminosity collected by CMS in each year, the $W \to \tau\nu$ production cross section and the related branching fractions, the finite MC sample statistics, the online and offline muon efficiency corrections, the trigger isolation correction, the NLO reweighting of the MC sample) and to the signal and background modeling (the normalization of the background pdf and the width of the signal pdf). These uncertainties are treated as nuisance parameters and profiled during the limit computation. The only sources of systematic uncertainties correlated among the analysis categories are the luminosity (independent between the two years) and the $W \to \tau\nu$ production cross section and the related branching fractions uncertainties (also correlated between the two years). When two sources of systematic uncertainty are correlated they are described by the same nuisance parameter.

The nuisance paramters related to the signal normalization are log-normal constrained, with a constraint strength equal to the associated systematic uncertainty.

To count for the uncertainty of categories with a low number of events in the sidebands, ran-

dom values are generated for nuisance parameters at each toy generation (this method is known as hybrid-bayesian method). In particular, background normalizations are randomized with a flat pdf defined in the 99% binomial confidence interval around the pre-fit value of the background yield.

Table 5 lists the systematic uncertainties of the W channel.

Table 5: Summary of systematic uncertainties for 2017 (2018). Where a single value is given, the systematic uncertainty is assumed correlated among the years.

Source of uncertainty	process	type	cat. A	cat. B	cat. C	correlated among cat.
luminosity	sig.	rate - lnN	2.3 (2.5)%	2.3 (2.5)%	2.3 (2.5)%	yes
$\sigma(\mathrm{pp} \to \mathrm{W})$	sig.	rate - lnN	3.7%	3.7%	3.7%	yes
$\mathcal{B}(W \to \tau\nu_\tau)$	sig.	rate - lnN	1.4%	1.4%	1.4%	yes
$\mathcal{B}(W \to \mu\nu_\mu)$	sig.	rate - lnN	1.8%	1.8%	1.8%	yes
finite MC statistics	sig.	rate - lnN	1.0 (0.7)%	0.7 (0.5)%	1.1 (1.2)%	no
muon identification	sig.	rate - lnN	1.3 (3.9)%	1.4 (4.7)%	1.5 (5.2)%	no
HLT muon efficiency	sig.	rate - lnN	1.9%(1%)	2.1%(1%)	2.2%(1%)	no
HLT tracker muon efficiency	sig.	rate - lnN	11%(8%)	10%(8%)	15%(9%)	no
HLT isolation correction	sig.	rate - lnN	12% (7%)	12% (7%)	12% (7%)	yes
NLO reweighing	sig.	rate - lnN	4%	4%	4%	yes
width of the signal	sig.	constr. param. - gauss	2% (2%)	6% (6%)	2% (2%)	no
background normalisation	bac.	unconst. rate param.	-	-	-	no

4 Combination of HF and W analyses on 2017 and 2018 datasets

This section documents the statistical interpretation of the combined results of the $W \to \tau\nu$, $\tau \to 3\mu$ (sometimes referred to as W channel) and $D/B \to \tau + X$, $\tau \to 3\mu$ (sometimes referred to as HF - heavy flavour - channel) analyses. The Higgs Combine package [7] has been used for the statistical interpretation of the results presented in this document.

The combined search consists of performing a simultaneous maximum likelihood fit to the invariant mass distribution of the 3μ signal candidates in the following analysis categories (per year):

- HF-3Global Muons: 9 categories according to the mass resolution (3 categories, labelled *A*, *B* and *C*) and to the BDT output (3 sub-categories, labelled *1*, *2* and *3*)
- HF-2Global and 1 Tracker Muons: 6 categories according to the mass resolution (3 categories, labelled *A*, *B* and *C*) and BDT output (2 sub-categories, labelled *1* and *2*)
- W: 3 categories according to the mass resolution (labelled *A*, *B* and *C*)

Thus, the simultaneous fit is perfomed across 18 categories per year.
The upper limits are coputed using the CLs method and the so-called LHC test statistics [8] [9], while the discovery significance is computed using the likelihood ratio test statistics of Eq. 8. Figure 18 shows the distribution of the test staitstics used for the upper limit computation of this analysis.

$$q_0 = -2\log\frac{\mathcal{L}(data|r=0,\theta_0)}{\mathcal{L}(data|r,\theta)} \tag{8}$$

The parameter of interest of the fit is the signal strength r, fully correlated among the categories. The nuisance parameters are incorporated in the likelihood function as constrained nuisance parameters. The full list of uncertainties and their constraints are reported in Sec. 2.3 for the HF analysis and in Sec. 3.3 for the W analysis.
The signal and background PDFs are both described by analytical functions as reported in Sec. 2.2 and Sec. 3.2.

The fit range of the 3μ invariant mass distribution is 1.62-2.00 GeV for both the signal and the background.
The systematic uncertainties are assumed to be uncorrelated between the two channels, as the two analyses are completely independent: all the theoretical uncertainties are different and the two analyses have different triggers. Notably, the uncertainty on the collected luminosity, which is accounted for in the W channel, is not included as a systematic for the HF as it cancels out as a consequence of the normalization strategy adopted by the HF analysis.

Within each channel, all categories are mutually exclusive. Data events selected by both the W and HF channels are removed from the HF dataset. This amounts to 3 events in 2017 and 6 events in 2018.

The acceptance of the W analysis to D/B signal events is negligible, whereas the heavy flavour analysis has non negligible acceptance to W signal events. W signal events that fail the W analysis selection (too low BDT scores) but enter the heavy flavour analysis selection amount to 5% of the D/B signal yield itself (7% in 2017, 4% in 2018). These W signal events are therefore

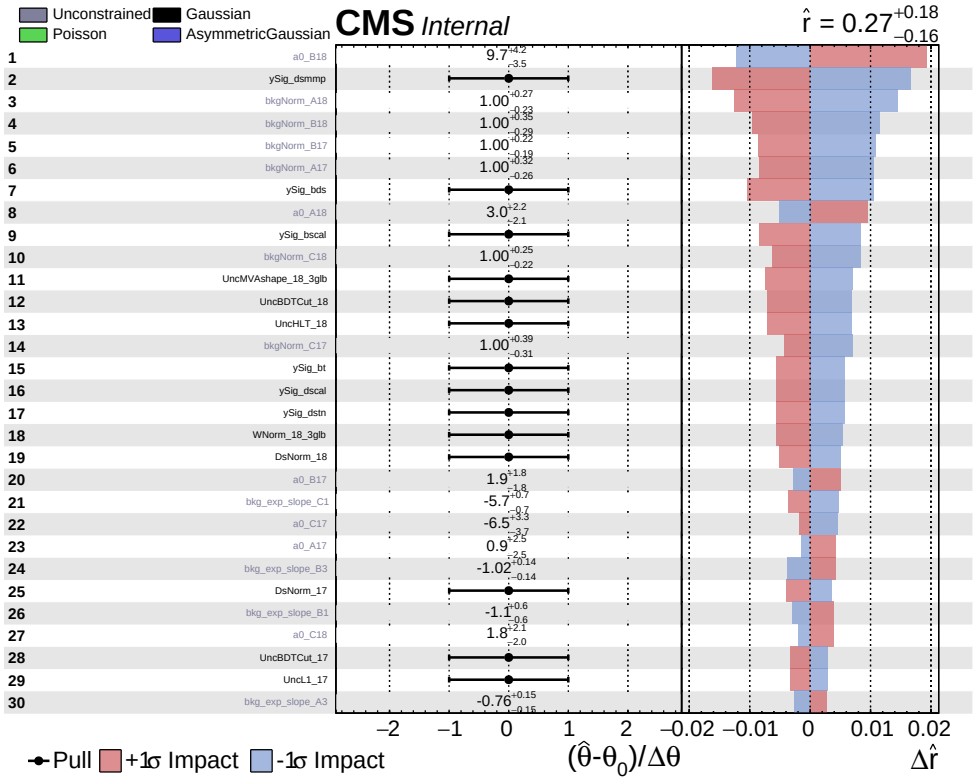


Figure 2: Impact of the nuisance parameters, computed on an Asimov dataset generated with signal strength equal to 0.27 (BR($\tau \to 3\mu$)=$2.7 \cdot 10^{-8}$) (page 1).

added to the D/B signal yield in the HF analysis.
A 100% systematic uncertainty is added on the additional yields to account for the missing data-to-MC correction for the W signal in the HF analysis. As shown by Fig. 17 and Fig. 16, the analysis is dominated by the statistical error.

Table 6 reports the number of expected and observed number of signal and background events, the sensitivity and the expected and observed upper limits of each category after the analysis selection.
The letters "A", "B", "C" refer to mass resolution categories, where category A is expected to have the best sensitivity. In some cases, the larger statistics in B compensates for the worse resolution with respect to A.

In the HF analysis, events are sub-categorised based on the BDT score. As described in Section 2.2, the sub-categories ("1", "2", "3") are defined to maximise the combined signal significance. Therefore, despite their names, the sub-categories are not sorted by sensitivity.

Table 7 reports the expected and observed upper limits of the 2017 and 2018 combination of the two channels.
Figures 2, 3, 4 show the impact and pulls of the nuisance parameters, computed on an Asimov dataset generated with signal strength equal to 0.27 (BR($\tau \to 3\mu$)=$2.7 \cdot 10^{-8}$). Figure 5 shows the correlation between the nuisance parameters following a fit to an Asimov dataset.

Table 6: Observed (expected) number of events in each category of the 2017 and 2018 analysis after the analysis selection, sensitivity and expected (observed) upper limits at 90% of confidence level for each category of the analysis. The signal yield is normalized to BR($\tau \to 3\mu$)=10^{-7} for the number of observed events, BR($\tau \to 3\mu$)=10^{-8} for the upper limits. The expected number of background events is estimated from the integral of the background model in the signal region. The observed number of background events is given by the observation in the signal region, after subtracting the integral of the post-fit signal model.

Category	Signal	observed (expected) n. of bkg events in signal region	observed signal significance	observed (expected) upper limit
HF channel - 3 global muons - 2017				
A1	1.04	1.5 (0.79)	0.45	35.3 (23.0)
B1	1.50	2.2 (2.49)	0.00	17.4 (21.9)
C1	1.99	9.5 (13.0)	0.17	36.7 (33.6)
A2	4.27	17.5 (18.9)	0.00	13.9 (17.0)
B2	2.88	10.6 (12.5)	0.00	14.5 (20.5)
C2	3.57	94.4 (107)	0.54	63.0 (47.7)
A3	7.88	121.8 (135)	0.02	23.4 (22.8)
B3	12.5	185.5 (243)	0.04	19.0 (18.8)
C3	2.79	280.1 (343)	0.92	163 (108)
HF channel - 3 global muons - 2018				
A1	4.60	8.8 (10.0)	0.06	12.9 (12.3)
B1	2.42	3.8 (4.94)	0.00	18.8 (17.2)
C1	2.60	6.0 (5.95)	0.47	24.4 (17.9)
A2	9.42	47.3 (59.7)	1.06	20.1 (12.9)
B2	12.8	54.6 (59.4)	0.00	07.8 (10.0)
C2	4.84	32.6 (37.3)	0.25	25.5 (21.3)
A3	16.7	331.0 (393)	0.00	14.3 (18.1)
B3	33.8	731.7 (944)	0.00	12.1 (13.5)
C3	10.7	270.4 (346)	0.51	35.5 (28.4)
HF channel - 2 global muons and 1 tracker muon - 2017				
A1	0.65	12.2 (2.08)	0.00	29.7 (48.8)
B1	1.28	10.8 (9.93)	0.00	50.4 (59.4)
C1	0.34	4.5 (6.30)	1.24	242 (143)
A2	1.36	24.3 (26.7)	0.00	36.8 (62.2)
B2	3.31	205.8 (224)	0.35	83.4 (72.2)
C2	0.50	41.3 (47.5)	0.66	316 (225)
HF channel - 2 global muons and 1 tracker muon - 2018				
A1	1.09	3.7 (6.49)	0.32	52.3 (35.5)
B1	4.61	39.6 (54.4)	0.00	24.2 (26.2)
C1	2.43	38.0 (42.0)	0.00	28.0 (45.6)
A2	1.63	27.1 (31.0)	0.20	59.4 (55.0)
B2	8.12	393.8 (487)	0.00	37.3 (45.5)
C2	4.12	272.3 (339)	0.32	85.5 (80.0)
W channel - 2017				
A	1.2	0.0 (0.5)	0.0	20.0 (20.5)
B	1.5	2.0 (1.7)	0.0	17.6 (23.0)
C	0.6	1.0 (1.7)	0.7	88.7 (71.8)
W channel - 2018				
A	2.2	2.0 (0.5)	1.5	18.6 (11.1)
B	1.6	1.0 (0.0)	NA	23.5 (16.6)
C	1.1	2.0 (2.3)	0.3	43.5 (43.0)

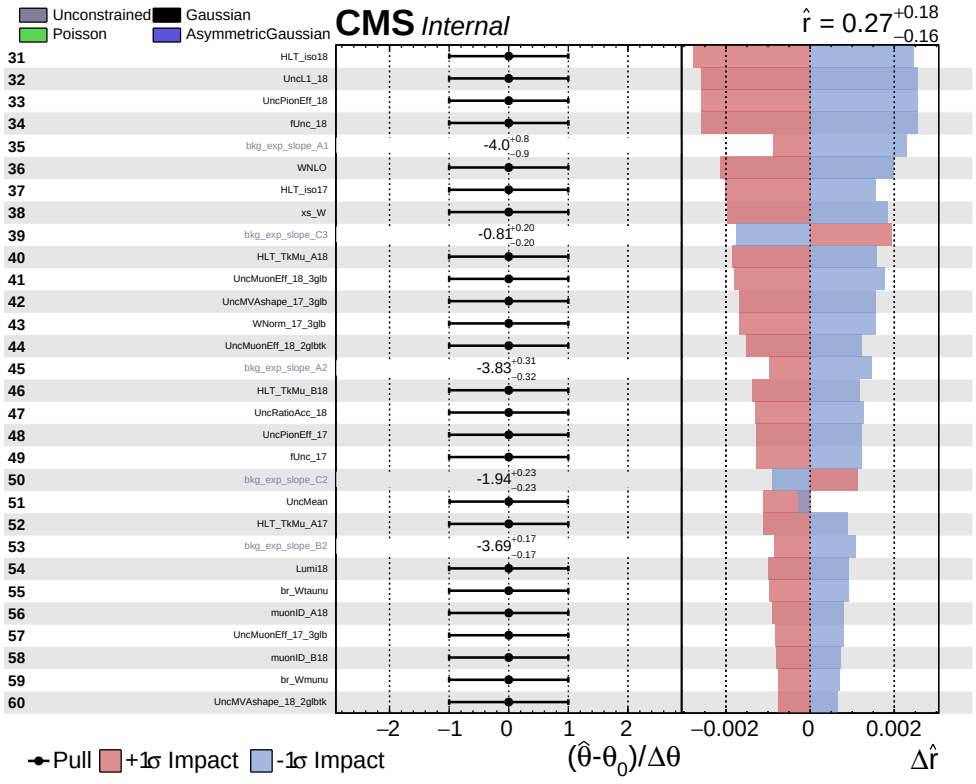


Figure 3: Impact of the nuisance parameters, computed on an Asimov dataset generated with signal strength equal to 0.27 (BR($\tau \to 3\mu$)=$2.7 \cdot 10^{-8}$) (page 2).

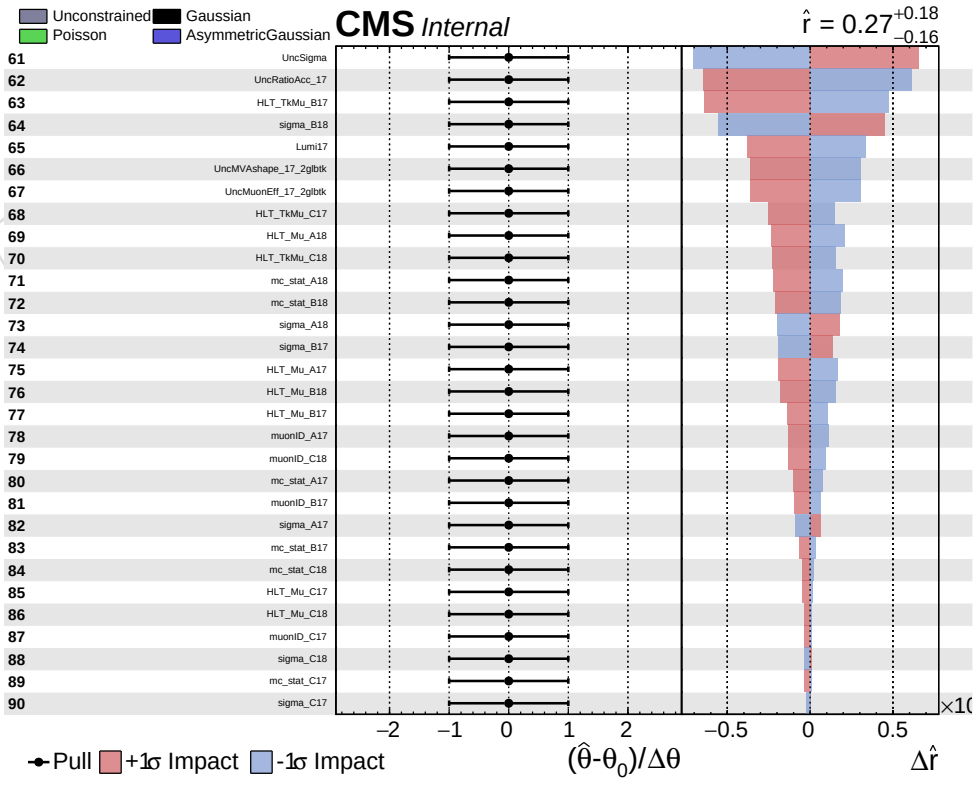


Figure 4: Impact of the nuisance parameters, computed on an Asimov dataset generated with signal strength equal to 0.27 (BR($\tau \to 3\mu$)=$2.7 \cdot 10^{-8}$) (page 3).

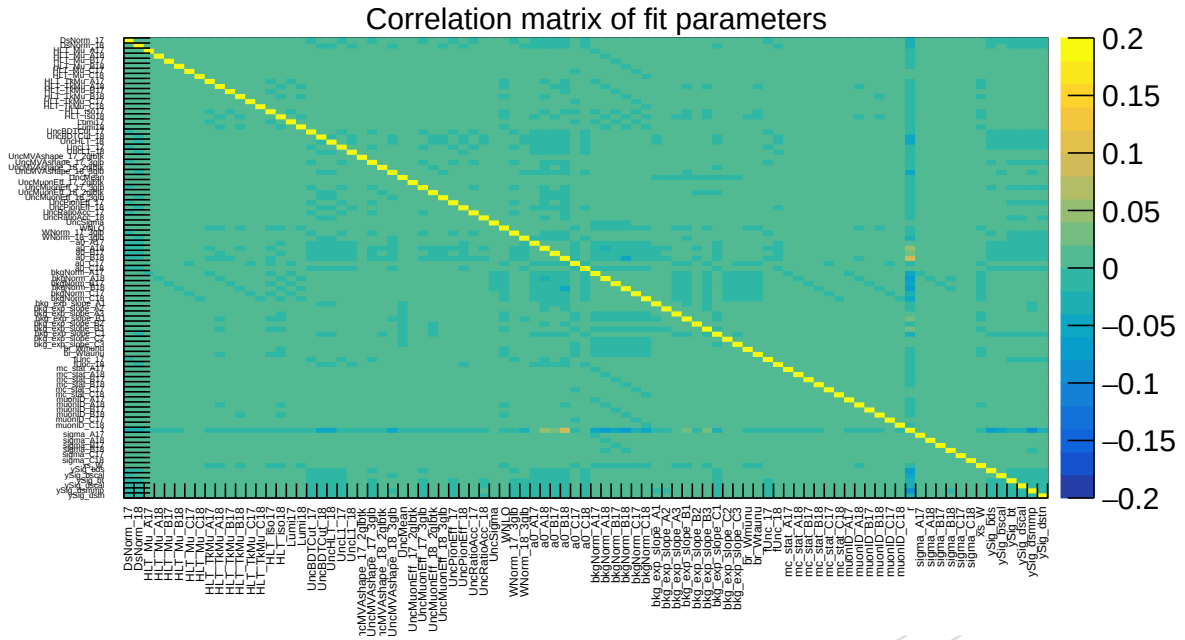


Figure 5: Covariance matrix showing the correlations between the nuisance parameters, computed while fitting an Asimov dataset generated with signal strength equal to 0.27 (BR($\tau \to 3\mu$)=$2.7 \cdot 10^{-8}$).

Figure 15 reports the pulls and impacts for the nuisance paramters of each channel.

The limit computation while fixing all the nuisance parameters in the fit provides the stat-only result. This check has been performed in the AsymptoticLimit approximation. Therefore, the limits are slightly different with respect to Table7 (computed with toys). The combined limit after freezeing the nuisances reads as follows:

Figure 7, 8, 9 and 6 report the unblinded signal-candidate invariant mass plots for each channel after the analysis selections.

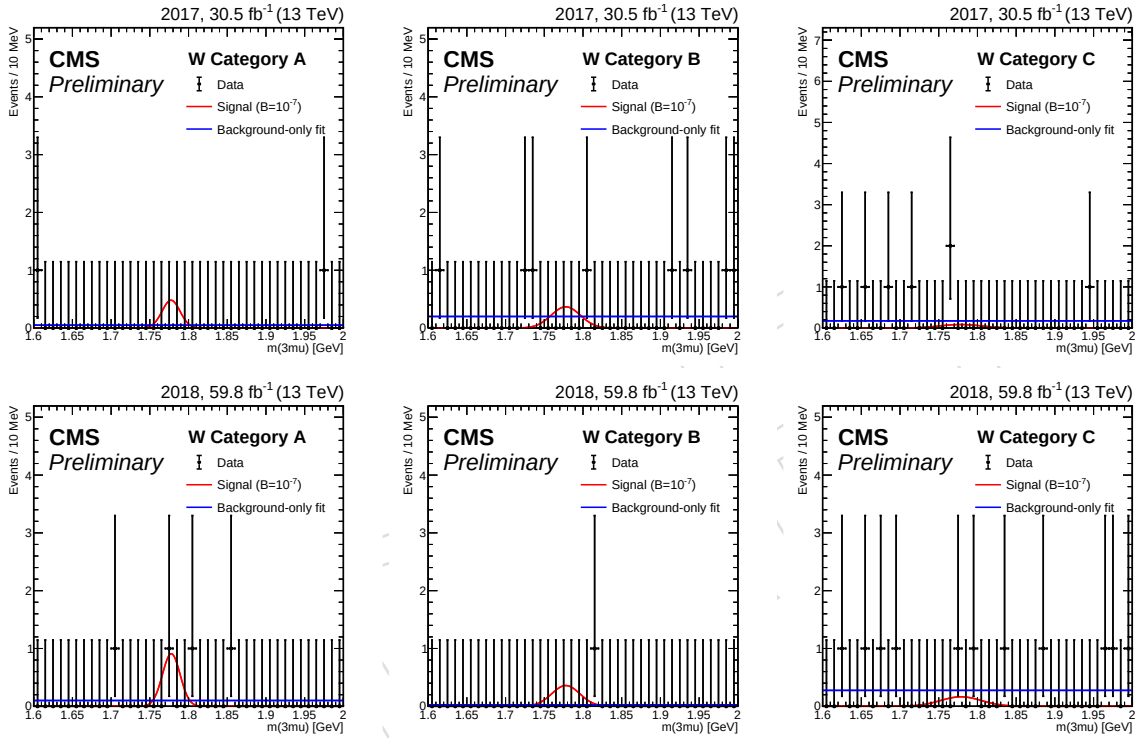


Figure 6: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 2017 (upper plots) and 2018 (lower plots). The signal-plus-background model fit to the data is shown as a blue line and its 68% confidence belt is shown as a yellow area. The background and signal components of the fit model are shown with a red and green line respectively.

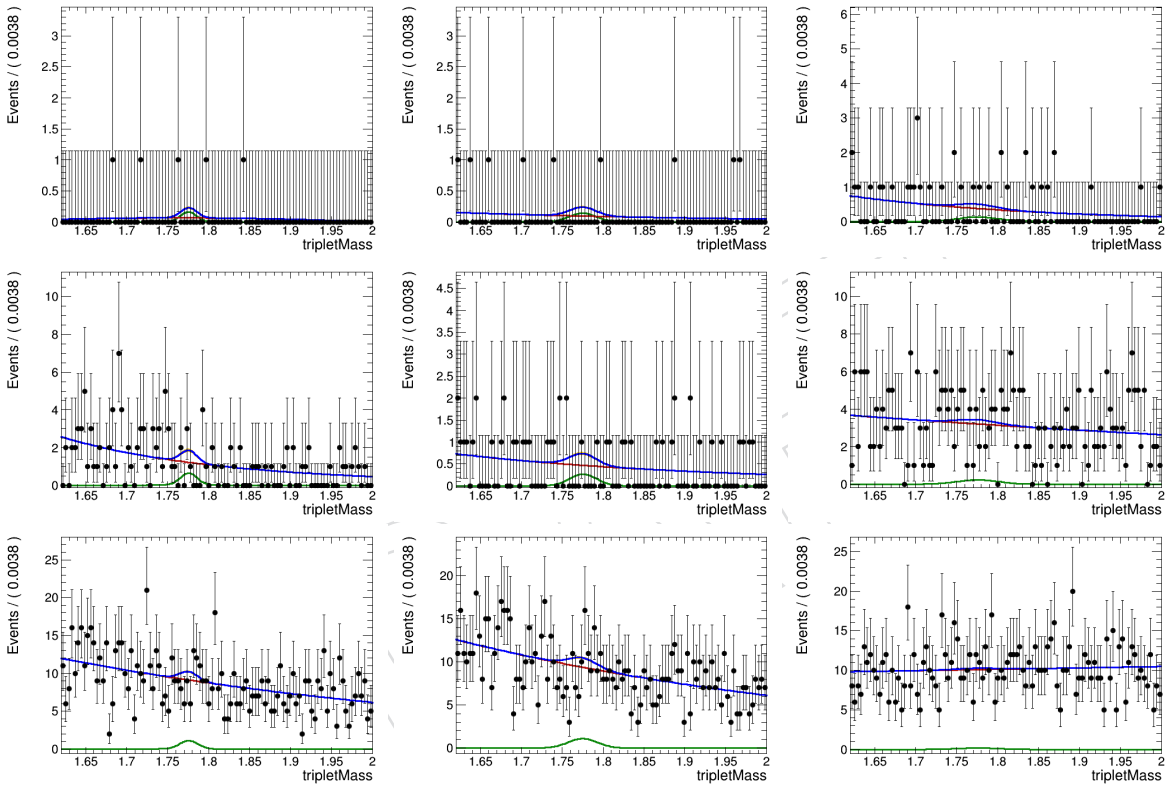


Figure 7: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 3Gl 2017, for the three different BDT-subcategories. The signal-plus-background pre-fit model is shown.

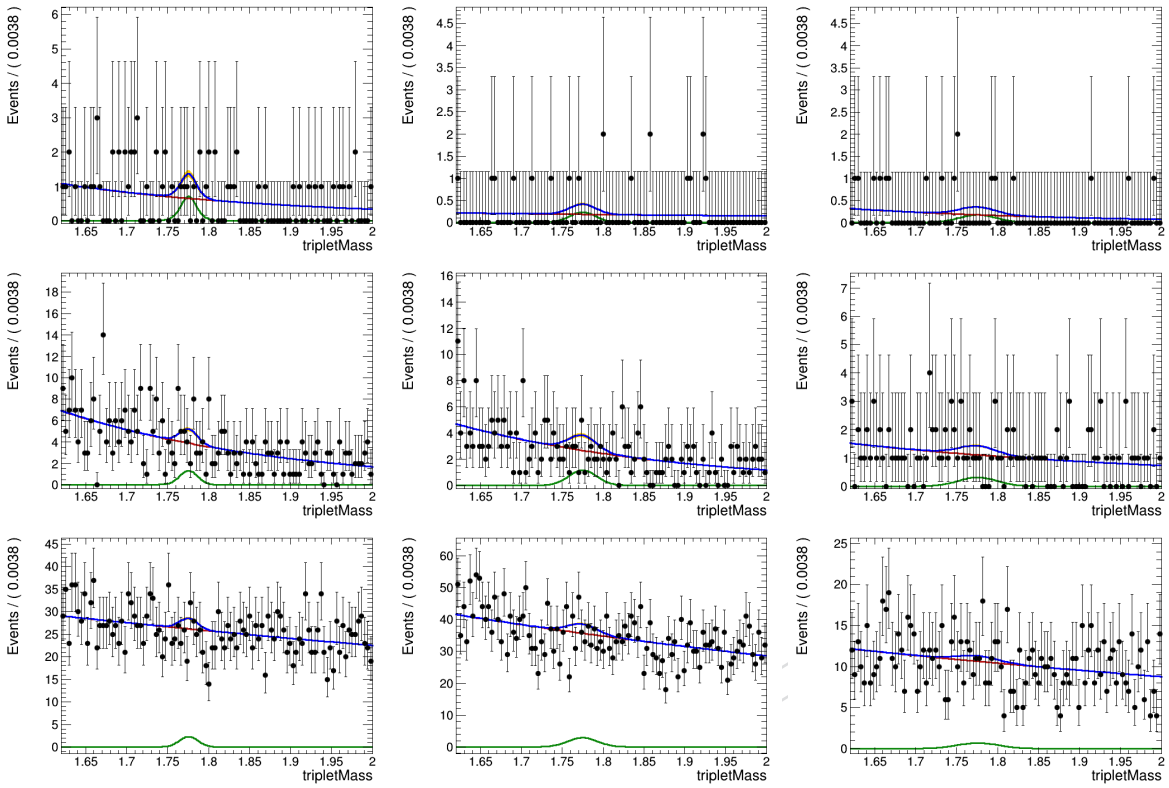


Figure 8: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 3Gl 2018, for the three different BDT-subcategories. The signal-plus-background pre-fit model is shown.

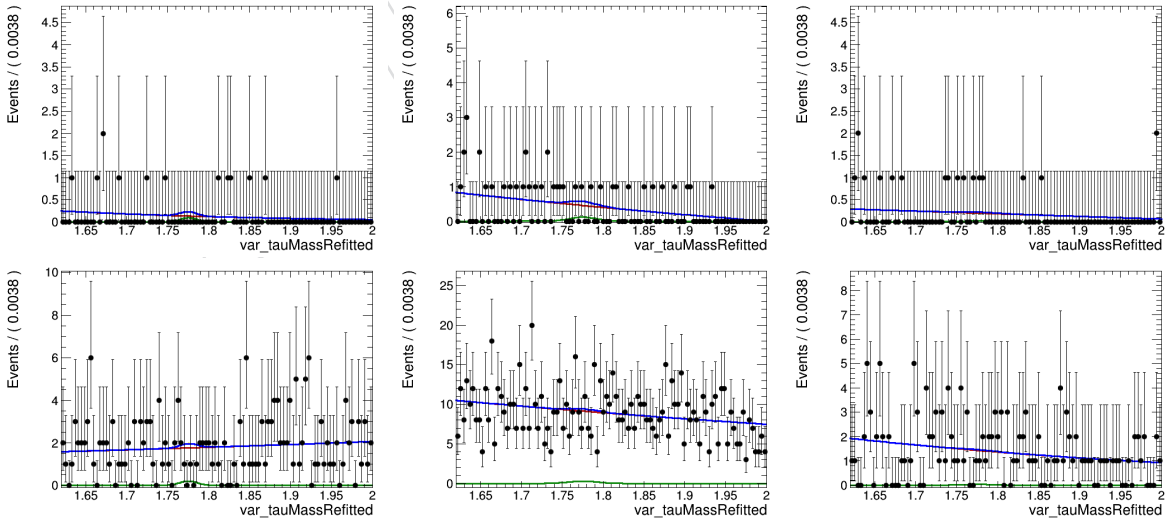


Figure 9: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 2Gl+Trk 2017, for the two different BDT-subcategories. The signal-plus-background pre-fit model is shown.

Table 7: The expected upper limit with its 68% confidence interval and the observed upper limit are reported for different combinations of the analysis categories. The upper limits are computed at 90% of confidence level and reported in units of 10^{-8}. The best fitted value of the signal strength (in units of 10^{-8}) and its significance is also reported for the main categories groups.

Category	-1 sigma	expected	+1 sigma	observed limit	best fit	significance
Three-global 2017	4.7	7.0	10.8	6.1	-0.26	0.0
Three-global 2018	3.1	4.6	7.0	5.2	+0.07	0.29
Two-global-one-tracker 2017	19.8	28.3	42.8	18.8	-1.91	0.0
Two-global-one-tracker 2018	10.6	15.9	24.4	13.9	-0.33	0.0
Heavy flavour channel	2.5	3.6	5.6	3.4	-0.04	0.0
W 2017	9.2	12.9	19.2	9.6	0.0	0.0
W 2018	5.7	7.0	10.5	14.0	0.55	1.4
W channel	4.2	5.6	8.4	8.0	+0.20	0.9
2017 combined	4.2	5.5	9.2	4.6	-0.62	0.0
2018 combined	2.2	3.3	5.2	5.0	+0.20	1.0
2017+2018 combined	1.8	2.7	4.2	3.1	+0.18	0.1

Table 8

Category	-1 sigma	expected	+1 sigma	observed limit	best fit	significance
2017+2018 combined, floating NP	1.84	2.74	4.20	3.20	+0.18	0.12
2017+2018 combined, freezed NP	1.80	2.71	4.10	3.12	+0.17	0.11

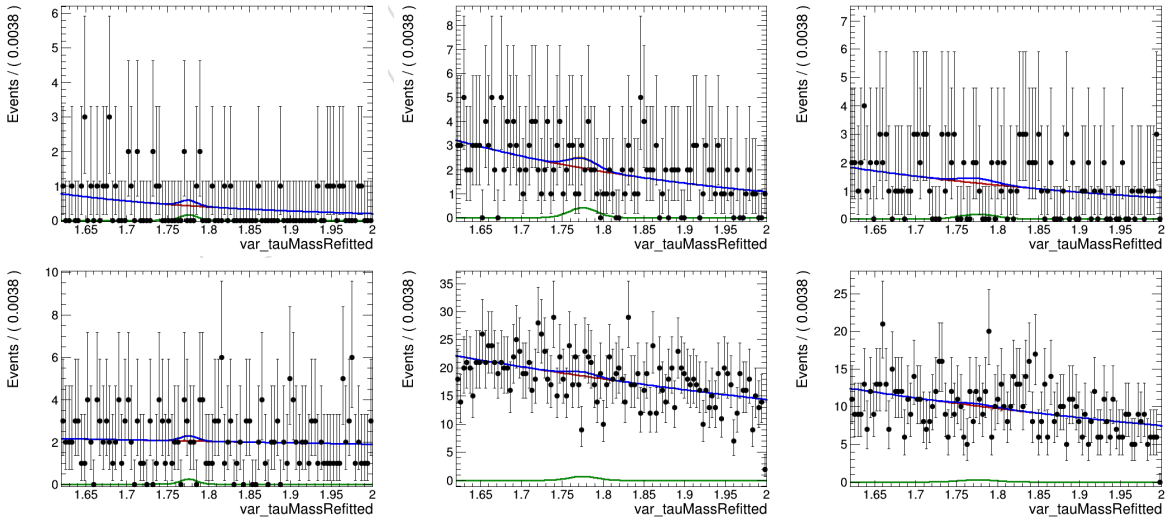


Figure 10: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 2Gl+Trk 2018, for the two different BDT-subcategories. The signal-plus-background pre-fit model is shown.

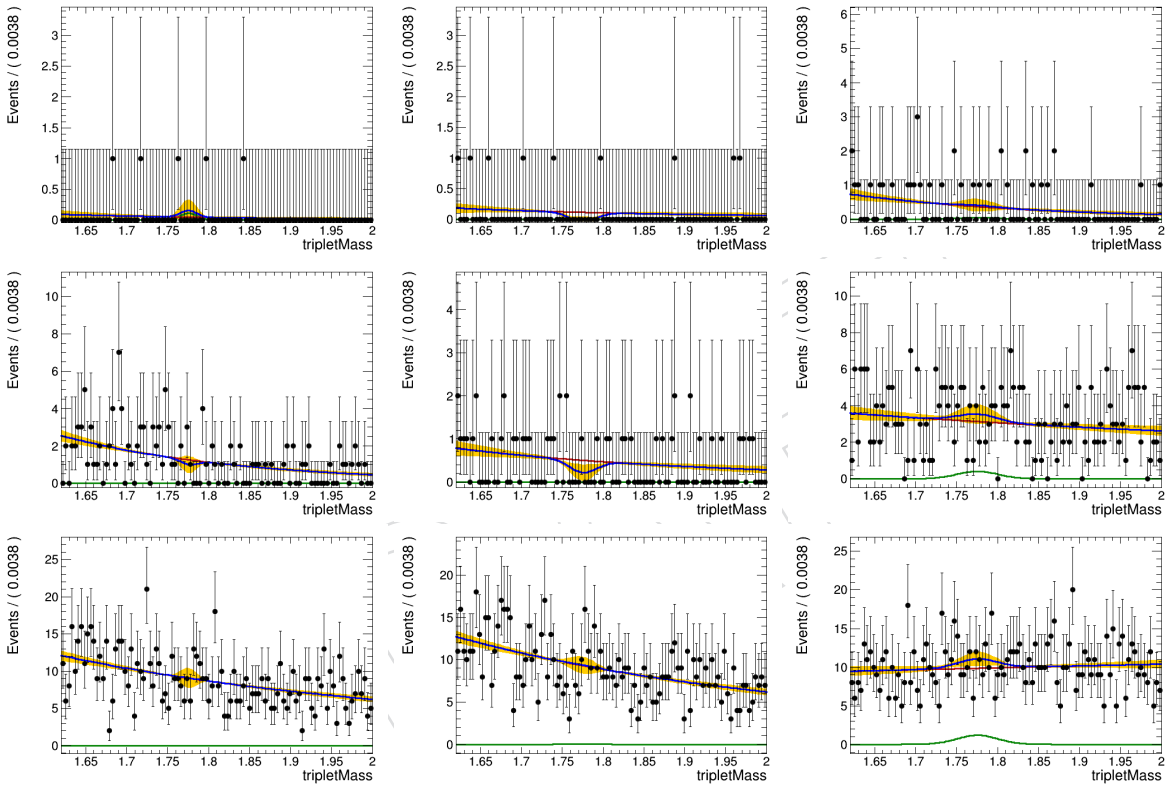


Figure 11: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 3Gl 2017, for the three different BDT-subcategories. The signal-plus-background post-fit model is shown.

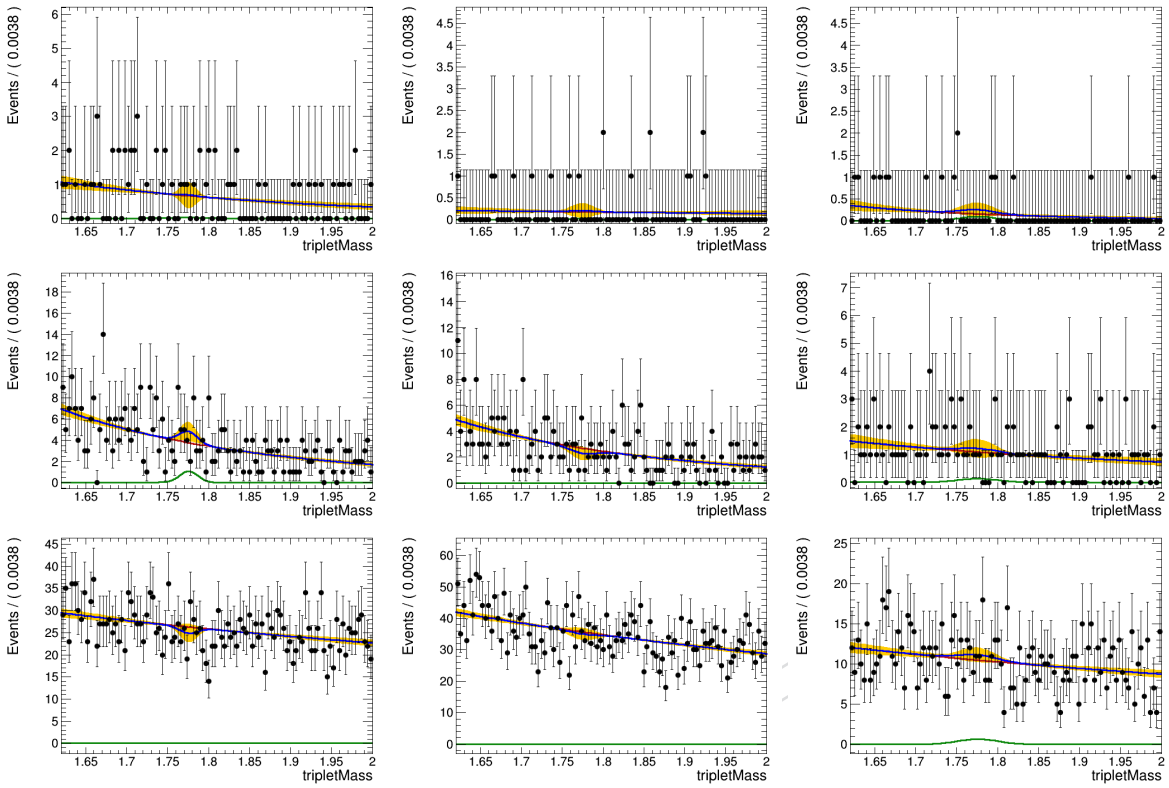


Figure 12: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 3Gl 2018, for the three different BDT-subcategories. The signal-plus-background post-fit model is shown.

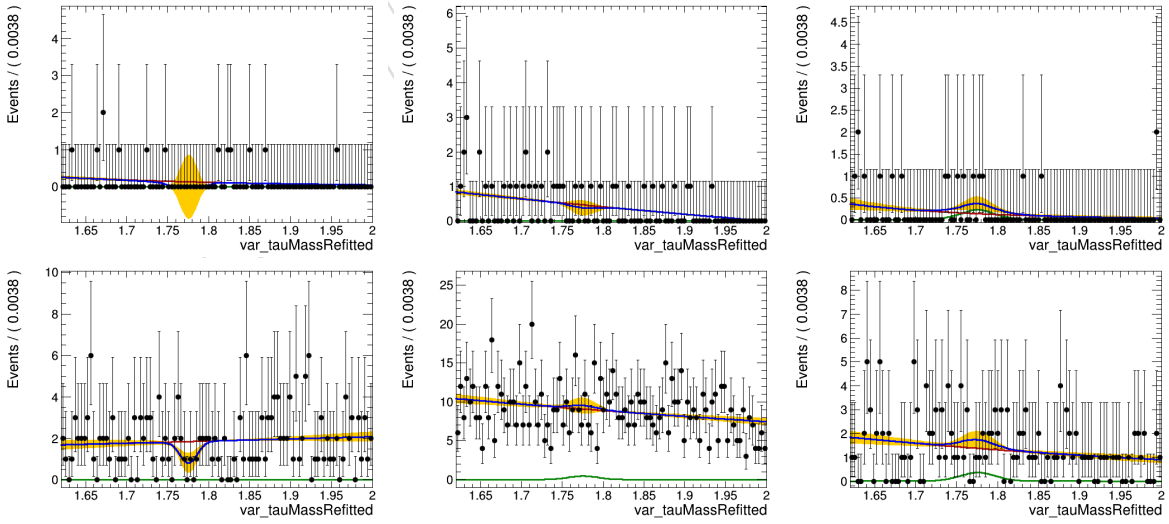


Figure 13: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 2Gl+Trk 2017, for the two different BDT-subcategories. The signal-plus-background post-fit model is shown.

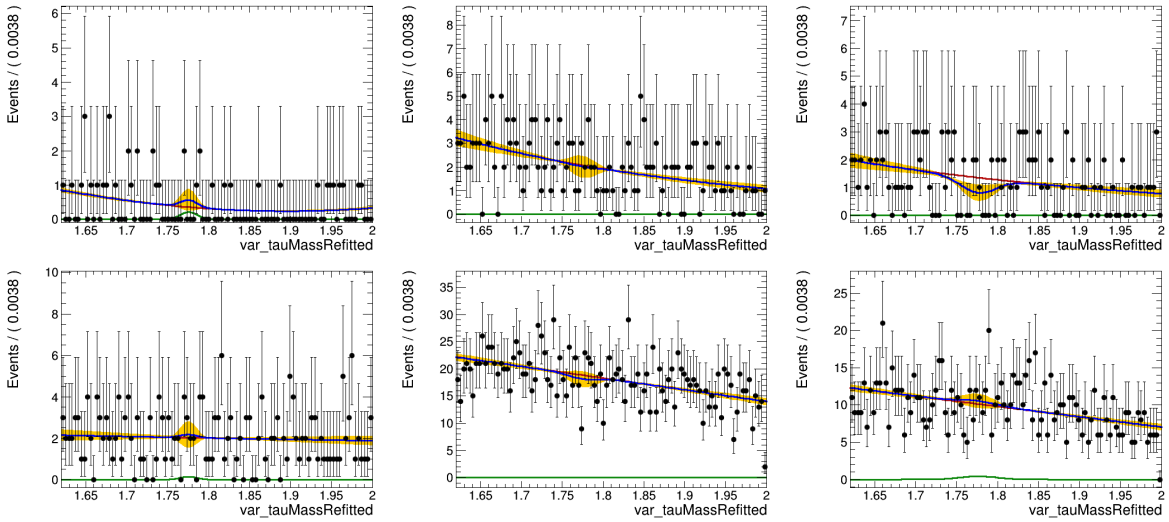


Figure 14: Three-muon mass distribution of evens passing the analysis selections for events in category A (left), B (central) and C (right) for 2GI+Trk 2018, for the two different BDT-subcategories. The signal-plus-background post-fit model is shown.

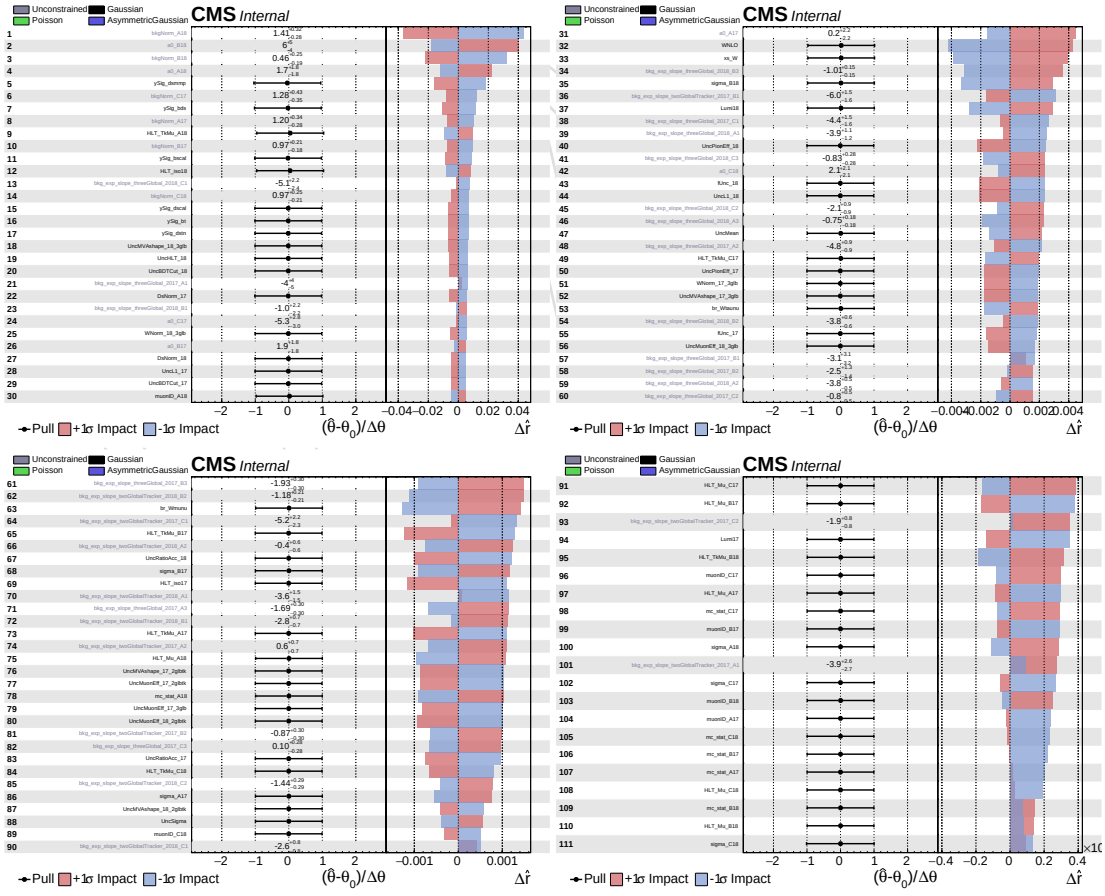


Figure 15: Pulls and impacts for the nuisance parameters of the analysis.

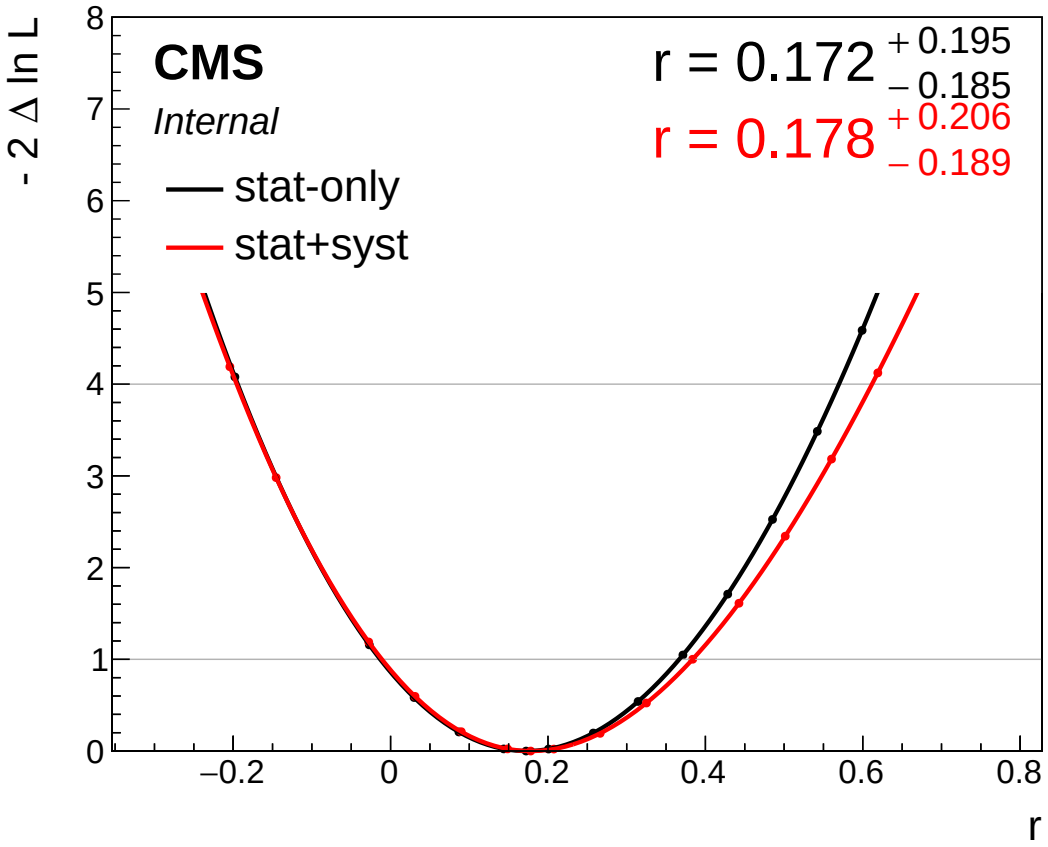


Figure 16: The plot shows likelihood scans in two scenarios: first all nuisances are floating and profiled, this is representative of the full model (black curve), second all constrained systematic uncertainties are frozen (red curve). The statistical uncertainty accounts for more than 95% of the total uncertainty (see uncertainties on fitted r). The scan has been run on 2017 and 2018 data of all the analysis categories.

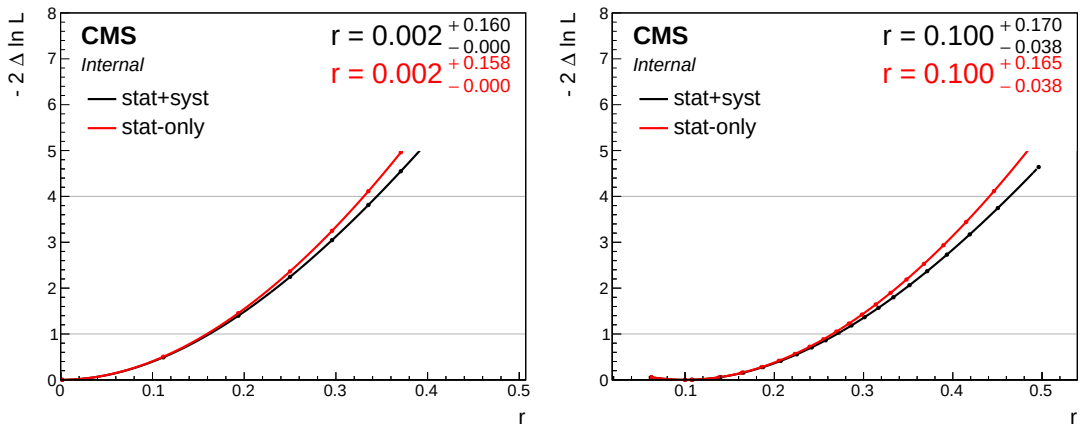


Figure 17: The plot shows likelihood scans in two scenarios: first all nuisances are floating and profiled, this is representative of the full model (black line), second all systematic uncertainties (except for the background normalizations) are frozen, and hence displaying only the statistical uncertainty (red line). The statistical uncertainty accounts for most of the total uncertainty. Two cases are shown: an asimov sample with r=0 (left) and an asimov sample with r=0.1 (BR($\tau \to 3\mu$)=10^{-8}) (right).

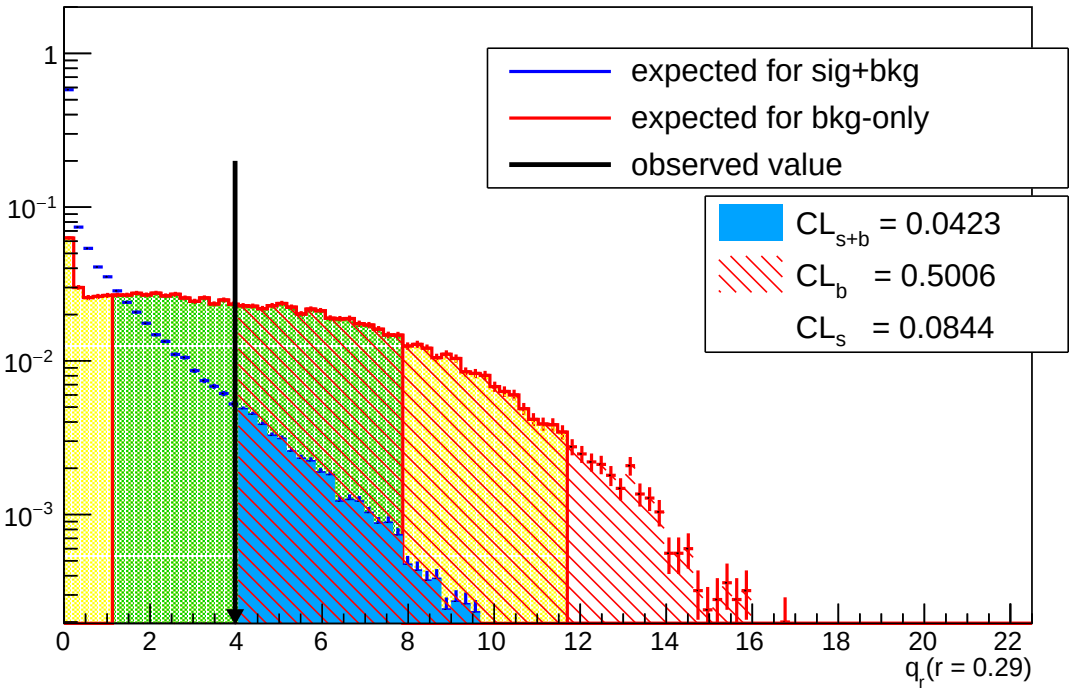


Figure 18: Toy based distributions of the test statistics q_r for the background only (red) and signal-plus-background (blue) hypotheses. The observed value is represented by a black arrow. The observed value is set to the median of the alternative distribution, as for the expected limits computation. The signal strength is arbitrarly set to 0.29 (BR($\tau \to 3\mu$) = $2.9 \cdot 10^{-8}$).

5 Full Run II combination

The results discussed in the previous chapters are combined with the result obtained on 2016 data, published in [3].
The combination is performed as described in Sec. 4. Table 9 lists the systematic uncertainties of the 2016 analysis; a detailed description is available in [4], [5] and [6].

Table 9: Summary of systematic uncertainties of the 2016 analysis.

Source of uncertainty	Process	Type	Barrel	Endcap	Correlated
W channel					
luminosity	signal	rate - logN	2.5 %	2.5 %	yes
$\sigma(pp \to W)$	signal	rate - logN	3.7 %	3.7 %	yes
$\mathcal{B}(W \to \tau\nu)$	signal	rate - logN	0.21%	0.21%	yes
$\mathcal{B}(W \to \mu\nu)$	signal	rate - logN	0.15%	0.15%	yes
finite MC statistics	signal	rate - logN	4.0 %	4.9 %	no
muon identification	signal	rate - logN	1.3 %	2.5 %	no
muon HLT efficiency	signal	rate - logN	0.7 %	2.3 %	no
track HLT efficiency	signal	rate - logN	0.8 %	0.8 %	no
track ineff. at HLT extrap.	signal	rate - logN	5.0 %	5.0 %	yes
EMTF ineff. correction	signal	rate - logN	4.0 %	30.0 %	yes
slope of the exponential	background	gaus param.	-	-	no
background normalisation	background	unconst. rate param.	-	-	no
HF channel					
D_s normalization	signal	rate - logN	10%		
$\mathcal{B}(D_s \to \tau\nu)$	signal	rate - logN	3%		
$\mathcal{B}(D_s \to \phi\pi \to \mu\mu\pi)$	signal	rate - logN	8%		
$\mathcal{B}(B \to D_s + ...)$	signal	rate - logN	5%		
$\mathcal{B}(B \to \tau + ...)$	signal	rate - logN	3%		
B/D ratio	signal	rate - logN	3%		
Signal from D^+ decay	signal	rate - logN	3%		
Signal from B_s decay	signal	rate - logN	4%		
Signal from trimuon trigger	signal	rate - logN	2%		
Ratio of acceptance $\mathcal{A}_{\text{sig}}/\mathcal{A}_{2\mu\pi}$	signal	rate - logN	1%		
Muon reconstruction efficiency	signal	rate - logN	1.5%		
π^+ reconstruction efficiency	signal	rate - logN	2.3%		
BDT cut efficiency	signal	rate - logN	5%		
Mass scale	signal	gaus param.	0.07%		
Mass resolution	signal	gaus param.	2.5%		

For the purpose of combination, the theoretical uncertainties (i.e. unc. on branching fractions) are taken as fully correlated across the years. The experimental uncertainties (i.e. normalisation, trigger, muon reconstruction) are computed separately for each year and thus are taken as uncorrelated.
Table 10 shows the summary of the systematic uncertainty for each channel of the $\tau \to 3\mu$ search on Run2 data.

Table 11 reports the upper limits of the combination of the full Run2 dataset.

Table 10: Systematic uncertainties (percent of the signal yield) for the $\tau \to 3\mu$ analysis on Run2 data. The systamatic uncertainties are reported for each category of the two channels for each year. Where one single value is given, it is the same for each category.

W channel											
source of uncertainty	**process**	**type**	**2016**		**2017**			**2018**			
			barrel	**endcap**	**A**	**B**	**C**	**A**	**B**	**C**	
luminosity	signal	rate - logN	2.5		2.3			2.5			
$\sigma(pp \to W)$	signal	rate - logN	3.7		3.7			3.7			
$\mathcal{B}(W \to \tau\nu)$	signal	rate - logN	0.21		1.4			1.4			
$\mathcal{B}(W \to \tau\nu)$	signal	rate - logN	0.15		1.8			1.8			
finite MC statistics	signal	rate - logN	4.0	4.9	1.0	0.7	1.1	0.7	0.5	1.2	
muon identification	signal	rate - logN	1.3	2.5	1.3	1.4	1.5	3.9	4.7	5.2	
muon HLT efficiency	signal	rate - logN	0.7	2.3	1.9	2.1	2.2	1.0	1.0	1.0	
HLT tracker muon efficiency	signal	rate - logN	-		11	10	15	8	8	9	
HLT track efficiency	signal	rate - logN	0.8	0.8	-			-			
track ineff. at HLT extrap	signal	rate - logN	5	5	-			-			
EMTF ineff. correction	signal	rate - logN	4	30	-			-			
HLT isolation correction	signal	rate - logN	-		12			7			
NLO reweighing	signal	rate - logN	-		4			4			
background normalization	background	parameter	unconstrained		unconstrained			unconstrained			
signal width	signal	parameter - gaussian	-		2	6	2	2	6	2	

Heavy Flavour channel							
source of uncertainty	**process**	**type**	**2016**	**2017**		**2018**	
				three-global	**two-global-one-tracker**	**three-global**	**two-global-one-tracker**
D_s normalisation	signal	rate - logN	10	6.2		4.0	
Uncertainty on W normalization	signal	rate - logN	-	3.0		5.0	
$\mathcal{B}(D_s \to \tau\nu)$	signal	rate - logN	3.0	3.0		3.0	
$\mathcal{B}(D_s \to \phi\pi \to \mu\mu\pi)$	signal	rate - logN	8.0	8.0		8.0	
$\mathcal{B}(B \to D_s + ...)$	signal	rate - logN	5.0	5.0		5.0	
$\mathcal{B}(B \to \tau + ...)$	signal	rate - logN	3.0	3.0		3.0	
Unc. on scaling D_s to include D^+	signal	rate - logN	3.0	3.0		3.0	
Unc. on scaling B^0 and $B^\pm$ to include B_s	signal	rate - logN	4.0	4.0		4.0	
B/D ratio	signal	rate - logN	3.0	2.0		2.0	
Signal from TripleMu L1	signal	rate - logN	2.0	5.0		2.0	
Ratio of acceptance $\mathcal{A}_{sig}/\mathcal{A}_{2\mu\pi}$	signal	rate - logN	1.0	1.0		1.0	
BDT selection efficiency	signal	rate - logN	5.0	20/5/0 (category-dependent)		15/7/0 (category-dependent)	
Muon reconstruction efficiency	signal	rate - logN	1.5	1.4	4.0	1.6	8.0
π^+ reconstruction efficiency	signal	rate - logN	2.3	2.2		2.1	
Shape correction on MVA muon ID	signal	rate - logN	-	3.0	4.0	6.0	4.0
Mass scale	signal	param. - gaussian	0.07	0.09		0.09	
Mass resolution	signal	param. - gaussian	2.5	2.0		2.0	
Index of bkg p.d.f	bkg	param. - discrete	-	-		-	

Table 11: The expected upper limit with its 68% confidence interval and the observed upper limit are reported for the Run2 analysis. The upper limits are computed at 90% and 95% of confidence level and reported in units of 10^{-8}. The best fitted value of the signal strength (in units of 10^{-8}) and its significance is also reported.

Category	**-1 sigma**	**expected**	**+1 sigma**	**observed limit**	**best fit**	**significance**
Run2 combination @ 90% CL	1.7	2.4	3.6	2.90	0.05	0.1
Run2 combination @ 95% CL	2.1	3.0	4.4	3.60		

6 Conclusion

The search for the $\tau \to 3\mu$ process at CMS has been carried out on Run2 data, corresponding to an integrated luminosity of 137 fb^{-1}. The analysis of the 2017 and 2018 datasets completes the preliminary analyis on 2016 data [3] (BR($\tau \to 3\mu$) $\leq$ 8.0 @ 90% of confidence level).
No evidence of signal was found and an upper limit was set on the branching fraction of the $\tau \to 3\mu$ process at 90% of confidence level [1].
The analysis was carried out using two different τ lepton production channel: the $W \to \tau\nu$ and the $D/B \to \tau X$ production channels. The two channels are combined to a single result.
The observed (expected) upper limit on the $\tau \to 3\mu$ branching fraction on 2017 and 2018 data alone is 3.1 (2.7) $\cdot 10^{-8}$ at 90% of confidence level. The observed value is compatible with the expected one within 1 sigma.
The observed (expected) upper limit on the $\tau \to 3\mu$ branching fraction on the full Run2 dataset is 2.9 (2.4) $\cdot 10^{-8}$ at 90% of confidence level. The observed (expected) upper limit on the $\tau \to 3\mu$ branching fraction on the full Run2 dataset is 3.6 (3.0) $\cdot 10^{-8}$ at 95% of confidence level.

7 Acknowledgments

[1] The 90% confidence level has been chosen to facilitate the comparison with the results already avalilable in the literature.

References

[1] V. C. et al, "Search for $\tau \to 3\mu$ decay with τ leptons produced in D and B decays using full Run II data", CMS Note CMS-AN-2020-102, 2020.

[2] S. M. L. Guzzi, "Search for Lepton Flavour Violating $\tau \to 3\mu$ decay in $W \to \tau\nu$ events with 2017 and 2018 CMS data", CMS Note CMS-AN-2020-127, 2020.

[3] CMS Collaboration, "Search for the lepton flavor violating decay $\tau \to 3\mu$ in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **01** (2021) 163, doi:10.1007/JHEP01(2021)163, arXiv:2007.05658.

[4] A. K. J. Wang, "Search for $\tau \to 3\mu$ decays using τ-leptons produced in D and B meson decays (2016 dataset)", CMS Note CMS-AN-2017-301, 2017.

[5] R. M. S. Fiorendi, L. Guzzi, "Search for Lepton Flavour Violating $\tau \to 3\mu$ decays in $W \to \tau\nu$ events with 2016 CMS data", CMS Note CMS-AN-2017-252, 2017.

[6] R. M. S. Fiorendi, "Combined search for LFV $\tau \to 3\mu$ decays using 2016 data", CMS Note CMS-AN-2020-024, 2020.

[7] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit.

[8] T. Junk, "Confidence level computation for combining searches with small statistics", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434** (1999), no. 2, 435–443.

[9] A. L. Read, "Modified frequentist analysis of search results (the CL_s method)", doi:10.5170/CERN-2000-005.81.