Contents

Chapter 1

Searches for New Physics in $\tau^+\tau^-$ Final States

The $\tau^+\tau^-$ final states are a powerful tool to search for new physics at collider experiments. As the heaviest lepton, they are sensitive to resonant production of new neutral particles where the couplings have mass hierarchy. They are also sensitive to non-resonant effects from new physics mediators. This chapter will detail the searches for two such areas of new physics: additional Higgs bosons and vector leptoquarks. These searches are split up into three sections:

i) A model independent search for single narrow spin-0 resonance, ϕ, produced via gluon fusion ($gg\phi$) or in association with a bottom quark ($bb\phi$). The SM Higgs boson is treated as a background. The Yukawa couplings that contribute to the gluon fusion loop are set to SM values.

ii) A search for the MSSM Higgs sector, in a number of benchmark scenarios. The benchmark scenarios are defined in Section 1.1.1. The production of SM Higgs boson is also used to constrain the available phase space.

iii) A search for the t-channel exchange of a U_1 vector leptoquark. Two scenarios are taken, based of the best fit to the b anomalies. These scenarios are detailed in Section 1.1.2.

These searches are performed with the full run-2 dataset (138 fb^{-1}) collected by the CMS experiment. The search for additional Higgs bosons had previously been performed with data collected in 2016 (39 fb^{-1}) and results were consistent with the SM background prediction.

1.1 Signal Modelling

1.1.1 Additional Higgs Bosons

Extended Higgs sectors, such as that of the MSSM, can be probed by direct searches for the additional bosons and further precise measurements of the Standard Model Higgs boson. This search for an extended Higgs sector is motivated by Type II 2HDMs, such as the MSSM. In these models $\tan\beta$ enhances couplings of additional Higgs bosons to bottom-like quarks and leptons, whilst top-like couplings are suppressed. This narrows down the most important production modes of the Higgs boson into two categories: Gluon fusion and production in association with a bottom quark. Examples of these are shown in Figure 1.1.

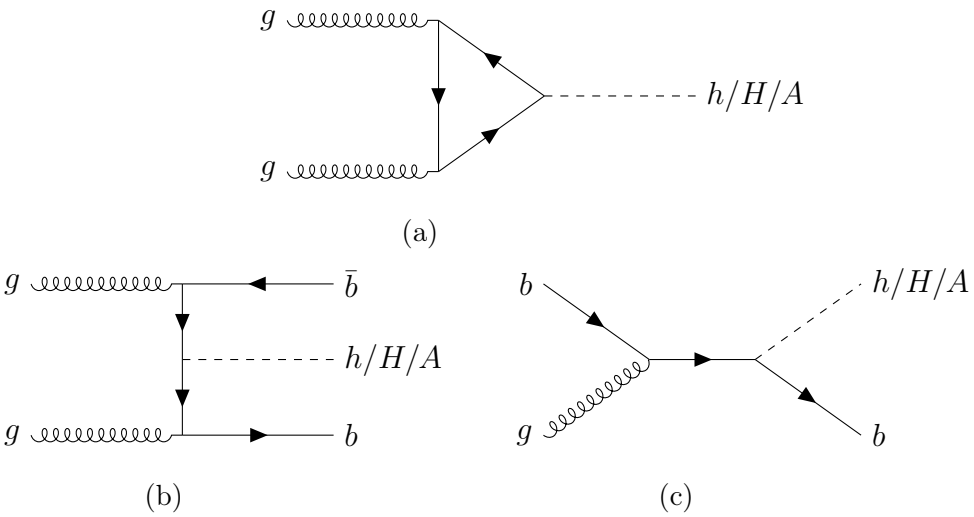


Figure 1.1: Diagram (a) shows the production of neutral Higgs bosons from gluon fusion. The dominant loop contributions to this diagrams are from top-only, bottom-only and top-bottom interference. Diagrams (b) and (c) show production in association with b quarks.

With the $\tan\beta$ enhancement, the decays of additional Higgs bosons to tau leptons and bottom quarks are most likely. Tau leptons are identified with a higher purity than bottom quarks a the CMS detector. It is also easier to separate $\tau^+\tau^-$ from the large QCD multijet background produced from the high energy proton-proton collisions. This hypothesis was tested with the 2016 dataset and although no deviations were observed, the strongest limits on the MSSM phase space was placed by the $\tau^+\tau^-$ final states.

For this analysis, the production of additional Higgs bosons over a mass range of 60 GeV to 3.5 TeV are generated. Gluon fusion is simulated at NLO precision

using the 2HDM implementation of POWHEG 2.0. The kinematic properties are highly dependent on the contributions to the loop, which vary dependent on the specific signal model. To account for the different loop contributions at the NLO plus parton shower prediction, weights based off the p_T spectra are calculated to split the contributions from the t quark only, b quark only, and tb-interference. Once individual templates have been determined for each contribution to the loop, the 2HDM samples can be scaled to the correct specific MSSM scenario prediction by the following formula.

$$\frac{d\sigma_{\text{MSSM}}}{dp_T} = \left(\frac{Y_{t,\text{MSSM}}}{Y_{t,\text{2HDM}}}\right)^2 \frac{d\sigma^t_{\text{2HDM}}}{dp_T}(Q_t) + \left(\frac{Y_{b,\text{MSSM}}}{Y_{b,\text{2HDM}}}\right)^2 \frac{d\sigma^b_{\text{2HDM}}}{dp_T}(Q_b) + \left(\frac{Y_{t,\text{MSSM}}}{Y_{t,\text{2HDM}}}\frac{Y_{b,\text{MSSM}}}{Y_{b,\text{2HDM}}}\right)\left\{\frac{d\sigma^{t+b}_{\text{2HDM}}}{dp_T}(Q_{tb}) - \frac{d\sigma^t_{\text{2HDM}}}{dp_T}(Q_{tb}) - \frac{d\sigma^b_{\text{2HDM}}}{dp_T}(Q_{tb})\right\},$$

where Q_i are resummation scales that depend on the mass of the additional Higgs boson. Further contributions from any Supersymmetric partners have been checked and account for less than a few percent and so are neglected. This is also done separately for the scalar and pseudoscalar additional Higgs bosons, as the p_T distributions can differ. The MSSM benchmark scenarios considered are detailed in Ref. []. The scenarios provide the relative Yukawa couplings (to calculate the cross sections) and branching fractions of the MSSM Higgs bosons. An example of the changes to gluon fusion production, in the MSSM M_h^{125} scenario with $m_A = 1600$ GeV and $\tan\beta$ varying is shown in Figure 1.2. The distributions peak at a higher p_T for the top quark loop, therefore at smaller $\tan\beta$, where the top quark contribution is dominant, an additional Higgs boson would be more boosted.

Production in association with bottom quarks is simulated at NLO precision using the corresponding POWHEG 2.0 implementation in the four-flavour scheme. All additional Higgs boson signal generation is performed using the parton distribution function (PDF) NNPDF3.1. Tau lepton decay, parton showering and hadronisation are all modelled with the PYTHIA event generator where the PU profile is matched to data. All events generated are passed through a GEANT4-based simulation of the CMS detector and reconstructed in the same way as data.

The model dependent search for the MSSM also looks to find differences from the observed SM Higgs boson and the predicted MSSM SM-like Higgs boson. In each MSSM benchmark scenario, an uncertainty of ± 3 GeV is given on the prediction for the SM Higgs boson mass. This uncertainty is to reflect the contribution from any

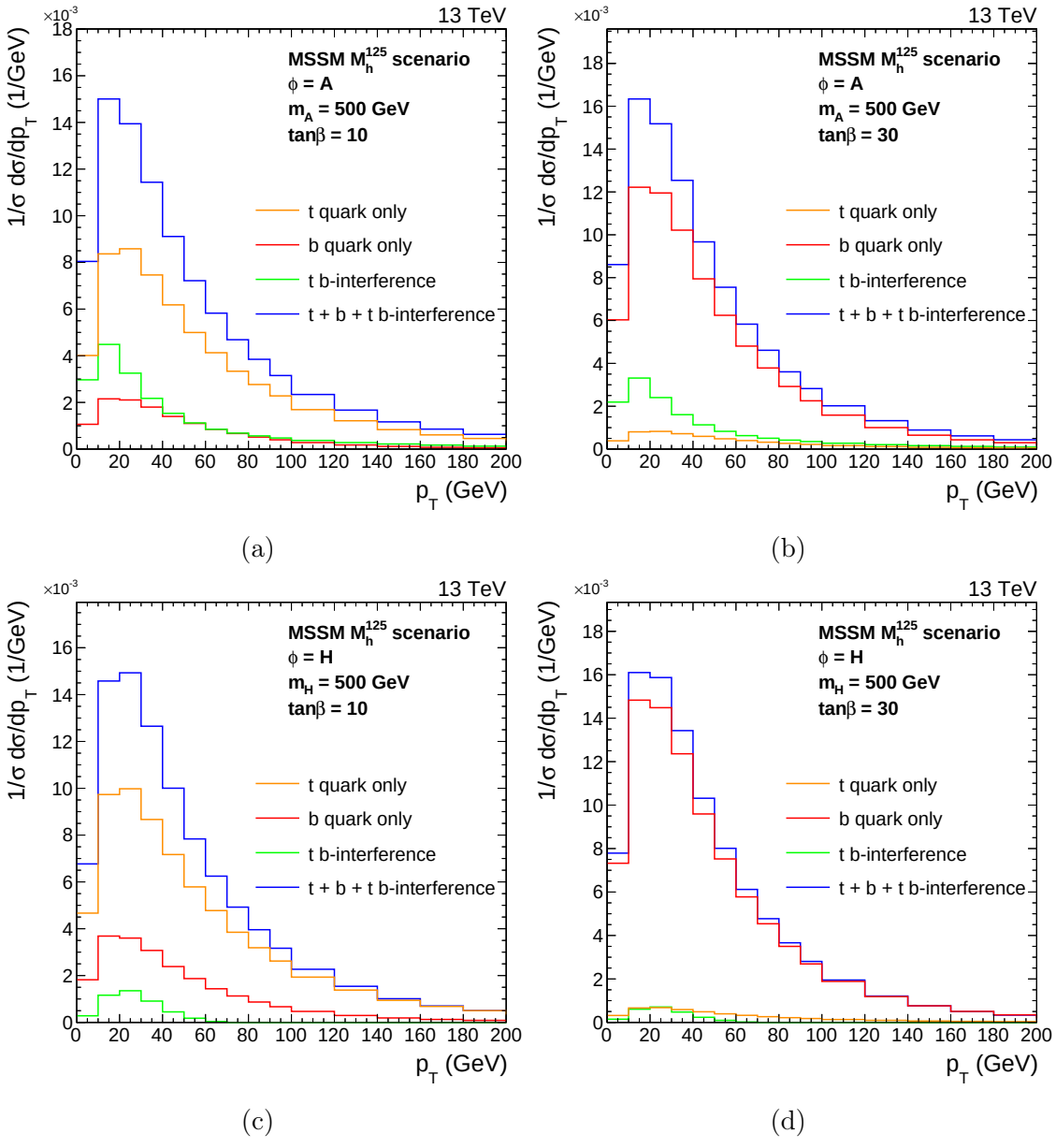


Figure 1.2: p_T density distributions of the A (top) and H (bottom) boson, with contributions to the gluon fusion loop displayed individually and summed. These are shown for $\tan\beta$ values of 10 (left) and 30 (right) where $m_A = 500$ GeV in the MSSM M_h^{125} scenario.

unknown higher-order corrections. The value of the mass is allowed to vary within this window, however the Yukawa couplings are rescaled the observed mass.

1.1.2 Vector Leptoquarks

The best fit in the vector leptoquark phase space to the B anomalies yielded large bottom quark and tau lepton couplings to the U_1 particle. The possible production modes of a $\tau^+\tau^-$ final state are shown in Figure 1.3.

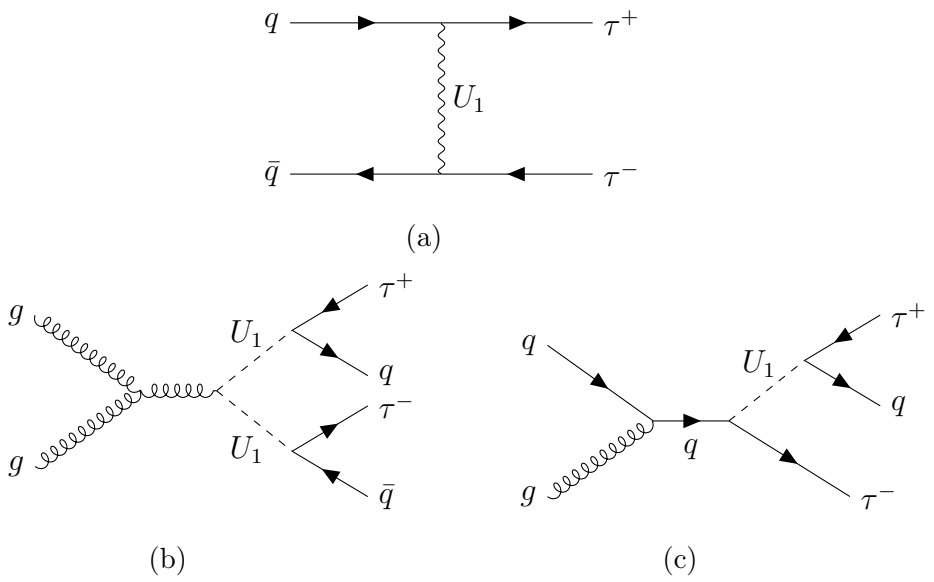


Figure 1.3: Feynman diagrams showing the contribution from U_1 vector leptoquarks to the final state with a pair of oppositely charged tau leptons. Diagram (a) shows t-channel, (b) pair and (c) single production of a vector leptoquark.

Pair and single production of a vector leptoquark is dependent on its strong coupling, which is highly model dependent. For large mass, m_U, the probability of producing an on-shell U_1 singlet or pair is heavily suppressed due to the momentum of the initial partons. These production processes are not discussed further in this search. Further studies have been used to search for single and pair production at the CMS experiment and no statistically significance derivation was observed.

The t-channel process contain two vertices with a U_1 vector leptoquark, a quark and a tau lepton, and hence the cross section will scale with g_U^4. From the best fit to B anomalies this vertex will be dominated by the b quark and hence the initial state will be mostly from $b\bar{b}$, with sub-dominant contributions from $b\bar{s}$, $s\bar{b}$ and $s\bar{s}$. Although there are no additional b quarks in the final state in the LO process, initial state radiation can lead to additional b quarks in the final state. In this search

the two scenarios discussed in Section **??** are considered. The only non negligible parameter for $\tau^+\tau^-$ final states from the fit in the m_U-g_U phase space is the $\beta_L^{s\tau}$ parameter. This is set to the best fit value.

The signal process of the U_1 t-channel exchange is simulated in the five-flavour scheme (5FS) at LO precision using the MadGraph5_aMC@NLO event generator, v2.6.5. Events are generated with one or fewer outgoing partons from the matrix element and the MLM prescription is then used for matching, with a scale set to 40 GeV. Negligible dependence of the U_1 decay width (Λ) is observed, for simulation this is chosen to approximately match the value predicted by the B anomaly fit. Samples with a mass between 1 and 5 TeV at $g_U = 1$ are generated.

The interference between the U_1 signal and $Z/\gamma^* \to \tau\tau$ production was checked. A large destructive affect is observed, with the magnitude dependent on g_U. To account for this, separate samples are produced for this interference, generated in the same way as the t-channel exchange. The interference samples are then split into two with a di-tau mass split in order to have a sufficient number of events in the high di-tau mass regions. The cross section of these interference samples scale with g_U^2. Examples of the generator level di-tau mass distributions are shown in Figure 1.4.

The t-channel signal produces a broad distribution in $m_{\tau\tau}$ due to its non-resonant nature. The interference is mostly a destructive effect (except for at small $m_{\tau\tau}$), with the yield becoming less negative at higher $m_{\tau\tau}$. The interference peaks negatively between 100 and 200 GeV and in this region the combined yield can be negative. Due to the difference in scaling of the two effects, at small g_U the interference is more dominant than the signal and hence the yield of the combined result is reduced.

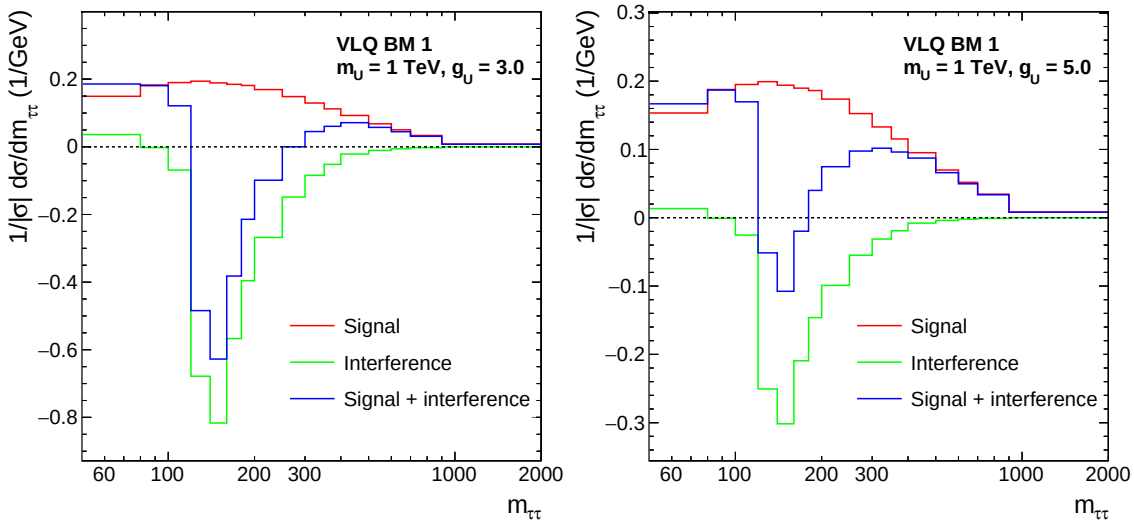
1.2 Event Selection

1.3 Signal Extraction

The variable D_ζ is defined as

$$D_\zeta = p_\zeta^{\text{miss}} - 0.85 p_\zeta^{\text{vis}}; \qquad p_\zeta^{\text{miss}} = \vec{p}_{\text{T}}^{\,\text{miss}} \cdot \hat{\zeta}; \qquad p_\zeta^{\text{vis}} = (\vec{p}_{\text{T}}^{\,e} + \vec{p}_{\text{T}}^{\,\mu}) \cdot \hat{\zeta} \tag{1.1}$$

where $\vec{p}_{\text{T}}^{\,e/\mu)}$ corresponds to the transverse momentum vector of the electron (muon) and $\hat{\zeta}$ to the bisectional direction between the electron and the muon in the trans-



(a) (b)

Figure 1.4: The generator level $m_{\tau\tau}$ density distributions of the t-channel vector leptoquark signal and the interference with Drell-Yan. This is shown in the VLQ BM 1 scenario for a leptoquark of mass 1 TeV for coupling strengths of $g_U = 3$ (a) and $g_U = 5$ (b).

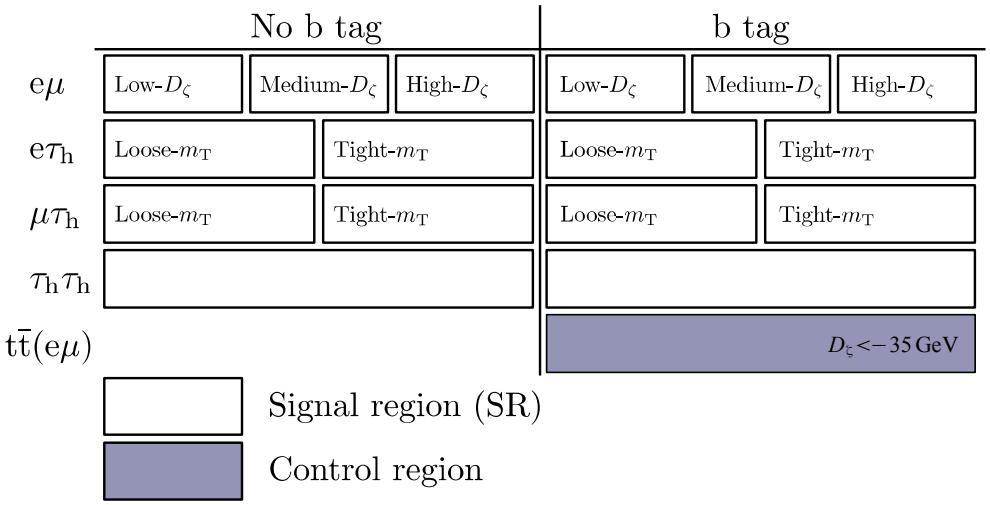


Figure 1.5: .

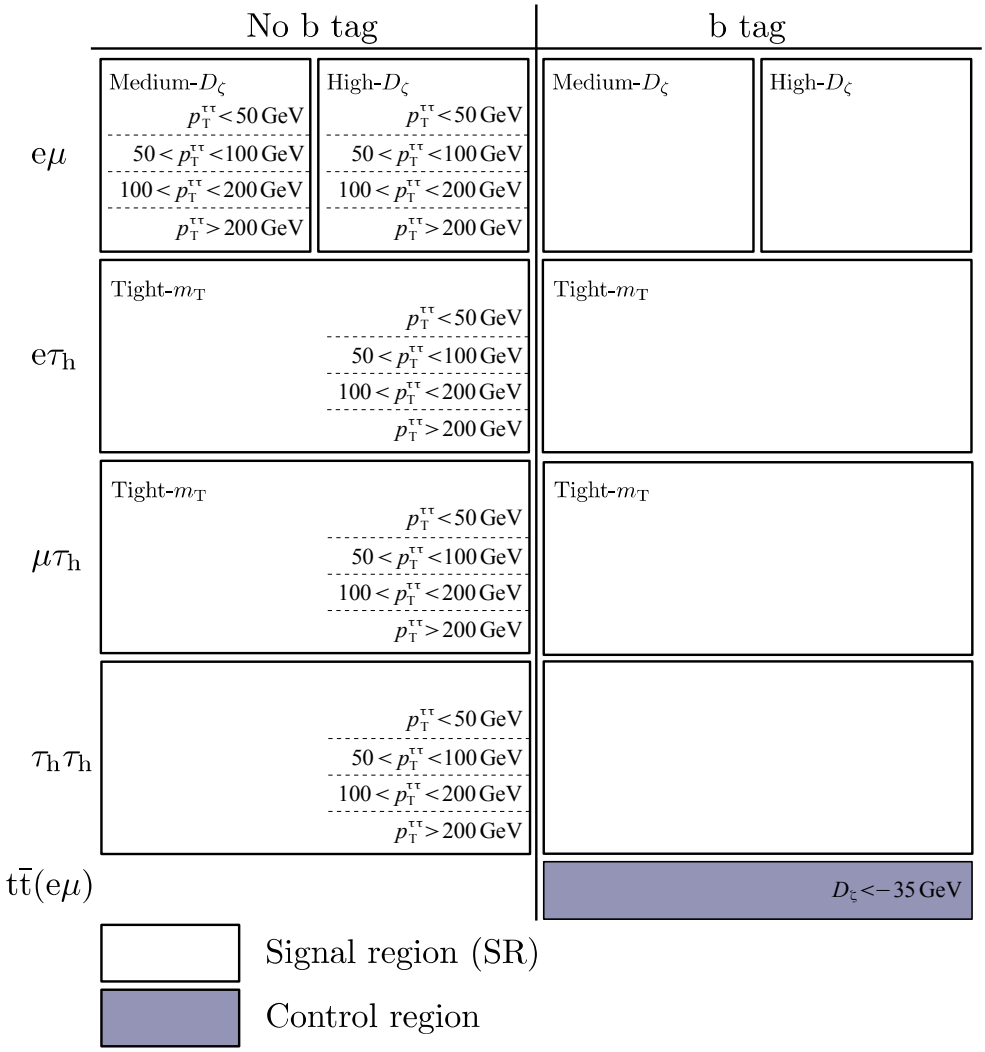

No b tag
b tag
eμ
Medium-D_ζ
$p_T^{\tau\tau} < 50$ GeV
$50 < p_T^{\tau\tau} < 100$ GeV
$100 < p_T^{\tau\tau} < 200$ GeV
$p_T^{\tau\tau} > 200$ GeV
High-D_ζ
$p_T^{\tau\tau} < 50$ GeV
$50 < p_T^{\tau\tau} < 100$ GeV
$100 < p_T^{\tau\tau} < 200$ GeV
$p_T^{\tau\tau} > 200$ GeV
Medium-D_ζ
High-D_ζ
eτh
Tight-m_T
$p_T^{\tau\tau} < 50$ GeV
$50 < p_T^{\tau\tau} < 100$ GeV
$100 < p_T^{\tau\tau} < 200$ GeV
$p_T^{\tau\tau} > 200$ GeV
Tight-m_T
μτh
Tight-m_T
$p_T^{\tau\tau} < 50$ GeV
$50 < p_T^{\tau\tau} < 100$ GeV
$100 < p_T^{\tau\tau} < 200$ GeV
$p_T^{\tau\tau} > 200$ GeV
Tight-m_T
τhτh
$p_T^{\tau\tau} < 50$ GeV
$50 < p_T^{\tau\tau} < 100$ GeV
$100 < p_T^{\tau\tau} < 200$ GeV
$p_T^{\tau\tau} > 200$ GeV
tt̄(eμ)
$D_\zeta < -35$ GeV
Signal region (SR)
Control region

Figure 1.6: .

Channel	Branching Fraction
$\tau_h\tau_h$	42.0%
$e\tau_h$	23.1%
$\mu\tau_h$	22.6%
$e\mu$	6.2%
ee	3.2%
$\mu\mu$	3.0%

Table 1.1: Branching fractions of the decays of two tau leptons.

verse plane [?].
In the $e\mu$ final state, three sub-categories are defined by

- `Low-`D_ζ: $-35 \leq D_\zeta[\text{GeV}] < -10$;
- `Medium-`D_ζ: $-10 \leq D_\zeta[\text{GeV}] < 30$;
- `High-`D_ζ: $D_\zeta[\text{GeV}] \geq 30$.

In this way sub-categories with different signal purities and $t\bar{t}$ fractions can be exploited during the statistical inference for the signal. The expected signal, for all masses tested, is mostly located in the `Medium-`D_ζ sub-category.
The transverse mass in the $e\tau_h$ and $\mu\tau_h$ final states is defined as:

$$m_{\mathrm{T}}^{e/\mu} = \sqrt{2\, p_T^{e/\mu}\, E_T\, (1 - \cos\Delta\phi)} \tag{1.2}$$

where $p_T^{e/\mu}$ refers to the p_T of the electron (muon) and $\Delta\phi$ to the azimuthal angle in the transverse plane between the electron/muon and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$.
The corresponding categories in $e\tau_h$ and $\mu\tau_h$ channels are defined by

- `Tight-`m_{T}: $m_{\mathrm{T}}^{e/\mu}[\text{GeV}] < 40$;
- `Loose-`m_{T}: $40 \leq m_{\mathrm{T}}^{e/\mu}[\text{GeV}] < 70$.

The bulk of the signal events, particularly for low mass hypotheses, lies in the `Tight-`m_{T} sub-category. The `Loose-`m_{T} category has been added to increase the signal acceptance for mass hypotheses of $m_{A,H} > 700$ GeV.

$$m_T^{\mathrm{tot}} = \sqrt{m_T^2(p_T^{\tau_1}, p_T^{\tau_2}) + m_T^2(p_T^{\tau_1}, p_T^{\tau_{\mathrm{miss}}}) + m_T^2(p_T^{\tau_2}, p_T^{\tau_{\mathrm{miss}}})} \tag{1.3}$$

1.4 Background Modelling Overview

The analysis considers several backgrounds including Drell-Yan, $t\bar{t}$, W+jets, QCD, di-boson, single-top, and electroweak W and Z bosons production. These are split into a five categories:

i) Events containing only genuine tau leptons.

ii) Events with a jet misidentified as a hadronic tau (jet$\rightarrow \tau_h$) in the $e\tau_h$, $\mu\tau_h$ or $\tau_h\tau_h$ channels.

iii) Events with jets faking both light leptons (jet$\rightarrow l$) in the $e\mu$ channel.

iv) Events from $t\bar{t}$ with a prompt light lepton (e or μ not from a τ decay) and the other object (if there are not two prompt light lepton) is from a genuine tau leptons.

v) Other events. This is a small contribution and hence why it is grouped.

- Non $t\bar{t}$ events with a prompt light lepton (e or μ not from a τ decay) and the other object (if there are not two prompt light lepton) is from a genuine tau leptons.
- Events with a light lepton faking a hadronic tau and the other object (if there are not two light leptons faking a hadronic tau) are reconstructed as prompt light lepton or from genuine tau leptons.
- Events with a jet faking a light lepton and the other object is from genuine tau leptons in the $e\tau_h$, $\mu\tau_h$ or $\tau_h\tau_h$ channels.
- Events with one jet faking a light lepton and the other object from a prompt light lepton in the $e\mu$ channel.

Backgrounds from (i) consists of largely $Z/\gamma^* \rightarrow \tau\tau$ events but there are also smaller contributions from other processes. This background is modelled by a data-simulation hybrid method called the embedding method and this is described in detail in Section 1.6. Group (ii) is dominated by QCD, W + jets and $t\bar{t}$ events with a jet$\rightarrow \tau_h$ misidentification. This is modelled from data by the fake factor method (F_F) and is explained in Section 1.7. Group (iii) is modelled from data to describe QCD multijet contribution to the background in the $e\mu$ channel. The method to obtain this background is described in Section 1.5. The data driven background estimations for (i), (ii) and (iii) contribute >98% of all expected background events in the $\tau_h\tau_h$ channel, >90% in $e\tau_h$ and $\mu\tau_h$ channels and >50% in the $e\mu$ channel.

The final groups, (iv) and (v), are modelled with MC. The $t\bar{t}$ process is separated due to its large contribution to the phase space where a b jet is required.

The W + jets and $Z \to ll$ processes are simulated at leading order (LO) using the MadGraph5_aMC@NLO 2.2.2 (2.4.2) event generator [?, ?] for the simulation of the data taken in 2016 (2017–2018). To increase the number of simulated events in regions of high signal purity, supplementary samples are generated with up to four outgoing partons in the hard interaction. For diboson production, MadGraph5_aMC@NLO is used at next-to-LO (NLO) precision. In each case, the FxFx [?] (MLM [?]) prescription is used to match the NLO (LO) matrix element calculation with the parton shower model. For $t\bar{t}$ [?] and (t-channel) single top quark production [?], samples are generated at NLO precision using POWHEG 2.0 [?, ?, ?, ?]. The POWHEG version 1.0 at NLO precision is used for single top quark production in association with a W boson (tw channel) [?].
When compared with data, W + jets, $Z \to ll$, $t\bar{t}$, and single top quark events in the tW channel are normalised to their cross sections at next-to-NLO (NNLO) precision [?, ?, ?]. Single top quark (t-channel) and diboson events are normalized to their cross sections at NLO precision or higher [?, ?, ?].

1.5 QCD Estimation in the $e\mu$ Channel

1.6 Embedding Method

Validation plots of it working

1.7 Fake Factor Method

Backgrounds in which a jet fakes a τ_h can be difficult to model using MC due to the poor description of the jet $\to \tau_h$ fake rate in simulation. In addition, the small probability of a jet being misidentified as a τ_h necessitates the production of high statistics MC samples at a significant computational expense. These shortcomings motivate the use of data-driven estimates for these processes. One such procedure is the fake factor (F_F) method.

The F_F method utilises regions in the data to model the jet $\to \tau_h$ background. Firstly, the determination regions (DR), which are jet $\to \tau_h$ enriched control regions orthogonal to the signal region (SR). It is used to calculate F_F by taking

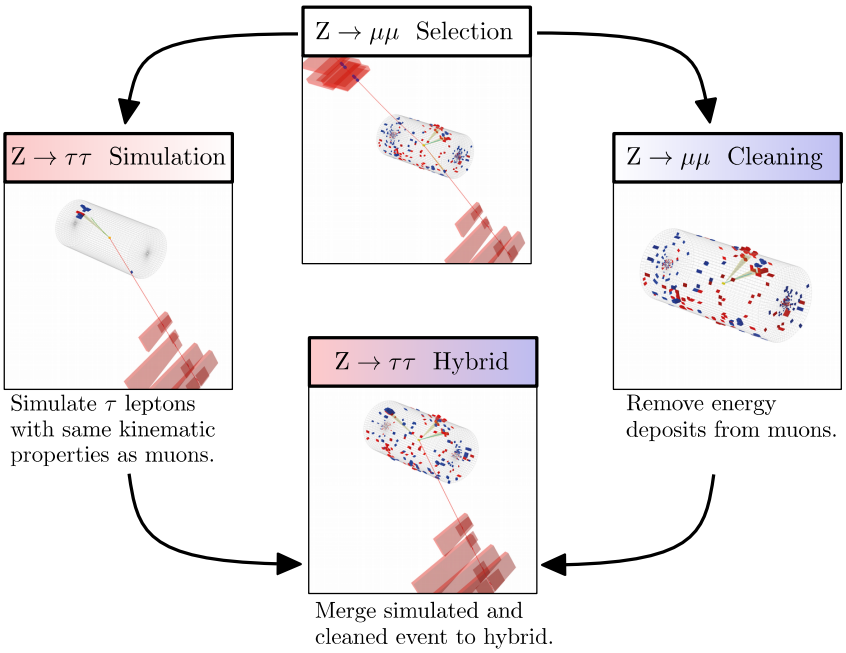


Figure 1.7: Schematic of the embedding method to model genuine di-tau backgrounds from di-muon events in data.

the ratio of number of jet fake events that pass the nominal hadronic tau ID requirement ($N(\texttt{Nominal})$), to the number of jet fake events that fail the nominal hadronic tau ID but pass a looser alternative hadronic tau ID requirement ($N(\texttt{Alternative \&\& !Nominal})$), as shown in Equation 1.4.

$$F_F = \frac{N(\texttt{Nominal})}{N(\texttt{Alternative \&\& !Nominal})}. \tag{1.4}$$

In the remaining text this numerator and denominator are referred to as the pass and fail regions. The derivation of this ratio is done differentially with respect to key parameters that differ in the two regions. Once F_F have been derived it is common to calculate corrections in other sideband regions (a region orthogonal to the signal region) and combine F_F measured from different processes. Finally, the F_F are applied to the application region (AR). This is defined as the SR but with the criteria that the jet fakes fail the nominal hadronic tau ID but pass the looser alternative tau ID requirement. This now models the background from jet $\rightarrow \tau_h$ events in SR.

The following Sections 1.7.1–1.7.5 detail the complexities of how this method is applied to this analysis. For these searches the nominal hadronc tau ID used is the `Medium` DeepTau Vs Jets ID working point and the alternative hadronic tau ID used is the `VVVLoose` DeepTau Vs Jets ID working point.

1.7.1 Determination Regions

The fake factors are measured separately in each year of data taking period (2016, 2017, 2018), in each channel containing hadronic taus ($e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$) and in enriched regions of dominant processes that contribute jet $\rightarrow \tau_h$ events. In the $e\tau_h$ and $\mu\tau_h$ channels F_F are measured for three processes: QCD, W + Jets and $t\bar{t}$. In the $\tau_h\tau_h$ channel F_F are measured only for the dominant QCD process. The QCD process is assumed to produce two jet fakes and so the fake factors is chosen to be calculated from leading p_T hadronic tau candidate only. Section 1.7.5 discusses how single jet fake events in the $\tau_h\tau_h$ channel are modelled.

Each separate measurement region is split into three sideband regions based off two cuts that surround the signal region. These regions are named the `Determination Region` (C), `Alternative Determination Region` (D) and `Correction Region` (B) and are schematically shown in Figure 1.8.

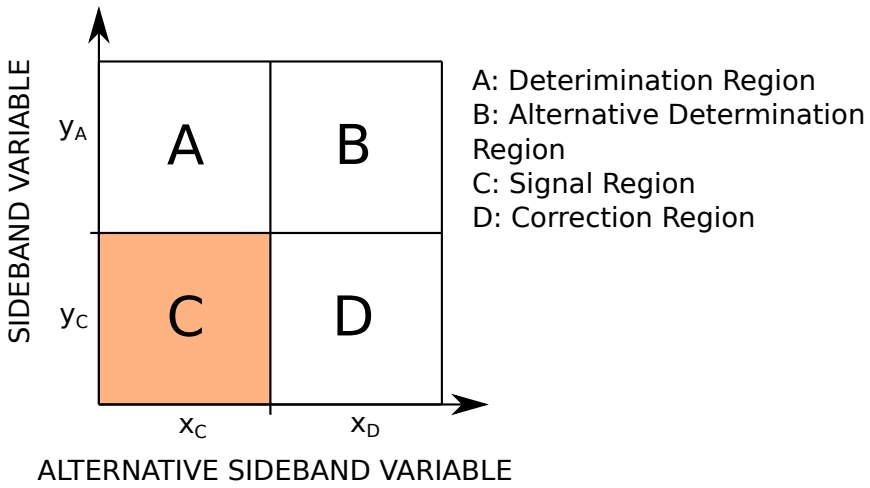


Figure 1.8: Schematic of the regions used for fake factor derivation.

Region A is used to measure and fit fake factors. Region B is an alternative region used to measure and fit fake factor to account for the difference in fake factors between A and C. These alternative fake factors are applied to the fail region in D and corrections are calculated comparing it to the pass region in D. The total fake factor per measurement region is calculated as the fake factors derived in region A multiplied by the correction calculated from region B to D.

The selection for x_C, x_D, y_C and y_A, as defined in Figure 1.8, in each separate

measurement region are shown below. These are chosen to balance the number of events and the purity of each background in the region.

i) $\tau_h\tau_h$ QCD
y_C: The τ_h candidates are required to have the opposite sign.
y_A: The τ_h candidates are required to have the same sign.
x_C: The subleading tau passes the `Medium` DeepTau Vs Jets working point.
x_D: The subleading tau fails the `VVLoose` DeepTau Vs Jets working point but passes the `VVVLoose` working point.

ii) $e\tau_h$ and $\mu\tau_h$ QCD
y_C: The e/μ and τ_h candidates are required to have the opposite sign.
y_A: The e/μ and τ_h candidates are required to have the same sign and the e/μ to have $I_{\text{rel}} > 0.05$.
x_C: The e/μ candidate is required to have $I_{\text{rel}} < 0.15$.
x_D: The e/μ candidate is required to have $0.25 < I_{\text{rel}} < 0.5$.

iii) $e\tau_h$ and $\mu\tau_h$ W + Jets
y_C: The m_T between the e/μ and the MET < 70 GeV.
y_A: The m_T between the e/μ and the MET > 70 GeV and no b jets in the event.
x_C: Data.
x_D: W + Jets MC.

iv) $e\tau_h$ and $\mu\tau_h$ $t\bar{t}$
y_C: Data.
y_A: MC ($t\bar{t}$ in B and W + Jets D).
x_C: $m_T < 70$ GeV.
x_D: $m_T > 70$ GeV and no b jets.

In the $\mu\tau_h$ and $e\tau_h$ channels QCD and W + Jets jet fake events are in general the most significant and contribute with approximately equal weights. $t\bar{t}$ inclusively is small but becomes more significant when searching for events with a b jet. The additional $I_{\text{rel}} > 0.05$ requirement in these channels for QCD is to reduce processes producing genuine leptons and the $N_{\text{b-jets}} = 0$ requirement for W + Jets is to reduce $t\bar{t}$ contamination. It is not possible to define a DR that is sufficiently pure in $t\bar{t}$ events to make a reasonable measurement of $F_F^{t\bar{t}}$ from data. Therefore $F_F^{t\bar{t}}$ are derived from MC. A comparison of the $F_F^{\text{W+jets}}$ measured in data and MC shows only $\sim$ 10–20% differences in the fake rates in data and MC. This observation coupled

with the fact that the $t\bar{t}$ contribution is small compared to the other processes means that any bias introduced by using $F_F^{t\bar{t}}$ measured in MC is small compared to the uncertainties on the fake factors, discussed in Section **??**.

1.7.2 Parametrisation

The F_F^i take into account dependencies on N_{jets} via the analysis tailed variable $N_{\text{pre b-jets}}$, the p_T of the τ_h candidate ($p_T^{\tau_h}$) and the p_T of the jet matched in ΔR to the τ_h (p_T^{jet}). $N_{\text{pre b-jets}}$ is defined to map the dependence of F_F^i on N_{jets} and describe the categorising variable $N_{\text{b-jets}}$ well. It is the number of jets in the event with $|\eta| < 2.4$ and $p_T > 20$. These are the same η and p_T thresholds required for a b-jet. They are also calculated separately for the three channels where a jet can fake a hadronic tau, $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$. The F_F^i are then measured separately for the dominant jet $\rightarrow \tau_h$ processes i, which for the $e\tau_h$ and $\mu\tau_h$ channels includes QCD, W+jets and $t\bar{t}r$, and for the $\tau_h\tau_h$ channel includes only QCD. The F_F^i are finally corrected to account for extrapolations from the DR to the SR, and for missing variable dependencies (via non-closure corrections). The F_F^i measured for the different processes are combined into an overall factor, F_F, using

$$F_F = \sum_i f_i \cdot F_F^i, \tag{1.5}$$

where the factor f_i is defined as

$$f_i = \frac{N_{\text{AR}}^i}{\sum_j N_{\text{AR}}^j}, \tag{1.6}$$

which is the fraction of events with a jet $\rightarrow \tau_h$ originating from process i over the total number of jet $\rightarrow \tau_h$ events for all processes in the AR.

1.7.3 Corrections

1.7.4 Application Region Fractions

1.7.5 Applying Fake Factors

using leading tau subtracting off rest w fakes in tt

1.8 Uncertainty Model

bulleted list of uncertainties used in the analysis.

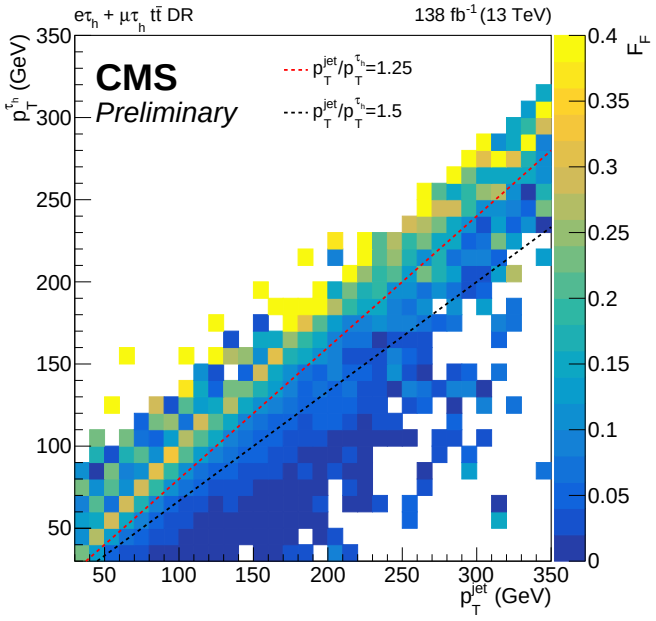
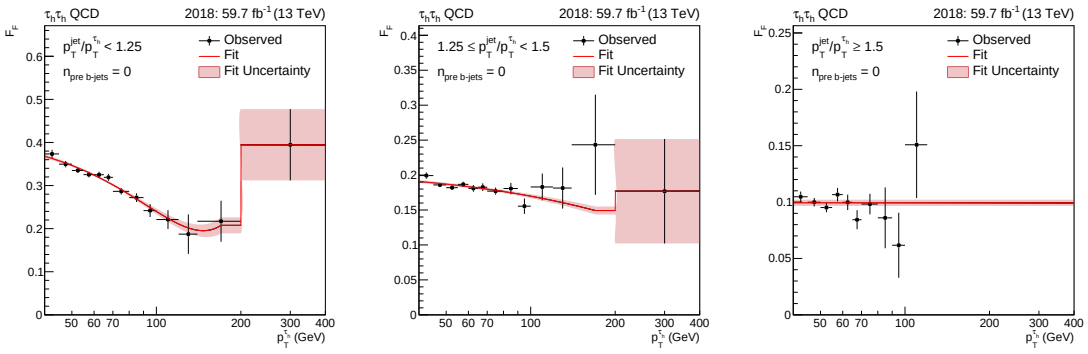
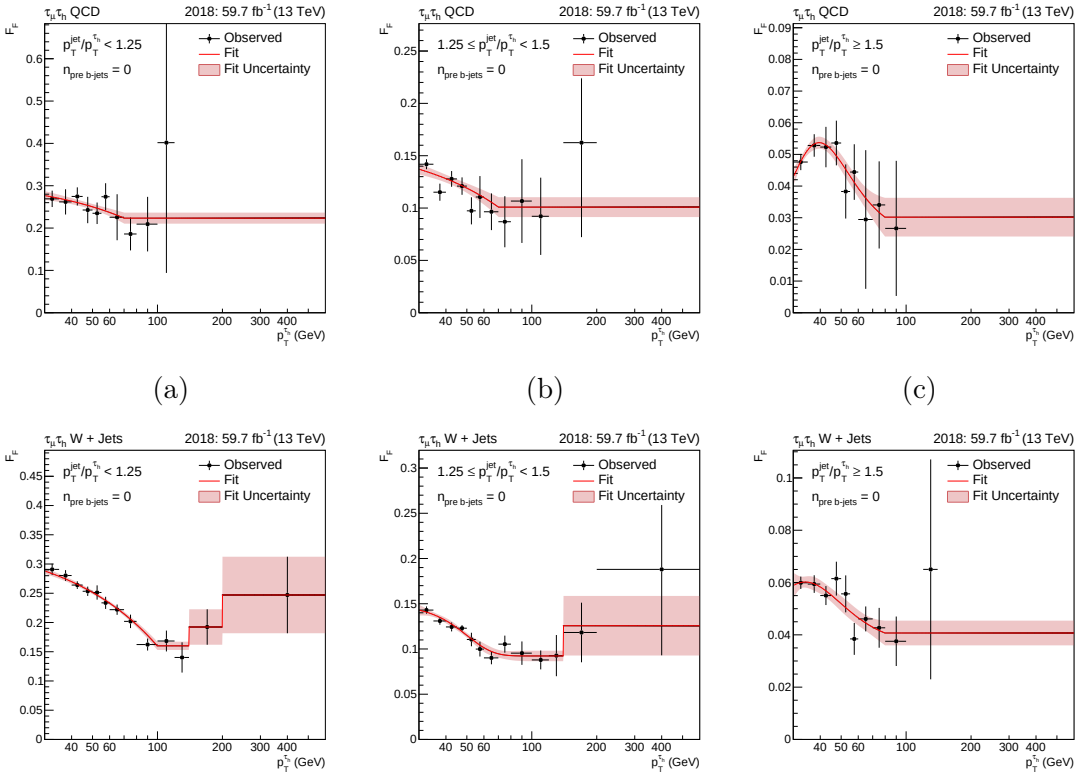


Figure 1.9: A 2D heat map of the fake factors determined from $t\bar{t}$ MC for the full run-2 dataset in the combined $e\tau_h$ and $\mu\tau_h$ channels. This is shown with respect to the hadronic tau p_T and the p_T of the jet matched to the hadronic tau. The ratio of jet to hadronic tau p_T categorisation used is shown split by the dashed lines.



(a) (b) (c)

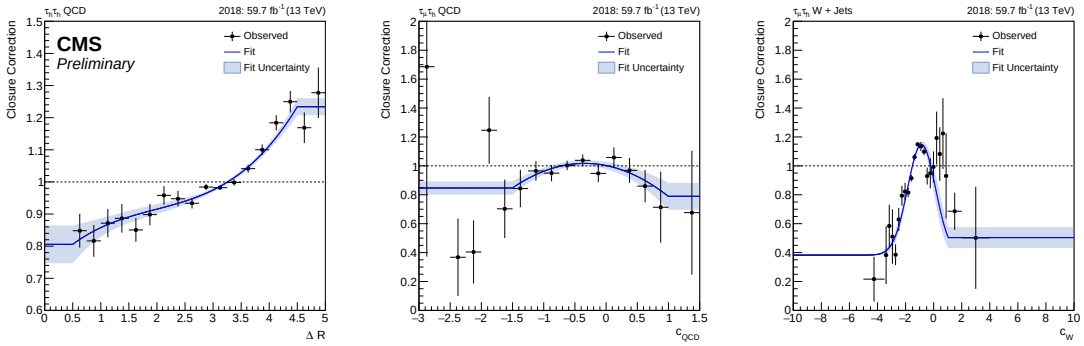
Figure 1.10: Fake factor fits in $\tau_h\tau_h$ channel for the QCD $n_{prebjets} = 0$ category with 2018 data. The three jet p_T to hadronic tau p_T categories are shown.



(a) (b) (c)

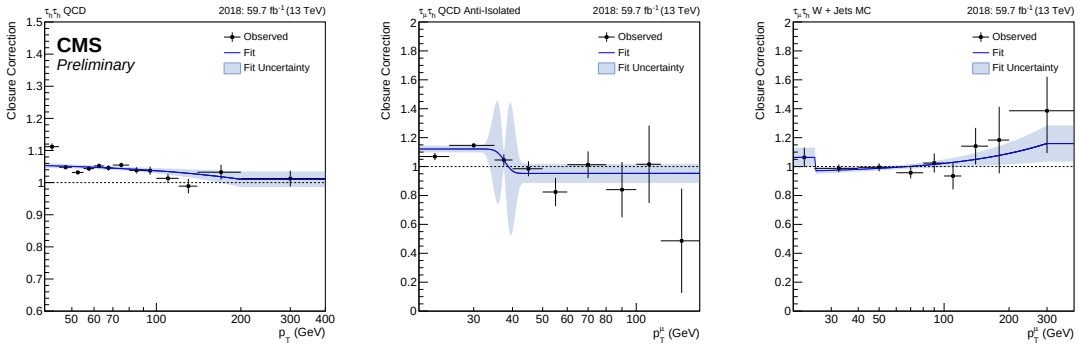
(d) (e) (f)

Figure 1.11: Fake factor fits in $\mu\tau_h$ channel for the QCD and W + Jets $n_{prebjets} = 0$ category with 2018 data. The three jet p_T to hadronic tau p_T categories are shown for each process.



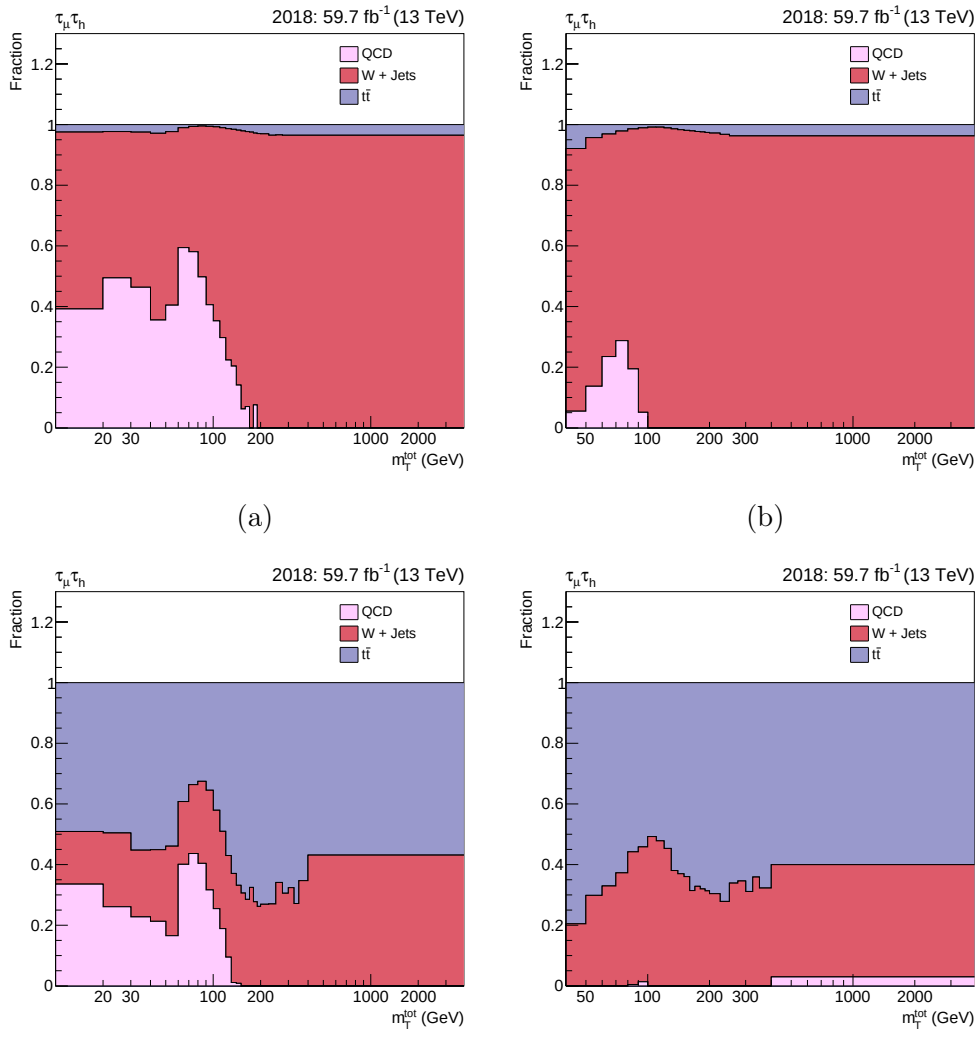
(a) (b) (c)

Figure 1.12: DR closures



(a) (b) (c)

Figure 1.13: DR to AR closures.

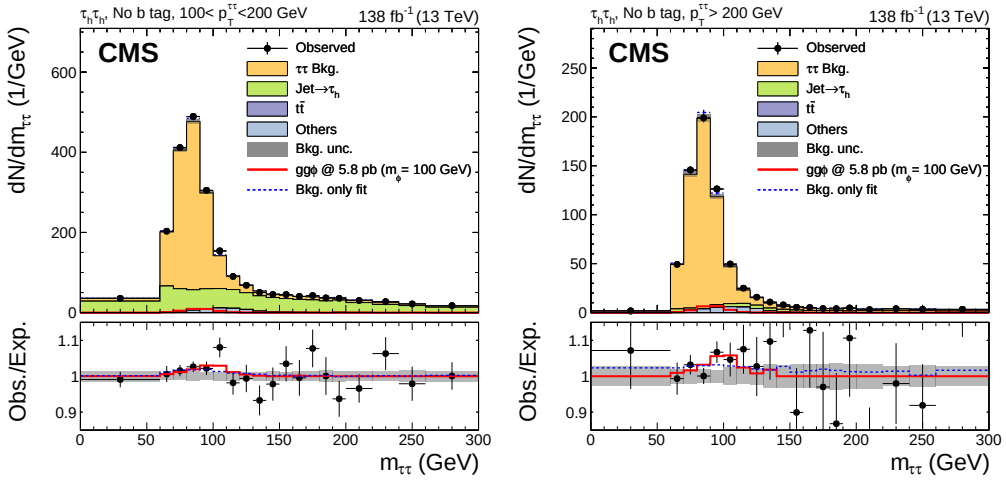


(a) (b)

(c) (d)

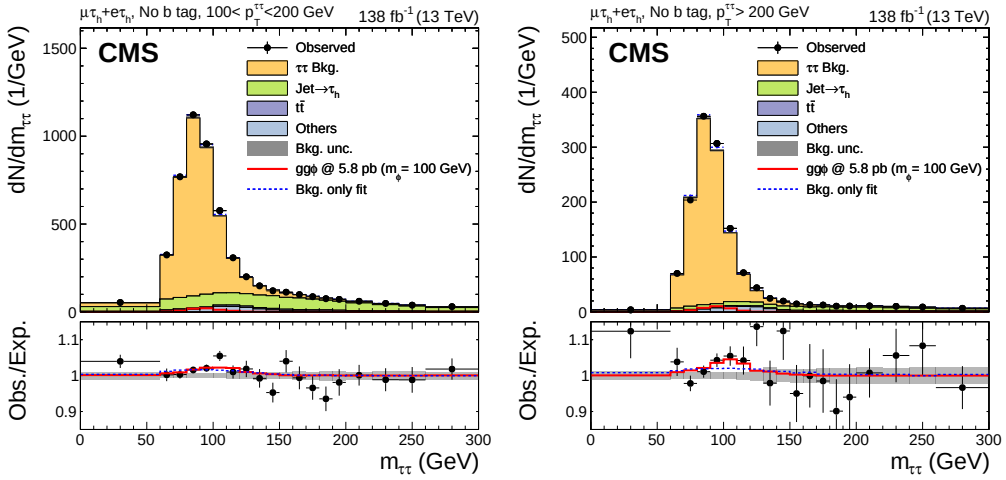
Figure 1.14: Fake factor fractions.

1.9 Postfit Plots



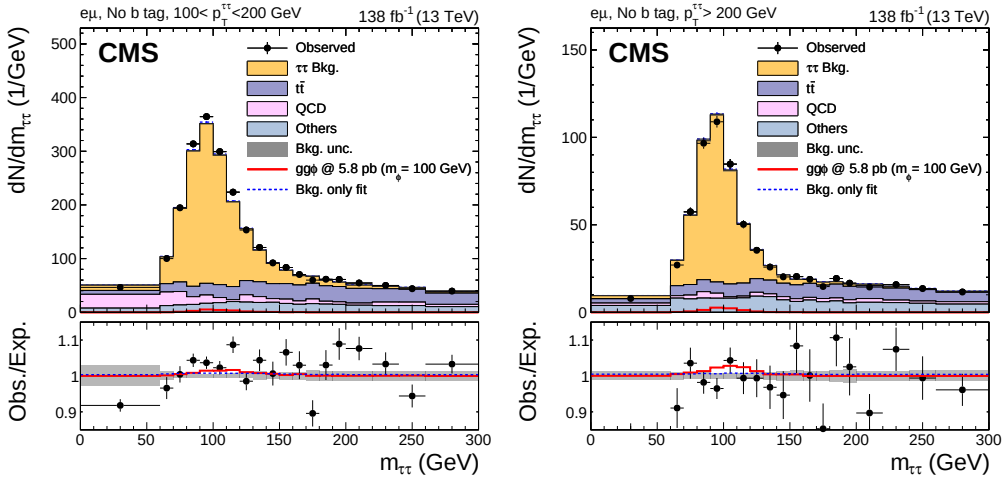
(a)

(b)



(c)

(d)



(e)

(f)

Figure 1.15: Low mass postfit.

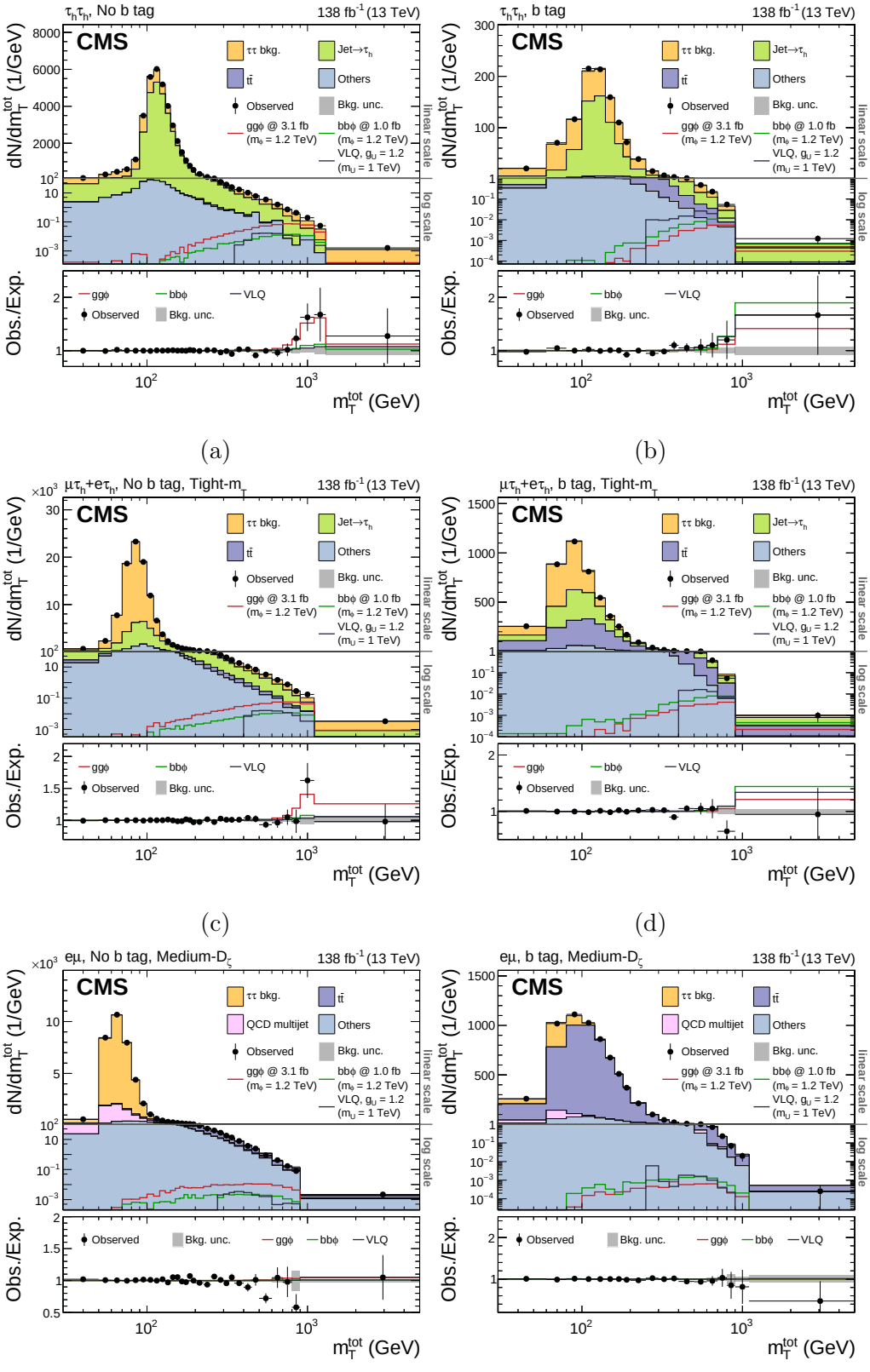


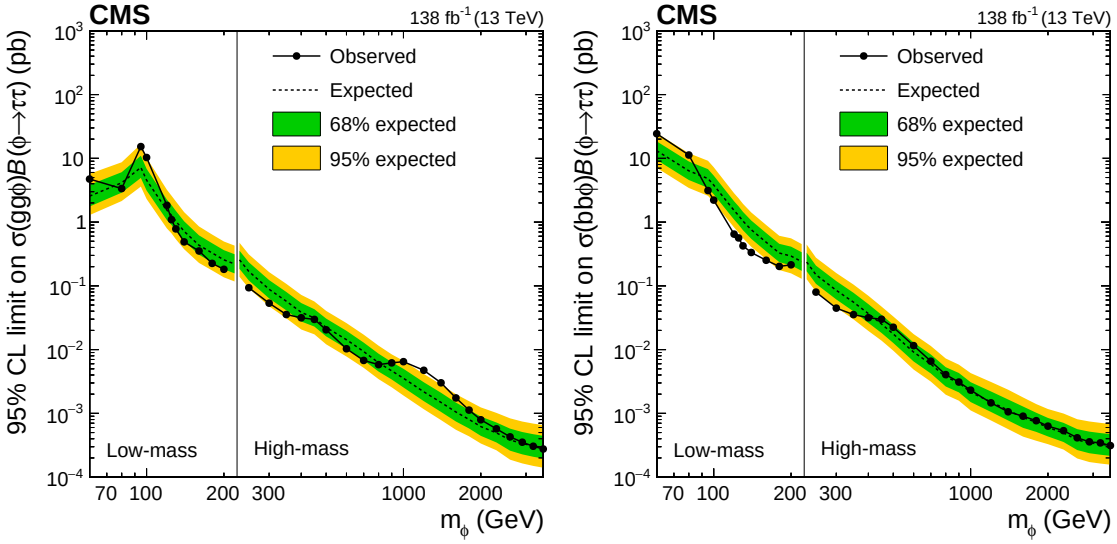
Figure 1.16: High mass postfit.

1.10 MC Corrections

1.11 Model Independent Results

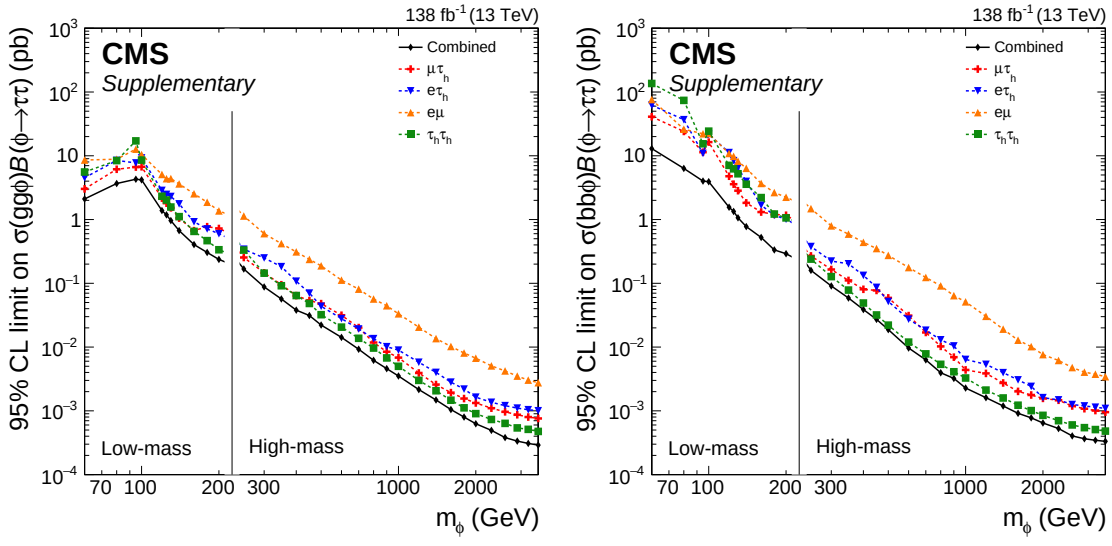
1.11.1 Limit Setting

$$q_\mu = -2\ln\left(\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta}_{\hat{\mu}})}\right), 0 \leq \hat{\mu} \leq \mu, \tag{1.7}$$



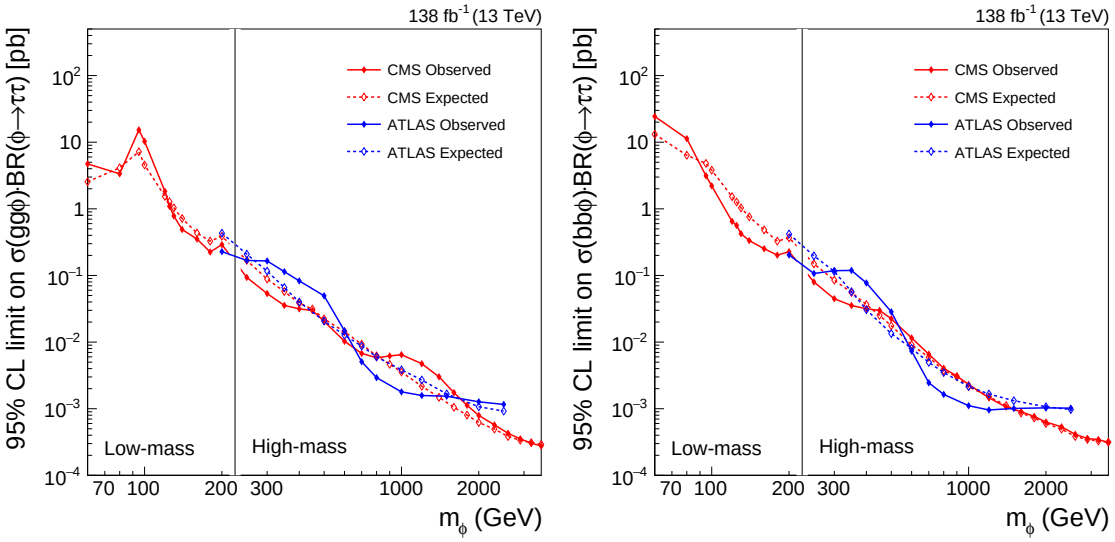
(a) (b)

Figure 1.17: Model independent limits.



(a) (b)

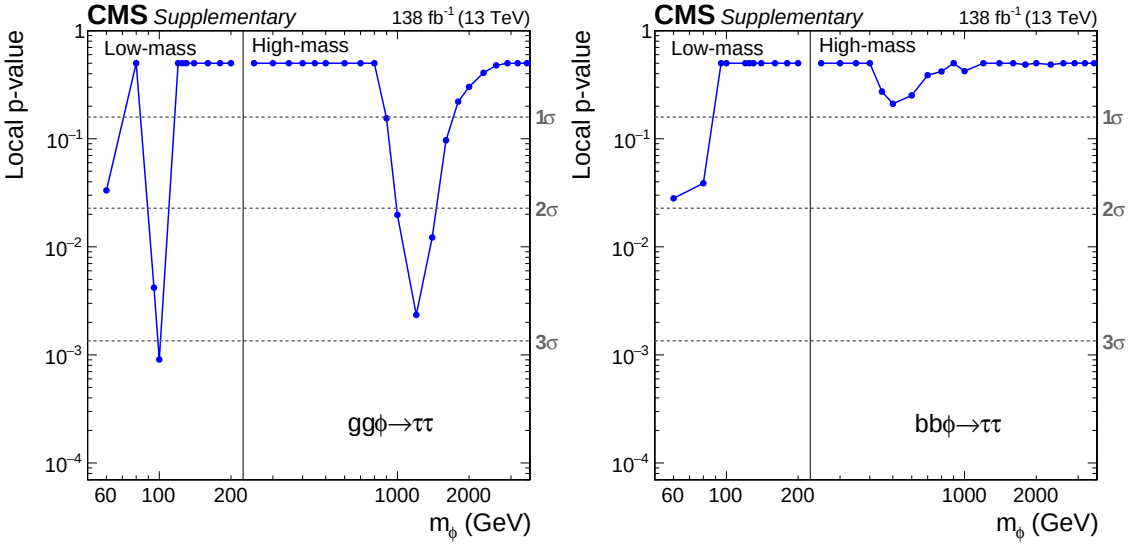
Figure 1.18: Model independent limits.



(a) (b)

Figure 1.19: Model independent limits.

1.11.2 Significance and Compatibility



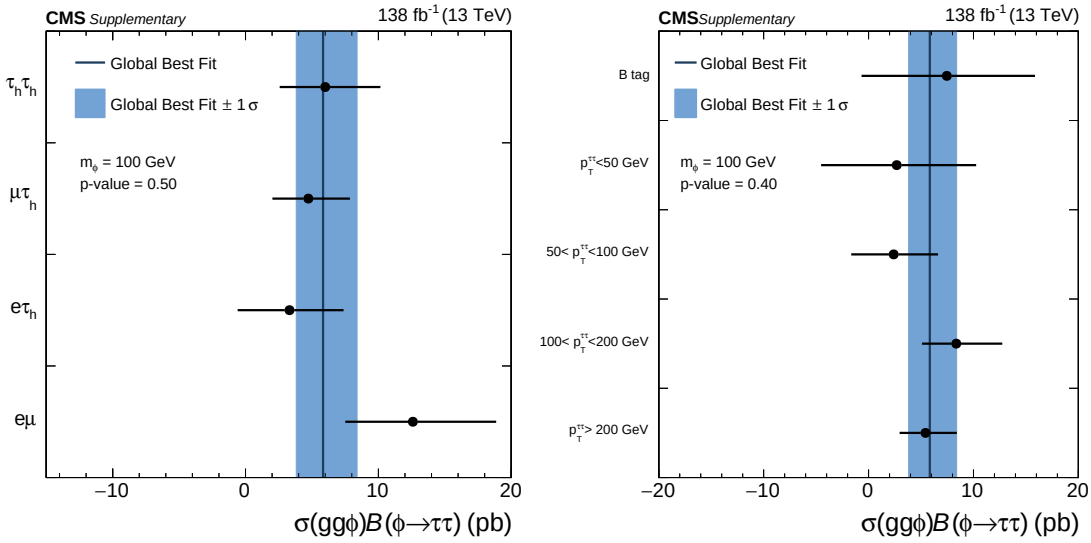
(a) (b)

Figure 1.20: Significance.

1.11.3 2D Likelihood Scans

1.12 Model Dependent Limits

1.12.1 Limit Setting



(a) (b)

Figure 1.21: Low mass compatibility.

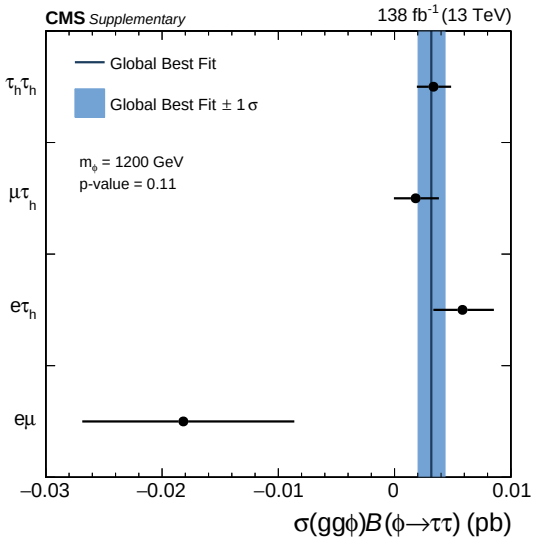


Figure 1.22: High mass compatibility.

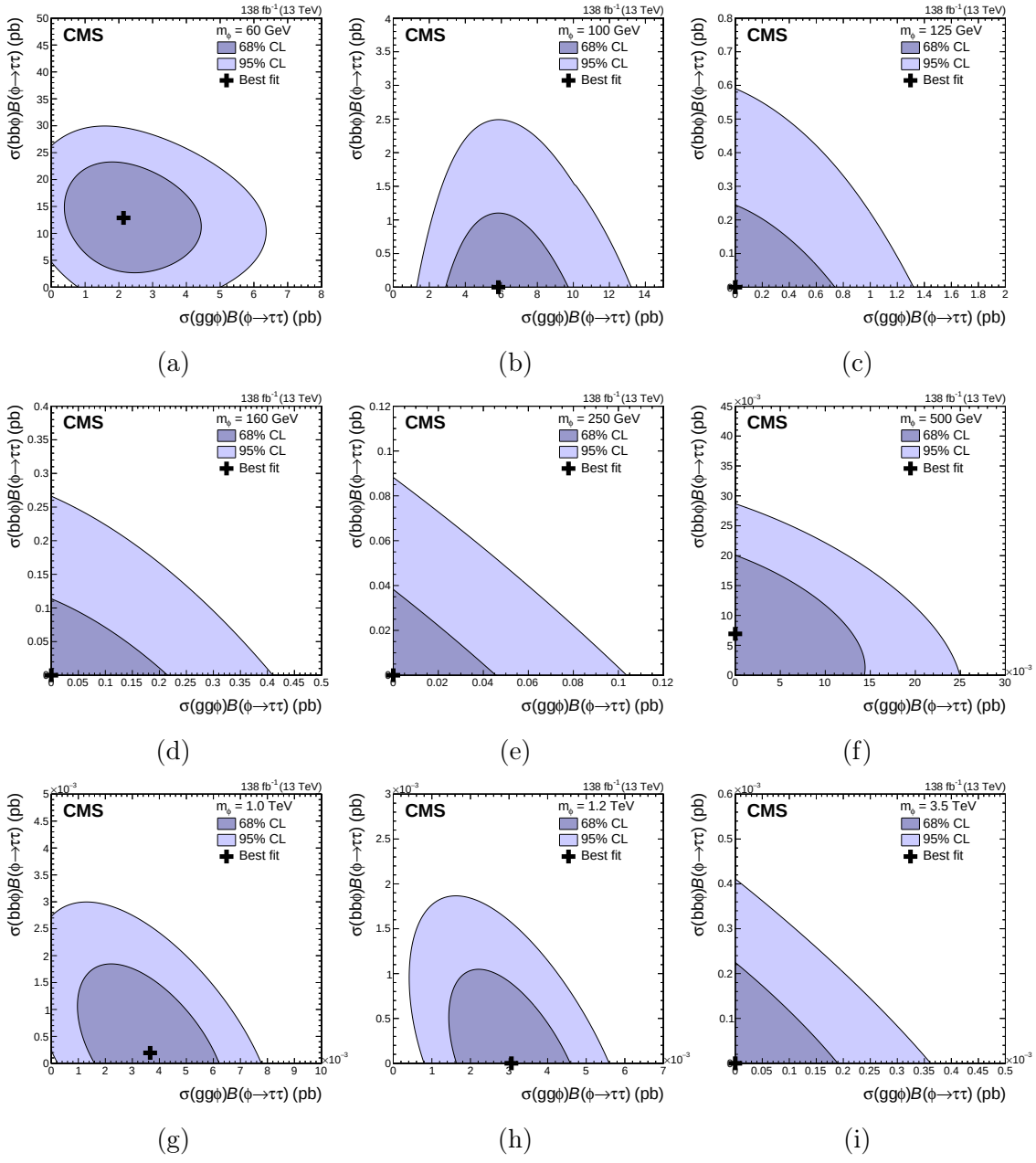
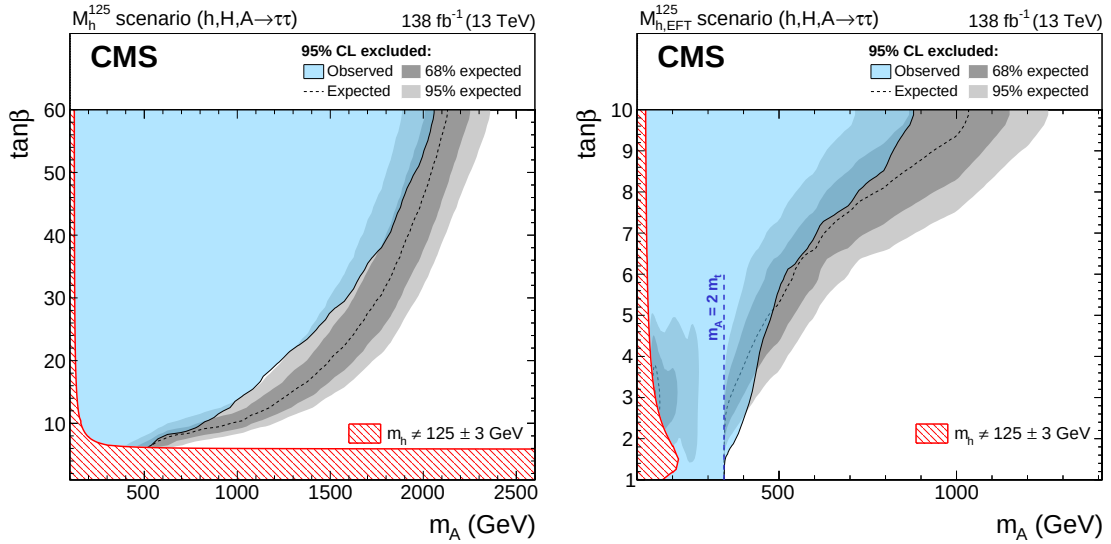
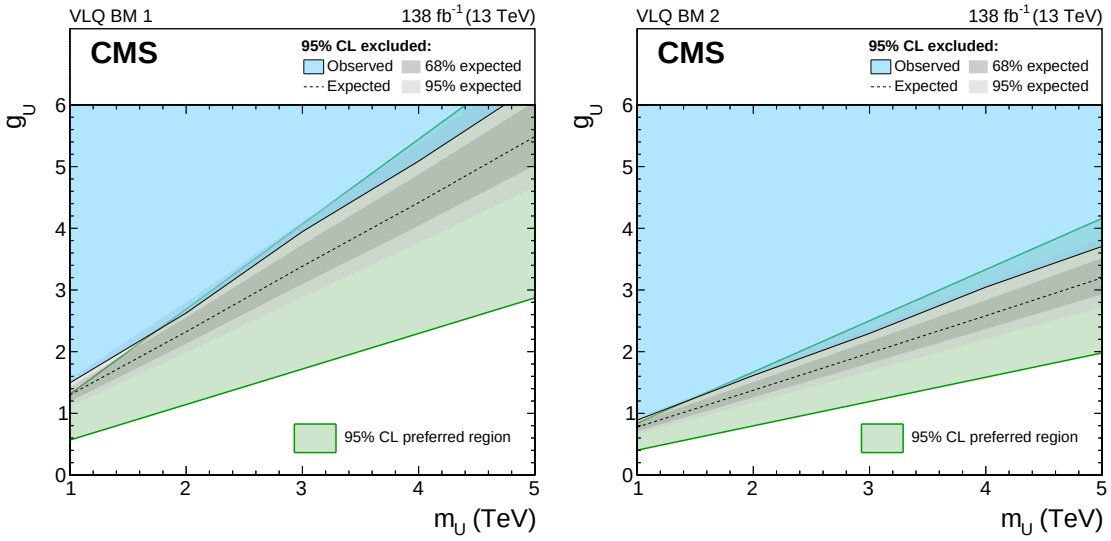


Figure 1.23: 2D Likelihood scans.



(a) (b)

Figure 1.24: MSSM limits.



(a) (b)

Figure 1.25: VLQ limits.