CHAPTER 8

Prospects for $H \to \tau\tau$

Prospects for the $H \to \tau\tau$ analysis in Run-II and at the HL-LHC are described. Discussion of prospects at the HL-LHC are drawn largely from the recent ATLAS documentation on the topic [111].

8.1 Run-II

In Run-II, the LHC is expected to collide protons with $\sqrt{s} = 13$ TeV, a peak instantaneous luminosity of approximately $1.6 \times 10^{34} \mathrm{cm}^{-2}\mathrm{s}^{-1}$, 25 nanosecond bunch spacing, and $\langle \mu \rangle \approx 40$. These data-taking conditions are much harsher than in 2012, as shown in Table 8.1.

The ATLAS L1 trigger rate would increase approximately five-fold if the 2012 trigger menu was ported unchanged to 2015 data-taking conditions [112]. However, the L1 bandwidth is expected to increase from 75 kHz to 100 kHz, much less than this rate increase. The L1 menu must then be adapted to accommodate this.

This is especially true of triggers which rely on τ_{had} since hadronic objects are challenging in the trigger, both in identification and calibration. An unsustainable increase of trigger rate as a function of instantaneous luminosity is shown in Fig. 8.1. Multiple avenues are explored to ensure tenable

Table 8.1: LHC data-taking conditions in 2011 and 2012 compared with the expected data-taking conditions in 2015.

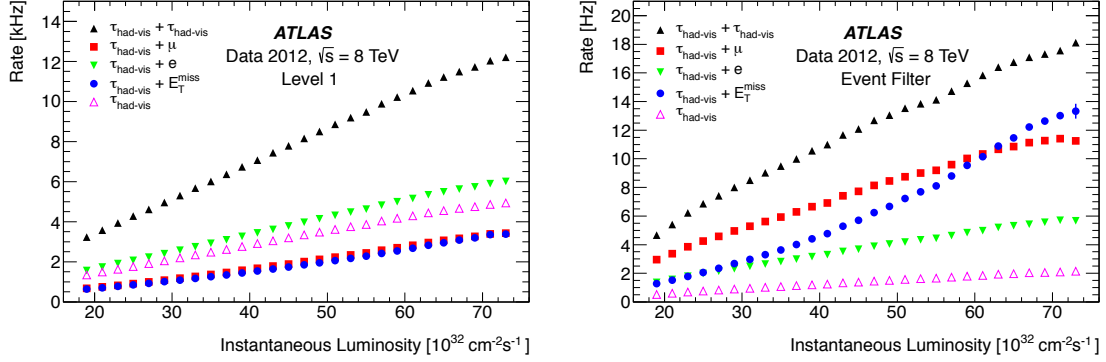| Year of operations	$\sqrt{s}$ [TeV]	peak lumi. $[\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	bunch spacing [ns]	$\langle \mu \rangle$
2011	7	0.4×10^{34}	50	≈ 10
2012	8	0.8×10^{34}	50	≈ 20
2015	13	1.6×10^{34}	25	≈ 40

Figure 8.1: Tau trigger rates in 2012 data-taking as a function of instantaneous luminosity for L1 (left) and HLT (right) [1].

trigger rates without significant loss of physics.

8.1.1 Run-I triggers for $H \to \tau\tau$

In Run-I, only the $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ analysis relies on τ_{had} triggers for physics. The $H \to \tau_\ell\tau_{\mathrm{had}}$ use single lepton triggers with an offline threshold of 26 GeV. $\ell + \tau_{\mathrm{had}}$ triggers are considered but ultimately dropped because they bring additional complication to the analysis without significant improvement in sensitivity. The $H \to \tau_\ell\tau_\ell$ analysis relies on single and di-lepton triggers. The list of triggers used in 2012 data-taking, and their expected 2015 versions, is shown in Table 8.2 and Table 8.3.

8.1.2 Run-II triggers

Trigger options for Run-II are most critical for the $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ since it relies entirely on τ_{had} triggers. The $H \to \tau_\ell\tau_\ell$ analysis will continue to use single and di-lepton triggers, which enjoy low

Table 8.2: L1 triggers used in the 2012 $H \to \tau\tau$ analysis, and their expected 2015 versions, grouped by $\tau\tau$ decay channel.

channel	L1, 2012	L1, 2015
$H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$	2TAU11I_TAU15	no di-τ_{had} item planned
$H \to \tau_e\tau_{\mathrm{had}}$	EM18VH	EM24VHI
$H \to \tau_\mu\tau_{\mathrm{had}}$	MU20	MU20
$H \to \tau_e\tau_e$	EM18VH || 2EM10VH	EM24VHI || 2EM15VH
$H \to \tau_\mu\tau_\mu$	2MU10	2MU10
$H \to \tau_e\tau_\mu$	EM10VH_MU6	EM15VH_MU10

thresholds and are among the most high profile triggers, and thus less likely to be cut in case of unexpectedly high rates. The $H \to \tau_\ell\tau_{\mathrm{had}}$ analysis will continue to use the single lepton trigger, but the potential benefit of recovering events with leptons below the single lepton trigger thresholds could be helpful. Accordingly, only the $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ and $H \to \tau_\ell\tau_{\mathrm{had}}$ analyses are considered for trigger optimizations, with emphasis on $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$.

8.1.2.1 Object thresholds

The first and simplest option is to raise object thresholds in the trigger. Since the Run-I signal regions only include event topologies with additional jets, triggering on one or two additional jets in the event is considered. To assess the impact of this, the Run-I analysis is re-run but with progressively higher thresholds on the final state objects, and the resulting sensitivity is derived. This is shown in Fig. 8.2.

For both $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ and $H \to \tau_\ell\tau_{\mathrm{had}}$, triggering on the lead jet in the VBF category is promising, and triggering on the sub-lead jet is not. The lowest realistic threshold for triggering on the sub-lead jet is $15-20$ GeV at L1 ($\approx 50-60$ GeV offline) costs significant sensitivity in both final states. Requiring a second jet would also be inefficient for the boosted category. To trigger on the lead jet, however, an offline threshold of ≈ 70 GeV (25 GeV at L1) does not cost significant sensitivity.

For $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$, raising the threshold on the lead τ_{had} is more promising then raising the threshold on the sub-lead τ_{had}. For $H \to \tau_\ell\tau_{\mathrm{had}}$, raising the threshold on the τ_{had} is more promising than raising the threshold on the lepton.

8.1.2.2 Topological requirements

A new feature of the L1 trigger in 2015 is the ability to make topological selections, whereas in Run-I, only object multiplicity selections could be made. This topological selection are implemented via the

Table 8.3: HLT triggers used in the 2012 $H \to \tau\tau$ analysis, and their expected 2015 versions, grouped by $\tau\tau$ decay channel.

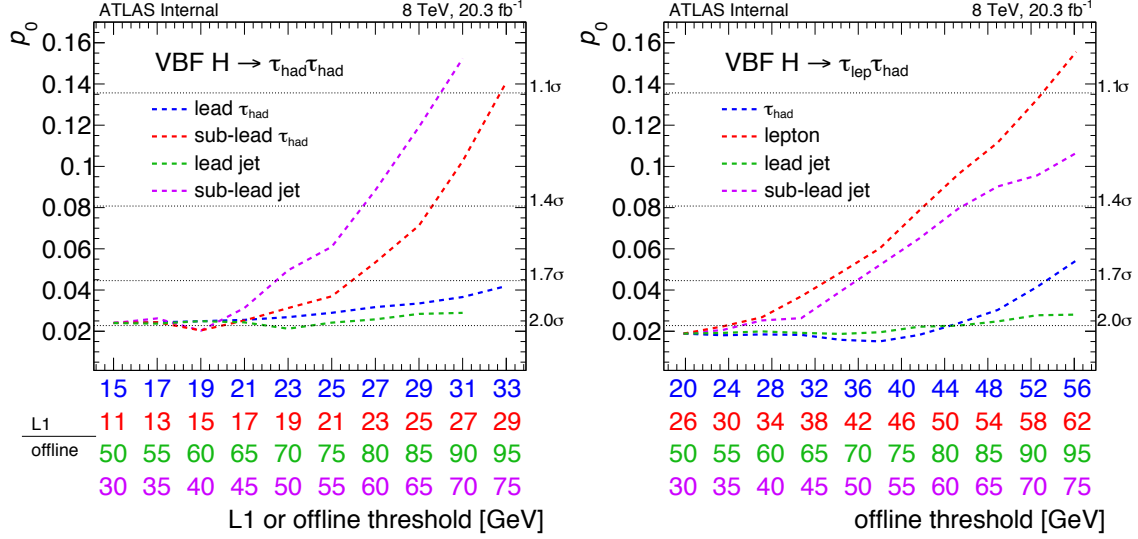| channel	HLT, 2012	HLT, 2015
$H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$	`tau29Ti_medium1_tau20Ti_medium1`	no di-τ_{had} item planned
$H \to \tau_e\tau_{\mathrm{had}}$	`e24vhi_medium1`	`e28i_tight`
$H \to \tau_\mu\tau_{\mathrm{had}}$	`mu24i_tight`	`mu26i_medium`
$H \to \tau_e\tau_e$	`e24vhi_medium1 || 2e12Tvh_loose1`	`e28i_tight || 2e17_loose`
$H \to \tau_\mu\tau_\mu$	`mu18_tight_mu8_EFFS`	`2mu14`
$H \to \tau_e\tau_\mu$	`e12Tvh_medium1_mu8`	`e17_medium_mu12`

Figure 8.2: Significance (p_0) of the $H \to \tau\tau$ VBF category in 2012 for the $H \to \tau_{\text{had}}\tau_{\text{had}}$ (left) and $H \to \tau_\ell\tau_{\text{had}}$ (right) analyses as a function of offline or L1 threshold for various objects.

new `L1topo` processor [113, 114].

At L1, the dominant background for τ_{had} triggers is QCD di-jet production. Topological selections can be used in various ways to suppress this process, especially in the $H \to \tau\tau$ signal regions where the $\tau\tau$ system tends to be boosted.

$\Delta\phi(\tau\tau) < X$:

QCD di-jets tend to be produced back-to-back in the transverse plane, and the $\tau\tau$ system tends to have smaller $\Delta\phi$ due to transverse boost of the Z/H.

$\Delta\eta(\tau\tau) < X$:

QCD di-jets tend to be produced broadly in η, and the $\tau\tau$ system tends to have smaller $\Delta\eta$ due to longitudal boost of the Z/H.

$\Delta R(\tau\tau) < X$:

This combines discriminating power of $\Delta\phi(\tau\tau)$ and $\Delta\eta(\tau\tau)$.

$p_{\text{T}}(\tau\tau) > X$:

QCD di-jet systems tend to be produced at rest in the transverse plane, and the $\tau\tau$ system is usually not due to transverse boost.

$m_{\tau\tau} > X$:

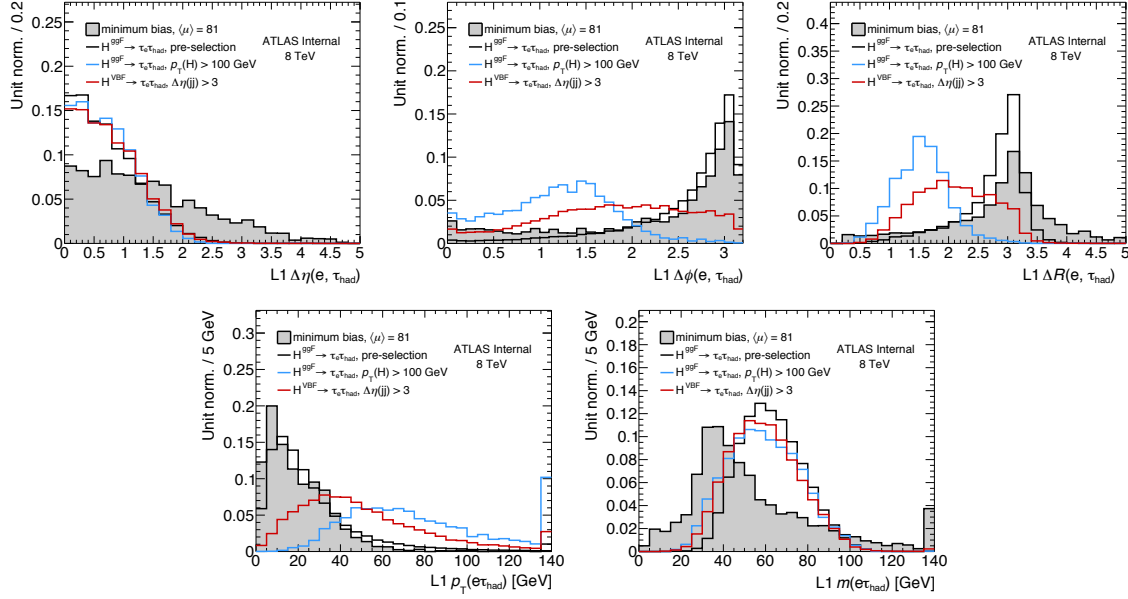QCD di-jet systems tend to be non-resonant and low-p_{T}, and the $\tau\tau$ system is from Z/H decays.

Figure 8.3: Topological distributions at L1 for $H \to \tau_e \tau_{\mathrm{had}}$ MC versus high-pileup ($\langle \mu \rangle = 81$) minimum bias MC.
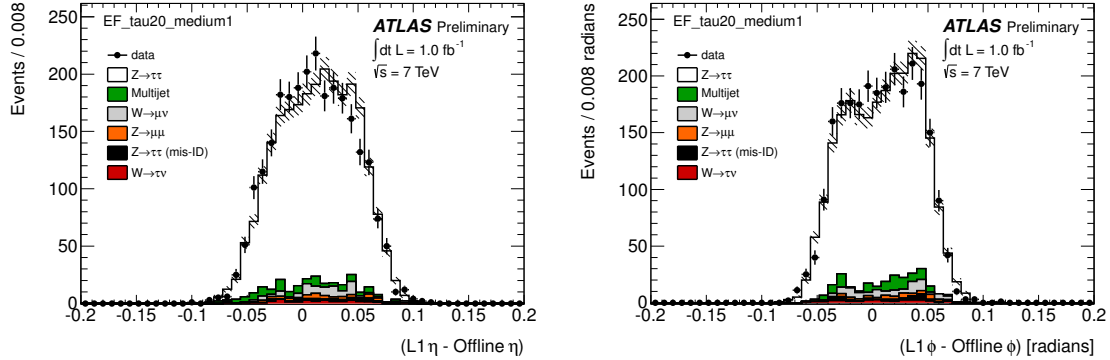

Figure 8.4: L1 angular resolution for τ_{had} in simulation and data [115].

The discriminating power of these variables is shown in Fig. 8.3 for $H \to \tau_e \tau_{\mathrm{had}}$ MC versus high-pileup minimum bias MC.

Of these options, angular discrimination is appealing because the angular resolution for τ_{had} at L1 is better than momentum resolution, as shown in Fig. 8.4 and Fig. 8.5. Sharper efficiency turn-ons can then be expected at HLT and offline. $\Delta R(\tau\tau) < 2.8$ is ultimately chosen for discrimination because it combines the discriminating power of $\Delta\phi(\tau\tau)$ and $\Delta\eta(\tau\tau)$.

The impact of requiring an additional jet and $\Delta R(\tau\tau) < 2.8$ on the predicted 2015 L1 trigger rate
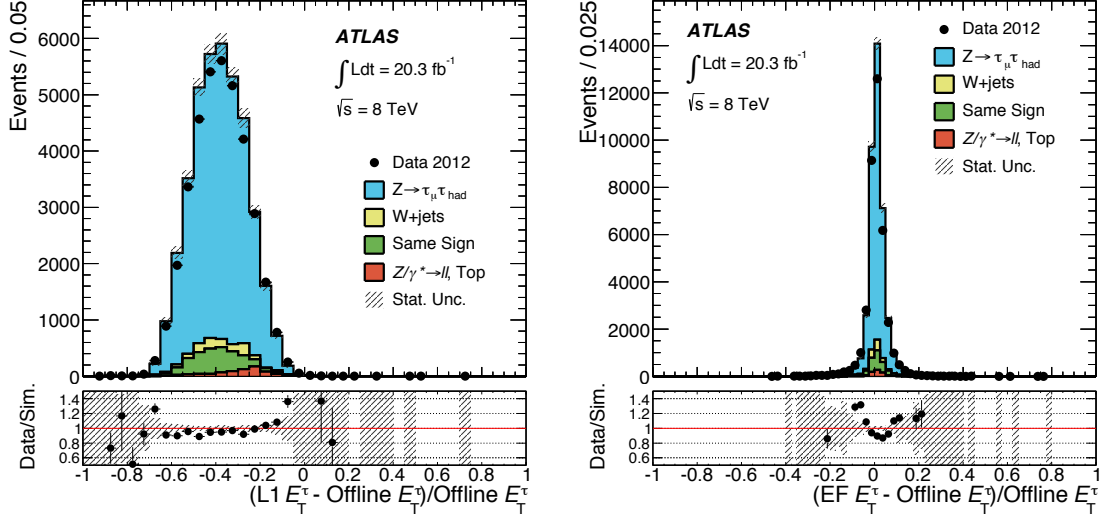
Figure 8.5: Momentum resolution for τ_{had} in simulation and data at L1 (left) and HLT (right). The resolution is significantly improved at HLT [1].

is shown in Table 8.4. A ten-fold reduction of the rate is achieved without signicant loss of sensitivity.

8.1.2.3 Gains with $\ell+\tau_{\text{had}}$ triggers

To assess the potential gain in sensitivity of $\ell+\tau_{\text{had}}$ triggers in 2015, the $H \to \tau_\ell\tau_{\text{had}}$ analysis is re-run in the regime below the single lepton trigger threshold with the $\ell+\tau_{\text{had}}$ triggers running in 2012, shown in Table 8.5. Kinematic distributions are shown in Fig. 8.6 and Fig. 8.7, including the final BDT discriminant.

A simple, analytic formula for the discovery significance Z_A [116] gives 0.7σ for the $\ell+\tau_{\text{had}}$ triggered

Table 8.4: L1 trigger items and rate predictions for 2015 data-taking. A baseline L1 menu is used for calculating the unique rate.

L1 item	rate [kHz]	unique rate [kHz]	notes
2TAU11	181	147	$H \to \tau_{\text{had}}\tau_{\text{had}}$, early 2011
2TAU11I	121	99	$H \to \tau_{\text{had}}\tau_{\text{had}}$, late 2011
TAU15I TAU11I	96	75	$H \to \tau_{\text{had}}\tau_{\text{had}}$, 2012
TAU20I TAU11I	69	48	Raise lead τ_{had} p_{T}
TAU20I TAU12I	61	42	Raise sub-lead τ_{had} p_{T}
TAU20I TAU12I J25	20	12	Additional jet
TAU20I TAU12I J25 DR2.8	7.8	4.3	$\Delta R(\tau_{\text{had}}\tau_{\text{had}}) < 2.8$

category. This is not a significant improvment over the single lepton triggered regime.

Table 8.5: L1 and HLT $\ell + \tau_{\mathrm{had}}$ trigger items operating in 2012.

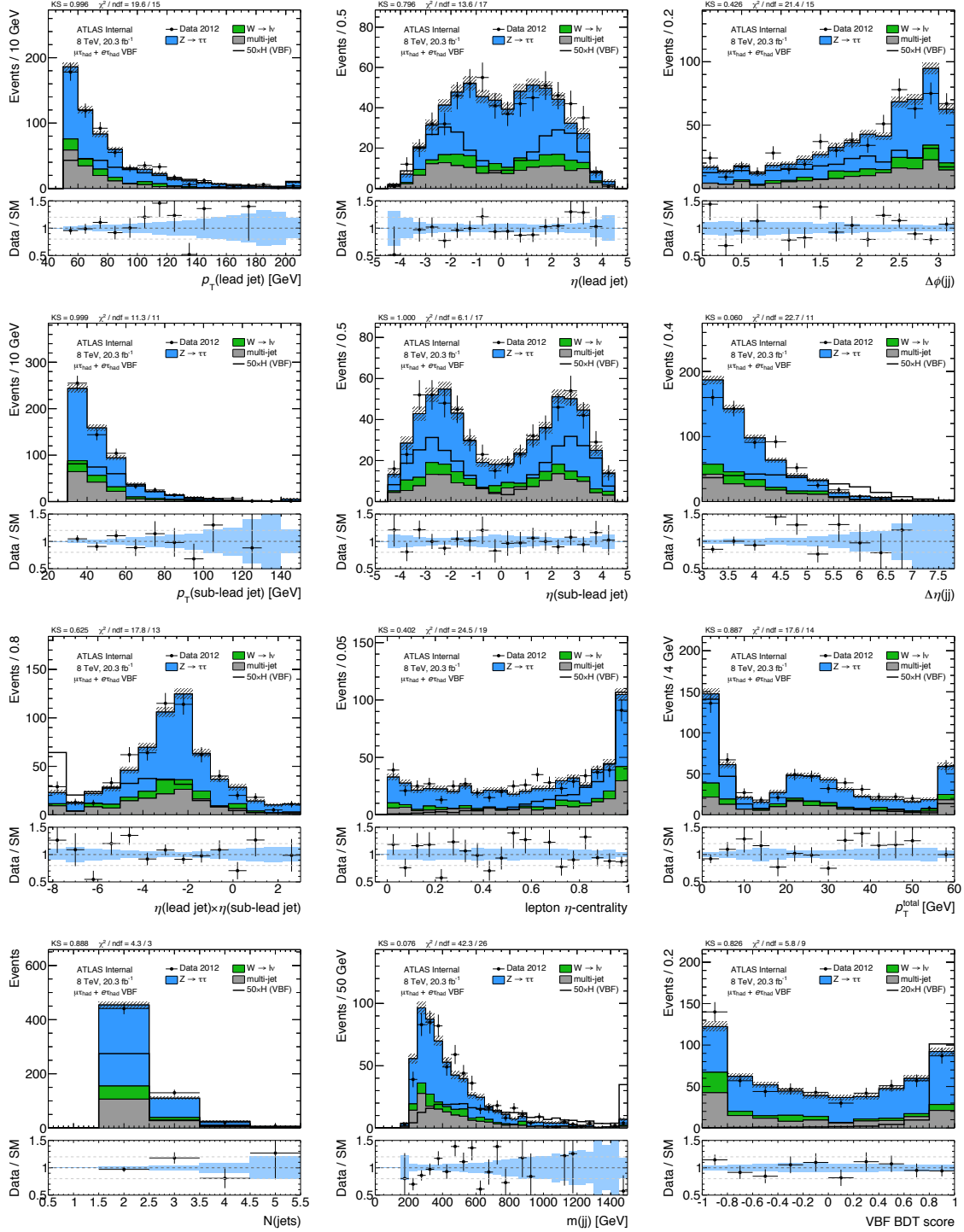| channel	L1	HLT
$H \to \tau_e \tau_{\mathrm{had}}$	2TAU11I_EM14VH	tau20Ti_medium1_e18vh_medium1
$H \to \tau_\mu \tau_{\mathrm{had}}$	TAU8_MU10	tau20_medium1_mu15
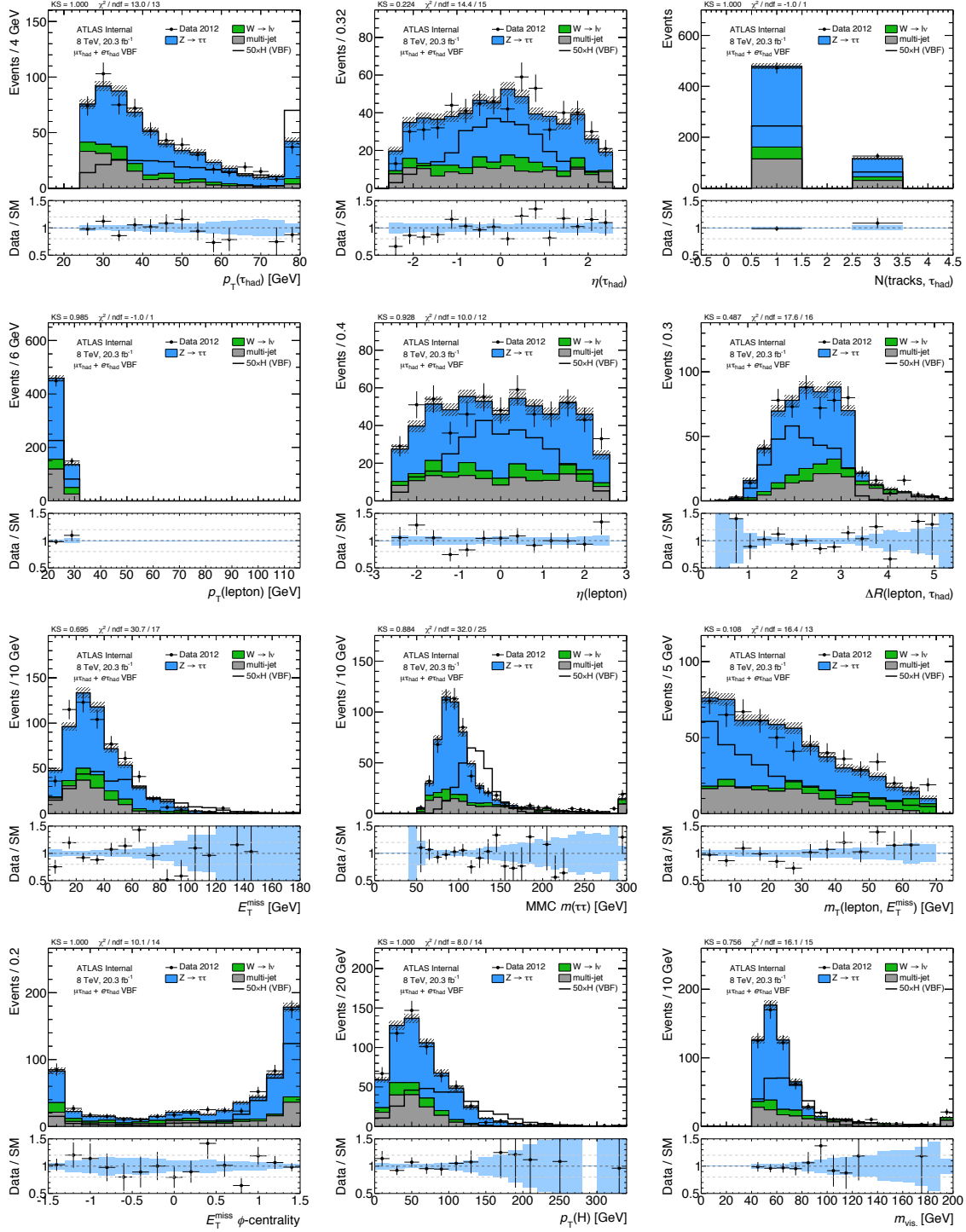
Figure 8.6: Kinematic distributions in the $\ell + \tau_{\text{had}}$ category of the 8 TeV VBF $H \to \tau_\ell \tau_{\text{had}}$ analysis.
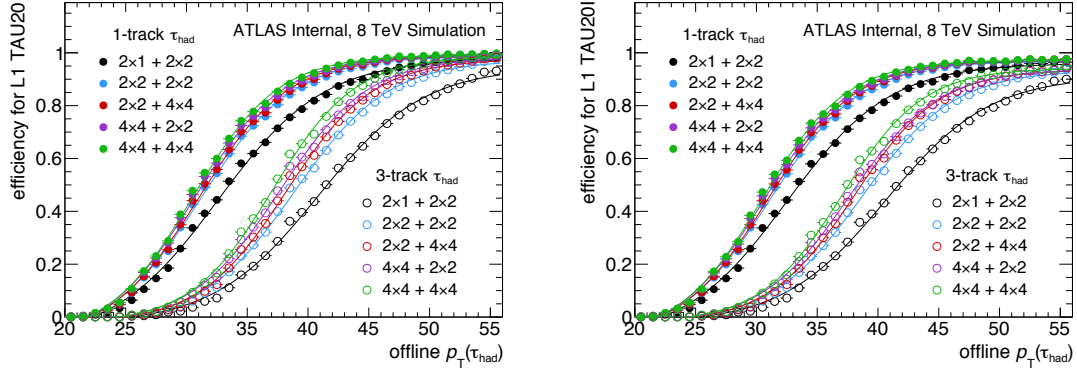
Figure 8.7: Kinematic distributions in the $\ell + \tau_{\mathrm{had}}$ category of the 8 TeV VBF $H \to \tau_\ell \tau_{\mathrm{had}}$ analysis.

Figure 8.8: Efficiency for firing the 20 GeV L1 τ_{had} trigger as a function of offline $p_{\mathrm{T}}(\tau_{\mathrm{had}})$ for no isolation requirement (left) and the 2012 isolation requirement (right) for various definitions of the L1 τ_{had} item. The current definition (2×1 EM, 2×2 had.) has the slowest efficiency turn-on. Fits are performed with a Fermi-Dirac distribution.

8.1.3 L1 τ_{had}

Updates to the L1 τ_{had} object are also considered. These are explored in hopes of improving the trigger efficiency for τ_{had} generally, without significant rate increases, which could assist all analyses which rely on τ_{had} triggers.

8.1.3.1 Size

The size of the L1 τ_{had} is made larger to check if a sharper efficiency turn-on curve can be gained without significant background contamination. This is especially important for 3-track τ_{had}, which are expected to be wider than 1-track τ_{had} and are observed to have a slower efficiency turn-on curve. The default L1 τ_{had} used in 2012 data-taking is 2×1 in the electromagnetic calorimeter and 2×2 in the hadronic calorimeter.

Efficiency turn-on curves for various L1 τ_{had} definitions are shown in Fig. 8.8 for simulated τ_{had} with and without isolation requirements. A sharper efficiency turn-on is achieved relative to the default definition, as shown in Table 8.6. Since background rates increase non-trivially with larger L1 τ_{had} objects, the default definition will be retained in 2015 data-taking.

8.1.3.2 Isolation

The definition of L1 isolation is also reconsidered for 2015 data-taking. In 2012, a flat cut of $p_{\mathrm{T}}^{\mathrm{L1,iso}} \leq 4$ GeV is used. In 2015, the option of $p_{\mathrm{T}}^{\mathrm{L1}}$-dependent isolation is available. A simple linear dependence

is explored:

$$\text{require } p_\text{T}^\text{L1,iso} \leq m \times p_\text{T}^\text{L1,iso} + b \text{ GeV} \tag{8.1}$$

With this definition, the 2012 isolation requirement is $m = 0$ and $b = 4$.

Removing the isolation requirement is considered above $p_\text{T}^\text{L1} = 60$ GeV. This is equivalent to a logical OR with the lowest unprescaled single τ_had trigger, TAU60, and thus will not cost any additional unique rate to the total L1 trigger menu.

The predicted relative rate is shown in Fig. 8.9, with and without isolation at high p_T. An inverse relationship in the rate is observed between offset and slope, as expected. The L1 τ_had efficiency is shown in Fig. 8.10 for the isolation requirements which give equal rate. Significant signal efficiency is recovered at high p_T.

8.1.4 Conclusions and contingencies

Despite significantly harsher data-taking conditions in 2015, triggers for the $H \to \tau_\text{had}\tau_\text{had}$ and $H \to \tau_\ell\tau_\text{had}$ analyses can be adapted to retain physics output with manageable rates and without major loss of generality. The most important adaptations for both channels are requiring an additional high-p_T jet and requiring low $\Delta R(\tau\tau)$. The L1 τ_had menu is shown in Table 8.7.

Contingency plans are considered due to the uncertainty of the rate prediction and the uncertainty of the LHC data-taking conditions. If the $H \to \tau\tau$ section of the τ_had L1 menu is asked to reduce its bandwidth consumption, options are available with minor physics losses, such as raising the τ_had

Table 8.6: Fits of the efficiency for firing the 20 GeV L1 τ_had trigger with a Fermi-Dirac distribution for various definitions of the L1 τ_had item. No isolation requirement is made.

τ_had	L1 size (EM, had.)	fitted p_T offset [GeV]	fitted sharpness
	2×1, 2×2	32.6	4.43
	2×2, 2×2	31.2	3.87
1-track	2×2, 4×4	31.0	3.72
	4×4, 2×2	30.7	3.67
	4×4, 4×4	30.5	3.52
	2×1, 2×2	41.0	4.47
	2×2, 2×2	38.9	3.93
3-track	2×2, 4×4	38.3	3.65
	4×4, 2×2	37.7	3.71
	4×4, 4×4	37.1	3.42
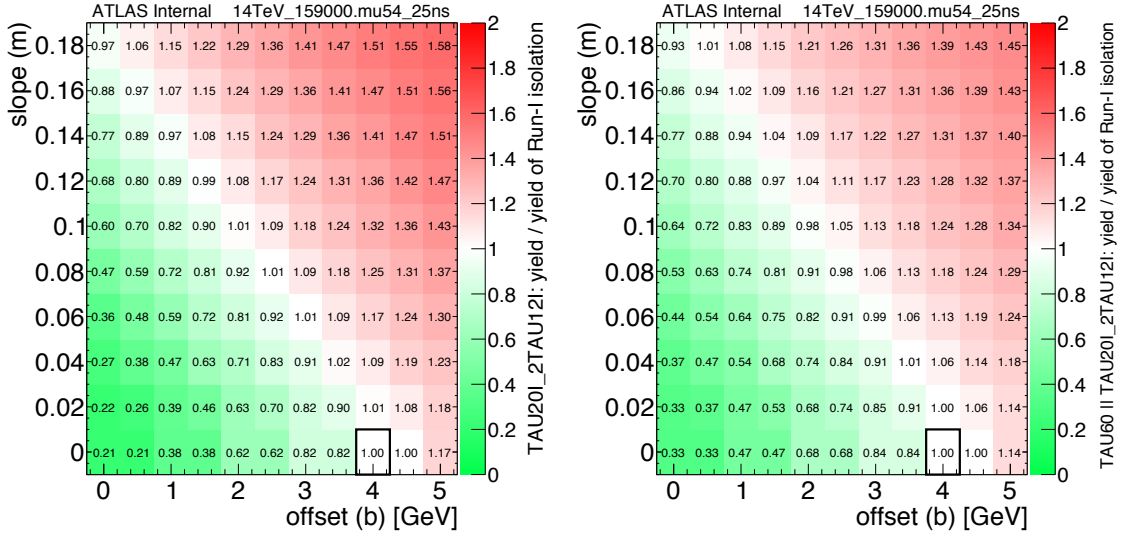
Figure 8.9: L1 rate for the di-τ_{had} trigger in 14 TeV minimum bias MC for various $p_{\mathrm{T}}^{\mathrm{L1}}$-dependent isolation definitions relative to the 2012 definition: $p_{\mathrm{T}}^{\mathrm{L1,iso}} \leq 4$ GeV. Many options give the same rate (white color). The rate is calculated irrespective of the lowest unprescaled single τ_{had} trigger (left) and with a logical OR of it (right).
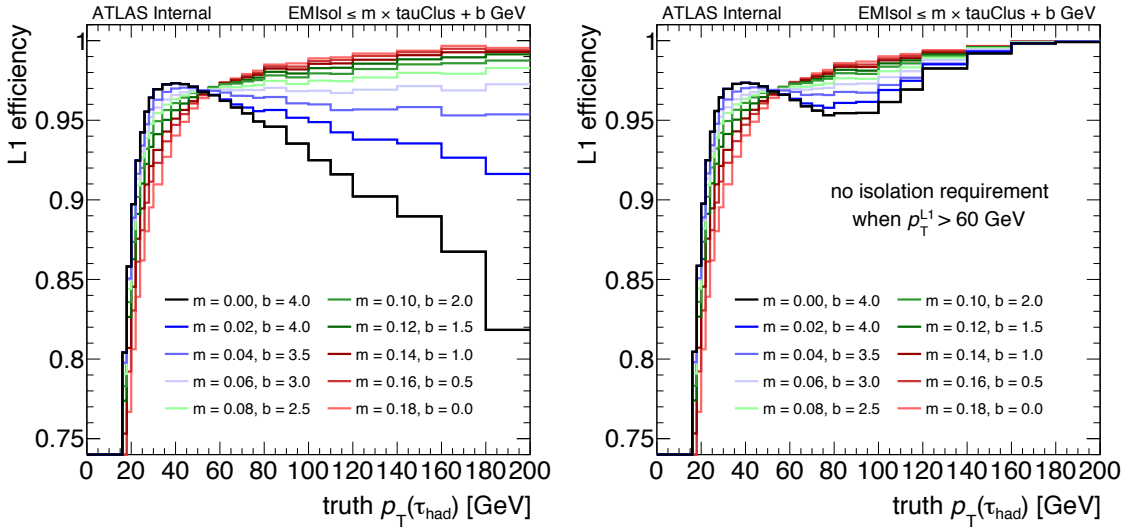
Figure 8.10: Efficiency for firing the L1 τ_{had} trigger for various $p_{\mathrm{T}}^{\mathrm{L1}}$-dependent isolation definitions which have similar rates, as derived from Fig. 8.9. The 2012 definition is the black line. The efficiency is calculated irrespective of the lowest unprescaled single τ_{had} trigger (left) and with a logical OR of it (right).

120

threshold of the $\ell + \tau_{\mathrm{had}}$ triggers or raising the jet threshold of all triggers. If additional bandwidth is available, the jet threshold or $\Delta R(\tau\tau)$ requirement can be relaxed. The contingencies are shown in Table 8.8.

8.2 HL-LHC

This section documents projections of the Standard Model $H \to \tau\tau$ analysis for the ECFA High Luminosity LHC Experiments Workshop 2014 [111]. The projection considers High Luminosity LHC (HL-LHC) running conditions with 14 TeV pp collisions, 3000 fb^{-1} delivered integrated luminosity, and an average number of overlapping pp collisions per bunch-crossing (pile-up) $\langle \mu \rangle = 140$. Only the VBF $\tau_{\ell}\tau_{\mathrm{had}}$ ($\ell = e, \mu$) analysis category is considered.

This projection is built from the existing Run-I analysis [2] by using the same data samples, Monte Carlo samples, data-driven background estimates, and multivariate analysis (MVA) techniques. It is projected to HL-LHC conditions by adding emulation of the harsher pile-up conditions and scaling the predictions by the ratios of cross-sections and integrated luminosity for HL-LHC versus 2012 conditions. The harsher pile-up conditions impacts jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ significantly.

The analysis also considers possible extensions of the tracking volume to investigate the impact of tracking-based rejection of pile-up jets.

Table 8.7: The τ_{had} L1 menu. A baseline L1 menu is used for calculating the unique rate.

signature	target	L1 item	unique rate [kHz]
single τ_{had}	exotics, SUSY	TAU60V	1.6
di-τ_{had}	Higgs, di-H	TAU12I_TAU20I_J25-DR28	3.7
$e + \tau_{\mathrm{had}}$	Higgs	EM15HI_TAU12I_J25-DR28	1.2
$e + \tau_{\mathrm{had}}$	exotics, SUSY	EM15HI_TAU40	0.2
$\mu + \tau_{\mathrm{had}}$	Higgs	TAU12I_MU10_J25-DR28	0.5
$\mu + \tau_{\mathrm{had}}$	exotics, SUSY	TAU20_MU10	0.5
$\tau_{\mathrm{had}} + E_{\mathrm{T}}^{\mathrm{miss}}$	SUSY	TAU20_XE45_J20	0.5
$\tau_{\mathrm{had}} + E_{\mathrm{T}}^{\mathrm{miss}}$	SUSY	TAU12I_TAU20I_XE35	2.1
$\tau_{\mathrm{had}} + E_{\mathrm{T}}^{\mathrm{miss}}$	SUSY	EM15HI_TAU12I_XE35	0.2
$\tau_{\mathrm{had}} + E_{\mathrm{T}}^{\mathrm{miss}}$	SUSY	TAU12I_MU10_XE35	0.0

8.2.1   Selection

The selection of a VBF-like sample for the $\tau_\ell\tau_{\mathrm{had}}$ final state discussed in this note relies on the identification of one lepton (electron or muon), one hadronic tau and at least two jets. Muons are selected if they have transverse momentum higher than 26 GeV and are in the region $|\eta| < 2.4$. Quality criteria on the inner detector track associated to the muon are also applied. Electron candidates are formed from a cluster in the electromagnetic calorimeters (transition regions, $1.37 < |\eta| < 1.52$, are excluded) that is matched with a track reconstructed within the inner detector, $|\eta| < 2.47$. Electrons with a transverse momentum higher than 26 GeV are selected and a `tightPP` working point is used. For both muons and electrons, calorimeter and track-based isolation criteria with similar combined selection efficiency as the LHC Run-I analysis are assumed.

Jets are reconstructed using the anti-k_t algorithm [64] with a radius parameter $R = 0.4$, taking topological clusters in the calorimeters as inputs. Only jets with $|\eta| < 4.5$ are selected in this analysis. The general p_{T} threshold for jets is 30 GeV, and specific selection of the VBF analysis is mentioned later. Track-based pile-up suppression with jet-vertex fraction (JVF) [66] is applied in the range of the tracking volume ($|\eta^{\mathrm{jet}}| < 2.4$). In the range $|\eta| < 2.5$, b-tagged jets are identified using the MV1 tagging algorithm based on the impact parameter information and on the reconstruction of the displaced vertices of the hadron decays inside the jets [67].

Tau candidates are seeded by anti-k_t [64], $R = 0.4$ jets with $p_{\mathrm{T}} > 20$ GeV whose calorimeter cluster and leading track must satisfy $|\eta| < 2.47$. The Boosted Decision Tree (BDT) tau identification method [1] is used, requiring that the tau candidate passes the `medium` tightness, corresponding to to approximately 55-60% efficiency A dedicated selection to reject fake tau candidates from electrons and muons is applied.

Table 8.8: Contingency options for the $H \to \tau\tau$ section of the τ_{had} L1 menu. The change in unique rate is with respect to the baseline menu. A baseline L1 menu is used for calculating the unique rate.

scenario		item change	motivation	Δ [kHz]
	3rd	`TAU20I` $\to$ `TAU15I`	slow τ_{had} turn-on	+3.2
add rate	2nd	$\Delta R(\tau\tau) < 2.8 \to 3.0$	ΔR inefficiency for di-H	+2.6
	1st	`J25` $\to$ `J20`	slow jet turn-on	+0.4
baseline	–	–	–	–
	1st	`TAU12I` $\to$ `TAU20I` ($\tau_\ell\tau_{\mathrm{had}}$)	small gain with $\ell+\tau_{\mathrm{had}}$ triggers	−1.7
cut rate	2nd	`J25` $\to$ `J30`	jet threshold is robust	−2.2
	3rd	`TAU20I` $\to$ `TAU25I`	τ_{had} threshold is robust	−3.8

When different objects selected according to the above criteria overlap with each other geometrically (within $\Delta R < 0.2$), only one of them is considered for further analysis. The overlap is resolved by selecting muon, electron, τ_{had} and jet candidates in this order of priority.

The signal events are characterized by true $E_{\text{T}}^{\text{miss}}$ due to the presence of the neutrinos from the tau decays. In this analysis, the $E_{\text{T}}^{\text{miss}}$ reconstruction uses reconstructed high-p_{T} physics objects (electrons, photons, τ_{had}, jets and muons) and a measurement of the soft term, which includes contributions from the underlying event, multi-parton interactions, and physics objects below analysis threshold.

Unless otherwise noted, the topological selection criteria are identical to the Run-I analysis. One lepton and one hadronically decaying tau are required, and there must be at least two jets with a significant separation in η as expected for VBF production. Some additional topological cuts are applied to suppress backgrounds while retaining most of the signal. The VBF category is intentionally defined loosely since the discrimination of signal from background is meant to be handled by the BDTs. The event selection is summarized in Table 8.9.

8.2.2 Emulation of High-Luminosity LHC conditions

The method of pile-up emulation follows the same procedure used by the $H \to WW^*$ projection for ECFA 2013 [117]. The approach is to port the existing Run-I analysis to HL-LHC conditions by overlaying pile-up jets on the 8 TeV samples, degrading the hard-scatter (HS) jet and $E_{\text{T}}^{\text{miss}}$ resolution, and propagating the impact of this to the analysis.

Table 8.9: Event selection and categorization criteria. The $m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}})$ requirement is relaxed to avoid signal loss due to the degradation of the $E_{\text{T}}^{\text{miss}}$ resolution at high $\langle\mu\rangle$.

Type	Selection
$\tau_\ell\tau_{\text{had}}$ preselection	exactly one identified and isolated lepton (e, μ)
	exactly one identified tau
	opposite sign lepton and tau
	no additional leptons passing loosened identification criteria
	no jets passing the b-tagging criteria
	$m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}}) < 100$ GeV
VBF categorization	leading jet with $p_{\text{T}} > 50$ GeV
	any additional jet with $p_{\text{T}} > 30$ GeV
	$|\Delta\eta(\text{lead jet}, \text{sub-lead jet})| > 3.0$
	$m_{\tau\tau}^{\text{vis.}} > 40$ GeV

8.2.2.1 Performance assumptions

The Run-I triggering thresholds and efficiency are assumed. The ATLAS trigger has many upgrades planned to mitigate the impact of higher instantaneous luminosity, including the New Small Wheel, hardware trackers, and finer L1 EM-calorimeter granularity. Improvements to the L1 granularity will especially improve triggers with electrons and taus, since shower shape variables can be built to discriminate against QCD jets better than existing isolation variables. It is then deemed unnecessary to consider scenarios of significant trigger efficiency loss in detail.

It is also assumed that the lepton and tau reconstruction and identification efficiencies are equivalent to that observed in the 2012 data. This is chosen because the detector upgrades for the HL-LHC aim for achieving a performance similar to Run-I despite the harsher pile-up conditions. Hard-scatter jets from the 8 TeV samples are smeared to emulate the reconstructed jet resolution at HL-LHC conditions [118]. The jet smearing is propagated to the $E_\mathrm{T}^\mathrm{miss}$ calculation. The soft-term resolution of the $E_\mathrm{T}^\mathrm{miss}$ is smeared: 33 MeV per unit of $\langle\mu\rangle$, which is derived from high pile-up Z/γ^* simulation.

8.2.2.2 Pileup emulation

Pileup jets are inserted into the event according to rates recommended by the ATLAS projections [118]. For $\langle\mu\rangle = 140$ and jet $p_\mathrm{T} \geq 30$ GeV, the rate is 2.4 additional pile-up jets per event. The kinematics of the inserted pile-up jets are derived from high pile-up Z/γ^* simulation. Templates are built for p_T and η and are then randomly sampled for each inserted pileup jet.

The hard-scattering jets are assumed to have a reconstruction efficiency of the track-confirmation algorithm recommended by the existing ATLAS projections. The pile-up jets are assumed to have the pile-up suppression efficiency of the 2012 JVF algorithm, which is 98% within the tracking volume [66]. No degradation of pile-up jet suppression is assumed because techniques already exist which outperform the JVF tagger. The pile-up suppression is assumed to not depend on p_T and η other than at the boundary of the tracker. The insertion of pileup jets is then propagated to the $E_\mathrm{T}^\mathrm{miss}$ calculation.

A new tracker in the forward region $|\eta| > 2.5$ is being considered for Phase-II ATLAS upgrades. To emulate the impact of pile-up jet suppression with a forward tracker, the analysis is re-run with forward pile-up jet rejection imposed by hand. Since the scope of the forward tracker is uncertain, a range of coverage and performance is considered: coverage of $|\eta| < 3.0$, 3.5, and 4.0 and pile-up jet suppression of 50%, 75%, and 90% with negligible signal (hard-scattering jets) efficiency loss is evaluated. For comparison, a pile-up jet suppression of 90% with negligible signal efficiency loss is comparable to the performance of the JVF tagger within the existing tracker [66].
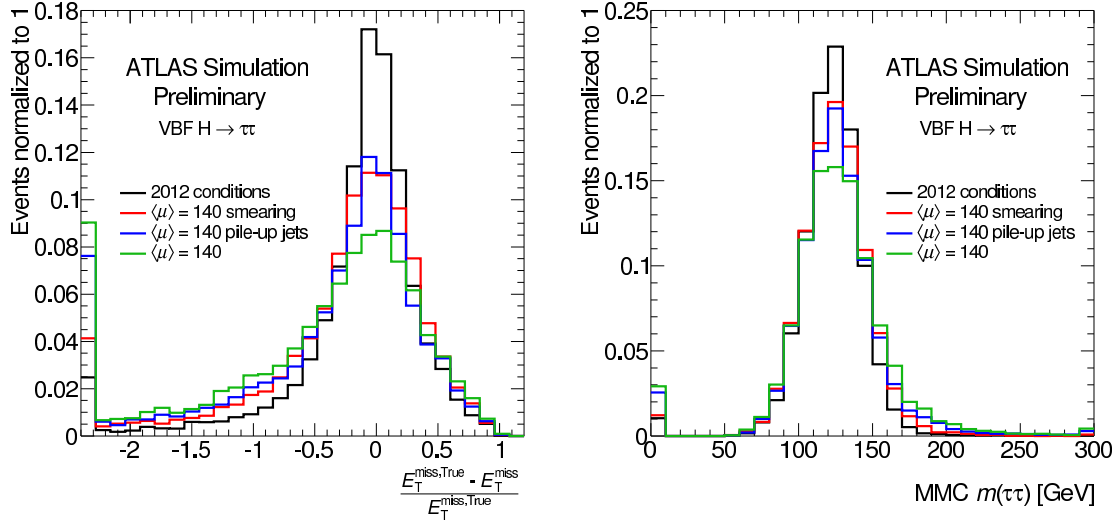
Figure 8.11: Degradation of $E_\mathrm{T}^\mathrm{miss}$-related observables at HL-LHC conditions for VBF $H \to \tau_\ell\tau_\mathrm{had}$: the $E_\mathrm{T}^\mathrm{miss}$ resolution (left) and reconstructed $m_{\tau\tau}^\mathrm{MMC}$ (right) [111]. The underflow of the $m_{\tau\tau}^\mathrm{MMC}$ shows the fraction of events which fail the mass reconstruction.

The method of randomly inserting pile-up jets into existing events does not consider correlations between pile-up jets. This could be problematic because the p_T imbalance of the event could be overstated, which would propagate to unphysical biases of quantities which are sensitive to balance like $E_\mathrm{T}^\mathrm{miss}$. The impact of this has been assessed by overlaying entire truth pile-up events onto the existing hard scatter events. The results are comparable.

8.2.2.3 Impact on observables

The VBF $H \to \tau\tau$ analysis relies on two jets with large $\Delta\eta$ to describe the VBF topology and on $E_\mathrm{T}^\mathrm{miss}$ to describe di-τ decays. The presence of forward pile-up jets can then be expected to degrade the sensitivity of the analysis because they will cause migration of background events into the VBF category, and because they will bias the $E_\mathrm{T}^\mathrm{miss}$ calculation.

The contamination of pile-up jets in the VBF category for the total background is significant. Most events (72%) have a sub-lead pile-up jet, and nearly half (42%) have a lead pile-up jet. These pile-up jets are especially problematic because they are typically forward, thus any event with pile-up lead and sub-lead jets in opposite hemispheres will have $|\eta_{j1} - \eta_{j2}| > 4.8$.

Degradation of the $E_\mathrm{T}^\mathrm{miss}$ and MMC $m(\tau\tau)$ under the high-luminosity conditions is shown in Fig. 8.11 for simulated VBF $H \to \tau\tau$. Both observables are degraded by jet and $E_\mathrm{T}^\mathrm{miss}$ smearing and by the presence of forward pile-up jets biasing the $E_\mathrm{T}^\mathrm{miss}$ calculation.

8.2.3   Analysis

8.2.3.1   Boosted decision tree training

A multi-variate analysis approach is used by training BDTs to discriminate signal from background. It is trained using all backgrounds scaled to their respective cross-sections against the total (ggF + VBF) signal shapes. The same training parameters and input variables as the Run-I analysis are used, and the input variables are listed in Table 8.10.

BDTs are trained for a variety of forward tracker coverages ($|\eta| < 3$, $|\eta| < 3.5$ and $|\eta| < 4$) and pile-up rejection values (50%, 75% and 90%). Fig. 8.12 shows the efficiency for rejecting the background versus the efficiency for selecting the signal for the scenario of 90% forward pile-up rejection. For a given signal efficiency, the background rejection improves with larger coverage.

8.2.3.2   Kinematic distributions

Predicted signal and background BDT input distributions as well as basic object kinematics are shown in Figures Fig. 8.13, Fig. 8.14 and Fig. 8.15. Signal and background are predicted with the same methods as the Run-I analysis [2], including the data-driven prediction of the dominant $Z \to \tau\tau$ and fake backgrounds. The resulting BDT score is presented in Figure Fig. 8.16.

8.2.4   Results

Yields for signal and background in the high BDT score bins are shown in Table 8.11. As in the Run-I analyses, $Z \to \tau\tau$ and fakes are the dominant backgrounds in the most signal-like regime. The binning of the BDT is optimized to maximize the expected sensitivity.

Table 8.10: Discriminating variables used for the BDT training.

| Variable	Definition
$\Delta R(\tau_{\text{had}}, \ell)$	Separation of the lepton and τ_{had}
$m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}})$	Transverse mass of the lepton and $E_{\text{T}}^{\text{miss}}$
$E_{\text{T}}^{\text{miss}}\phi$-centrality	Centrality of the $E_{\text{T}}^{\text{miss}}$ between the lepton and τ_{had}
MMC $m_{\tau\tau}$	$\tau\tau$ mass estimator
$m_{j1,j2}$	Invariant mass of the 2 leading jets
$\eta_{j1} \times \eta_{j2}$	Product of the ηs of the two leading jets
$|\eta_{j1} - \eta_{j2}|$	Absolute difference ηs of the two leading jets
ℓ η-centrality	Centrality of the lepton between the two leading jets
$p_{\text{T}}^{\text{Total}}$	$|\vec{p}_{\text{T}}^{\,\ell} + \vec{p}_{\text{T}}^{\,\tau_h} + \vec{p}_{\text{T}}^{\,j1} + \vec{p}_{\text{T}}^{\,j2} + \vec{E}_{\text{T}}^{\text{miss}}|$
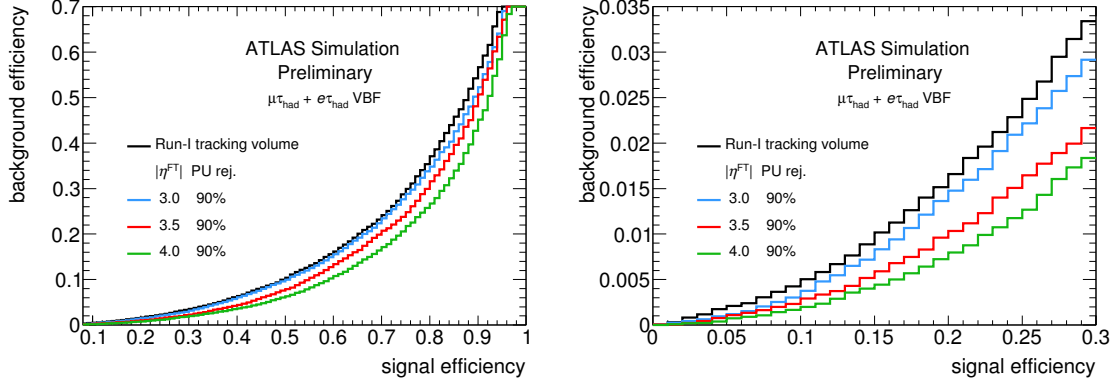
Figure 8.12: Signal efficiency versus background efficiency for scenarios of generic forward tracker coverage and rejection power (left) and zoomed in to lower signal efficiency (right) [111]. A BDT is trained in the VBF category for each scenario.

8.2.4.1 Uncertainties assumptions

When calculating the sensitivity of the analysis, three scenarios of background uncertainties and two scenarios of theory uncertainties are considered. The theory uncertainties are varied from no theory uncertainties to Run-I theory uncertainties, which are as large as 6% (30%) for the VBF (ggF) Higgs production modes [2]. The experimental signal uncertainty is fixed at 5% accounting for experimental sources such as jet energy scale uncertainties. The experimental background uncertainties are varied to 10% and 5% of the prediction, and they are treated as uncorrelated between backgrounds and between bins of the BDT score.

The projected sensitivity is shown in Table 8.12. The two scenarios of background uncertainties

Table 8.11: Yields for signal and background in the VBF category and in the most sensitive BDT bins, as shown in Fig. 8.16.

process	VBF category	third highest bin	second highest bin	highest bin
VBF $H \to \tau\tau$	8970	114	147	206
ggF $H \to \tau\tau$	16410	44	46	39
$Z \to \tau\tau$	1682400	875	720	514
fakes	2959800	205	190	155
$t\bar{t}$	191400	100	20	< 20
other	198600	< 20	< 20	< 20
signal	25380	158	193	245
background	5032200	1180	930	669
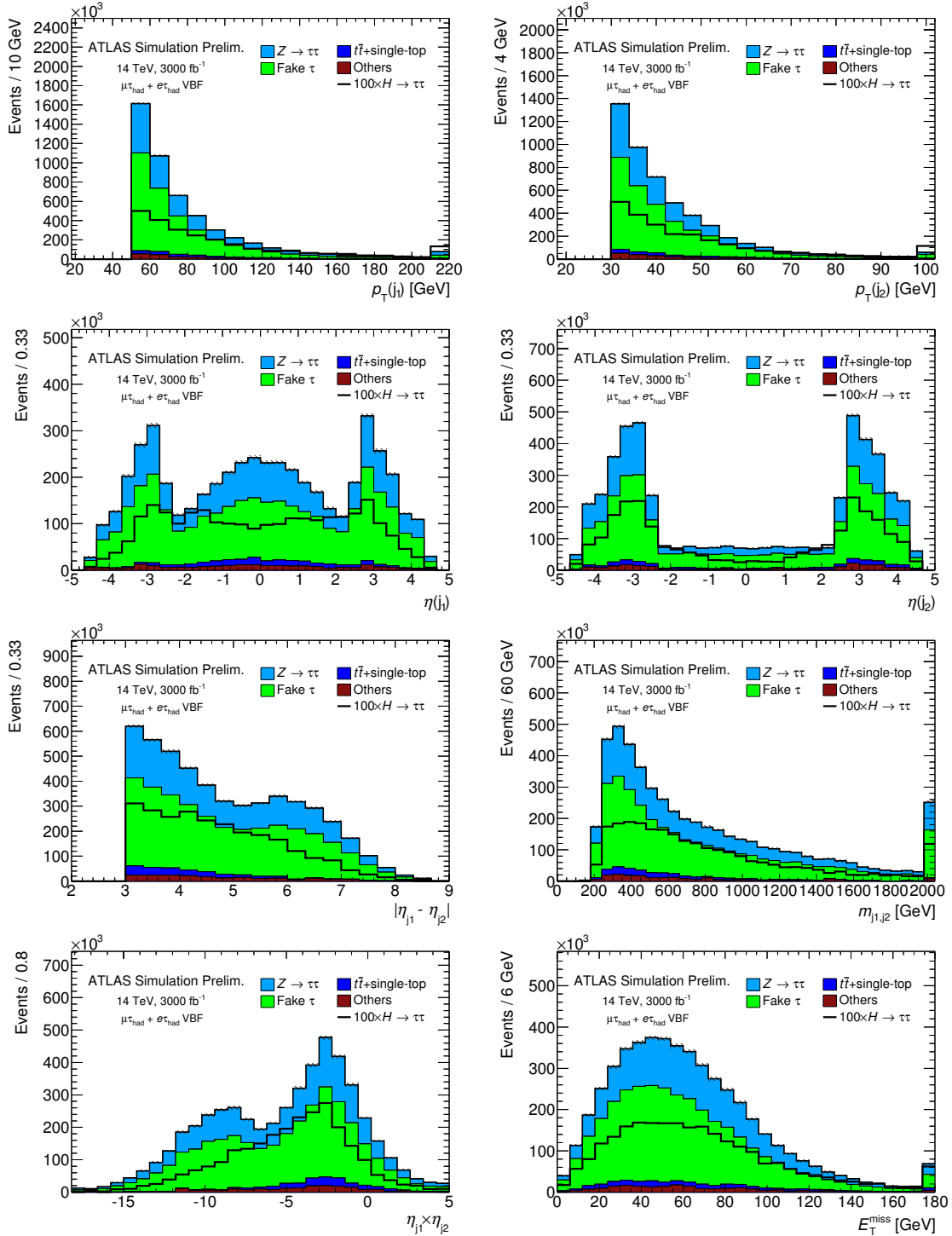
Figure 8.13: Signal and background HL-LHC predictions of (a) leading jet p_{T}, (b) sub-leading jet p_{T}, (c) leading jet η, (d) sub-leading jet η, (e) $\Delta\eta_{jj}$, (f) m_{jj}, (g) $\eta_{leadjet} \times \eta_{sub-leadjet}$ and (h) $E_{\mathrm{T}}^{\mathrm{miss}}$ [111]. The last bin contains the overflow events.
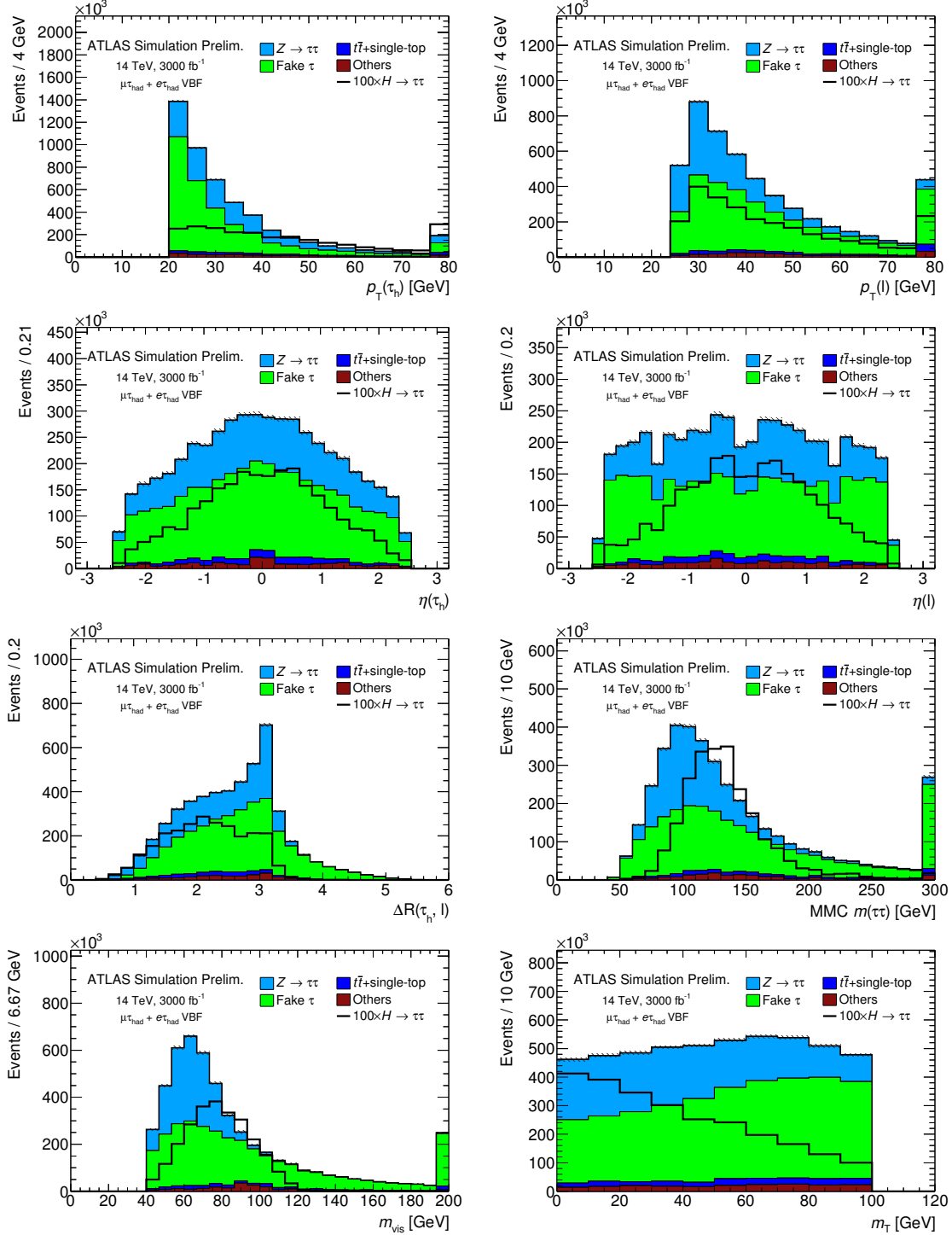
Figure 8.14: Signal and background HL-LHC predictions of (a) $p_{\mathrm{T}}(\tau_{\mathrm{had}})$, (b) $p_{\mathrm{T}}(\text{lepton})$, (c) $\eta(\tau_{\mathrm{had}})$, (d) $\eta(\text{lepton})$, (e) $\Delta R(\tau_{\mathrm{had}}, \text{lepton})$, (f) MMC (g) $m_{\tau\tau}^{\mathrm{vis}}$. and (h) $m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ [111]. The last bin contains the overflow events.
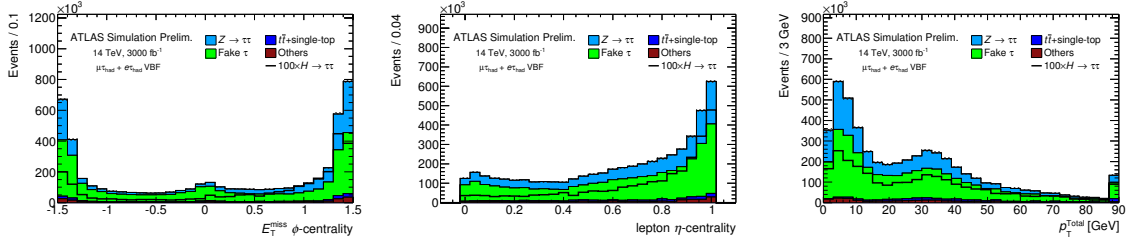
Figure 8.15: Signal and background HL-LHC predictions of (a) $E_T^{\text{miss}}\phi$-centrality, (b) lepton η-centrality and (c) p_T^{Total} [111]. The last bin contains the overflow events.
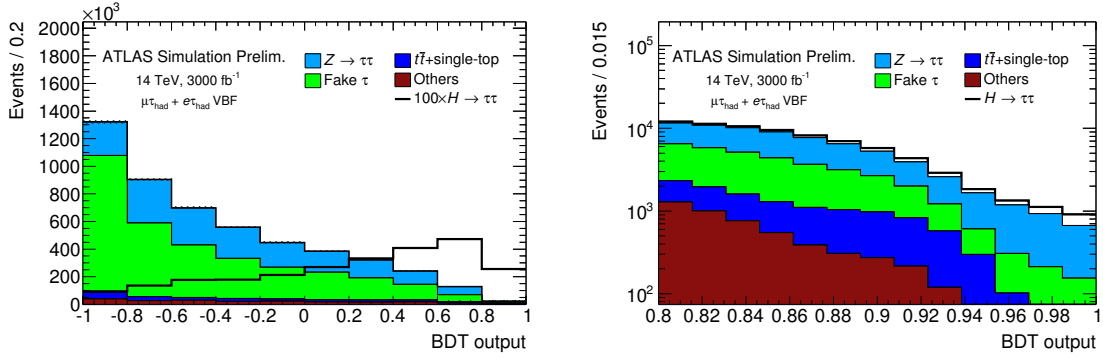

Figure 8.16: Signal and background HL-LHC predictions of the BDT spectrum in the (a) full range and (b) highest bins range [111]. Signal and background are overlaid in (a) and stacked in (b).

and two scenarios of theory uncertainties are shown. The sensitivity of the projection is driven by the uncertainty on the background prediction. For $\sigma_B^{\text{syst.}} = 10\%$, the projected uncertainty on μ with no signal theory uncertainties is 0.24. For $\sigma_B^{\text{syst.}} = 5\%$, this projected uncertainty is 0.13.

The impact of pile-up jet rejection in the forward region is also evaluated as an example of the impact of a forward tracker, and results are given in Table 8.13. Multiple scenarios of $|\eta|$ coverage and pile-up jet rejection are considered. For these scenarios, negligible loss of HS jets to forward pile-up jet rejection is assumed, a 10% systematic uncertainty is assumed for backgrounds, a 5% experimental

Table 8.12: Uncertainty on the signal strength ($\Delta\mu$) for different scenarios of background uncertainties and signal theory uncertainties.

		current $\sigma_S^{\text{theo.}}$	no $\sigma_S^{\text{theo.}}$
$\sigma_B^{\text{syst.}}$	$\sigma_S^{\text{syst.}}$	$\Delta\mu$	$\Delta\mu$
10%	5%	0.25	0.24
5%	5%	0.16	0.13

systematic uncertainty is assumed for signals, and theoretical uncertainties on signals are ignored.

8.2.5 Conclusions

The projection of the Standard Model $H \to \tau\tau$ analysis to the High Luminosity LHC (HL-LHC) running conditions with 14 TeV pp collisions, 3000 fb^{-1} delivered integrated luminosity, and an average number of overlapping pp collisions $\langle\mu\rangle = 140$ is performed. The VBF $\tau_\ell\tau_{\text{had}}$ $(\ell = e, \mu)$ analysis category is considered, and the uncertainty on the signal strength (μ) is projected to be 24% when theory uncertainties are ignored and 10% (5%) background (signal) uncertainties are assumed. The projected uncertainty could be reduced significantly if pile-up jets outside the current tracking volume could be rejected similar to pile-up jet rejection within the tracking volume in 2012. The uncertainty on μ is projected to be $8 - 18\%$ depending on the scenario of forward tracker coverage and pile-up jet rejection.

Table 8.13: Uncertainty on the signal strength (Δμ) for different scenarios of forward tracking.

forward pileup jet rejection	50%	75%	90%
forward tracker coverage		$\Delta\mu$	
Run-I tracking volume		0.24	
$|\eta| < 3.0$	0.18	0.15	0.14
$|\eta| < 3.5$	0.18	0.13	0.11
$|\eta| < 4.0$	0.16	0.12	0.08