A self-calibrating, double-ratio method to test tau lepton universality in W boson decays at the LHC

S. Dysch and T. R. Wyatt.

Particle Physics Group, Department of Physics and Astronomy, University of Manchester, UK.

Received: 26/10/2019

Abstract. Measurements in $W^+W^-$ events at LEP2 and in $B$ hadron semileptonic decays at B factories and LHCb provide intriguing hints of a violation of lepton universality in the charged current coupling of tau leptons relative to those for electrons and muons. We propose a novel, self-calibrating method to test tau lepton universality in $W$ boson decays at the LHC. We compare directly the ratio of the numbers of selected $ℓτ$had and $eτ$ final states in di-leptonic $tt$ events with that in $Z/γ^* → τ^+τ^-$ events. Here $ℓ = e$ or $μ$ and $τ$had is a candidate semi-hadronic tau decay. This “double-ratio” cancels to first order sensitivity to systematic uncertainties on the reconstruction of $e$, $μ$, and $τ$ leptons, thus improving very significantly the precision to which tau lepton universality can be tested in $W$ boson decay branching ratios at the LHC. Using particle-level Monte Carlo events, and a parameterised simulation of detector performance, we demonstrate the effectiveness of the method and estimate the most significant residual sources of uncertainty arising from experimental and phenomenological systematics. Our studies indicate that a single experiment precision on the tau lepton universality test of around 1.4% is achievable with a data set of $\int dt = 140 \text{ fb}^{-1}$ at $\sqrt{s} = 13 \text{ TeV}$. This would improve significantly upon the precision of 2.5% on the four-experiment combined LEP2 measurements. If the central value of the proposed new measurement were equal to the central value of the LEP2 measurement this would yield an observation of $2.5\sigma$ on the four-experiment combined LEP2 measurements. If the central value of the proposed new measurement were equal to the central value of the LEP2 measurement this would yield an observation of $2.5\sigma$ on the four-experiment combined LEP2 measurements. If the central value of the proposed new measurement were equal to the central value of the LEP2 measurement this would yield an observation of $2.5\sigma$ on the four-experiment combined LEP2 measurements.

1 Introduction

An important feature of the Standard Model (SM) of particle physics is Lepton Universality (LU): the idea that the Electroweak (EW) couplings of the leptons are identical in each fermion generation. For the Neutral Current (NC) interactions, mediated in the SM by the $γ$ and $Z$ bosons, the validity of LU for the three flavours of charged leptons ($e$, $μ$, $τ$) has been demonstrated at the scale of the $Z$ boson mass $m_Z$ at LEP1 and SLC to high precision [1]. For example, the ratios of the leptonic partial widths ($Γ$) or branching fractions ($B$) of the $Z$ boson are:

\begin{align}
\frac{Γ_{μμ}}{Γ_{ee}} &= \frac{B(Z/γ^* → μ^+μ^-)}{B(Z/γ^* → e^+e^-)} = 1.0009 ± 0.0028, \\
\frac{Γ_{ττ}}{Γ_{ee}} &= \frac{B(Z/γ^* → τ^+τ^-)}{B(Z/γ^* → e^+e^-)} = 1.0019 ± 0.0032,
\end{align}

which are consistent with LU to a precision of around three per mille [1]. Measurements of the leptonic asymmetry parameters $A_ℓ$, from forward-backward asymmetries, the left-right asymmetry, and the tau polarisation and its asymmetry:

\begin{align}
A_e &= 0.1514 ± 0.0019, \\
A_μ &= 0.1456 ± 0.0091,
A_τ &= 0.1449 ± 0.0040,
\end{align}

are also consistent with LU, albeit at the precision of a few percent [1].

In contrast, for the Charged Current (CC) interactions, mediated in the SM by the $W^±$ bosons, the experimental measurements of leptonic branching ratios are less precise. A direct test of LU at the scale of the $W$ boson mass $m_W$ at LEP2 can be made from the ratios of the leptonic branching fractions of the $W$ boson. The ratio [2]:

\begin{align}
\frac{B(W → μν)}{B(W → eν)} = 0.993 ± 0.019
\end{align}

is consistent with $e-μ$ universality to a precision of around two percent. However a hint at the possible violation of LU in the CC couplings of the $τ$ is present in the ratio [2]:

\begin{align}
R(W) &≡ \frac{B(W → τν)}{B(W → eν)} = 1.066 ± 0.025,
\end{align}

where $B(W → ℓν)$ is the average of the branching fractions for $W → eν$ and $W → μν$. This result deviates from
the assumption of $\tau$-$\ell$ universality by 2.6$\sigma$ (standard deviations).

At lower mass scales $e$-$\mu$ universality is tested very precisely, for example in leptonic $\tau$ decays the ratio of the muonic and electronic partial widths is measured to be \(3\):

\[
\frac{\Gamma(\tau^+ \rightarrow \mu^- \nu_\mu \bar{\nu}_\tau)}{\Gamma(\tau^+ \rightarrow e^- \nu_e \bar{\nu}_\tau)} = 0.9762 \pm 0.0028,
\]

which is consistent with the SM prediction including mass effects of 0.9726. $e$-$\mu$ universality is also tested in the decays of charged kaons \(3\):

\[
\frac{\Gamma(K^+ \rightarrow e^- \bar{\nu}_e)}{\Gamma(K^+ \rightarrow \mu^- \bar{\nu}_\mu)} = (2.488 \pm 0.009) \times 10^{-5},
\]

which is consistent with the SM prediction \(4\) of 2.477 $\times 10^{-5}$. In the decay of $B$ hadrons, $\tau$-$\ell$ universality can be tested by measurements of the branching fractions for the exclusive decays of $B^0 \rightarrow \tau^- \nu_\tau D^+$ and $B^+ \rightarrow \tau^- \nu_\tau D^{++}$ expressed as ratios to the branching fractions for the exclusive decays $B^0 \rightarrow \ell^- \bar{\nu}_\ell D^+$ and $B^+ \rightarrow \ell^- \bar{\nu}_\ell D^{++}$, respectively. Systematic uncertainties due to hadronic effects largely cancel in these ratios. A combination \(5\) of the results from the BaBar \(6,7\), Belle \(8,9,10,11,12\), and LHCb \(13,14,15\) experiments yields the results:

\[
R(D) \equiv \frac{B(B^0 \rightarrow \tau^- \nu_\tau D^+)}{B(B^0 \rightarrow \ell^- \bar{\nu}_\ell D^+)} = 0.340 \pm 0.027 \pm 0.013,
\]

\[
R(D^*) \equiv \frac{B(B^0 \rightarrow \tau^- \nu_\tau D^{++})}{B(B^+ \rightarrow \ell^- \bar{\nu}_\ell D^{++})} = 0.295 \pm 0.011 \pm 0.008.
\]

These measured values exceed the SM predictions (calculated assuming LU) of:

\[
R(D) = 0.299 \pm 0.011,
\]

\[
R(D^*) = 0.300 \pm 0.008,
\]

by 1.4$\sigma$ and 2.5$\sigma$ respectively (see \(5\) and the references contained therein). Taking into account correlations between the measurements, the combined discrepancy with regard to the SM predictions corresponds to 3.1$\sigma$ \(5\).

Possible Beyond the Standard Model (BSM) explanations have been proposed for the potential violation of LU in $R(D)$ and $R(D^*)$. A Leptoquark (LQ) that couples more strongly to the $\tau$ than to $e$ or $\mu$ could contribute at tree level to the decays $B \rightarrow \tau^- \nu_\tau D^{(*)+}$. For example, in \(16\) a charge 2/3 scalar LQ with $br$ and $cv$ Yukawa couplings is able to accommodate the measured central values of $R(D)$ and $R(D^*)$ without introducing an unacceptable level of flavour changing neutral current processes involving the first two generations of quarks and leptons. Interestingly, a possible link between the LEP2 measurement of $R(W)$ and the $B^0 \rightarrow \tau^- \nu_\tau D^{(*)+}$ excess has received little attention in the literature. We note that $R(W)$ might receive contributions at loop level from LQ that couples preferentially to the $\tau$. For example, if $cs$ and $st$ Yukawa couplings were added to the charge 2/3 LQ scenario of \(16\) it could produce an enhancement in $B(W \rightarrow \tau\nu)$. Such loop-level contributions might naturally lead to a smaller fractional deviation from LU in $R(W)$ than in $R(D)$ and $R(D^*)$, but this would depend, obviously, on the sizes of the assumed couplings.

Clearly it is important to improve on the precision of the measurements of $R(W)$ [currently 2.5$\sigma$], $R(D)$ [currently 11$\%$] and $R(D^*)$ [currently 5$\%$]. The precision of the LEP2 measurement of $R(W)$ was dominated by the limited number of available $W^+W^-$ events. A future high energy, high luminosity $e^+e^-$ collider, at which an improved $R(W)$ measurement could be performed, is likely to be decades away.

Measurements at hadron colliders that are sensitive to $R(W)$ usually rely on the identification of $\tau$ lepton decays in which the visible final state is hadronic ($\tau_{\text{had}}$). The current best published hadron collider measurements typically have small statistical uncertainties, but are dominated by large systematic uncertainties that render them uncompetitive with the LEP2 measurement of $R(W)$. For example, a measurement of the inclusive single $W \rightarrow \tau_{\text{had}}\nu$ cross section in $pp$ collisions at 7 TeV by ATLAS \(17\) has a relative uncertainty of 15$\%$, which is dominated by systematic uncertainties on the efficiency to trigger on and select events containing $\tau_{\text{had}}$ candidates.

Also of interest in this context are measurements at hadron colliders of the rate of events containing top quark pairs ($tt$), especially in the di-lepton final state. For the purpose of determining $R(W)$ events containing top quarks may be regarded as a convenient source of on-mass-shell $W$ bosons. In the absence of non-SM decay mechanisms for the top quark the branching fraction $B(t \rightarrow b\ell\nu)$ may be reinterpreted as the branching fraction $B(W \rightarrow \tau\nu)$. Relative to measurements of the inclusive single $W \rightarrow \tau_{\text{had}}\nu$ cross section, measurements using di-leptonic $tt \rightarrow b\ell\ell$ events eliminate systematic uncertainties associated with the trigger efficiencies for $\tau_{\text{had}}$ candidates, because single-$\ell$ ($e$ or $\mu$) triggers can be used. In addition, non-$tt$ backgrounds can be almost completely eliminated from the $tt \rightarrow b\ell\ell$ final state by employing $b$-jet tagging and the presence of the $\ell$ candidate. This means that the background from misidentified hadronic jets to the $\tau_{\text{had}}$ signature in the $tt \rightarrow b\ell\ell$ final state is likely to be significantly smaller than that in the inclusive single $W \rightarrow \tau_{\text{had}}\nu$ final state. However, systematic uncertainties associated with the $\tau_{\text{had}}$ candidate (identification efficiency, background and energy scale) still contribute directly to the measured rate of $tt \rightarrow b\ell\ell$ events. For example, in a measurement by ATLAS \(18\) in di-leptonic $tt$ events the systematic uncertainty on the branching fraction $B(t \rightarrow b\ell\nu)$ is around 7.5$\%$. In a measurement by CMS of the cross section for the $tt \rightarrow b\ell\ell$ final state \(19\), the systematic uncertainty is around 9.5$\%$, to which the combined contribution from the identification efficiency...
(6.0%), background (4.3%), and energy scale (2.5%) for \( \tau_{\text{had}} \) candidates was 7.8%. In the above-mentioned best currently published measurements from ATLAS [18] and CMS [19] based on the \( tt \rightarrow b\bar{b}\ell\tau_{\text{had}} \) final state the systematic uncertainties associated with the \( \tau_{\text{had}} \) candidate are around a factor of three greater than total uncertainty of 2.5% on the LEP2 measurement of \( R(W) \).

We propose here a novel, self-calibrating, “double-ratio” method that will allow \( R(W) \) to be measured using top quark pair \( (tt \rightarrow b\bar{b}W^+W^-) \) and \( Z/\gamma^* \rightarrow \tau^+\tau^- \) events at the LHC with a target precision of around 1%, which would improve significantly upon the LEP2 measurements. We define the ratio:

\[
R(bbWW) \equiv \frac{N(tt \rightarrow b\bar{b}\ell\tau_{\text{had}})}{N(tt \rightarrow b\bar{b}e\mu)},
\]

where \( N(tt \rightarrow b\bar{b}\ell\tau_{\text{had}}) \) and \( N(tt \rightarrow b\bar{b}e\mu) \) are the numbers of observed candidate events in the \( tt \rightarrow b\bar{b}\\ell\tau_{\text{had}} \) and \( tt \rightarrow b\bar{b}e\mu \) final states, respectively. We define also the ratio:

\[
R(Z) \equiv \frac{N(Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}})}{N(Z \rightarrow \tau\tau \rightarrow e\mu)},
\]

where \( N(Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}}) \) and \( N(Z \rightarrow \tau\tau \rightarrow e\mu) \) are the numbers of observed candidate events in the \( Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}} \) and \( Z \rightarrow \tau\tau \rightarrow e\mu \) final states, respectively. We then define the double ratio:

\[
R(WZ) \equiv \frac{R(bbWW)}{R(Z)} = \frac{N(tt \rightarrow b\bar{b}\ell\tau_{\text{had}}) \times N(Z \rightarrow \tau\tau \rightarrow e\mu)}{N(tt \rightarrow b\bar{b}e\mu) \times N(Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}})}.
\]

From an experimental perspective we note that the ratios \( R(bbWW) \) and \( R(Z) \) are designed to have approximately the same sensitivity to systematic uncertainties on the identification of e, \( \mu \) and \( \tau_{\text{had}} \) candidates, and on the efficiencies of the single-\( \ell \) triggers. Therefore, in the double ratio \( R(WZ) \) these systematic uncertainties cancel to first order. This cancellation is not necessarily perfect for the following reasons:

- The distributions in transverse momentum \( (p_T) \) and pseudorapidity \( (\eta) \) \([20]\) of the leptons \( (e, \mu \) and \( \tau \) are significantly different in the \( tt \rightarrow b\bar{b}W^+W^- \) and \( Z/\gamma^* \rightarrow \tau^+\tau^- \) signal samples. Systematic uncertainties on lepton identification and single-\( \ell \) trigger efficiencies are not necessarily fully correlated across all \( p_T \) and \( |\eta| \) bins.
- The levels of background from misidentified hadronic jets in the \( \tau_{\text{had}} \) candidates in the \((tt \rightarrow b\bar{b}\ell\tau_{\text{had}} \) and \( Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}} \) samples are not necessarily identical.
- The level and nature of the non-\( tt \) background in the \( tt \rightarrow b\bar{b}\ell\tau_{\text{had}} \) and \( tt \rightarrow b\bar{b}e\mu \) samples will not necessarily be identical. Similarly, the level and nature of the non-\( Z \) boson background in the \( Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}} \) and \( Z \rightarrow \tau\tau \rightarrow e\mu \) samples will not necessarily be identical.

From a phenomenological perspective the double ratio \( R(WZ) \) exploits the fact that LU has been precisely verified experimentally in the Z boson branching fractions (see equation 1). Therefore, any non-SM effects should affect \( R(WZ) \) primarily through \( R(bbWW) \), which is sensitive to \( R(W) \).

The rest of this paper is organised as follows. In section 2 we describe a Monte Carlo (MC) study of the proposed analysis method employing a simple parameterised simulation of detector performance. In section 3 we present our summary and conclusions.

2 A Monte Carlo study of the proposed double-ratio analysis method

Our study uses particle-level MC events for the various physics processes of relevance. In section 2.1 we describe the simple parameterised simulation of detector performance we use in the study of the proposed double-ratio analysis method. We describe also the variations in detector performance we consider in the study of experimental systematic uncertainties. In section 2.2 we describe the MC generators used and the potential sources of phenomenological systematic uncertainties that we have considered in our study. In section 2.3 we describe the candidate event selection criteria we employ for the four signal candidate event samples used in the double-ratio method. In the selection of \( Z/\gamma^* \rightarrow \tau^+\tau^- \) candidates we propose a novel selection variable, \( m_T^* \), that improves the discrimination power against the dominant backgrounds, such as \( tt \), diboson production, as well as events containing a leptonically decaying vector boson plus a QCD jet that is misidentified as a \( \tau_{\text{had}} \) candidate \((W+\text{jet})\). In section 2.4 we evaluate the size and composition of the four selected candidate event samples. The expected numbers of events are given for an integrated luminosity at \( \sqrt{s} = 13 \text{ TeV} \) of \( \int L dt = 140 \text{ fb}^{-1} \), which corresponds approximately to that available for physics analysis in ATLAS and CMS at the end of LHC run 2 \([21,22]\). In section 2.5 we evaluate the sensitivity of the measured double ratio \( R(WZ) \) to the physical quantity of interest \( R(W) \). In section 2.6 we present the effect of systematic uncertainties on the ratios \( R(bbWW), R(Z), R(WZ), \) and \( R(W) \). We thus demonstrate that in the double-ratio \( R(WZ) \) there is a high degree of cancellation in the experimental systematic uncertainties that have dominated previous related analyses at hadron colliders. e evaluate the residual systematic uncertainties on \( R(WZ) \) arising from the effects discussed in sections 2.1 and 2.2.
variations in detector performance we consider in the study of systematic uncertainties.

Clearly, our aim here is not to produce a completely accurate simulation of the data from either the ATLAS or CMS detectors. Nevertheless, we base our parameterisations of the detection of leptons and jets on published measurements of LHC detector performance and their associated uncertainties. This approach enables us to demonstrate some of the principle benefits of the proposed double-ratio method and allows us to investigate within a simple and controlled framework the principal sources of residual systematic uncertainty to which the method is sensitive.

The efficiencies associated with the reconstruction, identification, and triggering of high $p_T$, isolated lepton candidates are typically determined by the ATLAS and CMS collaborations using “tag and probe” measurements on $Z \rightarrow ee, \mu\mu, \tau\tau$ events in both MC simulations and the real data. Systematic uncertainties are usually quoted on the “scale factors” employed to correct MC simulations to provide an accurate description of the real data.

As noted in section 1, the ratios $R(bbWW)$ and $R(Z)$ are designed to have approximately the same sensitivity to systematic uncertainties on the identification of $e$, $\mu$ and $\tau_{\text{had}}$ candidates, and on the efficiencies of the single-\ell triggers. In the double ratio $R(WZ)$, these systematic uncertainties cancel to first order. Therefore, in our study we give particular attention to the $p_T$ and $\eta$ dependence of efficiencies and to any potential $p_T$ and $\eta$ dependence in the associated systematic uncertainties. In general we expect algorithms that are designed to have $p_T$- and $\eta$-independent efficiencies to have smaller $p_T$- and $\eta$-dependent systematic uncertainties.

Figure 1 shows as a function of $p_T$ and $|\eta|$ the overall efficiencies we assume for the reconstruction, identification, and isolation criteria for prompt $e, \mu$, and $\tau_{\text{had}}$ candidates. Table 1 gives a summary of the sources of systematic uncertainty on the reconstruction of leptons and jets considered in this study, stating separately the assumed size of the $p_T/\eta$-independent and $p_T/\eta$-dependent systematic uncertainties. The reasoning behind these choices is given below. Figure 2 shows the $p_T$-dependent systematic uncertainties on the efficiencies we assume for the reconstruction, identification, and isolation criteria for prompt $e, \mu$, and $\tau_{\text{had}}$ candidates. Table 2 gives a summary of the other sources of systematic uncertainty considered in this study.

### 2.1.1 Simulation of muon candidates

The efficiencies and systematic uncertainties associated with the reconstruction and identification of high $p_T$, isolated muon candidates have been presented by the ATLAS [23] and CMS [24] collaborations. In both experiments the muon reconstruction efficiency is around 99% and is independent of $p_T$ and $|\eta|$, except for some well-defined, poorly instrumented regions of both detectors that are usually excluded for precision measurements. We choose to simulate the efficiency for muon identification according to that given for the “Medium” category described in [23]. This gives an efficiency of around 96% for muons with $p_T > 20$ GeV. Systematic uncertainties on the efficiency are around 0.1% for muon $p_T$ around $m_Z/2$, increasing to around 0.5% for $p_T \approx 30$ GeV and $p_T \approx 100$ GeV. We choose to simulate the “Tight” lepton isolation requirement given in [23], which is measured to have an efficiency that is independent of $p_T$ and $|\eta|$ of around 96% with a systematic uncertainty at the per mille level. We assume the efficiency of the isolation criteria for sources of non-prompt muons to be 0.03, based on the range of values given in [23]. We consider a combined systematic uncertainty on the efficiency for muon reconstruction, identification, and isolation. We consider a $p_T/\eta$-independent relative systematic uncertainty of 2%. We generate a $p_T$-dependent systematic uncertainty by modifying the efficiency by a relative amount that varies.
Table 1. Sources of systematic uncertainty, together with the size (in percent, unless stated otherwise) of the $p_T/\eta$-independent and $p_T/\eta$-dependent systematic uncertainties considered in this study. In all expressions $p_T$ is given in units of GeV. See the text for further details.

| Source of systematic uncertainty | $p_T/\eta$-independent | $p_T/\eta$-dependent |
|---------------------------------|-------------------------|----------------------|
| Muon ID & isolation efficiency  | 2.0                     | 0.5 ($\frac{80 - p_T}{50}$) |
| Single-muon trigger efficiency  | 1.0                     | 1.0 (endcap only)     |
| Muon $p_T$ resolution          | 0.15                    | 0.15 (endcap only)   |
| Muon $p_T$ scale               | 0.2                     | 0.2 (endcap only)    |
| Electron ID & isolation efficiency | 2.0                 | 2.0 ($\frac{80 - p_T}{50}$) |
| Single-electron trigger efficiency | 1.0                   | 1 GeV change in $p_T$ threshold |
| Electron $p_T$ resolution      | $\frac{\sigma_{p_T}}{p_T} \rightarrow \frac{\sigma_{p_T}}{p_T} \oplus 0.002$ | vary resolution in endcap only |
| Electron $p_T$ scale           | 0.2                     | 0.2 (endcap only)    |
| $\tau_{\text{had}}$ efficiency | 5                      | 5 ($\frac{130 - p_T}{100}$) |
| $\tau_{\text{had}}$ $p_T$ scale | 1                      | 1 (endcap only)      |
| jet energy scale               | 1                      | —                    |
| $b$-tag efficiency for $b$-, $c$-, light-jets | 1.5, 4, 10 | — |

Table 2. Other sources of systematic uncertainty considered in this study (given in percent). In all expressions, $p_T$ and $m(t\bar{t})$ are given in units of GeV. See the text for further details.

| Source of systematic uncertainty | Size of fractional systematic uncertainty (in %) |
|---------------------------------|-----------------------------------------------|
| Misidentification rates for $\tau_{\text{had}}$ | 10 |
| MJ background in $Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}}$ sample | 5 |
| Ratio of diboson and $Z$ boson cross sections | 3 |
| Ratio of $t\bar{t}$ and $Z$ boson cross sections | 3 |
| $p_T(Z)$ reweighting | 0.5 for $50 < p_T(Z) < 150$ |
|                                 | 1.0 for $p_T(Z) > 150$ |
| $m(t\bar{t})$ reweighting | $10 + 0.018(m(t\bar{t}) - 1000)$ |
| $t\bar{t}$ modelling | Tested with alternative sample |

linearly between 0.5% at $p_T = 30$ GeV and 0% at $p_T = 80$ GeV and above.

We simulate the efficiency of the single muon trigger to be 65% (80%) in the barrel (endcap) region, as has been measured for ATLAS [25,26]. For offline $p_T$ at least 1 GeV above the trigger threshold the efficiency is almost independent of $p_T$. The trigger efficiency is assumed to be similar to the electron trigger efficiency of ATLAS during most of run 2 at $p_T = 26$ GeV [27]. We consider a $p_T/\eta$-independent relative systematic uncertainty of 1%. We generate an $\eta$-dependent systematic uncertainty by modifying the relative efficiency by 1% in the endcap region, whilst the efficiency in the barrel region remains unchanged.

We simulate the resolution in muon $p_T$ by means of a Gaussian with a width of $2.30 \pm 0.15\%$ in the barrel region and $2.90 \pm 0.15\%$ in the endcap region, independent of $p_T$ [23]. We consider a $p_T/\eta$-independent relative systematic uncertainty of 0.2% on the $p_T$ scale [23]. We generate $\eta$-dependent systematic uncertainties by modify-
ifying the resolution and $p_T$ scale only in the endcap region, whilst the values in the barrel region remain unchanged.

### 2.1.2 Simulation of electron candidates

The efficiencies and systematic uncertainties associated with the reconstruction and identification of high $p_T$ isolated electron candidates have been presented by the ATLAS [25] and CMS [29] collaborations. In both experiments the electron reconstruction efficiency is around 98% and is fairly independent of $p_T$ and $|\eta|$, except for some well-defined, poorly instrumented regions of both detectors [30]. A general feature in both ATLAS and CMS is that the efficiency of the commonly used electron identification algorithms tend to have a larger dependence on $p_T$ than is the case for muon identification. We choose to simulate approximately the efficiency for electron identification according to the “Tight likelihood” algorithm of [25]. The efficiency is around 75% at $p_T = 30$ GeV and rises to around 90% for $p_T > 80$ GeV and above [31]. The efficiency is determined with a systematic uncertainty of a few per mille for $p_T > 30$ GeV. We choose to simulate also for electrons the “Tight” lepton isolation requirement given in [25], with an efficiency that is independent of $p_T$ and $|\eta|$ of around 96%. We consider a combined systematic uncertainty on the efficiency for electron reconstruction, identification, and isolation. We consider a $p_T/\eta$-independent relative systematic uncertainty of 2%. We generate a $p_T$-dependent systematic uncertainty by modifying the position of the effective trigger threshold in $p_T$ by 1 GeV.

The efficiency of the single electron trigger has been measured in ATLAS to be around 90% for offline $p_T$ at least 1 GeV above the trigger threshold [32]. We consider a $p_T/\eta$-independent relative systematic uncertainty of 1%. We generate a $p_T$-dependent systematic uncertainty by modifying the position of the effective trigger threshold in $p_T$ by 1 GeV.

The resolution and $p_T$ scale for high $p_T$, isolated electron candidates and the associated systematic uncertainties have been measured by the ATLAS [33] and CMS [29] collaborations. We simulate the resolution in electron $p_T$ by means of a Gaussian with a width $\sigma_{p_T}$ given by $\sigma_{p_T} = 0.16 \sqrt{p_T}$. We consider a $p_T/\eta$-independent systematic uncertainty on the resolution by adding in quadrature a value 0.002 to $\sigma_{p_T}$ [34]. The $p_T$ scale is calibrated with a precision of around 0.2% in both ATLAS and CMS [33,29]. We consider a $p_T/\eta$-independent relative systematic uncertainty of 0.2% on the $p_T$ scale. We generate $\eta$-dependent systematic uncertainties by modifying the resolution and $p_T$ scale only in the endcap region, whilst the values in the barrel region remain unchanged.

Since we apply tight isolation requirements on the candidate leptons, we make the conservative assumption that the identification efficiency for non-prompt electrons from heavy flavour decays is the same as that for prompt electrons. In [35] around 2/3 of the background electrons arise from heavy flavour decays. The remaining background arises from photon conversions and misidentified hadrons, which are difficult to simulate in the context of our parameterised detector simulation. We therefore make the approximation that the total background to the sample of high $p_T$, isolated electrons in the simulated events is obtained by multiplying the number of electrons from heavy flavour decays by a factor of 1.5.

### 2.1.3 Simulation of $\tau$ hadron candidates

Of central importance to any measurement of $R(W)$ at a hadron collider will be an understanding of the efficiencies and backgrounds associated with the identification of $\tau$ candidates. A recent paper by the CMS Collaboration [36] describes the methods used to identify $\tau$ hadron candidates at $\sqrt{s} = 13$ TeV and details the methods used to determine identification efficiencies, background rates and energy scale, and their associated systematic uncertainties, using a data set corresponding to $\int L dt = 36$ fb$^{-1}$. We choose to simulate the $\tau$ hadron identification efficiency and fake probabilities corresponding to the “very-very-tight” operating point of [36]. The efficiency is around 30% and is independent of $p_T$. The misidentification rate for hadronic jets is around $2 \times 10^{-3}$ for $p_T \approx 25$ GeV decreasing to around $10^{-4}$ for $p_T \approx 100$ GeV. The misidentification rate is approximately independent of $|\eta|$. We simulate also the discriminants against electrons and muons described in [36]; the “Tight” discriminant against electrons has a $\tau$ hadron identification efficiency of 75% and a fake probability of $10^{-3}$; the “Tight” discriminant against muons has a $\tau$ hadron identification efficiency of 99% and a fake probability of $1.4 \times 10^{-3}$ [38].

In our study it is particularly important to assign realistic systematic uncertainties on the identification efficiency for $\tau$ hadron candidates. A tag and probe analysis of $Z \rightarrow \tau\tau \rightarrow \ell\ell\tau\tau$ events in [36] results in a relative uncertainty on the identification efficiency of $\tau$ hadron candidates of 5% for $\tau$ had $p_T > 60$ GeV. A sample of $t\bar{t}$ events is used in [39] to cross check the efficiency for $\tau$ had $p_T > 100$ GeV; a relative uncertainty of 7% for $60 < p_T < 100$ GeV is assigned. Implicitly this latter analysis assumes that $B(W\rightarrow\tau\nu)$ takes its SM value. Clearly, for our proposed test of LU using $t\bar{t}$ events it would not be legitimate to set data-MC scale factors for $\tau$ had using $t\rightarrow b\bar{b}\tau$ events; this means that it will be difficult to control the $p_T$ dependence of the identification efficiency for $\tau$ had candidates beyond the range in $\tau$ had $p_T$ covered by the $Z \rightarrow \tau\tau \rightarrow \ell\ell\tau\tau$ event sample. We consider a $p_T/\eta$-independent relative systematic uncertainty of 5%. We generate a $p_T$-dependent systematic uncertainty on the identification efficiency of $\tau$ had candidates by modifying the efficiency by a relative amount that varies linearly between 5% at $p_T = 30$ GeV and 0% at $p_T = 130$ GeV and above. The $p_T$ resolution for $\tau$ had candidates is given by $\frac{\Delta p_T}{p_T} = 0.16$ [39].

The relative uncertainties on the probabilities for electrons, muons, and hadronic jets to be misidentified as a $\tau$ had candidate are around 10% [40]. The uncertainty on
the $p_T$ scale for $\tau_{\text{had}}$ candidates is around 1% \cite{36}. We consider an $\eta$-dependent systematic uncertainty by modifying the $p_T$ scale by 1% only in the endcap region, whilst the scale in the barrel region remain unchanged. Preliminary systematic uncertainties of a similar magnitude have been assessed by the ATLAS Collaboration for the identification of $\tau_{\text{had}}$ candidates at $\sqrt{s} = 13\text{ TeV}$ \cite{41}, using the methods described in \cite{12}.

Backgrounds from QCD multijet (MJ) events (that is, events that do not contain any prompt leptons from $W$ or $Z$ boson decay) are typically estimated at hadron colliders using data driven methods. Such backgrounds cannot reliably be estimated from MC. Of the four signal samples, $\tau\tau$ to simulate the “very-very-tight” operating point of \cite{36} for $\tau_{\text{had}}$ identification, which has relatively low efficiency, but high rejection power, is to minimise the uncertainties arising from MJ backgrounds. From \cite{24} and \cite{53} we estimate the fraction of MJ background in the selected $Z \rightarrow \ell\ell_{\text{had}}$ sample to be around 5%, with a relative uncertainty of around 5% \cite{43}.

The probability to mis-measure the sign of the charge of lepton candidates is expected to be less than 1% over the range of $p_T$ of relevance to our study and is neglected in our simulation.

### 2.1.4 Simulation of hadronic jets, including $b$-tagging

Jet finding is performed at particle-level using the anti-$k_T$ algorithm \cite{44} with a distance parameter $R = 0.4$. Determinations of the jet energy scale (JES) and jet energy resolution (JER), along with their systematic uncertainties, have been presented by the ATLAS \cite{45} and CMS \cite{46} collaborations. We simulate the resolution in jet $p_T$ by means of a Gaussian with a width $\sigma_{p_T}$ given by $\frac{\sigma_{p_T}}{p_T} = \frac{0.7}{\sqrt{E_T}}$. We assume an uncertainty on the jet energy scale of 1%.

The efficiencies, backgrounds and systematic uncertainties associated with the $b$-tagging of hadronic jets have been presented by the ATLAS \cite{47} and CMS \cite{48} collaborations. We choose to simulate tag probabilities according to those given for the “DeepCSV Loose” category of \cite{48}.

Jets with $|y| < 2.5$ are flagged as $b$-tagged with a probability that depends on their flavour at truth level as follows: truth $b$ 85%, truth $c$ 40%, truth light quark or gluon 1%. These tag probabilities are approximately independent of $p_T$ and $|y|$. The relative uncertainty in the $b$-jet efficiency scale factors is around 1.5% and the uncertainties in the $c$-jet and light-jet mis-tag scale factors are around 4% and 10% respectively \cite{49}.

The missing transverse momentum ($E_T$) is calculated from the vector sum at particle level of the $p_T$ of all neutrinos in the event. Resolution in $E_T$ is taken into account by adding the difference between particle-level and detector level $p_T$ of each lepton and jet in the event. We have checked that this procedure reproduces approximately the $E_T$ resolutions given in \cite{50} and \cite{51}.

The effects of multiple proton-proton collisions or “pile-up” are not simulated in our study. In general, the lepton and jet identification algorithms employed by ATLAS and CMS are designed to have small pile-up dependence \cite{52}. We take the performance values we have implemented to represent averages over the pile-up conditions experienced at the LHC.

### 2.2 Monte Carlo generators and phenomenological systematic uncertainties

Events containing $W$ bosons, $Z$ bosons, top quark pairs, and the EW production of single top quarks are generated using POWHEG BOX \cite{53}, interfaced to PYTHIA \cite{54} for the simulation of parton showering and fragmentation. We thus ensure a consistent treatment of tau lepton decay in the principle sources of candidate events. EW diboson events are generated using SHERPA \cite{55}.

Significant sources of background in the selected $Z/\gamma^* \rightarrow \tau^+\tau^-$ samples arise from EW diboson and $t\bar{t}$ production. Cross sections at $\sqrt{s} = 13\text{ TeV}$ have been measured for EW diboson \cite{56} and $t\bar{t}$ \cite{57} production. Our estimates of the fractional composition of the selected $Z/\gamma^* \rightarrow \tau^+\tau^-$ samples are not sensitive to systematic uncertainties on the integrated luminosity or on the predicted absolute cross sections for $Z$ boson, $t\bar{t}$, or EW diboson production. They are, however, sensitive to systematic uncertainties on the ratio of the cross sections for $t\bar{t}$ and EW diboson production to that for $Z$ bosons. We assume an uncertainty of 3% on both of these cross section ratios \cite{58}.

In $Z/\gamma^* \rightarrow \tau^+\tau^-$ events the momentum distribution of electrons and muons produced in $\tau$ decay is softer than that for the visible $\tau_{\text{had}}$ systems. Changes in the distribution of the transverse momentum of the produced $Z$ bosons ($p_T(Z)$) can, therefore, affect the relative acceptance for $Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}}$ and $Z \rightarrow \tau\tau \rightarrow e\mu$ candidate events. Measurements of $p_T(Z)$ have been made at $\sqrt{s} = 8\text{ TeV}$ by ATLAS \cite{59}. In order to evaluate systematic uncertainties arising from $p_T(Z)$ we increase the weight of events satisfying 50 GeV $< p_T(Z) < 150$ GeV by 0.5% and the weight of events satisfying $p_T(Z) > 150$ GeV by 1.0% \cite{60}.

In $tt \rightarrow b\bar{b}\ell_{\text{had}}$ events, the majority of observed electrons and muons are from direct $W$ boson decays and therefore have a distribution in $p_T$ that is much harder than that for the visible $\tau_{\text{had}}$ systems. The relative acceptance for $tt \rightarrow b\bar{b}\ell_{\text{had}}$ and $tt \rightarrow b\bar{b}\ell_{\mu}$ final states may be sensitive to the details of the modelling of $t\bar{t}$ production. We investigate this sensitivity by using an alternative $t\bar{t}$ generator-level sample in which the QCD factorisation and renormalisation scales are changed by a factor of two relative to the default values. In addition, we evaluate systematic uncertainties arising from simulating the distribution of the mass of the $t\bar{t}$ system, $m(tt)$, by increasing the weights of events by an amount that varies linearly between 1% at $m(tt) = 500$ GeV to 10% at $m(tt) = 1000$ GeV \cite{61}.
2.3 Candidate event selection criteria

Candidate electrons and muons are considered in the analysis if after simulation of resolution and momentum scale they satisfy \( p_T > 27 \text{ GeV} \). Candidate \( \tau_{\text{had}} \) and hadronic jets are considered if after simulation of resolution and momentum scale they satisfy \( p_T > 25 \text{ GeV} \). Hadronic jets must satisfy \( |\eta| < 4.5 \). Candidate electrons, muons, \( \tau_{\text{had}} \), and \( b \)-tagged jets must satisfy \( |\eta| < 2.5 \).

2.3.1 Candidate event selection criteria on leptons

In order to maximise the cancellation of systematic uncertainties between \( tt \rightarrow b\bar{b}W^+W^- \) and \( Z/\gamma* \rightarrow \tau^+\tau^- \) event samples the same candidate event selection criteria on leptons are applied in the two event classes.

Candidate \( e_{\tau_{\text{had}}} \) events are required to contain:
- Exactly one \( e \) candidate.
- Exactly one \( \tau_{\text{had}} \) candidate of opposite sign to the \( e \) candidate.
- No \( \mu \) candidates.
- The \( e \) candidate must fire the single-\( e \) trigger.

Candidate \( \mu_{\tau_{\text{had}}} \) events are required to contain:
- Exactly one \( \mu \) candidate.
- Exactly one \( \tau_{\text{had}} \) candidate of opposite sign to the \( \mu \) candidate.
- No \( e \) candidates.
- The \( \mu \) candidate must fire the single-\( \mu \) trigger.

Candidate \( e\mu \) events are required to contain:
- Exactly one \( e \) candidate.
- Exactly one \( \mu \) candidate of opposite sign to the \( e \) candidate.
- The \( e \) candidate must fire the single-\( e \) trigger, or the \( \mu \) candidate must fire the single-\( \mu \) trigger.

2.3.2 \( tt \) candidate event selection criteria

In addition to the relevant criteria on leptons given in section 2.3.1 above, all candidate \( tt \rightarrow b\bar{b}W^+W^- \) events (\( tt \rightarrow b\bar{b}\tau_{\text{had}} \) as well as \( tt \rightarrow b\bar{b}\mu \)) are required to contain exactly two \( b \)-tagged jets. Because the same criterion is applied in the selection of the numerator \( tt \rightarrow b\bar{b}\tau_{\text{had}} \) and denominator \( tt \rightarrow b\bar{b}\mu \) events, we expect the ratio \( R(b\bar{b}W^W) \) to be largely insensitive to systematic uncertainties associated with jet reconstruction, JES, JER, and \( b \)-tagging. Requiring two \( b \)-tagged jets reduces the background in the selected \( tt \rightarrow b\bar{b}W^+W^- \) event samples from non-\( tt \) sources; it also reduces the background from \( t\bar{t} \) events in which a \( b \)-quark jet is misidentified as a prompt \( e, \mu, \tau_{\text{had}} \), or \( \tau_{\text{had}} \).

Candidate \( tt \rightarrow b\bar{b}\tau_{\text{had}} \) events are rejected if the invariant mass of the \( \tau_{\text{had}} \) candidate and the highest \( p_T \) non-\( b \)-tagged jet, \( m(j-\tau_{\text{had}}) \), satisfies 50 GeV < \( m(j-\tau_{\text{had}}) < 90 \text{ GeV} \). This criterion reduces the background from lepton+jet \( tt \) events in which the hadronically decaying \( W \) boson produces two reconstructed jets, one of which is misidentified as a \( \tau_{\text{had}} \) candidate. The effectiveness of this criterion is illustrated by Figure 3, which shows in the \( tt \rightarrow b\bar{b}\tau_{\text{had}} \) candidate event sample the distribution of \( m(j-\tau_{\text{had}}) \), having applied all other \( tt \rightarrow b\bar{b}\tau_{\text{had}} \) event selection criteria. The upper plot shows events in which the \( \tau_{\text{had}} \) candidate originates from a genuine \( \tau \) decay and the lower plot shows events in which the \( \tau_{\text{had}} \) candidate originates from a misidentified hadronic jet. A clear peak at around the mass of the \( W \) boson can be seen in the lower plot. In addition to helping reject background from misidentified hadronic jets, the distributions in Figure 3 offer the possibility to make a data-driven estimate of the background in the \( \tau_{\text{had}} \) candidate sample. This will be useful in reducing the systematic uncertainty on the background yield.

![Fig. 3. The distribution of \( m(j-\tau_{\text{had}}) \) in the \( tt \rightarrow b\bar{b}\tau_{\text{had}} \) candidate event sample. The event selection requirement on this quantity has not been applied. The upper plot shows events in which the \( \tau_{\text{had}} \) candidate originates from a genuine \( \tau \) decay and the lower plot shows events in which the \( \tau_{\text{had}} \) candidate originates from a misidentified hadronic jet.](image-url)
2.3.3 $Z/\gamma^* \rightarrow \tau^+\tau^-$ candidate event selection criteria

In addition to the relevant criteria on leptons given in section 2.3.1 above, all candidate $Z/\gamma^* \rightarrow \tau^+\tau^-$ events ($Z \rightarrow \ell\tau \rightarrow \ell\tau_{\text{had}}$ as well as $Z \rightarrow \tau\tau \rightarrow e\mu$) are required to satisfy the following criteria:

- Events should contain no $b$-tagged jets.
- $50 \text{ GeV} < m_{\ell\tau} < 100 \text{ GeV}$.
- $\Sigma \cos \Delta \phi > -0.1$.
- $\sigma_T < 60 \text{ GeV}$.

Distributions of each variable having applied all other selection criteria are shown in Figure 5 for $Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}}$ and in Figure 6 for $Z \rightarrow \tau\tau \rightarrow e\mu$.

As expected, the rejected event samples containing $b$-tagged jets are dominated by $tt$. Because the same criterion on $b$-tagged jets is applied in the selection of the numerator $Z \rightarrow \tau\tau \rightarrow \ell\tau_{\text{had}}$ and denominator $Z \rightarrow \tau\tau \rightarrow e\mu$ events, we expect the ratio $R(Z)$ to be largely insensitive to systematic uncertainties associated with jet reconstruction, JES, JER, and $b$-tagging. Requiring no $b$-tagged jets reduces the background in the selected $Z/\gamma^* \rightarrow \tau^+\tau^-$ event samples from $tt$ and also $W$ boson plus heavy flavour production in which a $b$-quark jet is misidentified as a prompt $e$, $\mu$, or $\tau_{\text{had}}$. It can be seen that the other selection criteria reject a large fraction of the remaining background, principally from $W+$jet production (in the $Z \rightarrow \tau\tau \rightarrow e\mu$ sample) and from $tt$ and diboson production (in the $Z \rightarrow \tau\tau \rightarrow e\mu$ sample).
Fig. 5. Distributions of variables in the selection of $Z \rightarrow \tau \tau \rightarrow \ell \nu_{\ell} \nu_{\tau}$ candidate events having applied all other selection criteria. The selection criteria are indicated by vertical lines.

Fig. 6. Distributions of variables in the selection of $Z \rightarrow \tau \tau \rightarrow e\mu$ candidate events having applied all other selection criteria. The selection criteria are indicated by vertical lines.
We propose here a novel selection variable, $m_3^2$:

$$m_3^2 \equiv \frac{m_T(\ell, \tau_{had}, E_T)}{\sin \theta^*_\eta}.$$  \hspace{1cm} (12)

Here $m_T(\ell, \tau_{had}, E_T)$ is the transverse mass of the 3-body system of $\ell$, $\tau_{had}$, and $E_T$, which may be defined by:

$$m_T(\ell, \tau_{had}, E_T)^2 = m_T(\ell, \tau_{had}, E_T)^2 + m_T(\tau_{had}, E_T)^2 \hspace{1cm} (13)$$

$\theta^*_\eta$ is an approximation to the scattering angle of the leptons relative to the beam direction in the dilepton rest frame. This variable is defined \[62\] solely using the measured track directions by:

$$\cos(\theta^*_\eta) \equiv \tanh \left( \frac{\eta^- - \eta^+}{2} \right), \hspace{1cm} (14)$$

where $\eta^-$ and $\eta^+$ are the pseudorapidities of the negatively and positively charge lepton, respectively. The division by $\sin \theta^*_\eta$ in the definition of $m_3^2$ takes into account the relative longitudinal motion of the two leptons and, therefore, $m_3^2$ is a more closely correlated with the $\tau^+\tau^-$ mass than is $m_T(\ell, \tau_{had}, E_T)$. In the selection of $Z/\gamma^* \rightarrow \tau^+\tau^-$ candidates this improves the discrimination power against the dominant backgrounds, e.g., $W$ plus jet and diboson events.

The variable $\Sigma \cos \Delta \phi$ \[63\]:

$$\Sigma \cos \Delta \phi \equiv \cos(\Delta \phi(\ell, E_T)) + \cos(\Delta \phi(\tau_{had}, E_T)),$$  \hspace{1cm} (15)

discriminates against background events containing leptonically decaying $W$ bosons. Here $\Delta \phi(\ell, E_T)$ is the azimuthal angle between the $\ell$ and the $E_T$, and $\Delta \phi(\tau_{had}, E_T)$ is the azimuthal angle between the $\tau_{had}$ and the $E_T$.

The variable $a_T$ \[64\] corresponds to the component of the $p_T$ of the dilepton system that is transverse to the dilepton thrust axis. This variable is well suited to the study of $\tau^+\tau^-$ final states, because it is less sensitive to any imbalance in the transverse momenta of the neutrinos produced in the tau decays than is $a_p$. \[44\], the component of the dilepton $p_T$ that is longitudinal to the dilepton thrust axis. The variable $a_T$ discriminates against background events containing leptonically decaying $W$ bosons.

Figure 2 shows the distributions of $p_T$ and $|\eta|$ of e, $\mu$ and $\tau_{had}$ in the selected $Z/\gamma^* \rightarrow \tau^+\tau^-$ candidate event samples. It can be seen that the $p_T$ distributions are considerably softer than those in Figure 4 for the selected $t\bar{t} \rightarrow b\bar{b}W^+W^-$ candidate event samples.

As can be seen from the lower plots in Figures 5 and 6, a cut of $a_T < 30$ GeV would be desirable to improve the suppression of backgrounds from $t\bar{t}$ and EW diboson events. However, a cut on $a_T$ suppresses also $Z/\gamma^* \rightarrow \tau^+\tau^-$ events that contain initial state parton radiation. Initial state radiation broadens the distributions of lepton candidate $p_T$ in $Z/\gamma^* \rightarrow \tau^+\tau^-$ events. A hard cut on $a_T$ would therefore suppress the high-$p_T$ regions in the distributions of lepton $p_T$ in the selected $Z/\gamma^* \rightarrow \tau^+\tau^-$ event samples, as is illustrated in Figure 8. The cut $a_T < 60$ GeV is chosen to reject background events from $t\bar{t}$ and diboson events, without unduly biasing the lepton $p_T$ distributions and further accentuating the differences between the lepton $p_T$ distributions seen in Figure 4 and Figure 7.

2.4 Size and composition of the selected event samples

The expected numbers of selected events and the composition of the four selected event samples for $\mathcal{L} dt = 140$ fb$^{-1}$ at $\sqrt{s} = 13$ TeV are given in Table 3. As a result of this study we estimate that the fractional statistical uncertainty on $R(WZ)$ for a single LHC experiment for an integrated luminosity of $\mathcal{L} dt = 140$ fb$^{-1}$ would be around 0.5%.

The selection requirements for $tt \rightarrow bb\tau_{had}$ and $tt \rightarrow b\bar{b}\mu$ can be seen to be very effective at removing non-$t\bar{t}$ sources of background, e.g., $W$ plus jet. The most significant source of non-$t\bar{t}$ events originates from the EW production of “single top” events in the associated production $tWb$ channel. Since these events contain two genuine leptonically decaying $W$ bosons they can effectively be considered as contributing to the signal sample, and not as background. They are listed as “$tWb$ true” in Table 4. Single top processes that do not contain a pair of $W$ bosons decaying to produce two correctly identified leptons are classified as background. Events which contain the associated production of either a $W$ or $Z$ boson with a $t$ pair ($tWb$) can be considered signal if they contain at least two $W$ bosons which decay to correctly identified leptons. Since the fractions of these events are small, < 0.1%, they are included in the $t\bar{t}$ categories in Table 5.

The most significant source of background for $tt \rightarrow b\bar{b}\tau_{had}$ originates from genuine $t\bar{t}$ events in which the $\tau_{had}$ candidate originates from a misidentified hadronic jet. The residual background from this source corresponds to about 2.5% of the selected sample of candidate $tt \rightarrow b\bar{b}\tau_{had}$ events. The $tt \rightarrow b\bar{b}\mu$ candidate event sample will be selected with entirely negligible levels of background.

The most significant source of background for $Z \rightarrow \tau\tau \rightarrow \tau_{had}$ is at the level of around 5% and originates from MJ events, as described in section 2.1.3 and \[43\]. In comparison, the MC-estimated backgrounds for $Z \rightarrow \tau\tau \rightarrow \tau_{had}$ from $t\bar{t}$ (0.3%) and $W$ plus jet (0.2%) are small. The most significant sources of background for $Z \rightarrow \tau\tau \rightarrow e\mu$ originate from $t\bar{t}$ (5.8%) and diboson (6.3%) production.

2.5 Sensitivity of $R(WZ)$ to $R(W)$

For definiteness we make the assumption in our study that the effective $W \rightarrow \tau\nu$ coupling relevant for on-mass-shell $W$ boson decays could be modified by some BSM effect, whilst all other $W$ boson couplings are maintained at their SM-predicted values. Under this assumption, if the branching ratio $B(W \rightarrow \tau\nu)$ is multiplied by a factor $X$ relative to its SM-predicted value $B(W \rightarrow \tau\nu)_{SM}$,

$$B(W \rightarrow \tau\nu) = X B(W \rightarrow \tau\nu)_{SM} \hspace{1cm} (16)$$

S. Dysch and T. R. Wyatt.: A self-calibrating, double-ratio method to test tau lepton universality in W boson decays 11
then all other $W$ boson branching fractions will be modified by a factor

$$F = \frac{1 - X_{\text{SM}}^{W\rightarrow\tau\nu}}{1 - X_{\text{SM}}^{W\rightarrow\tau\nu}}.$$  \hfill (17)

In our MC study we perform a “calibration” of the double-ratio method by reweighting every simulated event containing one or more $W$ boson decays by a factor $X^{W\rightarrow\tau\nu}$, where $n$ is the number of generator-level $W \rightarrow \tau\nu$ decays and $m$ is the number of other $W$ boson decays. This calibration procedure properly takes into account all events in the calculation of $R(bbWW)$ that contain pairs of leptonically decaying $W$ bosons with correctly identified decay products in the “numerator” $tt \rightarrow bb\ell_\text{had}$ and “denominator” $tt \rightarrow bb\ell_\mu$ samples. This includes, for example, the presence of events containing the cascade decay $W \rightarrow \tau\nu \rightarrow \ell\nu\nu$, whose presence in the “denominator” $tt \rightarrow bb\ell_\mu$ sample slightly decreases the correlation between $R(bbWW)$ and $R(W)$. The value of $R(Z)$ is designed to be independent of any change in $R(W)$. However, the backgrounds from $t\ell \rightarrow bbW^+W^-$ and $WW$ in the $Z \rightarrow \tau\tau \rightarrow \ell_\text{had}$ and $Z \rightarrow \tau\tau \rightarrow \ell_\mu$ events contain decays of $W$ bosons, which cause the expected background levels to alter with $R(W)$. A correction to the $R(Z)$ calculation from this effect is included. The result of this calibration is shown in Figure 8 which shows the variation of $R(bbWW)$, $R(Z)$, and $R(WZ)$ as a function of $R(W)$.

2.6 Evaluation of systematic uncertainties

Considering each source of systematic uncertainty described in sections 2.1 and 2.2 and summarised in Tables 1 and 2 we evaluate the resulting changes in the ratios $R(bbWW)$, $R(Z)$, $R(WZ)$, and $R(W)$.

The $p_T/\eta$-independent and $p_T/\eta$-dependent systematic variations on the reconstruction of leptons and jets, as summarised in Table 1 are considered separately. The resultant changes in the ratios are given in Table 4. It can be seen that the $p_T/\eta$-independent systematic uncertainties on $R(bbWW)$ and $R(Z)$ are large, but almost exactly
equal. Therefore, $p_T/\eta$-independent systematic uncertainties on the reconstruction of leptons and jets almost perfectly cancel in the double ratio $R(WZ)$. When considering $p_T/\eta$-dependent systematic uncertainties the cancellation is no longer perfect. The most significant sources of $p_T/\eta$-dependent systematic uncertainties are illustrated in Figure 10.

The alternative $tt$ sample with modified QCD factorisation and renormalisation scales leads to an uncertainty of 0.3% on $R(bbWW)$ and thus also on $R(W)$.

### 2.7 Considerations for future measurements

Measurements of $R(W)$ on the data from ATLAS and CMS using the double ratio technique proposed here will, clearly, require the use of fully simulated MC events and will employ the sophisticated procedures developed by the individual experiments to evaluate backgrounds and experimental systematic uncertainties. Our study is based on particle-level MC events and a simple parameterised simulation of detector performance. Nevertheless, we believe that demonstration that the dominant experimental systematic uncertainties cancel in the double ratio, as well as our estimates of the major residual systematic uncertainties, to be broadly realistic. We have based our simulation of lepton, jet and $E_\text{T}$ reconstruction on measurements of efficiencies, backgrounds and systematic uncertainties published by ATLAS and CMS, choosing identification algorithms whose performance is suited to the needs of our analysis. The above cited performance papers are for the most part based on around a quarter of the full run 2 data set of $\int \mathcal{L} \, dt = 140$ fb$^{-1}$ at $\sqrt{s} = 13$ TeV that is now available. It is, therefore, to be expected that the full run 2 data set will allow systematic uncertainties on lepton and jet identification efficiencies to be reduced by about a factor of two compared to the values we have assumed in our study. Of particular relevance to our study, it is to be hoped that the high statistics provided by the full run 2 data set will allow the $p_T$ dependence of the identification efficiency for $\tau_\text{had}$ candidates to be studied over the range $25 < p_T < 100$ GeV, without the need to use $tt$
$bb\ell\tau_{\text{had}}$ events and implicitly assume that $\mathcal{B}(W\rightarrow\tau\nu)$ takes its SM value. This could be achieved, e.g., by the selection of a dedicated $Z\rightarrow\tau\tau\rightarrow\ell\ell_{\text{had}}$ event sample in which there is a high transverse momentum initial state radiation.

Of course, if the measured value of $R(WZ)$ is found to disagree with that expected in the SM then further studies will be required to ascertain the nature of BSM physics that is responsible. For example, the decay $t\rightarrow bH^+$, where $H^+$ is a charged higgs boson, followed by $H^+\rightarrow\tau^+\nu$ would modify the effective $t\rightarrow b\tau\nu$ and $t\rightarrow bH^+$ branching ratios in a similar fashion to that discussed in the context of equations 16 and 17 above. Similarly, top decays via a neutral higgs boson $t\rightarrow qH$, $H\rightarrow\tau^+\tau^-$ will also increase the number of tau leptons in $t\ell$ events relative to the SM-expected value. Existing experimental searches for charged [65,67] and neutral [88] higgs bosons in top quark events suffer from the large systematic uncertainties associated with $\tau_{\text{had}}$ identification and will benefit from the double ratio method we propose in this paper to reduce experimental systematic uncertainties.

The large numbers of events from the EW production of diboson events ($WW$ and $WZ$) at the LHC provide alternative samples with which to make this novel measurement. Controlling systematic uncertainties on $R(W)$ in $WW$ events will be extremely challenging; we shall need extraordinarily good background rejection against fake $\tau$ candidates from misidentified jets in $W\rightarrow\ell\nu+\text{jet}$ events. One may also consider $ZW\rightarrow\ell^+\ell^-\tau\tau$ as an alternative sample with which to perform this measurement. This channel provides lower statistics because of the lower cross section times branching fraction. However, one can veto on $Z+\text{jet}$ backgrounds by removing events in which the $\ell^+\ell^-$ momentum is back to back with the $\tau$ candidate direction. One can calibrate the residual backgrounds by looking at the back-to-back events. If the experimental measurements proposed here observe a clear violation of LU then having three channels $t\ell\rightarrow bb\ell\tau_{\text{had}}, WW\rightarrow\ell\nu\tau\nu,$ and $ZW\rightarrow\ell^+\ell^-\tau\tau$ could increase the significance of the observation, and could help elucidate the underlying BSM origin of the effect.

### Tables

| Source of systematic uncertainty | Systematic uncertainty on measured ratios (%) |
|---------------------------------|-----------------------------------------------|
| $R(bbWW)$ | $R(Z)$ | $R(WZ)$ | $R(bbWW)$ | $R(Z)$ | $R(WZ)$ |
| Tau ID | 5.0 | 5.0 | < 0.1 | 3.9 | 4.7 | 0.8 |
| Electron ID | 1.1 | 1.3 | 0.2 | 0.5 | 1.2 | 0.7 |
| Electron trigger | 0.3 | 0.2 | 0.1 | 0.1 | 0.3 | 0.2 |
| Muon trigger | 0.4 | 0.5 | 0.1 | 0.4 | 0.5 | 0.1 |
| Muon ID | 1.0 | 0.7 | 0.3 | 0.1 | 0.2 | < 0.1 |
| b-jet ID | < 0.1 | < 0.1 | < 0.1 | - | - | - |
| Light jet mis-ID | 0.1 | 0.1 | < 0.1 | - | - | - |
| Tau $p_T$ scale | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 |
| Electron $p_T$ scale | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 |
| Muon $p_T$ scale | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 |
| Jet energy scale | < 0.1 | < 0.1 | < 0.1 | - | - | - |
| Electron $p_T$ resolution | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 |
| Muon $p_T$ resolution | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 | < 0.1 |

**Table 4.** Changes in the ratios $R(bbWW)$, $R(Z)$, and $R(WZ)$ resulting from the sources of systematic uncertainty described in Section 2.1 and summarised in Table 1. The $p_T/\eta$-independent and $p_T/\eta$-dependent systematic variations are considered separately.

| Source of systematic uncertainty | Systematic uncertainty on measured ratios (%) |
|---------------------------------|-----------------------------------------------|
| $R(bbWW)$ | $R(Z)$ | $R(WZ)$ |
| Fake $\tau_{\text{had}}$ background in $tt\rightarrow bb\ell\tau_{\text{had}}$ sample | 0.25 | 0 | 0.25 |
| MJ background in $Z\rightarrow\tau\tau\rightarrow\ell\ell_{\text{had}}$ sample | 0 | 0.25 | 0.25 |
| $t\ell\ell$ background in $Z\rightarrow\tau\tau\rightarrow e\mu$ sample | 0 | 0.2 | 0.2 |
| Diboson background in $Z\rightarrow\tau\tau\rightarrow e\mu$ sample | 0 | 0.2 | 0.2 |
| $p_T(Z)$ reweighting | 0 | < 0.1 | < 0.1 |
| $m(t\ell)$ reweighting | 0.2 | 0 | 0.2 |
| $t\ell$ modelling | 0.3 | 0 | 0.3 |

**Table 5.** Changes in the ratios $R(bbWW)$, $R(Z)$, $R(WZ)$, and $R(W)$ resulting from the sources of systematic uncertainty described in Section 2.2 and summarised in Table 2.

### Sections

**3 Summary and conclusions**

A measurement of $R(W)\equiv\mathcal{B}(W\rightarrow\tau\nu)/\mathcal{B}(W\rightarrow\ell\nu)$ ($\ell = e$ or $\mu$) represents a promising opportunity to discover a violation of lepton universality. We propose here a novel double-ratio method that will allow $R(W)$ to be measured using top quark pairs and $Z/\gamma^*\rightarrow\tau^+\tau^-$ events at the LHC.
We define $R(bbWW)$ in di-leptonic $t\bar{t}$ events to be the ratio of the numbers of $\ell\tau_{\text{had}}$ and $e\mu$ final states (equation 9). $R(bbWW)$ is sensitive to the value of $R(W)$, but also to systematic uncertainties on the reconstruction of $e$, $\mu$, and $\tau$ leptons. Similarly, we define $R(Z)$ in $Z/\gamma^* \to \tau^+\tau^-$ events to be the ratio of the numbers of $\ell\tau_{\text{had}}$ and $e\mu$ final states (equation 10). $R(Z)$ is similarly sensitive to systematic uncertainties on the reconstruction of $e$, $\mu$, and $\tau$ leptons, but is insensitive to the value of $R(W)$. The double ratio $R(WZ) \equiv R(bbWW)/R(Z)$ cancels to first order sensitivity to systematic uncertainties on the

Fig. 9. The variation of $R(bbWW)$ (upper), $R(Z)$ (middle), and $R(WZ)$ (lower), as a function of $R(W)$, obtained by reweighting events at generator level as described in the text. The displayed error bars illustrate the expected experimental statistical uncertainty on the relevant quantity, and are completely correlated point to point. MC statistical uncertainties on the point-to-point variation with $R(W)$ are negligible.

Fig. 10. Summary of the most significant changes in the ratios (a) $R(bbWW)$, (b) $R(Z)$, and (c) $R(WZ)$ arising from sources of $p_T/\eta$-dependent systematic uncertainty. From left to right in each sub-figure the sources are ordered in descending magnitude of the resulting systematic uncertainty.
reconstruction of $e$, $\mu$, and $\tau$ leptons, thus improving very significantly the precision to which $R(W)$ can be measured at a hadron collider.

We have performed a study of the double ratio $R(WZ)$ using particle-level MC events and a parameterised simulation of detector performance. We have based our simulation of jet, $E_T$ reconstruction on measurements of efficiencies, backgrounds and systematic uncertainties published by ATLAS and CMS [23–51], having chosen identification algorithms whose performance is suited to the needs of our analysis. For a data set of $fL dt = 140$ fb$^{-1}$ at $\sqrt{s} = 13$ TeV we estimate a statistical uncertainty on $R(W)$ of 0.5%. Our study confirms the almost perfect cancellation in $R(W)$ of systematic uncertainties on the reconstruction efficiencies of $e$, $\mu$, and $\tau$ leptons that are applied as constant factors. We find that the most significant residual sources of uncertainty on $R(W)$ arise from systematic uncertainties on the $p_T$ and $\eta$ dependence of the reconstruction efficiencies of $e$, $\mu$, and $\tau$ leptons, which total around 1.0%. We have evaluated also potential uncertainties arising from backgrounds to the selected event samples and from various phenomenological sources.

Our studies indicate that a single experiment precision on the measurement of $R(W)$ of around 1.4% is achievable with a data set of $fL dt = 140$ fb$^{-1}$ at $\sqrt{s} = 13$ TeV. This would improve significantly upon the precision of the LEP2 measurements of $R(W)$. If the central value of the new measurements were equal to the central value of the LEP2 measurements this would yield an observation of BSM physics at a significance level of around 5$\sigma$.

Final year undergraduate (MPhys) project students working in Manchester with T.W. have made some important contributions to the ideas presented in this paper. Jihyun Jeong (1993–2018) and Robin Upham made an early feasibility study for the measurement of $R(bWW)$ using $tt$ events at the LHC. Vilnis Cepaitis and Ricardo Wölker studied possible improvements to the selection of $Z/\gamma^* \rightarrow \tau^+\tau^-$ events at the LHC from which the variable $m_3$ arose. We are very grateful to our colleague Chris Parkes for his useful suggestions for the improvement of this paper.

References

1. ‘Precision electroweak measurements on the Z resonance’, ALEPH, DELPHI, L3, OPAL and SLD Collaborations, Phys. Rept. 427 (2006) 257.
2. ‘Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP’. The ALEPH, DELPHI, OPAL and L3 Collaborations, Physics Reports 532.4 (2013) 119–244.
3. ‘Review of Particle Physics’, The Particle Data Group, M. Tanabashi et al., Phys. Rev. D 98 (2018) 030001.
4. ‘Two-Loop Effective Theory Analysis of $\tau(K)\rightarrow e\pi^+\pi^-$ Branching Ratios’ V. Cirigliano and I. Rosell, Phys. Rev. Lett. 99 (2007).
5. ‘Average of $R(D)$ and $R(D^*)$ for Spring 2019’, Heavy Flavour Averaging Group (HFLAV), Y. Amhis et al., https://hflav-eos.web.cern.ch/hflav-eos/semi/spring19/html/RDsDstarr/RDRDstar.html using the methods described in Eur. Phys. J. C 77 (2017) no.12, 895.
6. ‘Evidence for an excess of $B\rightarrow D^{(*)}\tau^-\bar{\nu}_{\tau}$ decays’, BaBar Collaboration, J. P. Lees et al., Phys. Rev. Lett. 109 (2012) 101802.
7. ‘Evidence for an excess of $B\rightarrow D^{(*)}\tau^-\bar{\nu}_{\tau}$ decays and Implications for Charged Higgs Bosons’, BaBar Collaboration, J. P. Lees et al., Phys. Rev. D 88 (2013) 072012.
8. ‘Measurement of the branching ratio of $B\rightarrow D^{(*)}\tau^-\bar{\nu}_{\tau}$ relative to $B\rightarrow D^{(*)}\ell^-\bar{\nu}_{\ell}$ decays with hadronic tagging at Belle’, Belle Collaboration, M. Huschle et al., Phys. Rev. D 92 (2015) 072014.
9. ‘Measurement of the branching ratio of $B^0\rightarrow D^{(*)}\tau^-\bar{\nu}_{\tau}$ relative to $B^0\rightarrow D^{(*)}\ell^-\bar{\nu}_{\ell}$ decays with a semileptonic tagging method’, Belle Collaboration, Y. Sato et al. Phys. Rev. D 94 (2016) 072007.
10. ‘Measurement of the $\tau$ lepton polarization and $R(D^*)$ in the decay $B\rightarrow D^\tau\bar{\nu}_{\tau}$’. Belle Collaboration, S. Hirose et al. Phys. Rev. Lett. 118 (2017) 211801.
11. ‘Measurement of the $\tau$ lepton polarization and $R(D^*)$ in the decay $B\rightarrow D^\tau\bar{\nu}_{\tau}$ with one-prong hadronic $\tau$ decays at Belle’, Belle Collaboration, S. Hirose et al. Phys. Rev. D 97 (2018) 012004.
12. ‘Measurement of $R(D)$ and $R(D^*)$ with a semileptonic tagging method’, Belle Collaboration, A. Abdesselam et al. https://arxiv.org/abs/1904.08794 [hep-ex].
13. ‘Measurement of the ratio of branching fractions $B(\bar{B}\rightarrow D^{(*)}\tau^-\nu_{\tau})/B(\bar{B}\rightarrow D^{(*)}\mu^+\nu_{\mu})$', LHCb Collaboration, R. Aaij et al., Phys. Rev. Lett. 115 (2015) 111803. Erratum: Phys. Rev. Lett. 115 (2015) 159901.
14. ‘Measurement of the $R(B^{0})\rightarrow D^{(*)}\tau^-\nu_{\tau}$ branching fraction using three-prong $\tau$-lepton decays’, LHCb Collaboration, R. Aaij et al., Phys. Rev. Lett. 120 (2018) 171802.
15. ‘Test of lepton flavor universality by the measurement of the $B^{0}\rightarrow D^{(*)}\tau^-\nu_{\tau}$ branching fraction using three-prong $\tau$ decays’, LHCb Collaboration, R. Aaij et al., Phys. Rev. D 97 (2018) 072013.
16. ‘Testing lepton quark models in $B\rightarrow D^{(*)}\ell\bar{\nu}_{\ell}$’, Y. Sakaki, M. Tanaka, A. Tayduganov and R. Watanabe, Phys. Rev. D 88 (2013) 094012.
17. ‘Measurement of the $W\rightarrow \tau\nu$ cross section in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment’, ATLAS Collaboration, G. Aad et al., Phys. Lett. B 706 (2012) 276.
18. ‘Measurements of the top quark branching ratios into channels with leptons and quarks with the ATLAS detector’, ATLAS Collaboration, G. Aad et al., Phys. Rev. D 92 (2015) no.7, 072005.
19. ‘Measurement of the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 8$ TeV in dilepton final states containing one $\tau$ lepton’, CMS Collaboration, V. Khachatryan et al., Phys. Lett. B 739 (2014) 23.
20. The pseudorapidity, \( \eta \), is defined as \( \eta = -\ln[\tan(\frac{\theta}{2})] \), where \( \theta \) is the polar angle with respect to the beam direction, in the laboratory frame. The azimuthal angle around the beam direction is denoted by \( \phi \).

21. "Search for high-mass dilepton resonances using 139 fb\(^{-1}\) of pp collision data collected at \( \sqrt{s} = 13 \) TeV with the ATLAS detector", ATLAS Collaboration, G. Aad et al., arXiv:1903.06248 [hep-ex].

22. "Search for new physics in multilepton final states in pp collisions at \( \sqrt{s} = 13 \) TeV", CMS Collaboration, A. M. Sirunyan et al., CMS-PAS-EXO-19-002.

23. "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at \( \sqrt{s} = 13 \) TeV", ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C 76 (2016) 292.

24. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at \( \sqrt{s} = 13 \) TeV", CMS Collaboration, A. M. Sirunyan et al., JINST 13 (2018) P06015.

25. "Performance of the ATLAS trigger system in 2015", ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C 77 (2017) 317.

26. The efficiency of the single muon trigger has been measured in the barrel (endcap) region of CMS to be around 96% (90%) \(^{23}\) and to be almost independent of \( p_T \). In our simulation of the single muon trigger efficiency we choose to use the smaller values given in the text for ATLAS.

27. "Performance of electron and photon triggers in ATLAS during LHC Run 2", arXiv:1909.00761 [hep-ex].

28. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at \( \sqrt{s} = 13 \) TeV", ATLAS Collaboration, M. Aaboud et al., arXiv:1902.04655 [physics.ins-det].

29. "Performance of electron and photon identification in the ATLAS detector using proton-proton collisions at \( \sqrt{s} = 8 \) TeV", CMS Collaboration, A. M. Sirunyan et al., JINST 10 (2015) P06005.

30. Electron reconstruction and selection with the CMS detector in proton-proton collisions at \( \sqrt{s} = 8 \) TeV, CMS Collaboration, A. M. Sirunyan et al., JINST 10 (2015) P06005.

31. For example, the region \( 1.37 < |\eta| < 1.52 \) is usually excluded for precision measurements involving electrons or \( \tau \) hadrons. We simulate this gap in this detector parameterisation.

32. The efficiency for electron identification according for the "Tight likelihood" category is shown as functions of \( p_T \) and \( \eta \) in Figure 8 of \(^{28}\).

33. We simulate approximately the efficiency for the single-electron trigger according to that given as a function of \( p_T \) and \( \eta \) in Figure 26 (a) of \(^{24}\), except that the value of \( p_T \) on the abscissa is shifted up by 2 GeV to account for the fact that for most of the LHC run 2 the trigger threshold was set at \( p_T = 26 \) GeV \(^{27}\).

34. "Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data", ATLAS Collaboration, M. Aaboud et al., JINST 14 (2019) P03017.

35. The choice of 0.002 corresponds to a value in the middle of the range of values given in Table 2 of \(^{33}\).

36. "Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data", ATLAS Collaboration, M. Aaboud et al., Eur. Phys. J. C 77 (2017) 195. An estimate of various sources of backgrounds to the sample of high \( p_T \) isolated electron candidates in ATLAS is given in Table 3. The "Very tight likelihood" criterion in this paper has a very similar efficiency to "Tight likelihood" criterion in \(^{28}\). We, therefore, take the background probabilities given here for "Very tight likelihood" as being representative of those for the "Tight likelihood" in \(^{28}\).

37. The \( \tau \) hadron identification efficiencies and jet fake probabilities corresponding to various working points of the MVA-based \( \tau \) hadron isolation algorithm are shown as a function of \( p_T \) in Figure 4 of \(^{39}\).

38. The \( \tau \) hadron identification efficiencies and electron fake probabilities corresponding to various working points of the MVA-based electron discrimination algorithm are shown as a function of \( p_T \) in Figure 5 of \(^{40}\). The \( \tau \) hadron identification efficiencies and muon fake probabilities are given in section 5.4 of \(^{39}\).

39. See Figure 8 (right) of "Reconstruction and identification of \( \tau \) leptons decaying to hadrons and \( \nu_\tau \) at CMS", CMS Collaboration, V. Khachatryan et al., JINST 11, (2016) P01019.

40. The measurement in the CMS data of the probabilities for hadronic jets, electrons, and muons to be misidentified as \( \tau \) hadron candidates and comparison with MC simulations is described in section 10 of \(^{39}\).

41. "Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at \( \sqrt{s} = 13 \) TeV", ATLAS Collaboration, G. Aad et al., ATLAS-CONF-2017-029.

42. "Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment", ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C 76 (2016) no.5, 295.

43. In Figure 7 (left) of \(^{39}\) the fraction of MJ background in the signal region of visible mass in the selected \( Z \rightarrow \tau^+ \tau^- \rightarrow \tau^\tau\) hadron sample is seen to be around 20%, when using the "tight" operating point of the MVA-based \( \tau \) hadron identification algorithm and applying the "tight" isolation requirement to the muon. It can be seen from Figure 4 (left) of \(^{39}\) that the signal efficiency for \( \tau \) hadron decreases by a factor of around 1.5 between the "tight" and "very-very-tight" operating points and from Figure 4 (right) of \(^{39}\) that the background rejection improves by a factor of three. Therefore background to signal ratio improves by roughly a factor of two between the "tight" and "very-very-tight" \( \tau \) hadron operating points. Similarly, in section 6.2 of \(^{24}\) it is stated that the probability for a muon produced in a hadronic jet to satisfy tight isolation requirements is about 0.05 in the barrel, and goes up to about 0.15 in the endcap. This probability is around a factor of two larger than the probability of 0.03 that we have assumed for the probability for a muon produced in a hadronic jet. Taking these two factors of two into account we can, therefore, estimate the fraction of MJ background in the \( Z \rightarrow \tau^+ \tau^- \rightarrow \tau^\tau\) hadron sample selected using the "very-very-tight" operating point of the MVA-based \( \tau \) hadron identification algorithm and our chosen electron/muon isolation to be around 20%(2 x 2) = 5%. The relative uncertainty on the MJ background quoted in section 9.1 of \(^{39}\) is 5% and results from the limited size of the control samples in the \( \int L dt = 36 \) fb\(^{-1}\) of CMS data used in the measurement.

44. "The anti-\( k_t \) jet clustering algorithm", M. Cacciari, G. P. Salam, and G. Soyez, JHEP 4 (2008) 63.
45. ‘Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector’, ATLAS Collaboration, M. Aaboud et al., Phys. Rev. D 96 (2017) 072002.

46. ‘Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV’, CMS Collaboration, V. Khachatryan et al., JINST 12 (2017) P05014.

47. ‘Measurement of the $b$-jet identification efficiency at run-2 using $t\bar{t}$ events’, ATLAS Collaboration, G. Aad et al., JHEP 2018 (2018) 89.

48. ‘Measurement of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV’, CMS Collaboration, A. M. Sirunyan et al., JINST 13 (2018) P05011.

49. Tag probabilities for the DeepCSV Loose $b$-tagging algorithm of CMS has only a very moderate dependence on the number of pile-up interactions, as is shown in the top right-hand plot of Figure 17 of [48]. This means that the $b$-tagging requirements in the event selections will not introduce any significant difference in the pile-up distributions in the selected $t\bar{t}\rightarrow b\bar{b}W^{+}W^{-}$ and $Z/\gamma^{*}\rightarrow \tau^{+}\tau^{-}$ event samples.

50. ‘Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV’, ATLAS Collaboration, M. Aaboud et al., Eur. Phys. J. 78 (2018) 903.

51. ‘Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector’, CMS Collaboration, A. M. Sirunyan et al., arXiv:1903.06078 [hep-ex].

52. The assigned systematic uncertainties in our study correspond to the approximate size of the uncertainties on individual bins in $p_T(Z)$ in [59]. These are dominated by statistical uncertainties in the relevant region of $p_T(Z).

53. ‘A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX’, S. Alioli, P. Nason, C. Oleari and E. Re, JHEP 1006 (2010) 043.

54. ‘A brief introduction to PYTHIA 8.1’, T. Sjöstrand, S. Mrenna and P. Z. Skands, Comput. Phys. Commun. 178 (2008) 852.

55. ‘Event Generation with Sherpa 2.2’, E. Bothmann et al., arXiv:1905.09127 [hep-ph].

56. ‘Measurement of differential $W^{+}W^{-}$ production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector’, ATLAS Collaboration, M. Aaboud et al., arXiv:1905.02422 [hep-ex].

57. ‘Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV’, CMS Collaboration, A. M. Sirunyan et al., Eur. Phys. J. 79 (2019) 368.

58. Table 5 of [51] gives a number of predictions of the $W^{+}W^{-}$ fiducial cross-section, with systematic uncertainties in the range 2-5%. Given that $Z$ boson and EW diboson production are both dominated by $q\bar{q}$-initiated processes it is reasonable to assume that uncertainties due, e.g., to PDFs would at least partially cancel in the ratio. We assume an uncertainty of 3% on the ratio of the cross sections for $Z$ boson and EW diboson production. In [57] the $t\bar{t}$ cross section is measured with a precision of around 3%. Since $t\bar{t}$ production is dominated by gluon-initiated processes we do not assume any cancellation of systematic uncertainties in evaluating the ratio to the $Z$ boson cross section.

59. ‘Measurement of the transverse momentum and $\phi^{*}_{T}$ distributions of Drell-Yan lepton pairs in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector’, ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C 76 (2016) no.5, 291.

60. ‘Measurement of the $Z/\gamma^{*}\rightarrow \tau^{+}\tau^{-}$ cross section with the ATLAS detector’, ATLAS Collaboration, G. Aad et al., arXiv:1908.07305 [hep-ex].

61. ‘Relations of top-quark pair differential and double-differential cross-sections in the $t\bar{t}$+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector’, ATLAS Collaboration, G. Aad et al., arXiv:1905.09127 [hep-ph].