Reconstruction of the Higgs mass in events with Higgs bosons decaying into a pair of $\tau$ leptons using matrix element techniques

Lorenzo Bianchini a, Betty Calpas b, John Conway c, Andrew Fowlie d, Luca Marzola b,e, Lucia Perrini b, Christian Veelken f,*

a Institute for Particle Physics, ETH Zurich, 8093 Zurich, Switzerland
b National Institute for Chemical Physics and Biophysics, 10143 Tallinn, Estonia
c Department of Physics, University of California, Davis, CA 95616, USA
d ARC Centre of Excellence for Particle Physics at the Tera-scale, Monash University, Melbourne, Victoria 3800, Australia
e Institute of Particle Physics, University of Tartu, 50411 Tartu, Estonia
f CERN, 1211 Geneva, Switzerland

1. Introduction

A new boson of mass 125 GeV has been observed by the ATLAS and CMS collaborations [1,2]. The properties of the new particle are compatible with the predictions for the Standard Model (SM) Higgs ($H$) boson [3–8] within the present experimental uncertainties [9–12]. The observation of its decay into a pair of $\tau$ leptons, at a rate that is compatible with the SM expectation, has been reported recently [13–15]. The decay into a pair of $\tau$ leptons allows for the most precise measurement of the direct coupling of the SM H boson to fermions. Decays of heavy resonances into $\tau$ lepton pairs furthermore provide high sensitivity to search for models with an extended Higgs sector, constituting an important experimental signature at the LHC.

The sensitivity of the SM $H \rightarrow \tau\tau$ analysis critically depends on the capability to distinguish the signal from a large irreducible background, arising from $Z/\gamma^* \rightarrow \tau\tau$ Drell–Yan (DY) production. An important handle to separate the signal from the background is the mass of the $\tau$ lepton pair, which we denote by $m_\tau$. The signal is expected to show up as a small bump on the high mass tail of the $m_\tau$ distribution of the background (see e.g., Figs. 8, 9, and 11 of Ref. [13]).

The separation of the signal from the background improves if the mass distribution for the signal is narrow. The SM predicts the total width of the $H$ boson to be $\approx 4$ MeV. Present experimental upper limits on the total width amount to $\approx 10$ times the SM value [16,17]. These limits have been obtained by comparing the rates for off-shell versus on-shell $H$ boson production and depend on certain assumptions. Direct, model independent, upper limits on the total $H$ width, obtained by analyzing the mass spectra in $H \rightarrow ZZ \rightarrow 4\ell$ ($\ell = e, \mu$) and $H \rightarrow \gamma\gamma$ events, are $\approx 1$ GeV. In contrast, the width of the $m_\tau$ distribution reconstructed in SM $H \rightarrow \tau\tau$ events typically amounts to $\approx 20$ GeV and is dominated purely by the experimental resolution.

Different methods for the reconstruction of $m_\tau$ have been discussed in the literature [18–22]. The SVfit algorithm [23] has been used to reconstruct the $H$ boson mass in the SM $H \rightarrow \tau\tau$ analysis as well as in searches for further $H$ bosons predicted by models beyond the SM performed by the CMS collaboration during LHC Run 1 [24,25,13,26–31]. Compared to alternative mass variables, the usage of the SVfit algorithm has improved the sensitivity of the SM $H \rightarrow \tau\tau$ analysis for measuring the signal rate by $\approx 40\%$ [13]. The improvement in sensitivity corresponds to a gain by...
about a factor of two in integrated luminosity of the analysed dataset.

In this paper we report on the development of an improved version of the SVfit algorithm. Two variants of the improved algorithm have been implemented. The first variant allows to reconstruct the mass $m_{\tau}$ of the $\tau$ lepton pair. It has been developed within the paradigm of the matrix element (ME) method [32,33]. Whereas the algorithm described in Ref. [23] uses a likelihood function of arbitrary normalization, the improved algorithm is based on a proper normalization within the formalism of the ME method. The second variant of the improved algorithm uses a likelihood function of arbitrary normalization. The algorithm allows for the reconstruction of not only the mass $m_{\tau}$ of the $\tau$ lepton pair, but of any kinematic function of the two $\tau$ leptons, including the $p_{\tau}, \eta, \phi$, and transverse mass of the $\tau$ lepton pair. A further improvement concerns the extension of the algorithm to account for the experimental resolution on the reconstruction of hadrons that are produced in the $\tau$ decays.

The development of the formalism to handle $\tau$ decays in the ME method constitutes an important result of this paper, which has not been discussed in the literature so far. The formalism described in this paper allows one to extend the ME generated by automated tools such as CompHEP [34,35] or MadGraph [36], which treat $\tau$ leptons as stable particles, by the capability to handle $\tau$ decays.

The paper is organized as follows. In Section 2 we describe the first variant of the improved algorithm, in particular the formalism that we developed to handle $\tau$ decays in the ME method and our treatment of the experimental resolution on the reconstruction of hadrons that are produced in the $\tau$ decays. The second variant of the algorithm is presented in Section 3. The performance of both variants of the improved algorithm in terms of achieved $m_{\tau}$ resolution is compared to the previous version of the SVfit algorithm, used during LHC Run 1, and to selected alternative mass observables in Section 4. The results are discussed in Section 5. The paper concludes with a summary in Section 6.

2. The Matrix element method

In the ME method, an estimator for the unknown model parameter $\Theta$ is obtained by maximizing the probability density:

$$P(y|\Theta) = \frac{Q(y)}{\alpha(\Theta) R(\Theta)} \int dx_\alpha dx_\beta d\phi_{\alpha} \frac{f(x_\alpha) f(x_\beta)}{2 x_\alpha x_\beta}$$

$$\Phi(y|\Theta) = \int dy \left( \frac{Q(y)}{\alpha(\Theta) R(\Theta)} \right)$$

where $\Theta$ is the unknown model parameter, $\alpha(\Theta)$ and $R(\Theta)$ are the likelihood function and the acceptance, respectively, the energies and momenta of the two colliding protons, $\phi_{\alpha}$ represents the centre-of-mass energy, $x_\alpha$ and $x_\beta$ denote the Bjorken scaling variables [37] and $f(x_\alpha)$ and $f(x_\beta)$ the corresponding parton distribution functions (PDF). We use the MSTW 2008 LO PDF set [38] to evaluate $f(x_\alpha)$ and $f(x_\beta)$. We furthermore denote by $n$ the number of particles in the final state, by $p_i^0$ the momentum of the $i$-th final state particle and by $d\phi_{\alpha} = \prod_i \frac{d\phi_{i}^0}{2\pi r_{i}^{0}}$ the differential $n$-particle phase space element. Vector quantities are represented by bold letters. The symbol $|\mathcal{M}(p, \Theta)|^2$ represents the squared modulus of the ME for the process. The $\delta$-functions $\delta(x_\alpha E_\alpha + x_\beta E_\beta - \sum_{i} E_{i})$ and $\delta(x_\alpha p^\alpha + x_\beta p^\beta - \sum_{i} p^i)$ impose energy and momentum conservation. The set of observables measured in the detector is denoted by $y$. The function $W(y|\theta)$ represents the probability density to observe the measured values $y$, given a point $p$ in the $n$-particle phase space, and is referred to as "transfer function" (TF) in the literature, while the function $Q(y)$ is referred to as "indicator function" [39,40]. The value of the indicator function is 1 for events which pass the event selection criteria and 0 otherwise. The efficiency for an event originating at the phase space point $p$ to pass the event selection, i.e. to end up with measured observables $y$ for which $Q(y) = 1$, is denoted by $\epsilon(\theta, y)$. The symbol $\alpha(\Theta)$ corresponds to the inclusive cross section for the process under study, in the case considered in this paper the production, in $pp$ collisions, of a $H$ boson decaying into a pair of $\tau$ leptons with subsequent decay of the $\tau$ pair. The symbol $R(\Theta)$ represents the acceptance of the event selection, that is, the percentage of events which pass the event selection criteria. Division by $\alpha(\Theta) R(\Theta)$ ensures that the TF satisfy the normalization condition $\int dy \frac{Q(y)}{\alpha(\Theta) R(\Theta)} = 1$ for every $\Theta$.

The meaning of Eq. (1) is as follows. The best estimate for the unknown model parameter $\Theta$ is given by the value $\hat{\Theta}$ which maximizes the probability density to observe precisely the values $y$ that are measured in the detector. Within the scope of this paper, the unknown model parameter $\Theta$ corresponds to the mass $m_H$ of the $H$ boson or, equivalently, to the true mass $m_{\tau}$ of the $\tau$ lepton pair in a given event.

The individual terms of Eq. (1) are described in Sections 2.1–2.7. The ME for the process $pp \rightarrow H \rightarrow \tau \tau$ with subsequent decay of the $\tau$ leptons via $\tau \rightarrow e^\pm \nu_{e}, \tau \rightarrow \mu^\pm \nu_{\mu},$ or $\tau \rightarrow$ hadrons + $k_{s}$ is described in Section 2.1. A complication arises from the fact that we use a leading order (LO) ME to model the $H$ boson production process $pp \rightarrow H$. The LO ME strictly applies only to events in which the $H$ boson has zero $p_{T}$, while the production of $H$ bosons at the LHC typically proceeds in association with jets. The treatment of events with hadronic activity, in which the $H$ boson has non-zero $p_{T}$, is detailed in Section 2.2. The TPs are described in Section 2.3 and the integration over the $n$-particle phase space is described in Section 2.4. The computation of the cross section $\sigma(m_H)$ that is needed for a proper normalization of the probability density $P(y|m_H)$ in Eq. (1) is described in Section 2.5. The numerical maximization of the probability density $P(y|m_H)$ with respect to the mass $m_H$ of the H boson is described in Section 2.6. We denote this version of the SVfit algorithm by SVfitMEM.

We conclude this section with a description of an artificial term that we choose to add to the probability density $P(y|m_H)$ in order to reduce tails in the $m_{\tau}$ distribution reconstructed by the algorithm. The structure of this “regularization” term is described in Section 2.7. Distributions of $m_{\tau}$ in simulated events, reconstructed with and without this term, are presented in Section 4.

As a consequence of using LO ME to model the $H$ boson production process $pp \rightarrow H$, the efficiency $\epsilon(\theta, m_{\tau})$ and the acceptance $R(m_{\tau})$ cannot be determined reliably, because they depend on the $H$ boson $p_{T}$ spectrum. In particular for $H$ bosons of low mass $m_H$, the probability for the particles produced in the $\tau$ lepton decays to pass selection criteria on $p_{T}$ and $\eta$, which are necessitated by trigger requirements at the LHC, may vary significantly as function of $H$ boson $p_{T}$. For this reason, we will assume that $m_{\tau}$ is reconstructed before any event selection criteria are applied, i.e. $\Omega(y) = 1, \epsilon(\theta, m_{\tau}) = 1$, and $R(m_{\tau}) = 1$ for all evaluations of Eq. (1). We expect these assumptions to introduce a small bias on the reconstructed $m_{\tau}$ values and possibly a small degradation in $m_{\tau}$ resolution. The bias can be corrected with the Monte Carlo simulation, and we do not expect it to cause a problem in practical applications of our algorithm.
2.1. Matrix element

We decompose the squared modulus of the ME, \( |\mathcal{M}(p, m_{H})|^2 \), for the process \( pp \rightarrow H \rightarrow \tau \) with subsequent decay of the \( \tau \) leptons into electrons, muons, or hadrons into five parts:

\[
|\mathcal{M}(p, m_{H})|^2 = |\mathcal{M}_{pp-H-H\rightarrow\tau\tau}(p, m_{H})|^2 |\mathcal{W}(H)|^2.
\]

\[
|\mathcal{M}_{pp-H-H\rightarrow\tau\tau}(p, m_{H})|^2 = |\mathcal{M}_{12}(\tau\tau)(p)|^2 |\mathcal{M}_{22}(\tau\tau)(p)|^2 .
\]

where we use the superscripts (1) and (2) to refer to the \( \tau \) lepton of positive and negative charge, respectively. The first term, \( |\mathcal{M}_{pp-H-H\rightarrow\tau\tau}(p, m_{H})|^2 \), represents the squared modulus of the ME for H boson production with subsequent decay of the H boson into a pair of \( \tau \) leptons. This term can be computed by using automatized tools such as CompHEP or MadGraph or it can be taken from the literature.

We take \( |\mathcal{M}_{pp-H-H\rightarrow\tau\tau}(p, m_{H})|^2 \) from the literature and decompose it into a product of three factors:

\[
|\mathcal{M}_{pp-H-H\rightarrow\tau\tau}(p, m_{H})|^2 = |\mathcal{M}_{gg-H}\mathcal{W}(H)|^2 |\mathcal{M}_{H-H\rightarrow\tau\tau}|^2 .
\]

We model the H boson production using the LO ME for the gluon fusion process \( gg \rightarrow H \), which accounts for about 90% of the total H boson production rate at the LHC. The corresponding Feynman diagram is shown in Fig. 1. The squared modulus of the ME reads [41]:

\[
|\mathcal{M}_{gg-H}|^2 = \frac{\sqrt{2} G_F}{256 \pi^2} \alpha_s^2 \zeta^2 m_W^4 [1 + (1 - \zeta) f(\zeta)]^2 .
\]

(4)

with \( \zeta = 4 \frac{m_t^2}{m_H^2} \) and:

\[
f(\zeta) = \begin{cases} \arcsin \frac{1}{\sqrt{\zeta}} & \text{if } \zeta \geq 1 , \\ -1 + \left( \frac{1 + \sqrt{1 - \zeta}}{1 - \sqrt{1 - \zeta}} - i \zeta \right)^2 & \text{if } \zeta < 1 . \end{cases}
\]

(5)

The symbols \( G_F \) and \( \alpha_s \) denote, respectively, the Fermi and strong coupling constant. Their numerical values are:

\[
G_F = 1.166 \times 10^{-5} \text{ GeV}^{-2} \quad \text{and} \quad \alpha_s(m_Z) = 0.139384 .
\]

The squared modulus of the Breit-Wigner propagator \( |\mathcal{W}(H)|^2 \), given by:

\[
|\mathcal{W}(H)|^2 = \frac{1}{\left( \frac{m_H^2}{m_H^2} - m_W^2 \right) + m_H^2 f_H^2}
\]

associates H boson production and decay. In Eq. (7), the symbol \( \frac{m_H^2}{m_H^2} = (E_{H1} + E_{H2})^2 - (p_{H1}^2 + p_{H2}^2) \) denotes the mass of the H lepton pair. The squared modulus of the ME for the decay of the H boson into a \( \tau \) lepton pair, summed over the spin states of the two \( \tau \) leptons, is given by:

\[
|M_{H-H\rightarrow\tau\tau}|^2 = \frac{4 G_F m_W^2}{\sqrt{2} \pi} m_H^2 \left( 1 - \frac{m_t^2}{m_H^2} \right) .
\]

(8)

and is related to the branching ratio \( \mathcal{B}(H \rightarrow \tau \tau) \) by [42]:

\[
\mathcal{B}(H \rightarrow \tau \tau) = \frac{1}{16 \pi m_H f_H} \left( 1 - \frac{m_H^2}{m_H^2} \right) |M_{H-H\rightarrow\tau\tau}|^2.
\]

(9)

For a SM H boson, the branching ratio \( \mathcal{B}(H \rightarrow \tau \tau) \) becomes small and the total width \( f_H \) becomes large once the decay into a pair of W bosons is kinematically possible, i.e., for \( m_H \approx 2 m_W \). We remark that in theories beyond the SM, which motivate the search for additional heavy scalars, the branching ratio and total width may be very different from the SM values. In this paper, we assume \( \mathcal{B}(H \rightarrow \tau \tau) = 100\% \) and \( f_H = 10^{-2} m_W \), and we compute \( |M_{H-H\rightarrow\tau\tau}|^2 \) according to Eq. (9). The branching ratio for the decay \( H \rightarrow \tau \tau \) and the total width of the H boson actually have no effect on the value of \( m_{\tau} \) reconstructed by the algorithm, provided that \( \mathcal{B}(H \rightarrow \tau \tau) \) and \( f_H \) are treated consistently when computing the squared modulus of the ME \( |M_{H-H\rightarrow\tau\tau}|^2 \) and the normalization factor \( 1/\sigma(m_{\tau}) \) in Eq. (1).

Concerning the decay of the \( \tau \) leptons, we use the narrow-width approximation (NWA) and for each \( \tau \) lepton we take a separate and independent average over its spin states, effectively ignoring the correlation in spin orientations between the two \( \tau \) leptons. In the NWA, the squared modulus of the Breit-Wigner propagator \( |\mathcal{W}(\tau)|^2 \) that associates \( \tau \) lepton production and decay yields a \( \delta \)-function:

\[
|\mathcal{W}(\tau)|^2 = \frac{\pi}{m_\tau \Gamma_\tau} \delta(m_\tau^2 - m_\tau^2) \quad \text{with} \quad \Gamma_\tau = \frac{1}{\Delta \tau} = 2.267 \times 10^{-2} \text{ GeV} .
\]

(10)

where \( \Delta \tau = 290 \times 10^{-5} \) s denotes the lifetime of the \( \tau \) lepton [43]. The factor \( |\mathcal{M}_{\tau-H\rightarrow\tau\tau}|^2 \) in Eq. (2) represents the decay of the \( i \)-th \( \tau \) lepton. For the decays \( \tau \rightarrow e e \), \( \mu \mu \), and \( \tau \rightarrow \mu \ell_+ \ell_- \), which we refer to as ”lep-tonic” \( \tau \) decays, we take the ME from the literature. Taking the average over the spin states of the \( \tau \) lepton, the squared modulus of the ME is given by [44]:

\[
|M_{\tau-H\rightarrow\tau\tau}|^2 = 64 \mathcal{G}_0^2 (E_e, E_\mu) \mathcal{W}(E_e, E_\mu - p_\ell^2 + p_\tau^2) .
\]

(11)

The modelling of the decays \( \tau \rightarrow \mu \ell_+ \ell_- \) by matrix elements is more difficult, due to the fact that \( \tau \) leptons decay to a variety of hadronic final states, and because some of the decays proceed via intermediate meson resonances [43]. We refer to these decays as ”hadronic” \( \tau \) decays. The ME for the dominant hadronic \( \tau \) decay modes are discussed in the literature [45,46]. In this paper, we use a simplified formalism and treat \( \tau \rightarrow \mu \ell_+ \ell_- \) decays as two-body decays into a hadronic system \( 3\tau \) of momentum \( p_{3\tau} \) and mass \( m_{3\tau} \) plus a \( \ell \). The squared modulus of the ME for the decay is taken to be constant and denoted by \( |\mathcal{M}_{\tau-H\rightarrow\tau\tau}|^2 \). The value of \( |\mathcal{M}_{\tau-H\rightarrow\tau\tau}|^2 \) is chosen such that it reproduces the measured branching fraction for hadronic \( \tau \) decays. The following relation holds for the considered case of a two-body decay:

\[
\mathcal{B}(\tau \rightarrow \mu \ell_+ \ell_-) = \frac{1}{16 \pi m_\tau \Gamma_\tau} \frac{m_\tau^2 - m_{3\tau}^2}{m_\tau^2} |\mathcal{M}_{\tau-H\rightarrow\tau\tau}|^2 .
\]

(12)

from which it follows that:

\[
|M_{\tau-H\rightarrow\tau\tau}|^2 = 16 \pi m_\tau \Gamma_\tau m_\tau^2 - m_{3\tau}^2 \mathcal{B}(\tau \rightarrow \mu \ell_+ \ell_-) ,
\]

(13)

with \( \mathcal{B}(\tau \rightarrow \mu \ell_+ \ell_-) = 0.648 \) [43]. We have verified that the sum of all hadronic final states produced in \( \tau \) lepton decays is well reproduced by our simplified model. Fig. 2 shows the fraction of \( \tau \) lepton energy, in the laboratory frame, carried by the ”visible” \( \tau \) decay products:
We use the term “visible” $\tau$ decay products to refer to the sum of all hadrons produced in decays of the type $\tau \to \text{hadrons} + \nu_\tau$ as well as to the electron or muon produced in the decays $\tau \to e^- \nu_e$ and $\tau \to \mu^- \nu_\mu$, respectively.

### 2.2. Treatment of hadronic activity

The phase-space for QCD radiation is large at the centre-of-mass energies of the LHC and as a consequence, particles with masses up to a few hundred GeV are typically produced in association with a sizeable hadronic activity [47]. We refer to the vectorial sum of all particles in the event that do not originate from the $H$ boson decay as the “hadronic recoil” and denote the corresponding momentum by $p^\text{rec}$. CMS, as well as ATLAS, have developed sophisticated techniques to improve the reconstruction of the hadronic recoil [48,49].

The effect of the hadronic recoil is to modify the kinematics of the $H$ boson decay by boosting the $\tau$ leptons and their decay products, in the transverse plane and in longitudinal direction. The longitudinal component of $p^\text{rec}$ can be accounted for by a small adjustment of the Bjorken scaling variables $x_a$ and $x_b$ and does not cause a problem for modeling the $H$ boson production by LO ME. However, the components of $p^\text{rec}$ in the transverse plane need to be accounted for. The LO ME that we use to model the production of the $H$ boson is adequate for events in which the $H$ boson has zero $p_T$, implying the transverse components of $p^\text{rec}$ to be zero. In order to use the LO ME in analyses at the LHC, the momenta of the visible $\tau$ decay products and of the neutrinos produced in the $\tau$ decays need to be transformed into a frame in which the $H$ boson has zero $p_T$. The transformation is achieved by a Lorentz boost in the transverse plane, using \( \begin{pmatrix} \hat{E}^\text{rec} = \hat{p}_x^\text{rec} = \hat{p}_y^\text{rec} = \hat{p}_z^\text{rec} = 0 \end{pmatrix} \) as boost vector. We denote the true values of energies and momenta by symbols with a hat if they are given in the laboratory frame, and by symbols a tilde in case they are given in the zero transverse momentum frame of the $H$ boson. Symbols with neither hat nor tilde denote values measured in the laboratory frame.

The hadronic recoil is reconstructed with a typical resolution of $10 - 20$ GeV at the LHC [48,49]. It is demonstrated in Ref. [47] that the imperfect reconstruction of the hadronic recoil in the detector may cause a significant bias on the $H$ boson mass reconstructed by the ME method in case one ignores the experimental resolution on the hadronic recoil. We account for the experimental resolution by introduction of a TF $W(\hat{p}_x^\text{rec}, \hat{p}_y^\text{rec} | \hat{p}_x^\text{rec}, \hat{p}_y^\text{rec})$. We then marginalize Eq. (1) with respect to $\hat{p}_x^\text{rec}$ and $\hat{p}_y^\text{rec}$:

\[
\mathcal{P}(\mathbf{y}; \hat{p}_x^\text{rec}, \hat{p}_y^\text{rec} | m_H) = \frac{1}{\sigma(m_H)} \int dx_a dx_b d\hat{p}_x^\text{rec} d\hat{p}_y^\text{rec} \left\{ f(x_a) f(x_b) \right\} (2\pi)^4 \delta\left( \hat{E}^\text{rec} + \hat{E}_{n(1)} + \hat{E}_{n(2)} - \left( \hat{p}_x^\text{rec} + \hat{p}_y^\text{rec} + \hat{p}_z^\text{rec} \right) \right) W(\hat{p}_x^\text{rec}, \hat{p}_y^\text{rec} | \hat{p}_x^\text{rec}, \hat{p}_y^\text{rec}) \cdot
\]

The cross section $\sigma(m_H)$ needs to be replaced by:

\[
\sigma'(m_H) = \sigma(m_H) \int d\hat{p}_x^\text{rec} d\hat{p}_y^\text{rec}
\]

in order for $\mathcal{P}(\mathbf{y}; \hat{p}_x^\text{rec}, \hat{p}_y^\text{rec} | m_H)$ to have the correct dimensions. The symbol $\mathbf{y}$ refers to the measured momenta of the visible $\tau$ decay products, $p_x^\text{rec}$ and $p_y^\text{rec}$. Using the relations $\left( \hat{E}_{n(1)}, \hat{p}_x^\text{rec} \right) = \left( x_a, 0, 0, 1 \right)$ and $\left( \hat{E}_{n(2)}, \hat{p}_y^\text{rec} \right) = \left( x_b, 1, 0, 0, -1 \right)$ for the energies and momenta of the two colliding protons, the integral over the $\delta$-function that ensures the conservation of energy and momentum can be expressed by:

\[
\int dx_a dx_b d\hat{p}_x^\text{rec} d\hat{p}_y^\text{rec} \delta\left( \hat{E}_{n(1)} + \hat{E}_{n(2)} - \left( \hat{p}_x^\text{rec} + \hat{p}_y^\text{rec} + \hat{p}_z^\text{rec} \right) \right) = \frac{2}{s} \int d\hat{p}_x^\text{rec} d\hat{p}_y^\text{rec} \delta\left( \hat{p}_{x(1)} + \hat{p}_{y(1)} + \hat{p}_{z(1)} \right) \delta\left( \hat{p}_{x(2)} + \hat{p}_{y(2)} + \hat{p}_{z(2)} \right),
\]

with:

\[
x_a = \frac{1}{\sqrt{s}} \left( \hat{E}_{n(1)} + \hat{E}_{n(2)} - \left( \hat{p}_{x(1)} + \hat{p}_{y(1)} + \hat{p}_{z(1)} \right) \right),
\]

\[
x_b = \frac{1}{\sqrt{s}} \left( \hat{E}_{n(1)} + \hat{E}_{n(2)} - \left( \hat{p}_{x(2)} + \hat{p}_{y(2)} + \hat{p}_{z(2)} \right) \right).
\]

The integration over $d\hat{p}_x^\text{rec}$ and $d\hat{p}_y^\text{rec}$ removes the $\delta$-functions $\delta\left( \hat{p}_{x(1)} + \hat{p}_{y(1)} + \hat{p}_{z(1)} \right)$ and $\delta\left( \hat{p}_{x(2)} + \hat{p}_{y(2)} + \hat{p}_{z(2)} \right)$ in Eq. (17). Substituting Eq. (2) into Eq. (15), we obtain:
\[
\mathcal{P}(p^\text{vis(1)}, p^\text{vis(2)}, \hat{p}^\text{rec}_x, \hat{p}^\text{rec}_y | m_H) = \frac{1}{\sigma(m_H)} \frac{32\pi^4}{s} \int d\phi_n \frac{f(x_n)f(y_n)}{x_n y_n} |M_{pp \rightarrow H \eta \tau}(\hat{p}, m_H)|^2.
\]

\[
|\mathcal{B}|(p^\text{vis(1)}, (\hat{p}^2)|\mathcal{B}|(p^\text{vis(2)}, (\hat{p})^2), W(p^\text{vis(1)} | p^\text{vis(2)}) W(p^\text{vis(2)} | p^\text{vis(1)}) W^{\text{rec}}(\hat{p}^\text{rec}_x, \hat{p}^\text{rec}_y | \hat{p}^\text{rec}_x, \hat{p}^\text{rec}_y),
\]

with \(x_n\) and \(y_n\) given by Eq. (18) and:

\[
\hat{p}^\text{rec}_x = -\left(\frac{A^x}{\hat{p}^x} + \frac{A^{x(2)}}{\hat{p}^x}\right), \quad \hat{p}^\text{rec}_y = -\left(\frac{A^y}{\hat{p}^y} + \frac{A^{y(2)}}{\hat{p}^y}\right).
\]

The form of the TF for the hadronic recoil, \(W^{\text{rec}}(p^\text{rec}_x, p^\text{rec}_y | \hat{p}^\text{rec}_x, \hat{p}^\text{rec}_y)\), is discussed in Section 2.3.3.

In practice, the experimental resolution on the momenta of electrons, muons and also \(\eta\) is typically negligible compared to the resolution on the hadronic recoil. In good approximation, the resolution on the hadronic recoil is equivalent to the resolution on the vectorial sum of the momenta, in the transverse plane, of all particles reconstructed in the event. The latter is referred to as “missing transverse momentum” and denoted by \(p^\text{miss}\). The following relations hold for the components of the \(p^\text{miss}\) vector:

\[
p^\text{miss}_x = -\left(p^\text{rec}_x + p^{x(1)} + p^{x(2)}\right) \quad \text{and} \quad p^\text{miss}_y = -\left(p^\text{rec}_y + p^{y(1)} + p^{y(2)}\right).
\]

These relations are valid on reconstruction level as well as for the true values of the momenta, i.e. Eq. (21) is valid also in case all \(p^\text{rec}\) and \(p\) are replaced by \(\hat{p}^\text{rec}\) and \(\hat{p}\). Approximating the resolution on the hadronic recoil by the resolution on \(p^\text{rec}\) has the advantage that the resolutions on \(p^\text{miss}\) have been studied in detail and published by the ATLAS as well as CMS collaborations [48,49].

2.3. Transfer functions

The treatment of the experimental resolution on the momenta of electrons, muons and \(\eta\) produced in \(\tau\) decays, as well as of the hadronic recoil, are detailed in this section. The resolution on the momentum \(p^\text{vis}\) of the visible \(\tau\) decay products is parameterized as function of the true momentum \(\hat{p}\) and modeled by TF \(W(p^\text{vis} | \hat{p}^\text{vis})\). The case of hadronic and leptonic \(\tau\) decays is described in Sections 2.3.1 and 2.3.2, respectively. The neutrinos produced in the \(\tau\) decays are not included in the TF, but are handled separately, by the formalism detailed in Sections 7.1 and 7.2 of the appendix. The TF \(W^{\text{rec}}(p^\text{rec}_x, p^\text{rec}_y | \hat{p}^\text{rec}_x, \hat{p}^\text{rec}_y)\) that we use to model the experimental resolution on the hadronic recoil is discussed in Section 2.3.3.

2.3.1. Hadronic \(\tau\) decays

The energy, or equivalently \(p_T\), of the \(\eta\) is reconstructed with a resolution of 5–25% at the LHC [51,52]. The resolution typically varies as function of \(p_H\) and \(\eta\) and may additionally depend also on the multiplicity of the charged and neutral hadrons produced in the \(\tau\) decay. The resolution on the \(p_T\) of the \(\eta\) is of similar magnitude as the resolution on \(m_H\), that we aim to achieve and needs to be taken into account by suitable TF, denoted by \(W(p^\text{vis} | \hat{p}^\text{vis})\), when evaluating the integral in Eq. (19).

ATLAS as well as CMS report that the energy response for hadronic \(\tau\) decays may be asymmetric [53,54]. For the purpose of this paper, we assume the TF \(W(p^\text{vis} | \hat{p}^\text{vis})\) to follow a Gaussian distribution within the core region \(x = p_T^\text{vis}/p_T^\text{rec} \approx 1\) and to feature non-Gaussian tails, which follow power-law functions, on both sides of the Gaussian core. More specifically, we use the form:

\[
\frac{N}{2} \xi x \left(\frac{\alpha_1}{x} - x - \frac{x - \mu}{\sigma}\right) \quad \text{if} \quad x < x_1
\]

\[
N \exp\left(-\frac{1}{2} \left(\frac{x - \mu}{\sigma}\right)^2\right) \quad \text{if} \quad x_1 \leq x \leq x_2
\]

\[
\frac{N}{2} \xi_2 x \left(\frac{\alpha_2}{x} - x + \frac{x - \mu}{\sigma}\right) \quad \text{if} \quad x > x_2
\]

and we use the values \(\mu = 1.0, \sigma = 0.03, x_1 = 0.97, x_2 = 1.03,\) and \(\alpha_1 = 3.5\) for its parameters. The parameter values are chosen to approximately reproduce the \(\eta\) energy resolution expected for LHC Run 2 in case of the CMS experiment. The factors \(\xi_1\) and \(\xi_2\) are chosen such that the function \(W(p^\text{vis} | \hat{p}^\text{vis})\) is continuous at the points \(x = x_1\) and \(x = x_2\). The corresponding values are \(\xi_1 = \exp\left(-\frac{1}{2} \left(\frac{\alpha_1}{x} - x - \frac{x - \mu}{\sigma}\right)^2\right)\) and \(\xi_2 = \exp\left(-\frac{1}{2} \left(\frac{\alpha_2}{x} - x + \frac{x - \mu}{\sigma}\right)^2\right)\). The factor \(N\) is determined by the requirement that the function \(W(p^\text{vis} | \hat{p}^\text{vis})\) satisfies the normalization condition \(\int dp_T^\text{vis} W(p_T^\text{vis} | \hat{p}_T^\text{vis}) = 1\) for any given value of \(\hat{p}_T^\text{vis}\). The TF given by Eq. (22) is visualized in Fig. 3.

The resolution on the direction of the \(\eta\) is on the level of a few milliradians and is negligible in practice. We hence model the TF
for the momentum of the $\eta_{vis}$ by the product of Eq. (22) and two $\delta$-functions:

$$W_{\eta_{vis}}(p_{\eta_{vis}}^\text{vis}) = \frac{\sin^2\theta_{vis}}{(p_{\eta_{vis}}^\text{vis})^2} W_\eta(p_{\eta_{vis}}^\text{vis}, p_{\phi_{vis}}^\text{vis}) \delta(\eta_{vis} - \hat{\eta}_{vis}) \delta(\phi_{vis} - \hat{\phi}_{vis}).$$

(23)

The factor $\frac{\sin^2\theta_{vis}}{(p_{\eta_{vis}}^\text{vis})^2}$ is needed to ensure the correct normalization of the TF:

$$\int d^3p W_{\eta_{vis}}(p_{\eta_{vis}}^\text{vis}) = \int dp_{\eta_{vis}} dp_{\phi_{vis}} W_{\eta}(p_{\eta_{vis}}^\text{vis}, p_{\phi_{vis}}^\text{vis}) = 1.$$  

(24)

where the factor $(p_{\eta_{vis}}^\text{vis})^2/\sin^2\theta_{vis}$ corresponds to the Jacobian of the variable transformation from $(p_{\eta_{vis}}^\text{vis} = p_{\eta_{vis}}^\text{vis} \cos\theta_{vis}, p_{\phi_{vis}}^\text{vis} = p_{\phi_{vis}}^\text{vis} \sin\theta_{vis}$.

A few words of explanation are in order concerning the treatment of the mass $m_{\eta_{vis}}$ of the hadronic system $\eta_{vis}$ produced in $\tau \rightarrow\text{hadrons} + \nu$ decays. Within the scope of this paper, we will assume that $m_{\eta_{vis}}$ can be reconstructed with negligible experimental resolution. The $\eta_{vis}$ reconstruction algorithm of the CMS experiment during LHC Run 1 [51] does not allow to reconstruct the mass of $\eta_{vis}$ pair decays. Within the scope of this paper, we will assume that $m_{\eta_{vis}}$ can be reconstructed with negligible experimental resolution. The $\eta_{vis}$ reconstruction algorithm used by the ATLAS experiment during LHC Run 1 [51] does not allow to reconstruct the mass of $\eta_{vis}$. In case of semi-leptonic $\tau$ decays, the integral over $d\phi_{\eta_{vis}}$ is of dimension 12 in case of hadronic $\tau$ pair decays, 15 in case of semi-leptonic $\tau$ pair decays, and 18 in case of lepton $\tau$ pair decays. The differential space element reads:

$$d\phi_{\eta_{vis}} = \frac{d\phi_{\phi_{vis}}(\eta_{vis},\eta_{vis})}{2\pi \sqrt{W_{\eta_{vis}}}} \delta(\eta_{vis} - \hat{\eta}_{vis}) \delta(\phi_{vis} - \hat{\phi}_{vis}).$$

(25)

2.3.2. Leptonic $\tau$ decays

We assume that the $p_{\eta_{vis}}^\text{vis}$, $\eta_{vis}$, and $\phi_{vis}$ of the electrons and muons produced in $\tau$ decays are measured with negligible experimental resolution. Correspondingly, we model the TF by a product of three $\delta$-functions:

$$W_{\eta_{vis}}(p_{\eta_{vis}}^\text{vis}) = \frac{\sin^2\theta_{vis}}{(p_{\eta_{vis}}^\text{vis})^2} W_\eta(p_{\eta_{vis}}^\text{vis}, p_{\phi_{vis}}^\text{vis}) \delta(\eta_{vis} - \hat{\eta}_{vis}) \delta(\phi_{vis} - \hat{\phi}_{vis}).$$

(25)

2.3.3. Hadronic recoil

We use a two-dimensional normal distribution:

$$W_{\eta_{vis}}(p_{\eta_{vis}}^\text{vis}, p_{\phi_{vis}}^\text{vis}) = \frac{1}{2\pi \sqrt{W_{\eta_{vis}}}} \exp\left(-\frac{1}{2} \frac{\Delta p_{\eta_{vis}}^\text{rec}}{\Delta p_{\eta_{vis}}^\text{rec}} \cdot \Delta p_{\eta_{vis}}^\text{rec} \right),$$

(26)

to model the experimental resolution on the momentum, in the transverse plane, of the hadronic recoil. The symbols:

$$\Delta p_{\eta_{vis}}^\text{rec} = p_{\eta_{vis}}^\text{rec} - \hat{p}_{\eta_{vis}}^\text{rec} \quad \text{and} \quad \Delta p_{\eta_{vis}}^\text{rec} = p_{\phi_{vis}}^\text{rec} - \hat{p}_{\eta_{vis}}^\text{rec}$$

(27)

refer to the difference between reconstructed and true values of the momentum in $x$- and $y$-direction, and the covariance matrix:

$$\Delta\Sigma = \begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix}$$

(28)

accounts for the fact that the resolutions $\sigma_x$ and $\sigma_y$ in $x$- and $y$-direction may be correlated, with the correlation coefficient denoted by $\rho$. The symbol $|V|$ denotes the determinant of $V$. Two-dimensional normal distributions have been demonstrated to model well the resolution on $p_{\eta_{vis}}^\text{vis}$ in case of the CMS experiment.
and \( W(\mathbf{p}_{\text{vis}}^{\tau}(\mathbf{p}^\tau)) \). The \( \tau \) lepton energies and momenta need to be computed as function of the integration variables.

The dimension of the integration over the phase-space elements \( d^3\mathbf{p}(\mathbf{p}) \) and \( d^3\mathbf{p}(\mathbf{p}) \) can be reduced by means of analytic transformations. Two variables are sufficient to parametrize the kinematics of hadronic \( \tau \) decays. In case of leptonic \( \tau \) decays, three variables are sufficient. We choose to parametrize hadronic \( \tau \) decays by the variables \( z \) and \( \phi^{-\tau} \). The variable \( z \) represents the fraction of \( \tau \) lepton energy, in the laboratory frame, that is carried by the visible \( \tau \) decay products (cf. Eq. (14)). We denote the energy and momentum of the \( \tau \) neutrino produced in hadronic \( \tau \) decays as well as of the neutrino pair produced in leptonic \( \tau \) decays by the symbols \( E_{\nu\nu} \) and \( \mathbf{p}_{\nu\nu} \). The energy component \( E_{\nu\nu} \) is related to the variable \( z \) via:

\[
E_{\nu\nu} = \frac{1 - z}{z} E_{\text{vis}}.
\]

The angle \( \theta_{\text{inv}} \) between the \( \mathbf{p}_{\text{inv}} \) vector and the \( \mathbf{p}_{\text{vis}} \) vector is related to the variable \( z \) as well, and is given by Eq. (48) in case of hadronic \( \tau \) decays and by Eq. (67) in case of leptonic \( \tau \) decays. The variable \( \phi_{\nu\nu} \) specifies the orientation of the \( \mathbf{p}_{\nu\nu} \) vector relative to the direction of the visible \( \tau \) decay products. In case of hadronic \( \tau \) decays, the magnitude of the \( \mathbf{p}_{\text{inv}} \) vector is equal to \( E_{\nu\nu} \). We choose the mass \( m_{\nu\nu} \) of the neutrino pair as third variable to parametrize the kinematics of leptonic \( \tau \) decays, so that the magnitude of the \( \mathbf{p}_{\text{inv}} \) vector is given by

\[
\frac{(1 - z)^2}{z} E_{\text{vis}} - m_{\nu\nu}^2.
\]

With the convention that \( m_{\nu\nu} = 0 \) for hadronic \( \tau \) decays, Eqs. (48) and (67) can be expressed by a common form that is valid for hadronic as well as for leptonic \( \tau \) decays:

\[
\cos\theta_{\text{inv}} = \frac{1 - z}{z} E_{\text{vis}} - \frac{1}{2} \left( m_{\text{vis}}^2 - m_{\nu\nu}^2 \right).
\]

The \( \tau \) lepton momentum vector is given by the sum of the \( \mathbf{p}_{\text{vis}} \) and \( \mathbf{p}_{\text{inv}} \) vectors.

The angles \( \theta_{\text{inv}} \) and \( \phi_{\text{inv}} \) are illustrated in Fig. 4. The \( \mathbf{p}_{\text{vis}} \) vector is located on the surface of a cone, the axis of which is given by the \( \mathbf{p}_{\text{vis}} \) vector and the opening angle of which is given by Eq. (32). The variable \( \phi_{\text{inv}} \) represents the angle of rotation, in counter-clockwise direction, around the axis of the cone. The value \( \phi_{\text{inv}} = 0 \) is chosen to correspond to the case that the \( \mathbf{p}_{\text{inv}} \) vector is within the plane spanned by the \( \mathbf{p}_{\text{vis}} \) vector and the beam direction.

The parametrization of the \( \tau \) decay kinematics by \( \mathbf{p}_{\text{vis}} \) and the variables \( z \) and \( \phi_{\text{inv}} \), respectively by \( z \), \( \phi_{\nu\nu} \), and \( m_{\nu\nu} \), allows one to simplify the evaluation of the integral in Eq. (19) considerably. Expressions for the product of the differential phase space elements \( d^3\mathbf{p}(\mathbf{p}) \) and \( d^3\mathbf{p}(\mathbf{p}) \), with the squared modulus of the ME \( |\mathbf{B}|^2 \text{det} |\mathbf{M}^{(\tau)}(\mathbf{p})|^2 \) are derived in Sections 7.1 and 7.2 of the appendix. The results are:

\[
|\mathbf{B}|^2 \text{det} |\mathbf{M}^{(\tau)}(\mathbf{p})|^2 \left( \frac{d^3\mathbf{p}(\mathbf{p})}{2E_{\text{vis}}} \right) dz dm_{\text{vis}} dm_{\nu\nu} \left( d^3\mathbf{p}(\mathbf{p}) \right) dz dm_{\nu\nu} \left( d^3\mathbf{p}(\mathbf{p}) \right)
\]

The functions \( f_b \) and \( f_r \) are given by Eqs. (52) and (68).

Eq. (33) represents the quintessence of what is needed in order to extend the ME generated by automated tools such as ComPHEP or MadGraph by the capability to handle the \( \tau \) decays. Instead of performing an integration over \( d^3\mathbf{p}(\mathbf{p}) d^3\mathbf{p}(\mathbf{p}) \), which treats the \( \tau \) leptons as stable particles, one needs to perform the integration over \( |\mathbf{B}|^2 \text{det} |\mathbf{M}^{(\tau)}(\mathbf{p})|^2 \left( d^3\mathbf{p}(\mathbf{p}) \right) \) according to Eq. (33). The momenta of both \( \tau \) leptons need to be computed as function of the integration variables \( z_{1\tau}, \phi^{(1)}_{\nu\nu}, m_{\nu\nu}^{(1)} \), and \( z_{2\tau}, \phi^{(2)}_{\nu\nu}, m_{\nu\nu}^{(2)} \), using Eqs. (31) and (32), where \( m_{\nu\nu} \) is equal to zero in case the \( i \)-th \( \tau \) lepton decays hadronically. The \( \tau \) lepton momenta can then be used to evaluate the product of the PDF, the flux factor, and the squared modulus \( |\mathbf{M}_{\text{pp}}^{(\text{vis})}(\mathbf{p}, m_{\nu\nu}^2)|^2 \) of the ME for H boson production and subsequent decay into a \( \tau \) lepton pair in Eq. (19).

In order to improve the accuracy of the numerical integration, we perform a further variable transformation, replacing \( z_{2\tau} \) by the variable \( t_{2\tau} \), defined below. The transformation is executed in two steps. First, we replace \( z_{2\tau} \), by:

\[
q_{2\tau}^2 = \frac{q_{2\tau}^2}{z_{1\tau} z_{2\tau}},
\]

with \( q_{2\tau} \) denoting the mass of the visible decay products of both \( \tau \) leptons. Following Eq. (8) in Ref. [47], we then parametrize \( q_{2\tau}^2 \) by:

\[
q_{2\tau}^2 = m_{\tau}^2 + m_{\nu} l_{\tau} \tan t_{2\tau}.
\]

The form of the variable transformation in Eqs. (34) and (35) is chosen such that the Jacobi factor of the transformation from \( z_{2\tau} \) to \( t_{2\tau} \) is proportional to the inverse of \( |\mathbf{B}|^2 \), the squared modulus of the Breit-Wigner propagator of the H boson in Eq. (7). The Jacobi factor is given by:

\[
\frac{dz_{2\tau} dt_{2\tau}}{dz_{2\tau} dt_{2\tau}} = \frac{q_{2\tau}^2}{z_{1\tau} z_{2\tau}} \frac{m_{\tau}^2 + m_{\nu}^2}{m_{\tau} l_{\tau} m_{\nu} l_{\tau}}.
\]

Compared to Ref. [47] we differ by a factor \( \frac{1}{r} \) in the derivative \( \frac{dr}{r} \). We have verified that Eq. (36) is correct. This transformation improves the numerical precision of evaluating the integral in Eq. (19) by reducing the variance of the integrand.

2.5. Computation of \( \sigma(m_{\nu\nu}) \)

According to the paradigm of the ME method, the normalization factor \( 1/\sigma(m_{\nu\nu}) \) in Eq. (19) is to be computed by evaluating the integral:

\[
\sigma(m_{\nu\nu}) = \frac{32\pi^4}{s} \int d^3\mathbf{p}(\mathbf{p}) d^3\mathbf{p}(\mathbf{p}) dy_{\text{rec}} dy_{\text{rec}} dy_{\text{rec}} \left( f(x_f f(x_b)) \right)
\]

\[
|\mathbf{M}_{\text{pp}}^{(\text{vis})}(\mathbf{p}, m_{\nu\nu}^2)|^2 |\mathbf{B}|^2 \text{det} |\mathbf{M}^{(\tau)}(\mathbf{p})|^2 |\mathbf{B}|^2 \text{det} |\mathbf{M}^{(\tau)}(\mathbf{p})|^2.
\]

\[
W(\mathbf{p}^{(1)}(\mathbf{p}^{(2)}) | W(\mathbf{p}^{(1)}(\mathbf{p}^{(2)})) W_{\text{rec}}(\mathbf{p}^{(1)}(\mathbf{p}^{(2)})) \right). \]

Of the factors in the integrand of Eq. (37), only the TF depend on the measured momenta \( \mathbf{p}^{(1)}(\mathbf{p}^{(2)}) \) of the visible decay products of the two \( \tau \) leptons and on the measured momentum.
The values agree with the literature values for the LO gg → H cross section. Applying the LO ME to events in which the H boson recoils against a high pT jet. The scale of the ordinate is adjusted such that the probability density is equal to one for the value of mH\text{true} that maximizes \( \rho(p^{\text{vis}1}, p^{\text{vis}2}, p^{\text{rec}, \text{i}}, p^{\text{rec}, \text{jet}}) \).

The best estimate, \( m_\tau \), for the mass of the \( \tau \) lepton pair in a given event is obtained by computing the probability density \( \rho(p^{\text{vis}1}, p^{\text{vis}2}, p^{\text{rec}, \text{i}}, p^{\text{rec}, \text{jet}}) \) for a series of mass hypotheses \( m_{\text{H} \text{test}} \), using Eq. (19), and determining the value of \( m_\tau \) that maximizes this probability density.

The integral in Eq. (19) is evaluated numerically using the VAMP algorithm. For each mass hypotheses \( m_{\text{H} \text{test}} \), by inserting two \( \delta \)-functions to remove two integration variables analytically before evaluating the integral numerically. The values of \( \sigma(m_\tau) \) obtained in this way are shown as function of \( m_{\text{H} \text{test}} \) in Fig. 5. The values agree with the literature values for the LO gg → H cross section within ≈10%. The level of agreement is sufficient for our purposes.

2.6. Determination of \( m_\tau \)

The cross sections \( \sigma(m_\tau) \) computed in this way cannot be directly compared to literature values, due to the fact that we are applying the LO ME to events in which the H boson has non-zero \( p_T \). For the purpose of comparing \( \sigma(m_\tau) \) to literature values for the LO gg → H cross section, we compute the integral in Eq. (39) for the case that the H boson has zero \( p_T \), by inserting two \( \delta \)-functions, \( \delta(p_T^{\text{vis}1} + p_T^{\text{vis}2}) \) and \( \delta(p_T^{\text{rec}, \text{i}} + p_T^{\text{rec}, \text{jet}}) \), into the integrand. We then use the \( \delta \)-functions to remove two integration variables analytically before evaluating the integral numerically. The values of \( \sigma(m_\tau) \) obtained in this way are shown as function of \( m_{\text{H} \text{test}} \) in Fig. 6.
evaluated 20 000 times. The series of mass hypotheses is defined by a recursive relation:

$$m_H^{\text{test}(i+1)} = (1 + \delta) m_H^{\text{test}(i)} \quad \text{with} \quad m_H^{\text{test}(0)} = m_{\text{vis}}.$$  \hfill (40)

where $m_{\text{vis}}$ denotes the mass of the visible $\tau$ decay products. The step size $\delta$ is chosen such that it is small compared to the expected resolution on $m_H$, which typically amounts to 15–20% relative to the true mass of the $\tau$ lepton pair (cf. Section 4).

In order to reduce the computing time, the series is computed in two passes. The purpose of the first pass, which uses a step size $\delta = 0.10$, is to find an approximate value of $m_H$ that maximizes the probability density $P$. The series of mass hypotheses that is used in the first pass stops when $P$ falls below one per mille of the maximal probability density $P_{\text{max}}$ computed for any $m_H^{\text{test}}$ so far in a given event. In the second pass, which uses a step size of $\delta = 0.01$, further $P$ values are computed for mass hypotheses $m_H^{\text{test}}$ that are within a region around the maximum for which the probability density computed in the first pass exceeds 0.10 $P_{\text{max}}$. 

---

**Fig. 8.** Distributions in $m_\tau$ reconstructed by the CA method and different versions of the SVfit algorithm in simulated $Z/\gamma^* \to \tau \tau$ background events (a,c,e,g) and in SM $H \to \tau \tau$ signal events produced via the $gg \to H$ (b,d,f) and $q\bar{q} \to H$ (h) production processes in different event categories: 0-jet (a,b), 1-jet non-boosted (c,d), 1-jet boosted (e,f), and 2-jet VBF (g,h). The events are selected in the $\tau^\pm \tau^\mp$ decay channel. The axis of abscissae ranges from 0.4 to 8.
Finally, the graph of $\log(P \cdot \text{GeV}^8)$ versus $m_{H}^{\text{mass}}$ is fitted by a second order polynomial in the region around the maximum, and $m_{\tau}$ is taken to be the point at which the polynomial reaches its maximum.

Graphs of the probability density $P(m_{\tau}/m_{H}^{\text{mass}})$ as function of $m_{\tau}$, $m_{H}^{\text{mass}}$, and $m_{H}^{\text{true}}$ are shown in Fig. 6 for two exemplary events. The events are drawn from a simulated sample of $Z/\gamma^{*} \rightarrow \tau \tau$ background events, produced as described in Section 4. In the first event ("0-jet") the $Z$ boson has little $p_{T}$, while in the second event ("1-jet boosted") the $Z$ boson recoils against a high $p_{T}$ jet. The scale of the ordinate is adjusted such that the probability density is equal to one for the value of $m_{H}^{\text{mass}}$ that maximises $P$. The width of the graph reflects the experimental resolution on $m_{\tau}$. As will be explained in more detail in Section 4, the resolution on $m_{\tau}$ improves considerably in case the Z boson has high $p_{T}$. The graphs of the probability density are superimposed for two cases: for the case that the artificial regularization term described in Section 2.7 is used and for the case that it is not used. In the $Z/\gamma^{*} \rightarrow \tau \tau$ background event in which the $Z$ boson has little $p_{T}$ the probability density decreases slowly as function of the mass hypothesis in case no artificial regularization term is used with
the effect that such events have a sizeable probability for the reconstructed $m_\tau$, to populate, due to resolution effects, the region in which the Higgs boson signal is searched for.

2.7. Artificial regularization term

The SM $H \rightarrow \tau\tau$ signal is produced at a rate about three orders of magnitude smaller than the irreducible $Z/\gamma^* \rightarrow \tau\tau$ background. As illustrated in Fig. 5, the production rate for hypothetical heavy resonances, such as heavy pseudoscalar Higgs bosons or heavy spin 1 resonances, typically decreases steeply with mass. Physics analyses at the LHC will soon start to probe signal cross sections of order 1 fb. In order to maintain high sensitivity for the SM $H \rightarrow \tau\tau$ signal as well as for hypothetical heavy resonances it is imperative to reduce, as much as possible, high mass tails in the $m_\tau$ distribution reconstructed in $Z/\gamma^* \rightarrow \tau\tau$ background events, as tails that are on or below the per mille level may yet compete with or even dwarf potential signals.

Fig. 10. Distributions in $m_\tau$ reconstructed by the CA method and different versions of the SVfit algorithm in simulated $Z/\gamma^* \rightarrow \tau\tau$ background events (a,c,e,g) and in SM $H \rightarrow \tau\tau$ signal events produced via the $gg \rightarrow H$ (b,d,f) and $q \bar{q} \rightarrow H$ (h) production processes in different event categories: 0-jet (a,b), 1-jet non-boosted (c,d), 1-jet boosted (e,f), and 2-jet VBF (g,h). The events are selected in the $\mu e$ decay channel. The axis of abscissae ranges from 0.4 to 8.
As exemplified in Fig. 6, high mass tails in the $m_{\tau\tau}$ distribution for the $\gamma\tau\tau^* \rightarrow Z/\eta$ background predominantly arise from events in which the $Z$ boson has little $p_T$ and the probability density $P$ decreases slowly as function of $m_{\text{Htest}}$.

Penalized maximum likelihood (ML) estimation is an established method for circumventing problems that arise in the stability of parameter estimates in case the likelihood function is relatively flat [58]. Instead of maximising the log-likelihood function $\log L(\Theta|y)$, the penalized ML method finds the best estimate $\hat{\Theta}$ for the unknown model parameter $\Theta$ by maximising the sum $\log L(\Theta|y) + r(\Theta)$. The penalty function $r(\Theta)$ is added to the log-likelihood function in order to pull the best estimate $\hat{\Theta}$ for the parameter $\Theta$ towards a value that has some rationale as good guess for $\Theta$ in information outside of the likelihood function. The penalized ML approach can be viewed as a method for introducing some tolerable degree of bias in exchange for a reduction in the sampling variability of parameter estimates [59].

From a Bayesian perspective, the penalty function $r(\Theta)$ can be interpreted as prior distribution that describes the information that one has on the parameter $\Theta$ outside of any information conveyed by the measured observables $y$.

In the context of the SVfit algorithm, we choose a penalty function $r(m_{\text{Htest}})$ which gives preference to low values of $m_{\tau\tau}$ in case the probability density $P$ is relatively flat. Recall that the cross section for producing resonances decaying into $\tau$ lepton pairs decreases steeply with the mass of the resonance (cf. Fig. 5). The information that large mass values are less likely is not contained within the probability density defined by Eq. (19) and motivates the use of a penalized ML approach with a penalty function $r(m_{\text{Htest}})$ that increases (decreases) for small (large) $m_{\text{Htest}}$.

In previous applications of the SVfit algorithm for data analyses performed by the CMS collaboration, regularization functions of the following type were considered:

$$r(m_{\text{Htest}}) = -\kappa \cdot \log(m_{\text{Htest}} \text{ GeV}^{-1})$$

We will focus on this type of regularization functions in this paper, but remark that from a Bayesian perspective a well motivated alternative choice would be to use as penalty function the logarithm of the cross section as function of mass.

Our choice of the parameter $\kappa$ is performed with the objective of achieving an optimal compromise between reducing the high mass tail in the $m_{\tau\tau}$ distribution reconstructed for $Z/\eta$ background events on the one hand and causing no or at most a small bias on the $m_{\tau\tau}$ distribution reconstructed in signal events on the other hand. We find that the optimal value of $\kappa$ depends on the experimental resolution on the $p_T$ of $\tau$ and on $p_T^{\text{min}}$ and hence needs to be adjusted to the experimental conditions. Higher (lower) experimental resolution favors a small (large) value of $\kappa$. The optimal value of $\kappa$ may furthermore differ for events in which both $\tau$ leptons decay hadronically, events in which one $\tau$ lepton decays into hadrons and the other into an electron or muon, and events in which both $\tau$ leptons decay into electrons or muons, with...
larger (smaller) values of $\kappa$ being favored in case both (none) of the $\tau$ leptons decay hadronically.

The effect of adding a regularization term of the kind $i (m_{\text{vis}}(\text{test})) = -\kappa \log(m_{\text{vis}}(\text{test}))$ to the probability density $P$ given by Eq. (19) on the distribution in $m_{\tau\tau}$ reconstructed in $Z_{/\gamma^*} \to \tau\tau$ background events is visualized in Fig. 7. The effect on events in which the $Z$ boson recoils against a high $p_T$ jet, for which the probability density typically exhibits a narrow maximum, is small, while the addition of the regularization term effectively reduces the high mass tail in the $m_{\tau\tau}$ distribution reconstructed in events in which the $Z$ boson has little $p_T$, for which the probability density decreases more slowly as function of $m_{\text{vis}}(\text{test})$ (cf. Fig. 6).

3. “Classic” SVfit algorithm

A variant of the SVfit algorithm without proper normalization of the likelihood function is maintained for analyses of LHC Run 2 data. The algorithm differs from the SVfitMEM algorithm, described in Section 2, in two respects, altering the integrand in Eq. (1) as well as the procedure for computing the integral. The indicator function $\Omega(y)$, the PDF factor $f(x_a)f(x_b)$, the flux factor $\frac{1}{2} k_a k_b$, and the efficiency factor $\epsilon(\rho, \theta)$ of Eq. (1) are omitted, as is the normalization to inclusive cross section times acceptance, given by the factor $\frac{1}{2} m_N^2$. Instead of the complete ME $|M(p, m_\rho)|^2$ given by Eq. (2), only the terms $|BW_\tau^{(1)}|^2 |M_\tau^{(1)}(p)|^2$, $|BW_\tau^{(2)}|^2 |M_\tau^{(2)}(p)|^2$ for the decay of the $\tau$
leptons are retained. The equivalent to Eq. (19) reads:

\[
\mathcal{L}(p_{\text{vis}^{(1)}}, p_{\text{vis}^{(2)}}, p_x^{\text{rec}}, p_y^{\text{rec}} | m_t) = \frac{32\pi^4}{s} \int d\phi_n \cdot \\
|\mathbf{W}_1^{(1)}| |\mathcal{M}_1^{(1)}(\mathbf{p})|^2 |\mathbf{W}_2^{(1)}| |\mathcal{M}_2^{(1)}(\mathbf{p})|^2, \\
W(p_{\text{vis}^{(1)}} | p_{\text{vis}^{(1)}}) W(p_{\text{vis}^{(2)}} | p_{\text{vis}^{(2)}}) W_{\text{rec}}(p_x^{\text{rec}}, p_y^{\text{rec}} | p_x^{\text{rec}}, p_y^{\text{rec}}). 
\]

(42)

Instead of computing probability densities \(\mathcal{P}(p_{\text{vis}^{(1)}}, p_{\text{vis}^{(2)}}, p_x^{\text{rec}}, p_y^{\text{rec}} | m_t)\) for a series of mass hypotheses \(m_t^{\text{vis}^{(1)}}\) and determining the value of \(m_t\) that maximizes the probability density, the following integral is computed:

\[
\mathcal{L}(p_{\text{vis}^{(1)}}, p_{\text{vis}^{(2)}}, p_x^{\text{rec}}, p_y^{\text{rec}} | m_t) = \frac{32\pi^4}{s} \int d\phi_n \cdot \\
|\mathbf{W}_1^{(1)}| |\mathcal{M}_1^{(1)}(\mathbf{p})|^2 |\mathbf{W}_2^{(1)}| |\mathcal{M}_2^{(1)}(\mathbf{p})|^2, \\
W(p_{\text{vis}^{(1)}} | p_{\text{vis}^{(1)}}) W(p_{\text{vis}^{(2)}} | p_{\text{vis}^{(2)}}) W_{\text{rec}}(p_x^{\text{rec}}, p_y^{\text{rec}} | p_x^{\text{rec}}, p_y^{\text{rec}}) \mathcal{P}(\mathbf{p}). 
\]

The function \(\mathcal{P}(\mathbf{p})\) in the integrand may be an arbitrary function of the momenta \(\mathbf{p}^{(1)}\) and \(\mathbf{p}^{(2)}\) of the two \(\tau\) leptons. The integral is evaluated numerically, using a custom implementation of the Markov chain
Monte Carlo integration method with the Metropolis–Hastings algorithm [60]. The actual value $\mathcal{L}(y)$ of the integral is irrelevant. The reconstruction of the mass $m_{\tau}\ell$ of the $\tau$ lepton pair is based on choosing $\mathcal{F}(\mathbf{p}) \equiv \left( \hat{E}_{\tau(1)} + \hat{E}_{\tau(2)} \right)^2 - \left( \left( \vec{p}_{\tau(1)} + \vec{p}_{\tau(2)} \right)^2 \right) - \left( \left( \vec{p}_{\ell(1)} + \vec{p}_{\ell(2)} \right)^2 \right)$, recording the values of $\mathcal{F}(\mathbf{p})$ for each evaluation of the integrand in Eq. (43) by the Markov chain and taking the median of the series of $\mathcal{F}(\mathbf{p})$ values as the best estimate $m_{\tau}\ell$ for the mass of the $\tau$ lepton pair in a given event. The total number of evaluations of the integrand, referred to as Markov chain “states”, amounts to 100 000 per event. The first 10 000 evaluations of the integrand are used as “burn-in” period and are excluded from the computation of the median. The transitions between subsequent states of the Markov chain are computed for the case that $\mathcal{F}(\mathbf{p}) \equiv 1$. This choice has the advantage that the sequence of Markov chain states does not depend on $\mathcal{F}(\mathbf{p})$, which allows for $\mathcal{F}(\mathbf{p})$ to consist of multiple components that can be evaluated by the same Markov chain. Each component may be an arbitrary function of the momenta $\mathbf{p}^{(1)}$ and $\mathbf{p}^{(2)}$ of the two $\tau$ leptons. In the default implementation of the algorithm, the $p_{\tau}, \eta, \phi$ and transverse mass, $m_{\tau\ell} = \left( \vec{p}_{\tau(1)} + \vec{p}_{\tau(2)} \right)^2 - \left( \left( \vec{p}_{\ell(1)} + \vec{p}_{\ell(2)} \right)^2 \right)$, of the $\tau$ lepton pair are reconstructed in addition to the mass $m_{\tau\ell}$.

Table 3

| Sample | $m_{\tau\ell}$ | $\sigma_0/M$ | $\sigma_0/M$ | $s_s$ | $s_s$ |
|--------|----------------|--------------|--------------|-------|-------|
| Z $\rightarrow \tau\tau$ | | | | | |
| 0-jet | 48.1 | 0.196 | 0.274 | -- | 88.2 | 0.200 | 0.228 | -- |
| 1-jet non-boosted | 47.0 | 0.216 | 0.265 | -- | 89.5 | 0.170 | 0.192 | -- |
| 1-jet boosted | 38.2 | 0.304 | 0.456 | -- | 90.0 | 0.116 | 0.123 | -- |
| 2-jet VBF | 41.0 | 0.174 | 0.349 | -- | 89.2 | 0.167 | 0.096 | -- |
| SM $gg \rightarrow H, H \rightarrow \tau\tau$ | | | | | |
| 0-jet | 55.6 | 0.240 | 0.324 | 0.003 | 108.4 | 0.244 | 0.263 | 0.003 |
| 1-jet non-boosted | 55.2 | 0.250 | 0.314 | 0.006 | 117.3 | 0.209 | 0.191 | 0.010 |
| 1-jet boosted | 49.7 | 0.303 | 0.429 | 0.013 | 123.3 | 0.100 | 0.099 | 0.157 |
| SM $q\bar{q} \rightarrow H, H \rightarrow \tau\tau$ | | | | | |
| 2-jet VBF | 51.2 | 0.291 | 0.383 | 0.056 | 122.1 | 0.142 | 0.116 | 0.463 |

The cSVfit algorithm covers two use cases: data analyses use it either because of its capability to reconstruct kinematic
observables of the $\tau$ lepton pair other than the mass $m_{\tau}$, or because of its significantly reduced requirement on computing resources compared to the SVfitMEM algorithm. The $p_T$ of the $H$ boson was used for the purpose of categorizing events in the SM $H \rightarrow \tau\tau$ analysis performed by the CMS collaboration during LHC Run 1 \cite{13}. The $p_T$ was reconstructed by computing the vectorial sum of the momenta $p_{vis}$ of the visible $\tau$ decay products and of the missing transverse momentum $p_T^{miss}$. We will demonstrate in Section 4 that reconstructing the $p_T$ of the $H$ boson candidate by the cSVfit algorithm improves the resolution compared to taking the sum of the $p_{vis}$ and $p_T^{miss}$ vectors. The transverse mass $m_{\tau\tau}$ of the $H$ boson candidate has been used as observable to discriminate signal from background in the CMS search for heavy $H$ bosons in the first LHC Run 2 data \cite{61}, performed in the context of the minimal supersymmetric extension of the Standard Model (MSSM) \cite{62,63}.

Compared to the cSVfit algorithm, the version of the SVfit algorithm used for analyses of data recorded by the CMS experiment during LHC Run 1 used an incomplete expression for the product of the differential phase space element and for the ME modelling the $\tau$ lepton decays. In particular the factor $z_{1,2}$ in the functions $f_h$ and $f_\ell$, given by Eqs. (52) and (68), were missing by mistake. The effect of the
missing factors $\frac{1}{z^2}$ is equivalent to adding an artificial regularization term of the type described in Section 2.7 with $\kappa = 4$ to the likelihood function. This can be seen in the limit that the angles $\theta_{\text{inv}}$ between the $p_{\text{inv}}$ and $p_{\text{vis}}$ vectors are zero for both $\tau$ leptons: In this limit, $m_{\text{HH}} = \frac{m_{\text{vis}}}{\sqrt{\frac{m_{\text{vis}}^2}{2(2z^2)}}}$, with $m_{\text{vis}}$ denoting the measured mass of the visible $\tau$ decay products and $\log(p\cdot\text{GeV}) + 4\log\left(\frac{m_{\text{HH}}}{\sqrt{\frac{m_{\text{vis}}^2}{2(2z^2)}}}\right) \approx \log(p\cdot\text{GeV}) + \log(m_{\text{vis}}^4\cdot\text{GeV}^{-4}) - \log\left(\frac{m_{\text{HH}}^2}{m_{\text{vis}}^2} \cdot \frac{2}{1+\sqrt{2z^2}}\right)$. The term $\log(m_{\text{vis}}^4\cdot\text{GeV}^{-4})$ has been omitted from the sum in the last step. As the term $\log(m_{\text{vis}}^4\cdot\text{GeV}^{-4})$ does not depend on $m_{\text{HH}}$, it has no effect on the reconstruction of $m_{\tau\tau}$.

4. Performance

The performance of the $m_{\tau\tau}$ reconstruction is studied in simulated events. Samples of SM $H \to \tau\tau$ signal events are generated with the next-to-leading-order (NLO) program POWHEG v2 [64-66] for a H boson of mass $m_H = 125$ GeV and for the gluon fusion ($gg \to H$) and vector boson fusion ($q\bar{q} \to H$) production processes. We also study the $m_{\tau\tau}$ reconstruction in events containing heavy pseudoscalar Higgs bosons $A$ of mass $m_A = 200, 300, 500, 800, 1200, 1800$, and $2600$ GeV, produced via gluon fusion, and in events containing heavy spin 1 resonances, of mass $2500$ GeV, that decay into $\tau$ pairs. We denote the latter by the symbol $Z$. The $A \to \tau\tau$ and $Z \to \tau\tau$ signal samples are generated with the LO.
generator PYTHIA 8.2 [67]. The $Z/\gamma^* \rightarrow \tau\tau$ background sample is generated with the LO MadGraph program, in the version MadGraph_aMCatNLO 2.2.2 [68]. The sample contains $\tau$ lepton pairs of true mass $m_{\tau\tau}^{\text{true}} > 50$ GeV. Most of the $\tau$ lepton pairs have a mass near the $Z$ peak at $m_Z = 91.2$ GeV [43], but the sample also contains events of significantly higher mass. Drell-Yan events of mass $m_{\tau\tau}^{\text{true}} < 50$ GeV are not relevant for this study, because they do not pass the selection criteria on $p_T$ and $\eta$ that are applied on analysis level. All events are generated for proton-proton collisions at $\sqrt{s} = 13$ TeV centre-of-mass energy. The samples produced by MadGraph and POWHEG are generated with the NNPDF3.0 set of parton distribution functions, while the samples produced by PYTHIA use the NNPDF2.3LO set [69–71]. Parton shower and hadronization processes are modelled using the generator PYTHIA with the tune CUETP8M1 [72]. The latter is based on the Monash tune [73]. The decays of $\tau$ leptons, including polarization effects, are modelled by PYTHIA.

The samples are normalized according to cross section for the purpose of comparing different mass reconstruction algorithms in terms of signal-to-background separation. The cross section for the irreducible $Z/\gamma^* \rightarrow \tau\tau$ background is computed at NNLO accuracy and amounts to $1.92 \times 10^3$ pb [74]. The cross sections for the SM $H \rightarrow \tau\tau$ signal have been computed as detailed in Ref. [75], with the updates described in Ref. [76] included. The product of cross section times the branching fraction for the decay into a $\tau$ lepton pair amounts to 2.77 pb for SM H bosons produced via gluon fusion and to $2.37 \times 10^{-4}$ pb for SM H bosons produced via vector boson

![Fig. 14. Distributions in $m_{\tau\tau}$ reconstructed by the CA method and different versions of the SVfit algorithm in simulated $A \rightarrow \tau\tau$ signal events of different mass $m_A$: 200 GeV (a), 300 GeV (b), 500 GeV (c), 800 GeV (d), 1200 GeV (e), and 2600 GeV (f), as well as in $Z \rightarrow \tau\tau$ events of mass $m_Z = 2500$ GeV (g). The events are selected in the $\mu_{\tau\tau}$ decay channel. The axis of abscissae ranges from 0.15 to 8.](image-url)
The $\tau \rightarrow A$ and $\tau \rightarrow Z$ samples are scaled to a product of cross section times branching fraction, for the decay into a pair of $\tau$ leptons, of 1 pb. The events are studied on generator level. The $p_T$, $\eta$, and $\phi$ of the electrons and muons produced in the $\tau$ lepton decays are assumed to be reconstructed perfectly. The $\tau_h$ are built by adding the visible decay products of a given $\tau$ lepton, excluding the neutrinos. The experimental resolution on the $p_T$ of the $\tau_h$ is simulated by sampling from the TF described in Section 2.3.1, while the $\eta$, $\phi$, and $m_{\text{vis}}$ of the $\tau_h$ are assumed to be reconstructed perfectly. Jets are reconstructed by the anti-$k_T$ algorithm [77] with a distance parameter $R=0.4$, using all stable visible generator level particles as input. The components $p_T^{\text{rec}}$ and $p_T^{\text{rec}}$ of the hadronic recoil are computed according to Eq. (21): 

\[ p_x^{\text{rec}} = - (p_x^{\text{vis1}} + p_x^{\text{vis2}} + p_x^{\text{miss}}) \]

and 

\[ p_y^{\text{rec}} = - (p_y^{\text{vis1}} + p_y^{\text{vis2}} + p_y^{\text{miss}}) \]

where $p_x^{\text{miss}}$ and $p_y^{\text{miss}}$ correspond to, respectively, the $x$ and $y$ components of the sum of the momenta of all neutrinos produced in the $\tau$ lepton decays. The resolution on the hadronic recoil is simulated by sampling from the TF described in Section 2.3.3.

Distributions in $m_{\tau}$ are computed separately for events in which both $\tau$ leptons decay hadronically ($\tau_h\tau_h$), events in which one $\tau$ lepton decays hadronically and the other into a muon ($\mu\tau_h$), and events in which one $\tau$ lepton decays into a muon and the...

---

**Fig. 15.** Distributions in $m_{\tau_h}$ reconstructed by the CA method and different versions of the SVfit algorithm in simulated $A \rightarrow \tau\tau$ signal events of different mass $m_A$: 200 GeV (a), 300 GeV (b), 500 GeV (c), 800 GeV (d), 1200 GeV (e), and 2600 GeV (f), as well as in $Z \rightarrow \tau\tau$ events of mass $m_Z = 2500$ GeV (g). The events are selected in the $\mu$ decay channel. The axis of abscissae ranges from 0.15 to 8.
Table 4

| Sample | $m_{\text{vis}}$ | SVfitSA | $\mu_{\text{H}}$ |
|--------|-----------------|---------|-----------------|
| $A \rightarrow \tau\tau$ | $\mu_{H} = 200$ GeV | 148 | 0.172 | 0.137 | 0.111 |
| $m_{H} = 300$ GeV | 202 | 0.215 | 0.176 | 0.118 |
| $m_{Z} = 500$ GeV | 208 | 0.271 | 0.205 | 0.128 |
| $m_{\text{H}} = 800$ GeV | 487 | 0.302 | 0.205 | 0.118 |
| $m_{\text{H}} = 1200$ GeV | 724 | 0.318 | 0.206 | 0.116 |
| $m_{\text{H}} = 1800$ GeV | 1078 | 0.329 | 0.196 | 0.110 |
| $m_{\text{H}} = 2600$ GeV | 1553 | 0.339 | 0.198 | 0.103 |

| $Z \rightarrow \tau\tau$ | $m_{Z} = 2500$ GeV | 9486 | 0.340 | 0.245 | 0.147 |

| Sample | $m_{\text{vis}}$ | SVfitMEM, $\kappa = 0$ | SVfitMEM, $\kappa = 5$ | $\mu_{\text{H}}$ |
|--------|-----------------|-------------------|-------------------|-----------------|
| $A \rightarrow \tau\tau$ | $\mu_{H} = 200$ GeV | 207 | 0.116 | 0.140 | 0.113 |
| $m_{H} = 300$ GeV | 303 | 0.135 | 0.178 | 0.118 |
| $m_{Z} = 500$ GeV | 497 | 0.140 | 0.202 | 0.127 |
| $m_{\text{H}} = 800$ GeV | 788 | 0.138 | 0.205 | 0.116 |
| $m_{\text{H}} = 1200$ GeV | 1179 | 0.141 | 0.201 | 0.113 |
| $m_{\text{H}} = 1800$ GeV | 1769 | 0.134 | 0.193 | 0.107 |
| $m_{\text{H}} = 2600$ GeV | 2560 | 0.139 | 0.192 | 0.102 |

| $Z \rightarrow \tau\tau$ | $m_{Z} = 2500$ GeV | 2445 | 0.182 | 0.242 | 0.140 |

| Sample | $m_{\text{vis}}$ | SVfitMEM, $\kappa = 0$ | SVfitMEM, $\kappa = 5$ | $\mu_{\text{H}}$ |
|--------|-----------------|-------------------|-------------------|-----------------|
| $A \rightarrow \tau\tau$ | $\mu_{H} = 200$ GeV | 203 | 0.118 | 0.139 | 0.112 |
| $m_{H} = 300$ GeV | 298 | 0.146 | 0.179 | 0.124 |
| $m_{Z} = 500$ GeV | 495 | 0.168 | 0.231 | 0.132 |
| $m_{\text{H}} = 800$ GeV | 791 | 0.164 | 0.207 | 0.123 |
| $m_{\text{H}} = 1200$ GeV | 1166 | 0.151 | 0.208 | 0.118 |
| $m_{\text{H}} = 1800$ GeV | 1684 | 0.152 | 0.193 | 0.112 |
| $m_{\text{H}} = 2600$ GeV | 2399 | 0.183 | 0.205 | 0.117 |

| $Z \rightarrow \tau\tau$ | $m_{Z} = 2500$ GeV | 2222 | 0.220 | 0.243 | 0.152 |

Another into an electron ($e_{\nu}$). The visible $\tau$ decay products are required to pass selection criteria on $p_{T}$ and $E_{T}^{vis}$, which are motivated by the SM $H \rightarrow \tau\tau$ analysis performed by the CMS collaboration during LHC Run 1 [13]. Events in the $\tau_{\ell}\tau_{\ell}$ decay channel are required to contain two $E_{T}^{vis}$ of $p_{T} > 45$ GeV and $|\eta| < 2.1$. Events in the $\mu_{\ell}\tau_{\ell}$ channel are required to contain a muon of $p_{T} > 20$ GeV and $|\eta| < 2.1$ plus a $E_{T}^{vis}$ of $p_{T} > 30$ GeV and $|\eta| < 2.3$. Events selected in the $e_{\nu}\tau_{\ell}$ channel are required to contain a muon with $|\eta| < 2.1$ and an electron with $|\eta| < 2.4$. The lepton of higher $p_{T}$ (either the electron or the muon) is required to satisfy the condition $p_{T} > 20$ GeV, while the lepton of lower $p_{T}$ is required to satisfy $p_{T} > 10$ GeV. Similar selection criteria on $p_{T}$ and $E_{T}^{vis}$ of the visible $\tau$ decay products were applied in the $H \rightarrow \tau\tau$ analyses performed by the ATLAS collaboration during LHC Run 1 [78,79].

The $m_{\ell}\tau$ reconstruction in SM $H \rightarrow \tau\tau$ signal events is studied in event categories motivated by the SM $H \rightarrow \tau\tau$ analysis performed by the CMS collaboration during LHC Run 1 [13]:

- 0-Jet: Events containing no jets of $p_{T} > 30$ GeV and $|\eta| < 4.5$.
- 1-Jet non-booster: Events containing one or more jets of $p_{T} > 30$ GeV and $|\eta| < 4.5$ in which the $H$ (respectively $Z$) boson satisfies $p_{T} < 100$ GeV. Events in the 1-Jet non-booster category are required not to be selected in the 2-Jet VBF category.
- 1-Jet boosted: Events containing one or more jets of $p_{T} > 30$ GeV and $|\eta| < 4.5$ in which the $H$ (respectively $Z$) boson satisfies $p_{T} > 100$ GeV. Events in the 1-Jet boosted category are required not to be selected in the 2-Jet VBF category.
The table below presents the median $M$ and resolutions $σ_M/M$ and $σ_μ/M$ of the distributions in $m_{vis}$ and in $m_{rec}$, reconstructed by different versions of SVfit algorithm, in simulated signal events containing either heavy pseudoscalar Higgs bosons $A$ or heavy spin 1 resonances $Z'$ and in simulated $Zl^+l^− → ττ$ background events. The events are selected in the $ττ$ decay channel. The difference between the two resolutions $σ_μ/M$ and $σ_μ/M$ is explained in the text.

| $m_{vis}$ | $σ_M/M$ | $σ_μ/M$ | $m_{rec}$ | $σ_M/M$ | $σ_μ/M$ |
|----------|--------|--------|----------|--------|--------|
| $A → ττ$: | | | | | |
| $m_A = 200$ GeV | 77 | 0.318 | 0.426 | 181 | 0.303 | 0.187 |
| $m_A = 300$ GeV | 104 | 0.372 | 0.514 | 272 | 0.355 | 0.176 |
| $m_A = 500$ GeV | 162 | 0.427 | 0.564 | 464 | 0.351 | 0.136 |
| $m_A = 800$ GeV | 246 | 0.463 | 0.620 | 744 | 0.348 | 0.133 |
| $m_A = 1200$ GeV | 357 | 0.487 | 0.661 | 1310 | 0.353 | 0.116 |
| $m_A = 1800$ GeV | 538 | 0.517 | 0.659 | 1727 | 0.320 | 0.094 |
| $m_A = 2500$ GeV | 748 | 0.528 | 0.716 | 2501 | 0.312 | 0.086 |

| Sample | $m_{vis}$ | $σ_M/M$ | $σ_μ/M$ | $m_{rec}$ | $σ_M/M$ | $σ_μ/M$ |
|--------|--------|--------|--------|----------|--------|--------|
| $A → ττ$: | | | | | |
| $m_κ = 200$ GeV | 223 | 0.162 | 0.463 | 186 | 0.272 | 0.191 |
| $m_κ = 300$ GeV | 321 | 0.155 | 0.444 | 279 | 0.325 | 0.167 |
| $m_κ = 500$ GeV | 518 | 0.136 | 0.369 | 469 | 0.322 | 0.133 |
| $m_κ = 800$ GeV | 815 | 0.130 | 0.310 | 751 | 0.319 | 0.128 |
| $m_κ = 1200$ GeV | 1214 | 0.126 | 0.263 | 1141 | 0.317 | 0.111 |
| $m_κ = 1800$ GeV | 1812 | 0.115 | 0.202 | 1732 | 0.267 | 0.097 |
| $m_κ = 2600$ GeV | 2607 | 0.119 | 0.172 | 2518 | 0.268 | 0.089 |

- 2-jet VBF: Events containing two or more jets of $p_T > 30$ GeV and $|η| < 4.5$, with at least one pair of jets satisfying $m_{jj} > 500$ GeV and $Δφ_{jj} > 3.5$.

The event categorization is based on generator level quantities. We do not expect that migrations between the event categories due to resolution effects significantly affect the conclusions of the $m_{rec}$ resolution studies.

The $m_{rec}$ distributions reconstructed by the SVfitMEM and cSVfit algorithms in these event categories in the $ττ$, $μτ$, and $eτ$ decay channels are shown in Figs. 8–10. They are compared to the distributions in $m_{vis}$, reconstructed by the "collinear-approximation" (CA) method [18] and by the previous version of the SVfit algorithm, described in Ref. [23], to which we refer to as the SVfit "standalone" (SVfitSA) algorithm. The SVfitMEM and cSVfit algorithms are utilized with and without the artificial regularization term described in Section 2.7. The value of $κ$ used for each channel is chosen such that the optimal resolution on $m_{rec}$, quantified in terms of the ratio of the root-mean-square (RMS) to the median $M$ of the distribution, is attained. We find that the choice of $κ = 5$ for the $ν_τ$ channel, $κ = 4$ for the $μτ$ channel, and $κ = 3$ for the $eτ$ channel performs well for the SVfitMEM as well as for the cSVfit algorithm.

The distributions in $m_{rec}$ reconstructed by the cSVfit and SVfitMEM algorithms with artificial regularization term and by the SVfitSA algorithm are very similar. The SVfitSA algorithm performs well without adding an artificial regularization term of the type described in Section 2.7 to its likelihood function. This is because the effect of the missing factor $Δϕ$ in the likelihood function used by the SVfitSA algorithm (cf. Section 3) is equivalent to using an artificial regularization term with $κ = 4$. Regardless of the choice of $κ$, the peaks of the $m_{rec}$ distributions reconstructed by the different versions of the SVfit algorithm are close to the true value of the mass of the $τ$ lepton pair in the 1-jet boosted and 2-jet VBF categories. In the 0-jet and 1-jet non-boosted categories, the distributions in $m_{rec}$ reconstructed by the cSVfit and SVfitMEM algorithms with $κ = 0$ exhibit a tendency to overestimate the mass of the $τ$ lepton pair, resulting in pronounced high mass tails. The choice of a small positive $κ$ reduces this tendency, makes the distributions in $m_{rec}$ peak close to the true mass of the $τ$ lepton pair and significantly reduce the high mass tails.

The effect of adding the artificial regularization term is largest for events reconstructed in the 0-jet category, which are the most difficult to reconstruct. This is because in events in which the $H$ or $Z$ boson has low $p_T$, the $τ$ leptons are typically “back-to-back” in the transverse plane ($\Delta φ ≈ π$), with the effect that the neutrinos produced in the $τ$ lepton decays are emitted in opposite hemispheres and their contribution to $p_T^{vis}$ cancels. The cancellation of neutrino momenta causes mass hypotheses of low $m_{rec}^{vis} ≃ m_{vis}$ and of high $m_{rec}^{vis} >> m_{vis}$ to be degenerate in terms of the probability density $p$, computed according to Eq. (19). The best estimate for the mass of the $τ$ lepton pair in a given event may fluctuate due to mismeasurements, within the experimental resolution, of the components $p_κ^{rec}$ and $p_κ^{rec}$ of the hadronic recoil or, to a lesser extent, of the $p_T$ of $ν_τ$, degrading the resolution on $m_{rec}$.

The resolution on $m_{rec}$ is significantly higher in events selected in the 1-jet and 2-jet VBF categories, in which the $τ$ lepton pair typically recoils against high $p_T$ jets. As the mass of the H respectively $Z$ boson increases, the angle $\Delta φ$ between the $τ$ leptons decreases, due to the Lorentz boost in direction of the $H$ or $Z$ boson. The momenta of the neutrinos produced in the $τ$ lepton decays add constructively in this case, with the effect that the mass of the $τ$ lepton pair is constrained by the measured value of $p_T^{vis}$. The correlation between the $p_T$ of the $τ$ lepton pair and the angle $Δφ$ is visualized in Fig. 11 for SM $H → ττ$ Signal and for $Zl^+l^− → ττ$ background events. The events are selected in the $ν_τ$ decay channel. The distributions for events selected in the $ν_τ$ and $eτ$ decay channels are similar.

The variation of the resolution on $m_{vis}$ across event categories is most pronounced in the $eτ$ and least pronounced in the $ν_τ$ decay channel. This is because the fraction of $τ$ lepton energy carried by the visible $τ$ decay products is typically high for hadronic $τ$ decays and typically low for leptonic $τ$ decays, cf. Fig. 2. The consequence is that for events in the $ν_τ$ decay channel, regardless of the angle $Δφ$ between the $τ$ leptons, the best estimate for the mass of the $τ$ lepton pair is typically not much higher than $m_{vis}$, as the energies of the neutrinos produced in hadronic $τ$ decays are known to be most likely small. For events in the $eτ$ decay channel on the other
hand, lepton decays with high energetic neutrinos are known to be likely and, provided they are compatible with the measured value of $p_T^\text{mis}$, mass hypotheses of low $\approx()$ $\mu$ $\approx()$ $\tau$ $\approx()$ $\nu$, corresponding to the case $\approx()$ $\mu$ $\approx()$ $\tau$, and of high $\approx()$ $\mu$ $\approx()$ $\tau$, corresponding to the case $\approx()$ $\mu$ $\approx()$ $\tau$, are often degenerate in terms of the probability density $\mathcal{P}$.

Numerical values for the resolution in $m_\tau$ achieved by the different algorithms are given in Tables 1–3. The mass of the visible $\tau$ decay products, $m_\text{vis}$, is given in the tables for comparison. The resolution is quantified separately for the low mass and for the high mass tail of the mass distribution reconstructed by a given algorithm. The symbol $\sigma/M$ (the symbol $\sigma/M$) denotes the ratio of the difference between the median and the 16 quantile (between the 84 quantile and the median) relative to the median of the mass distribution. The capability of each algorithm to separate the SM $\tau \to H$ signal from the irreducible $\gamma \tau \to \gamma \tau$ background is additionally quantified in terms of $\mathcal{S}/(\mathcal{S} + \mathcal{B})$, the ratio of the number of SM $H \to \tau \tau$ signal events to the sum of signal plus $\gamma \tau \to \gamma \tau$ background events. The yields $\mathcal{S}$ of signal and $\mathcal{B}$ of background events are computed within a mass window containing 68% of signal events, with 16% of signal events on either side of the mass window, and for SM signal and background cross sections.

The choice of quantifying the resolution in terms of the ratios $\sigma/M$ and $\sigma_0/M$ is motivated by the fact that the distributions in $m_\tau$, $m_\text{vis}$, and $m_\mu$ are...
and in $m_{\text{vis}}$ may be shifted with respect to $m_{\tau\tau}^{\text{true}}$. While the distributions in $m_{\tau\tau}$ reconstructed by the different versions of the SVfit algorithm, as well as by the CA method, peak close to $m_{\tau\tau}^{\text{true}}$, the distributions in $m_{\text{vis}}$ exhibit significant shifts towards lower mass. The shift of the $m_{\text{vis}}$ distribution is in general highest for events in the $\mu\tau_h$ decay channel and lowest for events in the $\tau\tau_{hh}$ decay channel, reflecting the fact that the visible $\tau$ decay products typically carry a lower fraction of $\tau$ lepton energy in leptonic compared to hadronic $\tau$ decays. The shift is more pronounced in events selected in the 1-jet and 2-jet VBF categories and less pronounced in the 0-jet category. In particular in the $\tau\tau_{hh}$ channel, the $p_T$ cuts that are applied on the $p_T$ remove most of the $Z/\gamma^* \rightarrow \tau\tau$ events in the 0-jet category, except for a few events with large $m_{\text{vis}}$. The events selected in the 1-jet and 2-jet VBF categories typically have smaller $m_{\text{vis}}$, as $m_{\text{vis}}$ decreases proportional to the cosine of the angle between the $\tau$ leptons: $m_{\text{vis}} \approx p_{T,1}^{\text{vis}} \cos \theta_{1,2}^{\text{vis}} / p_T^{\text{vis}} \cos \theta_{1,2}^{\text{vis}}$. All algorithms can be trivially calibrated such that the median of each mass distribution coincides with the true mass of the $\tau$ lepton pair, by scaling the output of the algorithm by a suitably chosen constant. The advantage of quantifying the resolution in terms of the ratios $\sigma/M$ is that these ratios are invariant under such scaling.

The ratio $S/(S+B)$ of the number of SM $H \rightarrow \tau\tau$ signal to the sum of signal plus $Z/\gamma^* \rightarrow \tau\tau$ background events in general increases with jet multiplicity and with the $p_T$ of the $\tau$ lepton pair ($p_T^{\tau_1}$ respectively $p_T^{\tau_2}$). The categories with the highest signal-to-background ratio, the 1-jet boosted and the 2-jet VBF category, in fact provide most of the sensitivity of the SM $H \rightarrow \tau\tau$ analysis. The improvement in mass resolution provided by the SVfit algorithm is...
Table 7
CPU time, in seconds per event, needed to reconstruct $m_\tau$, by the SVFitMEM, cSVFit, and SVFitSA algorithms in simulated $Z/\gamma^* \to \tau\tau$ background and SM $H \to \tau\tau$ signal events in the decay channels $\tau_1\tau_2$, $p_{\tau\tau}$, $\rho_{\tau\tau}$, and $\phi_{\tau\tau}$.

| Algorithm | $Z/\gamma^* \to \tau\tau$ | $H \to \tau\tau$ | $m_\tau$ decay channel | $Z/\gamma^* \to \tau\tau$ | $H \to \tau\tau$ | $\mu_{\tau\tau}$ decay channel | $Z/\gamma^* \to \tau\tau$ | $H \to \tau\tau$ |
|-----------|-----------------------------|-----------------------------|-------------------------|-----------------------------|-----------------------------|-------------------------|-----------------------------|-----------------------------|
|           | Mean [s] | RMS [s] | Mean [s] | RMS [s] | Mean [s] | RMS [s] | Mean [s] | RMS [s] |
| SVFitSA   | 3.51    | 5.24    | 3.75    | 5.87    |   |   |   |   |
| cSVFit    | $\kappa = 0$ | 0.72 | 0.17 | 0.73 | 0.18 |   |   |   |   |
| $\kappa = 5$ | 0.77 | 0.20 | 0.78 | 0.20 |   |   |   |   |
| SVFitMEM  | $\kappa = 0$ | 13.07 | 6.27 | 14.03 | 6.89 |   |   |   |   |
| $\kappa = 5$ | 10.72 | 4.72 | 11.44 | 5.02 |   |   |   |   |
| $\mu_{\tau\tau}$ decay channel | $Z/\gamma^* \to \tau\tau$ | $H \to \tau\tau$ | $m_\tau$ decay channel | $Z/\gamma^* \to \tau\tau$ | $H \to \tau\tau$ | $\phi_{\tau\tau}$ decay channel | $Z/\gamma^* \to \tau\tau$ | $H \to \tau\tau$ |
|           | Mean [s] | RMS [s] | Mean [s] | RMS [s] | Mean [s] | RMS [s] | Mean [s] | RMS [s] |
| SVFitSA   | 5.82    | 9.22    | 5.02    | 7.91    |   |   |   |   |
| cSVFit    | $\kappa = 0$ | 0.62 | 0.11 | 0.56 | 0.12 |   |   |   |   |
| $\kappa = 4$ | 0.73 | 0.17 | 0.61 | 0.16 |   |   |   |   |
| SVFitMEM  | $\kappa = 0$ | 22.64 | 12.99 | 18.45 | 10.15 |   |   |   |   |
| $\kappa = 4$ | 18.38 | 7.49 | 15.83 | 6.68 |   |   |   |   |

largest in these categories, enhancing the separation of the SM $H \to \tau\tau$ signal from the $Z/\gamma^* \to \tau\tau$ background where it matters most. In the SM $H \to \tau\tau$ analysis performed by the CMS collaboration during LHC Run 1, the use of $m_\tau$, reconstructed by the SVFit algorithm, has increased the expected significance for observing a signal by ≈40% compared to $m_{\text{vis}}$ [13], corresponding to a gain of a factor two in luminosity.

The reconstruction of $m_\tau$ by the CA method performs sub-optimal compared to SVFit. In particular the pronounced high mass tails in the $m_\tau$ distribution, arising from resolution effects, significantly reduce the sensitivity for observing a signal, as they cause a sizeable fraction of $Z/\gamma^* \to \tau\tau$ background events to be reconstructed near the signal region $m_\tau \approx 125$ GeV. A further disadvantage of the CA method is that it fails to yield a physical solution for approximately half of the events, whereas the SVFit algorithm provides a physical solution for every event.

The performance of the cSVFit algorithm to reconstruct $p_{\tau\tau}$, $\eta_\tau$, and $\phi_\tau$ of the $\tau$ lepton pair is studied in SM $H \to \tau\tau$ signal events produced via the gluon fusion process, separately for the decay channels $\eta_\tau$, $p_{\tau\tau}$, and $\phi_\tau$. Distributions of the difference between reconstructed and true $p_\tau$, $\eta_\tau$, and $\phi_\tau$ of the H boson are shown in Fig. 12. Compared to the case that the H boson $p_\tau$ and $\phi_\tau$ are reconstructed by taking the vectorial sum of the momenta of the visible decay products and of $p_{\tau\tau}^{\text{true}}$, the cSVFit algorithm improves the resolution by 10–20%. The pseudo-rapidity $\eta_\tau$ of the H boson can only be reconstructed with the cSVFit algorithm. Typical resolutions on the $p_\tau$, $\eta_\tau$, and $\phi_\tau$ of the H boson, reconstructed by the cSVFit algorithm, amount to 10 GeV, 0.4 rad, and 0.8 rad, respectively.

Distributions in $m_\tau$, reconstructed by different versions of the SVFit algorithm and by the CA method, as well as in $m_{\text{vis}}$, in events containing hypothetical heavy pseudoscalar Higgs bosons and heavy spin 1 resonances are shown in Figs. 13–15. The distributions in $m_\tau$, reconstructed by the cSVFit and SVFitMEM algorithms are shown for the case that the artificial regularization term described in Section 2.7 is used and for the case that it is not used.

The SVFit algorithm significantly improves the separation of the heavy $A$ and $Z$ boson signals from the irreducible $Z/\gamma^* \to \tau\tau$ background in all three decay channels. Numerical values for the median $M$ as well as for the resolutions $s/M$ and $t/M$ are given in Tables 4–6.

We interpret the fact that the performance of the SVFitMEM algorithm is similar for an $A \to \tau\tau$ signal of mass 2600 GeV and for a $Z \to \tau\tau$ signal of mass 2500 GeV as evidence that the usage of a LO ME for the gluon fusion process $gg \to H$ in the SVFitMEM algorithm represents no limitation for using the SVFitMEM algorithm in data analyses of $\tau$ lepton pair production other than studies of Higgs boson production.

The improvement in signal-to-background separation provided by the SVFit algorithm in searches for high mass resonances decaying to $\tau$ lepton pairs is illustrated in Fig. 16.

We conclude the discussion of the performance of the SVFit algorithm with a comparison of the computing time requirements of the SVFitMEM, cSVFit, and SVFitSA algorithms. The computing time is dominated by the evaluation of the integrand during the numeric integration of Eqs. (19) and (43), and of Eq. (2) of Ref. [23], respectively. The computing time requirement is expected to be highest for the SVFitMEM algorithm, due to the time needed for evaluation of the PDF and of the TF in Eq. (19).

The time for reconstructing $m_\tau$ in a single event scales approximately linearly with the number of evaluations of the integrand. The SVFitMEM (SVFitSA) algorithm performs 20 000 (10 000) evaluations of the integrand per mass hypothesis $m_{\text{true}}^{\text{scan}}$. The number of evaluations of the integrand per mass hypothesis $m_{\text{true}}^{\text{scan}}$ is chosen such that the resolution on $m_\tau$ is within 1% compared to the resolution obtained in case an infinite (very large) number of evaluations of the integrand is performed. The number of evaluations is higher in case of the SVFitMEM algorithm, accounting for the fact that the additional PDF and TF factors raise the variation of the integrand, such that a higher number of evaluations is necessary in order to reach a given precision on the value of the integral.

The number of mass hypotheses tested for an event varies depending on the conditions described in Section 2.6. For the series of $m_{\text{true}}^{\text{scan}}$ values defined by Eq. (40), the number of mass hypotheses for which the integral $\mathcal{P}$ gets computed according to Eq. (19) increases approximately logarithmically with the true mass of the $\tau$ lepton pair. In signal events, the integral $\mathcal{P}$ typically gets computed for a series of 20–30 mass hypotheses. The average number of mass hypotheses varies with the decay channel. The series is typically the longest in the $\eta_\tau$ and the shortest in the $p_{\tau\tau}$ channel, as the value of $m_{\text{true}}$ that is used to initialize the series is farther away respectively less far away from $m_{\text{true}}$ at which the value of $\mathcal{P}$ is expected to reach its maximum $\mathcal{P}_{\text{true}}$. A larger number of mass hypotheses may be required for background events, in particular for backgrounds containing neutrinos that do
not originate from \( \tau \) lepton decays, such as \( W\) boson production and the production of top quark pairs, as the values of \( \mathcal{P} \) may be almost degenerate over a large range in \( m_{\ell\ell}^{\text{miss}} \).

The numeric integration is performed by the VAMP algorithm in case of the SVfitMEM algorithm (cf. Section 2.6) and by the VEGAS algorithm in case of the SVfitSA algorithm. In case of the cSVfit algorithm, the numeric integration is performed by a custom implementation of the Markov chain Monte Carlo method, as described in Section 3. In the latter case, the integrand is evaluated 100,000 times per event and there is no iteration over a sequence of mass hypotheses, cf. Eq. (43).

The computing time is measured by CPU time, using a single core of a 2.30 GHz Intel\textsuperscript{®} Xeon\textsuperscript{®} E5-2695 v3 processor. The distribution of the CPU time, in units of seconds per event, spent by the SVfitMEM, cSVfit and SVfitSA algorithms to compute \( m_\ell \) in SM \( H \to \tau\tau \) signal events and in \( Z/\gamma^* \to \tau\tau \) background events is shown in Fig. 17. Numerical values of the mean and RMS of the distributions are given in Table 7. The cSVfit algorithm requires about 0.5 s of CPU time per event to reconstruct the mass \( m_\ell \) of the \( \tau \) lepton pair, as well as its \( p_\ell, \eta_T, \phi \), and transverse mass \( m_{\tau\tau} \). The CPU time is reduced by a factor 5–30 compared to the SVfitSA algorithm, owing to the more efficient Markov chain Monte Carlo integration method compared to computing the probability density \( \mathcal{P} \) for a series of mass hypotheses. The computing time requirement of the SVfitMEM algorithm is higher by a factor 3–4 compared to the SVfitSA algorithm.

5. Discussion

Considering that the resolution on \( m_{\tau\tau} \), quantified by \( \sigma/M \) and \( \sigma_\eta/M \), achieved by the SVfitMEM and cSVfit algorithms is almost identical, we find that the cSVfit algorithm represents the best compromise between physics performance and computing time requirements in practical applications of the SVfit algorithm.

The optimal resolution is achieved in case an artificial regularization term of the type described in Section 2.7, with small positive \( \kappa \), is added to the likelihood function. We expect the optimal choice of \( \kappa \) to depend on the experimental resolution as well as on the rates of signal and background processes, and we recommend to perform a reoptimization of \( \kappa \) in practical applications of our algorithm.

The merit of the SVfitMEM algorithm is that the formalism to treat \( \tau \) lepton decays in the ME method, developed for the SVfitMEM algorithm, can be used in future applications of the ME method to data analyses with \( \tau \) leptons in the final state. An example for such an application is the analysis of SM H boson production in association with a pair of top quarks (tH) in final states with a \( \tau \) lepton [80], in which the existence of neutrinos from \( W \to \ell\nu \) decays preclude the reconstruction of the H boson mass by the SVfit algorithm.

6. Summary

An algorithm for reconstruction of the H boson mass in events in which the H boson decays into a pair of \( \tau \) leptons has been presented. The relative resolution on the H boson mass amounts to typically 15–20%. The algorithm has been used in data analyses performed by the CMS collaboration during LHC Run 1. It improves the sensitivity of the SM \( H \to \tau\tau \) analysis by \( \approx 40\% \), corresponding to a gain in luminosity by about a factor of two.

An improved version of the algorithm has been developed in preparation for LHC Run 2. Two variants of the improved algorithm have been implemented. The first variant, SVfitMEM, has been rigorously developed within the formalism of the ME method. It is based on proper normalization of the probability density \( \mathcal{P} \), given by Eq. (19). The second variant of the improved algorithm uses a likelihood function of arbitrary normalization. It allows to compute, on an event-by-event basis, any kinematic function of the two \( \tau \) leptons and provides a substantial reduction in computing time requirements. A further improvement concerns the modelling of the experimental resolution on the \( p_\ell \) of \( \eta_T \) via TF, described in Section 2.3.1.

The performance of the algorithm has been studied in simulated SM \( H \to \tau\tau \) signal and \( Z/\gamma^* \to \tau\tau \) background events, as well as in simulated samples of heavy pseudoscalar Higgs bosons and heavy spin 1 resonances. The SVfit algorithm is found to perform well in all event categories and over the full range in true mass of the \( \tau \) lepton pair that we studied.

The development of the formalism to handle \( \tau \) lepton decays in the ME method constitutes an important result of this paper. The formalism allows one to extend the matrix elements generated by automatized tools such as CompHEP or MadGraph by the capability to handle hadronic as well as leptonic \( \tau \) decays. We expect that the formalism will be very useful for future applications of the ME method to data analyses with \( \tau \) leptons in the final state.

Acknowledgements

JC and CV wish to acknowledge the collaboration with our former colleague Evan Friis, who made vital contributions to the early stage of this project. LM acknowledges the Estonian Research Council for supporting his work with the grant PUTJD110.
Appendix A. Appendix

We derive here the relations for the product of the squared moduli of the ME and the phase space elements \( d\Omega^{\nu}_{\nu, \nu} \) and \( d\Omega^{\nu}_{\nu, \nu} \) for, respectively, the decays \( \tau \rightarrow \) hadrons + \( \nu \) and \( \tau \rightarrow \ell^- \nu \), given by Eq. (33). We start with the simpler case of hadronic \( \tau \) decays in Section A.1 and turn to the more complex case of leptonic \( \tau \) decays in Section A.2. For clarity of notation, we omit the hat symbol in this section and use the convention that all symbols refer to true values in the laboratory frame, unless indicated explicitly otherwise.

A.1. The decay \( \tau \rightarrow \) hadrons + \( \nu \)

We treat hadronic \( \tau \) decays as a two-body decay into a hadronic system \( \tau_h \) plus a \( \nu \), as explained in Section 2.1, and take the squared modulus of the ME to be a constant, which we denote by \( |\mathcal{M}_{\tau_h \nu}^{\mathrm{eff}}|^2 \). We further denote the momentum of the neutrino produced in the \( \tau \) decay by \( p^{\nu} \) and its energy by \( E^{\nu} \) (cf. Section 2.4). For reasons that will become clear later, we allow the neutrino to have non-zero mass \( m^{\nu} \).

The product of the squared modulus of the ME and of the phase space element \( d\Omega^{\nu}_{\nu, \nu} \) reads:

\[
\frac{|\mathcal{B}|^2 |\mathcal{M}_{\tau \nu}^{\mathrm{eff}}(p)|^2}{(2\pi)^3 2E_{\nu, \nu}} \frac{d^3p^{\nu}}{(2\pi)^3 2E_{\nu, \nu}} = \frac{|\mathcal{B}|^2 |\mathcal{M}_{\tau \nu}^{\mathrm{eff}}(p)|^2}{(2\pi)^3 2E_{\nu, \nu}} \frac{d^3p^{\nu}}{(2\pi)^3 2E_{\nu, \nu}} \frac{d^3p^{\nu}}{(2\pi)^3 2E_{\nu, \nu}} \frac{d^3q}{(2\pi)^3 2E_{\nu, \nu}}.
\]

\[
\frac{E_{\nu, \nu} dE_{\nu, \nu} d\cos\theta_{\nu, \nu} d\phi_{\nu, \nu}}{2E_{\nu, \nu}}.
\]

where we have used the formula for recursive phase space generation, given by Eq. (46.12) in Ref. [43], for transforming the first line into the second and the identity:

\[
d^3p^{\nu} = |p^{\nu}| d|p^{\nu}| d\cos\theta_{\nu} d\phi_{\nu} = \frac{1}{2E_{\nu, \nu}} \frac{dE_{\nu, \nu} d\cos\theta_{\nu} d\phi_{\nu}}{2E_{\nu, \nu}}.
\]

for rewriting the third line by the fourth. The factor \( \frac{|\mathcal{B}|^2}{2E_{\nu, \nu}} \delta(q^2 - m^2) \) removes the integration over \( dq^2 \), enforcing the \( \tau \) lepton energy and momentum to be related by \( E_{\nu, \nu} = \sqrt{|p^{\nu}|^2 + m^2} \). The symbol \( E_{\nu, \nu}(p^{\nu}, p^{\nu}) \) indicates that \( E_{\nu, \nu} \) is a function of \( p^{\nu} \) and \( p^{\nu} \), as is necessary to satisfy the \( \delta \)-function \( \delta(|p^{\nu}|^2 + m^2) \).

We define \( z = E_{\nu, \nu}/E_{\nu} \) according to Eq. (14) and replace the integration over \( dE_{\nu, \nu} \) by an integration over \( z \). The Jacobi factor related to this transformation is:

\[
E_{\nu, \nu} = E_{\nu} - E_{\nu} = (1 - z)E_{\nu} = \frac{1}{z}E_{\nu} \
\Rightarrow \quad dE_{\nu, \nu} = \left| \frac{dE_{\nu}}{dz} \right| dz = \frac{E_{\nu}}{z} dE_{\nu}.
\]

We then perform the integration over \( d\cos\theta_{\nu} \). Following the convention that we introduced in Section 2.4, we choose the coordinate system such that \( \cos\theta_{\nu} \) is equal to the angle between the \( p^{\nu} \) and \( p^{\nu} \) vectors. The \( \delta \)-function \( \delta(E_{\nu}(p^{\nu}, p^{\nu}) - E_{\nu} - E_{\nu}) \) depends on \( \cos\theta_{\nu} \) via:

\[
E_{\nu}(p^{\nu}, p^{\nu}) = \sqrt{|p^{\nu}|^2 + m^2} = \sqrt{(p^{\nu} + p^{\nu})^2 + m^2} = \sqrt{|p^{\nu}|^2 + |p^{\nu}|^2 + 2 p^{\nu} p^{\nu} + m^2} = \sqrt{|p^{\nu}|^2 + |p^{\nu}|^2 + 2 |p^{\nu}| |p^{\nu}| \cos\theta + m^2}.
\]

\[
The \delta \)-function argument vanishes if \( E_{\nu}(p^{\nu}, p^{\nu}) - E_{\nu} - E_{\nu} = 0 \). This yields:

\[
\cos\theta_{\nu} = \frac{E_{\nu} E_{\nu} - \frac{1}{2}(m^2 - (m^2 + m^2))}{|p^{\nu}|^2 |p^{\nu}|}.
\]

When substituting the expressions of Eqs. (46) and (48) into Eq. (44), we need to account for the \( \delta \)-function rule:
\[ \delta(g(x)) = \sum_k \frac{\delta(x - x_k)}{|g'(x_k)|}. \]  \tag{49} 

We identify:

\[ g(\cos \theta_{\text{inv}}) = \sqrt{\left| p^{\text{vis}}_\nu \right|^2 + \left| p^{\text{vis}}_\mu \right|^2 + 2 \left| p^{\text{vis}}_\nu \right| \left| p^{\text{vis}}_\mu \right| \cos \theta_{\text{inv}} + m^2 - E_{\text{vis}} - E_{\text{inv}} \]  \tag{50} 

and obtain \( g'(x_0) = \left| p^{\text{vis}}_\nu \right| \left| p^{\text{vis}}_\mu \right| \left( E_{\text{vis}} + E_{\text{inv}} \right) = \left| p^{\text{vis}}_\nu \right| \left| p^{\text{vis}}_\mu \right| E_r. \)

This yields:

\[ \boxed{\text{BW}^{(i)}_{\tau \nu} \left| M_{\tau \nu} \right|^2 d\phi_{\text{inv}} = \frac{M_{\text{eff}}^{(i)}}{256\pi^4 m^2 \Gamma r} \frac{E_{\text{vis}}}{\left| p^{\text{vis}}_\nu \right| \left| p^{\text{vis}}_\mu \right|^2} d^3 p^{\text{vis}} dz d\phi_{\text{inv}}.} \]  \tag{51} 

We define:

\[ f_h(p^{\text{vis}}, m_{\text{vis}}, p^{\text{inv}}) = \frac{M_{\text{eff}}^{(i)}}{256\pi^4 m^2 \Gamma r} \frac{E_{\text{vis}}}{\left| p^{\text{vis}}_\nu \right| \left| p^{\text{vis}}_\mu \right|^2} \]  \tag{52} 

and obtain:

\[ \boxed{\text{BW}^{(i)}_{\tau \nu} \left| M_{\tau \nu} \right|^2 d\phi_{\text{inv}} = \frac{M_{\text{eff}}^{(i)}}{256\pi^4 m^2 \Gamma r} \frac{E_{\text{vis}}}{\left| p^{\text{vis}}_\nu \right| \left| p^{\text{vis}}_\mu \right|} f_h(p^{\text{vis}}, m_{\text{vis}}, p^{\text{inv}}) d^3 p^{\text{vis}} dz d\phi_{\text{inv}},} \]  \tag{53} 

which is the result that we quote in Eq. (33).

A.2. The decays \( \tau \to e\nu\bar{\nu}_e \) and \( \tau \to \mu\nu\bar{\nu}_\mu \)

We treat leptonic \( \tau \) decays as three-body decays and do account for the ME. Assuming the taus to be unpolarized, the squared modulus of the ME is given by [44]:

\[ \left| M_{\tau \nu} \right|^2 = 64 G_F^2 \left( E_{\nu} - p^\mu - p^\nu \right) \left( E_{\tau} - p^\mu - p^\nu \right), \]  \tag{54} 

where \( G_F \) denotes the Fermi constant, given by Eq. (6).

The product of the squared modulus of the ME and the phase space element \( d\phi_{\nu\bar{\nu}_e} \) reads:

\[ \boxed{\text{BW}^{(i)}_{\tau \nu} \left| M_{\tau \nu} \right|^2 d\phi_{\nu\bar{\nu}_e} = \text{BW}^{(i)}_{\tau \nu} \left| M_{\tau \nu} \right|^2 \frac{d^3 p^{\text{vis}}}{(2\pi)^4 2E_{\text{vis}}} \frac{d^3 p^{\text{vis}}}{(2\pi)^4 2E_{\text{vis}}} \frac{d^3 p^{\text{vis}}}{(2\pi)^4 2E_{\text{vis}}} \frac{d^3 p^{\text{vis}}}{(2\pi)^4 2E_{\text{vis}}} d^2 q^2,} \]  \tag{55} 

The equality \( \left| M_{\tau \nu} \right|^2 = \left| M_{\tau \nu} \right|^2 \) follows from the fact that the matrix element is a Lorentz invariant quantity and as such has the same value in any frame. We have used the identity:

\[ \int dE d^3 p \delta(E^2 - p^2 - m^2) = \int \frac{d^3 p}{2E}. \]  \tag{56} 

for expressing the second line by the third. Eq. (56) follows from the \( \delta \)-function rule Eq. (49). We assume that the mass of the \( \nu \) as well as the mass of the \( \bar{\nu}_e \) is zero.

We perform a variable transformation from \( (E_\nu, p^\nu) \) and \( (E_\mu, p^\mu) \) to:

\[ (u_0, u) = \frac{1}{\sqrt{2}} (E_\nu + E_\mu, p^\nu + p^\mu), \quad (v_0, v) = \frac{1}{\sqrt{2}} (E_\nu - E_\mu, p^\nu - p^\mu). \]

The variables \( u_0 \) and \( u \) represent the energy and momentum of the neutrino pair. The magnitude of the determinant of the Jacobi matrix for the transformation from \( (E_\nu, p^\nu; E_\mu, p^\mu) \) to \( (u_0, u; v_0, v) \) equals unity. Expressed in the new variables, the energy and momentum of the \( \nu \) and of the \( \bar{\nu}_e \) produced in the tau decay are given by:
\[ (E, \mathbf{p}) = \frac{1}{\sqrt{2}}(u_0 + v_0, \mathbf{u} + \mathbf{v}) \quad \text{and} \quad (E', \mathbf{p}') = \frac{1}{\sqrt{2}}(u_0 - v_0, \mathbf{u} - \mathbf{v}). \]

The product of the squared modules of the ME and the phase space element can then be expressed by:

\[
|\mathbf{B}|^2 \left| M_{1 \to f} \right|^2 \frac{d^4 p}{(2\pi)^3 2E_1} \delta(E_1 - E_{\text{vis}}) \delta(E_1 - E_0) \delta(p' - p^{\text{vis}} - p - p')
\]

\[
\frac{d^4 p}{(2\pi)^3 2E_1} \frac{d^4 p^{\text{vis}}}{(2\pi)^3 2E_{\text{vis}}} \frac{d\mathbf{u}_1}{(2\pi)^3} \frac{d\mathbf{u}}{(2\pi)^3} \frac{d\mathbf{u}_0}{(2\pi)^3} \frac{d\mathbf{v}}{(2\pi)^3}
\]

\[
M_{1 \to f} \frac{d\mathbf{u}_1}{(2\pi)^3} \frac{d\mathbf{u}}{(2\pi)^3} \frac{d\mathbf{u}_0}{(2\pi)^3} \frac{d\mathbf{v}}{(2\pi)^3}
\]

\[
\theta(u_0 - v_0) \delta \left( \frac{u_0^2 - |\mathbf{u}|^2 + v_0^2 - |\mathbf{v}|^2}{2} + u_0 v_0 - \mathbf{u} \cdot \mathbf{v} \right).
\]

We define:

\[
l_{\text{inv}} = \left| M_{1 \to f} \right|^2 \frac{d\mathbf{u}_1}{(2\pi)^3} \frac{d\mathbf{u}}{(2\pi)^3} \frac{d\mathbf{u}_0}{(2\pi)^3} \frac{d\mathbf{v}}{(2\pi)^3}
\]

\[
\theta(u_0 - v_0) \delta \left( \frac{u_0^2 - |\mathbf{u}|^2 + v_0^2 - |\mathbf{v}|^2}{2} + u_0 v_0 - \mathbf{u} \cdot \mathbf{v} \right).
\]

The quantity \(l_{\text{inv}}\) is a Lorentz invariant quantity. As such, it can be computed in any frame and will yield the same value as in the laboratory frame. We choose to evaluate it in the rest frame of the neutrino pair. In this frame, the energy is given by \(u_0 = m_{\text{inv}}\) and the momentum by \(\mathbf{u} = (0, 0, 0)\), with \(m_{\text{inv}}\) denoting the mass of the neutrino pair. Hence \(u_0 \mathbf{v}_0 - \mathbf{u} \mathbf{v} = m_{\text{inv}} \mathbf{v}_0\) in this frame. Performing the integration over \(v_0\), we obtain:

\[
l_{\text{inv}} = \left| M_{1 \to f} \right|^2 \frac{1}{2} \frac{d\mathbf{v}}{(2\pi)^3} \theta(u_0 + v_0) \delta \left( \frac{u_0^2 - |\mathbf{u}|^2 + v_0^2 - |\mathbf{v}|^2}{2} + u_0 v_0 - \mathbf{u} \cdot \mathbf{v} \right)
\]

\[
= \frac{1}{2m_{\text{inv}}} \theta(u_0) \left| M_{1 \to f} \right|^2 \frac{d\mathbf{v}}{(2\pi)^3} \delta \left( \frac{m_{\text{inv}}^2 - |\mathbf{v}|^2}{2} \right)
\]

\[
= \frac{1}{2m_{\text{inv}}} \theta(u_0) \int \frac{d^3 \mathbf{v}}{(2\pi)^3} \left| M_{1 \to f} \right|^2 \delta \left( \frac{|\mathbf{v}|^2 - m_{\text{inv}}^2}{2} \right)
\]

\[
= \frac{1}{2m_{\text{inv}}} \theta(u_0) \int \frac{d^3 \mathbf{v}}{(2\pi)^3} \left| M_{1 \to f} \right|^2.
\]

We have used the relation \(\delta(a + b) \delta(a - b) = \delta(2b) \delta(a - b) = \frac{1}{2} \delta(b) \delta(a)\) to express Eq. (58) by the first line in Eq. (59). The identify \(\frac{v_0^2 - |\mathbf{v}|^2}{2} = -\frac{1}{2} |\mathbf{v}|^2\) follows from the presence of the \(\delta(v_0)\) in the first line.

In the rest frame of the neutrino pair:

\[
(E, \mathbf{p}) = (E_0, 0, 0, |\mathbf{p}|)
\]

\[
(E_{\text{vis}}, \mathbf{p}^{\text{vis}}) = (E_{\text{vis}}, 0, 0, |\mathbf{p}^{\text{vis}}|)
\]

\[
(E, \mathbf{p}) = \frac{m_{\text{inv}}}{2} (1, 0, \sin \theta, \cos \theta)
\]

\[
(E, \mathbf{p}) = \frac{m_{\text{inv}}}{2} (1, 0, -\sin \theta, -\cos \theta),
\]

where we have chosen the polar axis such that it is parallel to \(\mathbf{p}^{\text{vis}}\).

The ME given by Eq. (54) evaluates to:

\[
\left| M_{1 \to f} \right|^2 = 64 G^2 F \left( E, E_{\text{vis}} - \mathbf{p}^{\text{vis}} \cdot \mathbf{p} \right) \delta \left( E_{\text{vis}} - E - \mathbf{p}^{\text{vis}} \cdot \mathbf{p} \right)
\]

\[
= 64 G^2 \frac{m_{\text{inv}}^2}{2} (E + |\mathbf{p}^{\text{vis}}| \cos \theta \left( E_{\text{vis}} - |\mathbf{p}^{\text{vis}}| \cos \theta \right)
\]

\[
= 16 G^2 \frac{m_{\text{inv}}^2}{2} (E + |\mathbf{p}| \cos \theta \left( E_{\text{vis}} - |\mathbf{p}| \cos \theta \right)
\]

(60)
in the rest frame of the neutrino pair.

The energy of the electron or muon is related by:

\[ m_{\text{vis}}^2 = E_{\text{vis}}^2 - \left| \mathbf{p}_{\text{vis}} \right|^2 = (E_\tau - u_0)^2 - (\mathbf{p} - \mathbf{u})^2 \]

\[ = m_\tau^2 + m_{\text{inv}}^2 - 2(E_\tau u_0 - \mathbf{p} \cdot \mathbf{u}) = m_\tau^2 + m_{\text{inv}}^2 - 2m_{\text{inv}}E_\tau , \]

from which it follows that:

\[ E_\tau = \frac{m_\tau^2 + m_{\text{inv}}^2 - m_{\text{vis}}^2}{2m_{\text{inv}}} \quad \text{and} \quad E_{\text{vis}} = E_\tau - m_{\text{inv}} = \frac{m_\tau^2 - m_{\text{inv}}^2 - m_{\text{vis}}^2}{2m_{\text{inv}}} . \]

Substituting Eq. (60) into Eq. (59) yields:

\[ I_{\text{inv}} = \frac{1}{2} \theta(u_0) \int \frac{d\Omega_{\ell}}{(2\pi)^3} \left| \mathcal{M}_{\ell \rightarrow \nu \ell \nu} \right|^2 \]

\[ = 8 G^2 F_\pi^2 m_{\text{inv}}^2 \theta(u_0) \int \frac{d\cos \theta}{2\pi} \left( E_\tau, E_{\text{vis}} - \mathbf{p} \cdot \mathbf{u} \right) \]

\[ \left( \frac{E_\tau}{2} E_{\text{vis}} - \frac{1}{2} \sqrt{E_\tau^2 - m_{\text{vis}}^2} \right) \]

\[ = \frac{2 G^2}{\pi} m_{\text{inv}}^2 \theta(u_0) \left( 2 E_\tau E_{\text{vis}} - \frac{1}{2} \sqrt{E_\tau^2 - m_{\text{vis}}^2} \right) , \]

with \( E_\tau \) and \( E_{\text{vis}} \) given by Eq. (62). Note that \( I_{\text{inv}} \) solely depends on a single kinematic variable, \( m_{\text{inv}} \), as \( m_\tau \) and \( m_{\text{vis}} \) are constants.

Before substituting Eq. (63) into Eq. (57), we perform a variable transformation from \( (u_0, u_0, u_1, u_1) \) to \( (m_{\text{inv}}^2, u_0, u_0, u_1) = (2(u_0^2 - u_0^2 - u_0^2), u_0, u_0, u_1) \). The magnitude of the determinant of the Jacobi matrix for this transformation is \( |J| = 4u_0 \), from which it follows that:

\[ du_0 d^2u = \frac{1}{|J|} dm_{\text{inv}}^2 d^2u \;
\]

Substituting Eqs. (63) and (64) into Eq. (57), we then obtain:

\[ |\mathcal{B}|^2 \left| \mathcal{M}_{\tau \rightarrow \nu \tau \nu} \right|^2 \left( \phi_1 \phi_2 \right) = \frac{8 \pi^3}{m_\tau^2} \delta(E_\tau - E_{\text{vis}} - u_0) \delta \left( \mathbf{p} - \mathbf{p}_{\text{vis}} - \mathbf{u} \right) \]

\[ \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{d^3\mathbf{p}_{\text{vis}}}{(2\pi)^3} \frac{d^2u}{(2\pi)^3} \frac{I_{\text{inv}}}{2u_0} \frac{dm_{\text{inv}}^2}{2} \]

\[ = \frac{I_{\text{inv}}}{512 \pi^2 m_\tau^2 \left| \mathbf{p}_{\text{vis}} \right|} \frac{E_{\text{vis}}}{2E_{\text{vis}}} \frac{d^3\mathbf{p}_{\text{vis}}}{dz} \frac{dm_{\text{inv}}^2}{d\phi_{\text{inv}}} . \]

where \( \phi_{\text{inv}} \) specifies the orientation of the neutrino pair momentum vector \( \mathbf{p}_{\text{inv}} \) with respect to the momentum vector \( \mathbf{p}_{\text{vis}} \) of the electron or muon.

The opening angle between the vector \( \mathbf{p}_{\text{inv}} \) and the direction of the electron respectively muon is given in analogy to Eq. (48) by:

\[ \cos \theta_{\text{inv}} = \frac{E_{\text{vis}} - \frac{1}{2} \left( m_\tau^2 - (m_{\text{vis}}^2 + m_{\text{inv}}^2) \right)}{\left| \mathbf{p}_{\text{vis}} \right| \left| \mathbf{p}_{\text{inv}} \right|} . \]

We define:
to obtain:

$$\frac{\mathcal{B}(\tilde{W}_R^0 \tilde{W}_L^0)}{\mathcal{B}(\tau^+ \tau^-)} d\phi_{\tilde{W}_R^0} = \frac{m_{\tilde{W}_R^0}}{m_{\tilde{W}_L^0}} \frac{d\phi_{W_R^0}}{2 \pi} \frac{d\phi_{\tau^+}}{2 \pi} = \mathcal{D}_{W_R^0} d\phi_{\tilde{W}_R^0} + d\phi_{\tau^+},$$

(69)

the result that we quote in Eq. (33).
[74] K. Melnikov, F. Petriello, Electroweak gauge boson production at hadron colliders through OXIALPS*2, Phys. Rev. D 74 (2006) 114017 (doi:10.1103/PhysRevD.74.114017).
[75] S. Dittmaier, et al., Handbook of LHC Higgs cross-sections: 1. inclusive observables arXiv:1101.0593 http://dx.doi.org/10.1017/CERN-2011-002.
[76] J.R. Andersen, et al., Handbook of LHC Higgs cross-sections: 3. Higgs properties arXiv:1307.1347 http://dx.doi.org/10.3170/CERN-2013-004.
[77] M. Cacciari, G.P. Salam, G. Soyez, The anti-kT jet clustering algorithm, JHEP 04 (2008) 063 (http://arXiv:0802.1189).
[78] G. Aad, et al., Evidence for the Higgs boson Yukawa coupling to leptons with the ATLAS detector, JHEP 04 (2015) 117, http://dx.doi.org/10.1007/JHEP04(2015)117 (http://arXiv:1501.04943).
[79] G. Aad, et al., Search for neutral Higgs bosons of the minimal supersymmetric standard model in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector, JHEP 11 (2014) 056, http://dx.doi.org/10.1007/JHEP11(2014)056.
[80] CMS Collaboration, Search for the associated production of a Higgs boson with a top quark pair in final states with a $\tau$ lepton at $\sqrt{s} = 13$ TeV, CMS Physics Analysis Summary CMS-PAS-HIG-17-003 (2017) (http://cds.cern.ch/record/2257067).