

# Application of the matrix element method to the analysis of $HH \rightarrow b\bar{b}W^+W^-$ in the dilepton decay channel

Karl Ehatäht<sup>1</sup> and Christian Veelken<sup>1</sup>

<sup>1</sup>National Institute for Chemical Physics and Biophysics, 10143 Tallinn, Estonia

December 7, 2018

## Abstract

We present the development of a discriminant for the analysis of Higgs boson pair (HH) production in the decay channel  $HH \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell^+\nu\ell^-\bar{\nu}$ , which separates the HH signal from the large irreducible background that arises from top quark pair production in this channel. The discriminant is based on matrix element method.

# 1 The Matrix element method

The matrix element method (MEM) computes probability densities  $w_i(\mathbf{y}|\hat{\mathbf{y}})$  for obtaining the measured observables  $\mathbf{y}$ , assuming the event has been produced by the process  $i$ . The probability densities  $w_i(\mathbf{y}|\hat{\mathbf{y}})$  are interpreted as quantifying the compatibility of the measured observables  $\mathbf{y}$  with the signal ( $i = 0$ ) and background ( $i = 1$ ) hypothesis. In the analysis of Higgs boson pair (HH) production in the decay channel  $\text{HH} \rightarrow \text{b}\bar{\text{b}} \text{W}^+ \text{W}^- \rightarrow \text{b}\bar{\text{b}} \ell^+ \nu \ell^- \bar{\nu}$  the observables  $\mathbf{y}$  refer to the measured momenta of the two b-jets, the two charged leptons, and the measured missing transverse momentum ( $\mathbf{p}_T^{\text{miss}}$ ) in the event. We use symbols with a hat to denote the true values of energies and momenta. Bold letters denote vector quantities. The vector  $\hat{\mathbf{y}}$  denotes the true values of the b-jet and charged lepton momenta and the true values of the momenta of the two neutrinos produced in the W boson decays. Note that  $\text{HH} \rightarrow \text{b}\bar{\text{b}} \text{W}^+ \text{W}^- \rightarrow \text{b}\bar{\text{b}} \ell^+ \nu \ell^- \bar{\nu}$  signal events contain the same number of b-jets, charged leptons, and neutrinos as the dominant background, arising from top quark pair production ( $\text{t}\bar{\text{t}} \rightarrow \text{b}\text{W}^+ \bar{\text{b}}\text{W}^- \rightarrow \text{b}\ell^+ \nu \bar{\text{b}}\ell^- \bar{\nu}$ ), in this channel. The separation of the HH signal from the irreducible  $\text{t}\bar{\text{t}}$  background is based on the difference in event kinematics, causing the probability densities  $w_0(\mathbf{y}|\hat{\mathbf{y}})$  ( $w_1(\mathbf{y}|\hat{\mathbf{y}})$ ) to be in general higher when evaluated on signal (background) events and lower when evaluated on background (signal) events. Given the probability densities  $w_0(\mathbf{y}|\hat{\mathbf{y}})$  and  $w_1(\mathbf{y}|\hat{\mathbf{y}})$  for the signal and background hypotheses, the Neyman lemma [1] postulates that the ratio:

$$P(\mathbf{y}) = \frac{w_0(\mathbf{y}|\hat{\mathbf{y}})}{w_0(\mathbf{y}|\hat{\mathbf{y}}) + w_1(\mathbf{y}|\hat{\mathbf{y}})}$$

then provides the optimal separation of the HH signal from the irreducible  $\text{t}\bar{\text{t}}$  background.

Different nomenclatures and conventions for the MEM exist in the literature. In this note, we follow the nomenclature and conventions introduced in Ref. [2]. The probability densities  $w_i(\mathbf{y}|\hat{\mathbf{y}})$  are given by the integral:

$$w_i(\mathbf{y}|\hat{\mathbf{y}}) = \frac{\Omega(\mathbf{y})}{\sigma \mathcal{A}} \int d\hat{\mathbf{x}}_a d\hat{\mathbf{x}}_b d\Phi_n \frac{f(\hat{\mathbf{x}}_a) f(\hat{\mathbf{x}}_b)}{2 \hat{\mathbf{x}}_a \hat{\mathbf{x}}_b s} (2\pi)^4 \delta \left( \hat{\mathbf{x}}_a \hat{E}_a + \hat{\mathbf{x}}_b \hat{E}_b - \sum_i^n \hat{E}_{(i)} \right) \cdot \delta^3(\hat{\mathbf{x}}_a \hat{\mathbf{p}}^a + \hat{\mathbf{x}}_b \hat{\mathbf{p}}^b - \sum_i^n \hat{\mathbf{p}}^{(i)}) |\mathcal{M}_i(\hat{\mathbf{y}})|^2 W(\mathbf{y}|\hat{\mathbf{y}}) \epsilon(\hat{\mathbf{y}}). \quad (1)$$

The symbol  $|\mathcal{M}_i(\hat{\mathbf{y}})|^2$  denotes the squared modulus of the matrix element (ME) for either the signal ( $i = 0$ ) or for the background ( $i = 1$ ) hypothesis. The symbols  $\hat{E}_a$  and  $\hat{E}_b$  ( $\hat{\mathbf{p}}^a$  and  $\hat{\mathbf{p}}^b$ ) denote the energies (momenta) of the two colliding protons,  $\sqrt{s}$  their center-of-mass energy,  $\hat{\mathbf{x}}_a$  and  $\hat{\mathbf{x}}_b$  the Bjorken scaling variables [3], and  $f(\hat{\mathbf{x}}_a)$  and  $f(\hat{\mathbf{x}}_b)$  the corresponding parton distribution functions (PDF). We use the MSTW 2008 LO PDF set [4] to evaluate  $f(\hat{\mathbf{x}}_a)$  and  $f(\hat{\mathbf{x}}_b)$ . We denote by  $n$  the number of particles in the final state, by  $\mathbf{p}^{(i)}$  ( $\hat{\mathbf{p}}^{(i)}$ ) the measured (true) momentum of the  $i$ -th final state particle, and by  $d\Phi_n = \prod_i^n \frac{d^3 \hat{\mathbf{p}}^{(i)}}{(2\pi)^3 2\hat{E}_{(i)}}$  the differential  $n$ -particle phase space (PS) element. For the HH signal as well as for the  $\text{t}\bar{\text{t}}$  background,  $n = 6$ , as the final state in both cases contains a pair of b and  $\bar{\text{b}}$  quarks, two charged leptons, and two neutrinos. The  $\delta$ -functions  $\delta(\hat{\mathbf{x}}_a \hat{E}_a + \hat{\mathbf{x}}_b \hat{E}_b - \sum_i^n \hat{E}_{(i)})$  and  $\delta^3(\hat{\mathbf{x}}_a \hat{\mathbf{p}}^a + \hat{\mathbf{x}}_b \hat{\mathbf{p}}^b - \sum_i^n \hat{\mathbf{p}}^{(i)})$  impose conservation of energy and momentum. The functions  $W(\mathbf{y}|\hat{\mathbf{y}})$  are referred to as “transfer functions” (TF) in the literature. They represent the probability density to observe the measured values  $\mathbf{y}$ , given the true values  $\hat{\mathbf{y}}$ . The vector  $\hat{\mathbf{y}}$  corresponds to a point  $\hat{\mathbf{p}}$  in the  $n$ -particle PS. The function  $\Omega(\mathbf{y})$  is referred to as “indicator function” in the literature [5, 2]. It attains the value 1 in case the event represented by the measured observables  $\mathbf{y}$  passes the event selection criteria and otherwise attains the value 0. The efficiency for an event originating at the PS point  $\hat{\mathbf{y}}$  to pass the event selection, i.e.

to end up with measured observables  $\mathbf{y}$  for which  $\Omega(\mathbf{y}) = 1$ , is denoted by  $\epsilon(\mathbf{y})$ . Finally, the symbol  $\mathcal{A}$  denotes the acceptance of the event selection, that is, the percentage of events which pass the event selection criteria, while  $\sigma$  denotes the total (inclusive) cross section. Division of the right-hand-side (RHS) of Eq. (1) by the product  $\sigma \mathcal{A}$  ensures that  $w_i(\mathbf{y}|\hat{\mathbf{y}})$  has the correct normalization required for a probability density, i.e.  $\int d\mathbf{y} w_i(\mathbf{y}|\hat{\mathbf{y}}) = 1$ , provided that the TF satisfy the normalization condition  $\int d\mathbf{y} \Omega(\mathbf{y}) W(\mathbf{y}|\hat{\mathbf{y}}) = 1$  for every  $\mathbf{y}$ .

We express the PS element  $d\Phi_6$  in terms of the energies  $\hat{E}_{(i)}$ , the polar angles  $\hat{\theta}_{(i)}$ , and the azimuthal angles  $\hat{\phi}_{(i)}$  of the pair of  $b$  and  $\bar{b}$  quarks, of the two charged leptons, and of the two neutrinos:

$$\begin{aligned}
d\Phi_6 &= \prod_i^6 \frac{d^3 \hat{\mathbf{p}}^{(i)}}{(2\pi)^3 2\hat{E}_{(i)}} = \frac{1}{2^{24} \pi^{18}} \prod_i^6 \frac{d^3 \hat{\mathbf{p}}^{(i)}}{\hat{E}_{(i)}} \\
&= \frac{1}{2^{24} \pi^{18}} \prod_i^6 \frac{d\hat{E}_{(i)} d\hat{\theta}_{(i)} d\hat{\phi}_{(i)} |\hat{\mathbf{p}}^{(i)}| \hat{E}_{(i)} \sin \hat{\theta}_{(i)}}{\hat{E}_{(i)}} \\
&= \frac{1}{2^{24} \pi^{18}} \prod_i^6 d\hat{E}_{(i)} d\hat{\theta}_{(i)} d\hat{\phi}_{(i)} \hat{\beta}_{(i)} \hat{E}_{(i)} \sin \hat{\theta}_{(i)}. \tag{2}
\end{aligned}$$

All energies  $\hat{E}_{(i)}$  as well as the angles  $\hat{\theta}_{(i)}$  and  $\hat{\phi}_{(i)}$  refer to the laboratory (detector) frame. The velocity  $\hat{\beta}_{(i)}$  of particle  $i$ , given by  $\hat{\beta}_{(i)} \equiv \frac{|\hat{\mathbf{p}}^{(i)}|}{\hat{E}_{(i)}}$ , has been used to simplify the expression for  $d\Phi_6$  in the last step. Note that the velocity  $\hat{\beta}_{(i)}$  is a function of energy  $\hat{E}_{(i)}$  and hence cannot be treated as constant when evaluating the integral over  $d\hat{E}_{(i)}$ . Similarly, the magnitude of the momentum  $|\hat{\mathbf{p}}^{(i)}|$  is a function of the energy  $\hat{E}_{(i)}$ . In the following, we use the identities  $|\hat{\mathbf{p}}^{(i)}| = \sqrt{\hat{E}_{(i)}^2 - m_{(i)}^2}$  and  $\hat{\beta}_{(i)} = \frac{\sqrt{\hat{E}_{(i)}^2 - m_{(i)}^2}}{\hat{E}_{(i)}}$  to make the dependency on the energy  $\hat{E}_{(i)}$  explicit.

The form of Eq. (2) is useful, as it allows to trivially perform the integration over the angles  $\hat{\theta}_{(i)}$  and  $\hat{\phi}_{(i)}$  for the pair of  $b$  and  $\bar{b}$  quarks and for the two charged leptons, taking advantage of the fact that the directions of quarks (jets) and charged leptons are measured with negligible experimental resolution. With the further assumption that also the energy of charged leptons can be measured with negligible experimental resolution, the integration over  $d\hat{E}_{\ell^+}$  and  $d\hat{E}_{\ell^-}$  can be carried out trivially too. We shall only consider events that pass the event selection criteria, i.e. for which the indicator function  $\Omega(\mathbf{y})$  is equal to 1. For simplicity, we moreover neglect the effect of the efficiency  $\epsilon(\hat{\mathbf{y}})$  and of the acceptance  $\mathcal{A}$ . With these assumptions and upon inserting the expressions for the TF given by Eqs. (14), (15), and (16)

in the appendix into Eq. (2), we obtain:

$$\begin{aligned}
w_i(\mathbf{y}|\hat{\mathbf{y}}) &= \frac{1}{2^{24} \pi^{18} \sigma E_{\ell^+} E_{\ell^-}} \int d\hat{x}_a d\hat{x}_b \frac{d\hat{E}_b}{\hat{E}_b} \frac{d\hat{E}_{\bar{b}}}{\hat{E}_{\bar{b}}} \frac{d^3\hat{\mathbf{p}}_\nu}{\hat{E}_\nu} \frac{d^3\hat{\mathbf{p}}_{\bar{\nu}}}{\hat{E}_{\bar{\nu}}} \frac{f(\hat{x}_a) f(\hat{x}_b)}{2\hat{x}_a \hat{x}_b s} \\
&\cdot (2\pi)^4 \delta(\hat{x}_a \hat{E}_a + \hat{x}_b \hat{E}_b - \sum_i^6 \hat{E}_{(i)}) \delta^3(\hat{x}_a \hat{\mathbf{p}}^a + \hat{x}_b \hat{\mathbf{p}}^b - \sum_i^6 \hat{\mathbf{p}}^{(i)}) \\
&\cdot |\mathcal{M}_i(\hat{\mathbf{p}})|^2 \frac{\hat{\beta}_b \hat{E}_b^2}{\beta_b E_b^2} W(E_b|\hat{E}_b) \frac{\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}^2}{\beta_{\bar{b}} E_{\bar{b}}^2} W(E_{\bar{b}}|\hat{E}_{\bar{b}}) \\
&= \frac{1}{2^{21} \pi^{14} \sigma E_{\ell^+} E_{\ell^-} E_b E_{\bar{b}}} \int d\hat{x}_a d\hat{x}_b d\hat{E}_b d\hat{E}_{\bar{b}} \frac{d^3\hat{\mathbf{p}}_\nu}{\hat{E}_\nu} \frac{d^3\hat{\mathbf{p}}_{\bar{\nu}}}{\hat{E}_{\bar{\nu}}} \frac{f(\hat{x}_a) f(\hat{x}_b)}{\hat{x}_a \hat{x}_b s} \\
&\cdot \delta(\hat{x}_a \hat{E}_a + \hat{x}_b \hat{E}_b - \sum_i^6 \hat{E}_{(i)}) \delta^3(\hat{x}_a \hat{\mathbf{p}}^a + \hat{x}_b \hat{\mathbf{p}}^b - \sum_i^6 \hat{\mathbf{p}}^{(i)}) \\
&\cdot |\mathcal{M}_i(\hat{\mathbf{p}})|^2 \frac{\hat{\beta}_b \hat{E}_b}{\beta_b E_b} W(E_b|\hat{E}_b) \frac{\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}}{\beta_{\bar{b}} E_{\bar{b}}} W(E_{\bar{b}}|\hat{E}_{\bar{b}}). \tag{3}
\end{aligned}$$

The terms  $\frac{\hat{\beta}_b \hat{E}_b}{\beta_b E_b}$  and  $\frac{\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}}{\beta_{\bar{b}} E_{\bar{b}}}$  arise because the integration over the PS elements  $d^3\hat{\mathbf{p}}$  of the  $b$  and  $\bar{b}$  quarks yields a factor  $\hat{\beta} \hat{E}^2 \sin \hat{\theta}$ , while the normalization of the TF yields a factor  $\frac{1}{\beta E^2 \sin \theta}$ , cf. Eq. (16). The terms  $\sin \hat{\theta}$  and  $\frac{1}{\sin \theta}$  cancel, due to the presence of the  $\delta(\theta - \hat{\theta})$  in the integrand, cf. Eq. (15). No similar terms arise for the charged leptons, as the TF for charged leptons demand  $\hat{\beta} = \beta$ ,  $\hat{E} = E$ , and  $\hat{\theta} = \theta$ , cf. Eq. (14).

Concerning the four-dimensional  $\delta$ -function  $\delta(\hat{x}_a \hat{E}_a + \hat{x}_b \hat{E}_b - \sum_i^6 \hat{E}_{(i)}) \cdot \delta^3(\hat{x}_a \hat{\mathbf{p}}^a + \hat{x}_b \hat{\mathbf{p}}^b - \sum_i^6 \hat{\mathbf{p}}^{(i)})$ , which ensures the conservation of energy and momentum, we first eliminate the energy and longitudinal momentum components by solving for the Bjorken scaling variables  $\hat{x}_a$  and  $\hat{x}_b$ . This yields [6]:

$$\hat{x}_a = \frac{1}{\sqrt{s}} \sum_i^6 \hat{E}_{(i)} + \hat{p}_z^{(i)} \quad \text{and} \quad \hat{x}_b = \frac{1}{\sqrt{s}} \sum_i^6 \hat{E}_{(i)} - \hat{p}_z^{(i)}. \tag{4}$$

For the purpose of eliminating the transverse momentum components of the four-dimensional  $\delta$ -function, we follow the approach of Ref. [6]. The approach is based on introducing the “hadronic recoil”, denoted by the symbol  $\rho$ , as a means to account for QCD radiation, which causes additional jets to be produced besides the two  $b$ -jets originating from the decay of the  $H$  boson (in signal events) or from the decay of the two top quarks (in background events). As detailed in Ref. [7], significant amounts of QCD radiation, in particular initial-state radiation (ISR), are a typical feature of most signal and background processes at the LHC. The longitudinal momentum of the additional jets produced by QCD radiation alters the relations for  $\hat{x}_a$  and  $\hat{x}_b$  somewhat, compared to the values given by Eq. (4). We expect the effect of QCD radiation on the energy and longitudinal momentum components to be small, however, and thus neglect it. The effect on the transverse momentum balance is important, however, as QCD radiation distorts the kinematic relations that would be expected to hold in the absence of such radiation. As a consequence, the  $\delta$ -functions that ensure the conservation of momentum in the transverse plane need to be modified. Their modified form reads:  $\delta(\hat{p}_x^\rho + \sum_i^6 \hat{p}_x^{(i)})$  and  $\delta(\hat{p}_y^\rho + \sum_i^6 \hat{p}_y^{(i)})$ , where  $\hat{p}_x^\rho$  ( $\hat{p}_y^\rho$ ) denotes the true value of the momentum of the hadronic recoil  $\rho$  in  $x$  ( $y$ ) direction. The  $\delta$ -functions hold on the level of measured and on the level of true values. For the measured values, they imply the relations:

$$p_x^\rho = - \left( p_x^b + p_x^{\bar{b}} + p_x^{\ell^+} + p_x^{\ell^-} + p_x^{\text{miss}} \right) \quad \text{and} \quad p_y^\rho = - \left( p_y^b + p_y^{\bar{b}} + p_y^{\ell^+} + p_y^{\ell^-} + p_y^{\text{miss}} \right),$$

while the corresponding relations for the true momenta read:

$$\hat{p}_x^\rho = -\left(\hat{p}_x^b + \hat{p}_x^{\bar{b}} + \hat{p}_x^{\ell^+} + \hat{p}_x^\nu + \hat{p}_x^{\ell^-} + \hat{p}_x^{\bar{\nu}}\right) \quad \text{and} \quad \hat{p}_y^\rho = -\left(\hat{p}_y^b + \hat{p}_y^{\bar{b}} + \hat{p}_y^{\ell^+} + \hat{p}_y^\nu + \hat{p}_y^{\ell^-} + \hat{p}_y^{\bar{\nu}}\right). \quad (5)$$

The experimental resolution on  $p_x^\rho$  and  $p_y^\rho$  is modeled by introducing a TF for the hadronic recoil into the integrand of Eq. (3). Following Ref. [6], we assume that the resolution on the transverse momentum components of  $\rho$  follows a two-dimensional normal distribution:

$$W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho) = \frac{1}{2\pi \sqrt{|V|}} \exp \left( -\frac{1}{2} \begin{pmatrix} p_x^\rho - \hat{p}_x^\rho \\ p_y^\rho - \hat{p}_y^\rho \end{pmatrix}^T \cdot V^{-1} \cdot \begin{pmatrix} p_x^\rho - \hat{p}_x^\rho \\ p_y^\rho - \hat{p}_y^\rho \end{pmatrix} \right). \quad (6)$$

The matrix  $V$  is computed event-by-event, using the  $p_T^{\text{miss}}$ -significance algorithm [8]. Following Ref. [6], we replace the  $\delta$ -functions  $\delta(\hat{p}_x^\rho + \sum_i^6 \hat{p}_x^{(i)})$  and  $\delta(\hat{p}_y^\rho + \sum_i^6 \hat{p}_y^{(i)})$  that ensure the momentum conservation in the transverse plane by the TF for the hadronic recoil. A remaining issue is that we use leading order (LO) ME  $\mathcal{M}_i(\hat{\mathbf{p}})$  for the HH signal and for the  $t\bar{t}$  background in Eq. (3). The LO ME for the signal (background) requires that the HH ( $t\bar{t}$ ) system has zero  $p_T$ , a condition that (in the laboratory frame) only holds in case the hadronic recoil has zero  $p_T$ , which rarely holds at the LHC, due to the abundance of ISR (as previously discussed). The issue that the LO ME is only well-defined for events with zero ISR is resolved by evaluating the ME  $\mathcal{M}_i(\hat{\mathbf{y}})$  in a frame in which the HH ( $t\bar{t}$ ) system has zero  $p_T$ , to which we refer as the zero-transverse-momentum (ZTM) frame. The Lorentz transformation of the energy  $\hat{E}_{(i)}$  and momenta  $\hat{\mathbf{p}}^{(i)}$  in Eq. (3) from the laboratory to the ZTM frame is performed using the vector  $(-\frac{p_x^\rho}{p_T^\rho}, -\frac{p_y^\rho}{p_T^\rho}, 0)$  as the boost vector.

Eliminating the energy and longitudinal momentum components of the (modified) four-dimensional  $\delta$ -function  $\delta(\hat{x}_a \hat{E}_a + \hat{x}_b \hat{E}_b - \sum_i^6 \hat{E}_{(i)}) \cdot \delta(\hat{x}_a \hat{p}_z^a + \hat{x}_b \hat{p}_z^b - \sum_i^6 \hat{p}_z^{(i)}) \cdot \delta(\hat{p}_x^\rho + \sum_i^6 \hat{p}_x^{(i)}) \cdot \delta(\hat{p}_y^\rho + \sum_i^6 \hat{p}_y^{(i)})$  by means of Eq. (4) and replacing its transverse momentum components by the TF  $W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho)$  for the hadronic recoil  $\rho$ , the expression for the probability density  $w_i(\mathbf{y}|\hat{\mathbf{y}})$  becomes:

$$\begin{aligned} w_i(\mathbf{y}|\hat{\mathbf{y}}) &= \frac{1}{2^{21} \pi^{14} \sigma E_{\ell^+} E_{\ell^-} E_b E_{\bar{b}}} \int d\hat{E}_b d\hat{E}_{\bar{b}} \frac{d^3 \hat{\mathbf{p}}_\nu}{\hat{E}_\nu} \frac{d^3 \hat{\mathbf{p}}_{\bar{\nu}}}{\hat{E}_{\bar{\nu}}} \frac{f(\hat{x}_a) f(\hat{x}_b)}{\hat{x}_a \hat{x}_b s} \\ &\quad \cdot |\mathcal{M}_i(\hat{\mathbf{p}})|^2 \frac{\hat{\beta}_b \hat{E}_b}{\beta_b E_b} W(E_b | \hat{E}_b) \frac{\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}}{\beta_{\bar{b}} E_{\bar{b}}} W(E_{\bar{b}} | \hat{E}_{\bar{b}}) W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho) \\ &= \frac{1}{2^{21} \pi^{14} \sigma E_{\ell^+} E_{\ell^-} E_b E_{\bar{b}}} \int d\hat{E}_b d\hat{E}_{\bar{b}} d\hat{E}_\nu d\hat{\theta}_\nu d\hat{\phi}_\nu d\hat{E}_{\bar{\nu}} d\hat{\theta}_{\bar{\nu}} d\hat{\phi}_{\bar{\nu}} \\ &\quad \cdot \hat{\beta}_\nu \hat{E}_\nu \sin \hat{\theta}_\nu \hat{\beta}_{\bar{\nu}} \hat{E}_{\bar{\nu}} \sin \hat{\theta}_{\bar{\nu}} \frac{f(\hat{x}_a) f(\hat{x}_b)}{\hat{x}_a \hat{x}_b s} \\ &\quad \cdot |\mathcal{M}_i(\hat{\mathbf{p}})|^2 \frac{\hat{\beta}_b \hat{E}_b}{\beta_b E_b} W(E_b | \hat{E}_b) \frac{\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}}{\beta_{\bar{b}} E_{\bar{b}}} W(E_{\bar{b}} | \hat{E}_{\bar{b}}) W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho). \quad (7) \end{aligned}$$

The expression in Eq. (7) concludes our discussion of the analytic transformations that we apply to  $w_i(\mathbf{y}|\hat{\mathbf{y}})$  in case of the signal ( $i = 0$ ) as well as the background ( $i = 1$ ) hypothesis.

A few more transformations are necessary to handle the presence of Breit-Wigner (BW) propagators in the ME  $\mathcal{M}_i(\hat{\mathbf{p}})$  while performing the numeric integration. The effect of the BW propagators is that only narrow slices in the 6-particle PS yield sizeable contribution to the integral, namely the regions where the 6 final state particles (the  $b$  and  $\bar{b}$  quarks, the two charged leptons, and the two neutrinos) satisfy certain mass constraints. The mass constraints reflect the presence of (on-shell) H bosons, W bosons, and top quarks in the decay

chains  $HH \rightarrow b\bar{b} W^+ W^- \rightarrow b\bar{b} \ell^+ \nu \ell^- \bar{\nu}$  and  $t\bar{t} \rightarrow b W^+ \bar{b} W^- \rightarrow b \ell^+ \nu \bar{b} \ell^- \bar{\nu}$ . The presence of these narrow slices in the 6-particle PS renders the numeric integration inefficient, unless the mass constraints are treated analytically. We use the narrow-width approximation [9] to treat the mass constraints and replace the BW propagators by  $\delta$ -functions, which has the effect of restricting the numerical integration to the regions in PS where the mass constraints are satisfied and the ME  $\mathcal{M}_i(\hat{\mathbf{p}})$  yields a sizeable contribution to the integral. The analytic transformations specific to the signal hypothesis are detailed in Section 2, while those specific to the background hypothesis are presented in Section 3.

The final numeric integration is performed using the VAMP algorithm [10], a variant of the popular VEGAS algorithm [11], which has been optimized for the case of integrating multimodal functions, typically appearing in the integration of ME over regions in PS.

## 2 Analytic transformations specific to the signal hypothesis

In  $HH \rightarrow b\bar{b} W^+ W^- \rightarrow b\bar{b} \ell^+ \nu \ell^- \bar{\nu}$  signal events, either the  $W^+$  or the  $W^-$  boson (but not both) are on mass shell. We denote the off-shell W boson in signal events by an asterisk (\*), and similarly use an asterisk when referring to the charged lepton and neutrino originating from the decay of this W boson. When evaluating the integrand in Eq. (7) for the signal hypothesis, a sizeable contribution to the value of the integral  $w_0(\mathbf{y}|\hat{\mathbf{y}})$  is obtained only for those points  $\hat{\mathbf{p}}$  in the 6-particle PS for which the mass of the system of b plus  $\bar{b}$  quark equals  $m_H = 125.1$  GeV [12], the mass of the system of charged lepton plus neutrino that originates from the decay of the on-shell W boson equals  $m_W = 80.4$  GeV [13], and the mass of the system of two charged leptons plus the two neutrinos equals  $m_H$ .

We formally introduce these mass constraints by inserting three  $\delta(g(x))$  into the integrand of Eq. (7). The procedure is explained in Section 4.2 of the appendix. More specifically, we insert one  $\delta$ -function of the type  $g(\hat{E}_{\bar{b}})$  given by Eq. (19), one of the type  $g(\hat{E}_{\nu})$  given by Eq. (22), and one of the type  $g(\hat{E}_{\ell^*})$  given by Eq. (25). We denote the charged leptons and neutrinos in signal events by the labels  $\ell$  and  $\nu$  ( $\ell^*$  and  $\nu^*$ ) in order to indicate which charged lepton and neutrino originates from the decay of the on-shell (off-shell) W boson and is therefore subject (not subject) to the W mass constraint.

After solving for the  $\delta$ -functions analytically, as detailed in Sections 4.2.1, 4.2.2, and 4.2.3 of the appendix, the resulting expression for the probability density  $w_0(\mathbf{y}|\hat{\mathbf{y}})$  for the signal

hypothesis reads:

$$\begin{aligned}
w_0(\mathbf{y}|\hat{\mathbf{y}}) &= \frac{(m_H \Gamma_H)^2 m_W \Gamma_W}{2^{21} \pi^{14} \sigma E_\ell E_{\ell^*} E_b E_{\bar{b}}} \int d\hat{E}_b d\hat{\theta}_\nu d\hat{\phi}_\nu d\hat{\theta}_{\nu^*} d\hat{\phi}_{\nu^*} \\
&\cdot \underbrace{\hat{\beta}_\nu}_{=1} \underbrace{\hat{E}_\nu \sin \hat{\theta}_\nu}_{=\hat{p}_T^\nu} \underbrace{\hat{\beta}_{\nu^*}}_{=1} \underbrace{\hat{E}_{\nu^*} \sin \hat{\theta}_{\nu^*}}_{=\hat{p}_T^{\nu^*}} \frac{f(\hat{x}_a) f(\hat{x}_b)}{\hat{x}_a \hat{x}_b s} \\
&\cdot |\mathcal{M}_0(\hat{\mathbf{p}})|^2 \frac{\hat{\beta}_b \hat{E}_b}{\beta_b E_b} W(E_b|\hat{E}_b) \frac{\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}}{\beta_{\bar{b}} E_{\bar{b}}} W(E_{\bar{b}}|\hat{E}_{\bar{b}}) W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho) \\
&\cdot \frac{1}{\hat{E}_b \left(1 - \frac{\hat{\beta}_b}{\beta_b} \cos \angle(\hat{\mathbf{e}}_b, \hat{\mathbf{e}}_{\bar{b}})\right)} \\
&\cdot \frac{1}{4 \hat{E}_\ell \sin^2 \left(\frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2}\right)} \frac{1}{\hat{E}_{\ell\nu\ell^*} \left(1 - \hat{\beta}_{\ell\nu\ell^*} \cos \angle(\hat{\mathbf{e}}_{\ell\nu\ell^*}, \hat{\mathbf{e}}_{\nu^*})\right)} \\
&= \frac{(m_H \Gamma_H)^2 m_W \Gamma_W}{2^{23} \pi^{14} \sigma s E_\ell E_{\ell^*} \beta_b E_b^2 \beta_{\bar{b}} E_{\bar{b}}^2} \int d\hat{E}_b d\hat{\theta}_\nu d\hat{\phi}_\nu d\hat{\theta}_{\nu^*} d\hat{\phi}_{\nu^*} \\
&\cdot \hat{p}_T^\nu \hat{p}_T^{\nu^*} \frac{f(\hat{x}_a) f(\hat{x}_b)}{\hat{x}_a \hat{x}_b} \\
&\cdot |\mathcal{M}_0(\hat{\mathbf{p}})|^2 \hat{\beta}_b \hat{E}_b W(E_b|\hat{E}_b) \hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}} W(E_{\bar{b}}|\hat{E}_{\bar{b}}) W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho) \\
&\cdot \frac{1}{\hat{E}_b \left(1 - \frac{\hat{\beta}_b}{\beta_b} \cos \angle(\hat{\mathbf{e}}_b, \hat{\mathbf{e}}_{\bar{b}})\right)} \\
&\cdot \frac{1}{\hat{E}_\ell \sin^2 \left(\frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2}\right)} \frac{1}{\hat{E}_{\ell\nu\ell^*} \left(1 - \hat{\beta}_{\ell\nu\ell^*} \cos \angle(\hat{\mathbf{e}}_{\ell\nu\ell^*}, \hat{\mathbf{e}}_{\nu^*})\right)}, \tag{8}
\end{aligned}$$

with:

$$\begin{aligned}
\hat{E}_{\bar{b}} &= \frac{a \Delta_{m_H} + |b| \sqrt{\Delta_{m_H}^2 - (a^2 - b^2) m_b^2}}{a^2 - b^2} \\
\hat{E}_\nu &= \frac{m_W^2}{4 \hat{E}_\ell \sin^2 \left(\frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2}\right)} \\
\hat{E}_{\nu^*} &= \frac{m_H^2 - m_{\ell\nu\ell^*}^2}{2 \hat{E}_{\ell\nu\ell^*} \left(1 - \hat{\beta}_{\ell\nu\ell^*} \cos \angle(\hat{\mathbf{e}}_{\ell\nu\ell^*}, \hat{\mathbf{e}}_{\nu^*})\right)}, \tag{9}
\end{aligned}$$

90 where:

$$\begin{aligned}
\Delta_{m_H} &= \frac{m_H^2}{2} - m_b^2 \\
a &= \hat{E}_b \\
b &= \underbrace{\sqrt{\hat{E}_b^2 - m_b^2}}_{=\hat{\beta}_b \hat{E}_b} \underbrace{\hat{\mathbf{e}}_b \cdot \hat{\mathbf{e}}_{\bar{b}}}_{=\cos \angle(\hat{\mathbf{e}}_b, \hat{\mathbf{e}}_{\bar{b}})} = \hat{\beta}_b \hat{E}_b \cos \angle(\hat{\mathbf{e}}_b, \hat{\mathbf{e}}_{\bar{b}}). \tag{10}
\end{aligned}$$

The integral given by Eq. (8) extends over 5 remaining variables, which are integrated numerically:  $\hat{E}_b$ ,  $\hat{\theta}_\nu$ ,  $\hat{\phi}_\nu$ ,  $\hat{\theta}_{\nu^*}$ , and  $\hat{\phi}_{\nu^*}$ .

There is one further aspect that needs to be taken into account when evaluating the compatibility of a given event with the signal hypothesis, and that is that there exists a fourfold ambiguity in associating the two measured b-jets to the b and  $\bar{b}$  quarks and in associating the two measured charged leptons to the on-shell and off-shell W bosons. We deal with the fourfold ambiguity by evaluating the integral  $w_0(\mathbf{y}|\hat{\mathbf{y}})$  given by Eq. (8) four times, each time permuting the association of the measured b-jets to the b and  $\bar{b}$  quarks and of the measured charged leptons to the on-shell and off-shell W bosons.

### 3 Analytic transformations specific to the background hypothesis

In  $t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow b\ell^+\nu\bar{b}\ell^-\bar{\nu}$  background events, both W bosons are on-shell and sizeable contributions to the value of the integral  $w_1(\mathbf{y}|\hat{\mathbf{y}})$  are thus obtained only for those points  $\hat{\mathbf{p}}$  in the 6-particle PS for which the masses of the  $\ell^+\nu$  as well as of the  $\ell^-\bar{\nu}$  system are equal to  $m_W = 80.4$  GeV [13], and the masses of the  $b\ell^+\nu$  and  $\bar{b}\ell^-\bar{\nu}$  systems are equal to the top quark mass of  $m_t = 173.0$  GeV [13].

We account for these mass constraints by inserting four  $\delta(g(x))$  into the integrand of Eq. (7): two  $\delta$ -functions of the type  $g(\hat{E}_\nu)$  given by Eq. (22) and two  $\delta$ -functions of the type  $g(\hat{E}_b)$  given by Eq. (28). Note that the first (second)  $\delta$ -function of each type refers to the decay of the top (anti-top) quark and hence carries the labels  $\nu$  and b (and hence the labels  $\nu$  and b need to be replaced by new labels  $\bar{\nu}$  and  $\bar{b}$ ).

After solving for the  $\delta$ -functions analytically, following Sections 4.2.2 and 4.2.4 of the appendix, the resulting expression for the integral  $w_1(\mathbf{y}|\hat{\mathbf{y}})$  for the background hypothesis is



given by:

$$\begin{aligned}
w_1(\mathbf{y}|\hat{\mathbf{y}}) &= \frac{(m_t \Gamma_t)^2 (m_W \Gamma_W)^2}{2^{21} \pi^{14} \sigma E_{\ell^+} E_{\ell^-} E_b E_{\bar{b}}} \int d\hat{\theta}_\nu d\hat{\phi}_\nu d\hat{\theta}_{\bar{\nu}} d\hat{\phi}_{\bar{\nu}} \\
&\cdot \underbrace{\hat{\beta}_\nu}_{=1} \underbrace{\hat{E}_\nu \sin \hat{\theta}_\nu}_{=\hat{p}_T^\nu} \underbrace{\hat{\beta}_{\bar{\nu}}}_{=1} \underbrace{\hat{E}_{\bar{\nu}} \sin \hat{\theta}_{\bar{\nu}}}_{=\hat{p}_T^{\bar{\nu}}} \frac{f(\hat{\mathbf{x}}_a) f(\hat{\mathbf{x}}_b)}{\hat{\mathbf{x}}_a \hat{\mathbf{x}}_b s} \\
&\cdot |\mathcal{M}_1(\hat{\mathbf{p}})|^2 \frac{\hat{\beta}_b \hat{E}_b}{\beta_b E_b} W(E_b|\hat{E}_b) \frac{\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}}{\beta_{\bar{b}} E_{\bar{b}}} W(E_{\bar{b}}|\hat{E}_{\bar{b}}) W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho) \\
&\cdot \frac{1}{4 \hat{E}_{\ell^+} \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_{\ell^+}, \hat{\mathbf{e}}_\nu)}{2} \right)} \frac{1}{4 \hat{E}_{\ell^-} \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_{\ell^-}, \hat{\mathbf{e}}_{\bar{\nu}})}{2} \right)} \\
&\cdot \frac{1}{\hat{E}_{\ell^+ \nu} \left( 1 - \frac{\hat{\beta}_{\ell^+ \nu}}{\hat{\beta}_b} \cos \angle(\hat{\mathbf{e}}_{\ell^+ \nu}, \hat{\mathbf{e}}_b) \right)} \frac{1}{\hat{E}_{\ell^- \bar{\nu}} \left( 1 - \frac{\hat{\beta}_{\ell^- \bar{\nu}}}{\hat{\beta}_{\bar{b}}} \cos \angle(\hat{\mathbf{e}}_{\ell^- \bar{\nu}}, \hat{\mathbf{e}}_{\bar{b}}) \right)} \\
&= \frac{(m_t \Gamma_t)^2 (m_W \Gamma_W)^2}{2^{25} \pi^{14} \sigma s E_{\ell^+} E_{\ell^-} \beta_b E_b^2 \beta_{\bar{b}} E_{\bar{b}}^2} \int d\hat{\theta}_\nu d\hat{\phi}_\nu d\hat{\theta}_{\bar{\nu}} d\hat{\phi}_{\bar{\nu}} \\
&\cdot \hat{p}_T^\nu \hat{p}_T^{\bar{\nu}} \frac{f(\hat{\mathbf{x}}_a) f(\hat{\mathbf{x}}_b)}{\hat{\mathbf{x}}_a \hat{\mathbf{x}}_b} \\
&\cdot |\mathcal{M}_1(\hat{\mathbf{p}})|^2 \hat{\beta}_b \hat{E}_b W(E_b|\hat{E}_b) \hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}} W(E_{\bar{b}}|\hat{E}_{\bar{b}}) W_\rho(p_x^\rho, p_y^\rho | \hat{p}_x^\rho, \hat{p}_y^\rho) \\
&\cdot \frac{1}{\hat{E}_{\ell^+} \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_{\ell^+}, \hat{\mathbf{e}}_\nu)}{2} \right)} \frac{1}{\hat{E}_{\ell^-} \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_{\ell^-}, \hat{\mathbf{e}}_{\bar{\nu}})}{2} \right)} \\
&\cdot \frac{1}{\hat{E}_{\ell^+ \nu} \left( 1 - \frac{\hat{\beta}_{\ell^+ \nu}}{\hat{\beta}_b} \cos \angle(\hat{\mathbf{e}}_{\ell^+ \nu}, \hat{\mathbf{e}}_b) \right)} \frac{1}{\hat{E}_{\ell^- \bar{\nu}} \left( 1 - \frac{\hat{\beta}_{\ell^- \bar{\nu}}}{\hat{\beta}_{\bar{b}}} \cos \angle(\hat{\mathbf{e}}_{\ell^- \bar{\nu}}, \hat{\mathbf{e}}_{\bar{b}}) \right)}, \quad (11)
\end{aligned}$$

with:

$$\begin{aligned}
\hat{E}_b &= \frac{a_t \Delta_{m_t} + |b_t| \sqrt{\Delta_{m_t}^2 - (a_t^2 - b_t^2) m_b^2}}{a_t^2 - b_t^2} \\
\hat{E}_{\bar{b}} &= \frac{a_{\bar{t}} \Delta_{m_t} + |b_{\bar{t}}| \sqrt{\Delta_{m_t}^2 - (a_{\bar{t}}^2 - b_{\bar{t}}^2) m_b^2}}{a_{\bar{t}}^2 - b_{\bar{t}}^2} \\
\hat{E}_\nu &= \frac{m_W^2}{4 \hat{E}_{\ell^+} \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_{\ell^+}, \hat{\mathbf{e}}_\nu)}{2} \right)} \\
\hat{E}_{\bar{\nu}} &= \frac{m_W^2}{4 \hat{E}_{\ell^-} \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_{\ell^-}, \hat{\mathbf{e}}_{\bar{\nu}})}{2} \right)}, \quad (12)
\end{aligned}$$

where:

$$\begin{aligned}
\Delta_{m_t} &= \frac{m_t^2 - m_b^2 - m_W^2}{2} \\
a_t &= \hat{E}_{\ell^+ \nu} \\
b_t &= \sqrt{\hat{E}_{\ell^+ \nu}^2 - m_W^2} \underbrace{\hat{\mathbf{e}}_{\ell^+ \nu} \cdot \hat{\mathbf{e}}_b}_{=\cos \angle(\hat{\mathbf{e}}_{\ell^+ \nu}, \hat{\mathbf{e}}_b)} = \sqrt{\hat{E}_{\ell^+ \nu}^2 - m_W^2} \cos \angle(\hat{\mathbf{e}}_{\ell^+ \nu}, \hat{\mathbf{e}}_b) \\
a_{\bar{t}} &= \hat{E}_{\ell^- \bar{\nu}} \\
b_{\bar{t}} &= \sqrt{\hat{E}_{\ell^- \bar{\nu}}^2 - m_W^2} \underbrace{\hat{\mathbf{e}}_{\ell^- \bar{\nu}} \cdot \hat{\mathbf{e}}_{\bar{b}}}_{=\cos \angle(\hat{\mathbf{e}}_{\ell^- \bar{\nu}}, \hat{\mathbf{e}}_{\bar{b}})} = \sqrt{\hat{E}_{\ell^- \bar{\nu}}^2 - m_W^2} \cos \angle(\hat{\mathbf{e}}_{\ell^- \bar{\nu}}, \hat{\mathbf{e}}_{\bar{b}}). \quad (13)
\end{aligned}$$

The integral given by Eq. (11) extends over 4 remaining variables, which are integrated numerically:  $\hat{\theta}_\nu$ ,  $\hat{\phi}_\nu$ ,  $\hat{\theta}_{\bar{\nu}}$ , and  $\hat{\phi}_{\bar{\nu}}$ .

When evaluating the compatibility of a given event with the background hypothesis, there exists a twofold ambiguity in associating the two measured b-jets to the b and  $\bar{b}$  quarks. We deal with this ambiguity by evaluating the integral  $w_1(\mathbf{y}|\hat{\mathbf{y}})$  given by Eq. (8) two times, each time permuting the association of the measured b-jets to the b and  $\bar{b}$  quarks. Note that in case of the background hypothesis, there is no ambiguity in associating the two measured leptons to the  $W^+$  and  $W^-$  bosons, as both W bosons are on-shell and the measurement of the lepton charge allows for a unique association of each charged lepton to either the  $W^+$  or the  $W^-$  boson, as conservation of electric charge in the decays  $W^+ \rightarrow \ell^+ \nu$  and  $W^- \rightarrow \ell^- \bar{\nu}$  implies that the charge of each lepton  $\ell$  matches the charge of the W boson from the decay of which it originates.

## 4 Appendix

In this section, we derive a few useful relations that allow us to simplify the expression for the probability density  $w_i(\mathbf{y}|\hat{\mathbf{y}})$  in Eq. (1). We start by deriving relations for the TF of charged leptons and b-jets in Section 4.1. In Section 4.2, we will then derive expressions that constrain the energy of the b and  $\bar{b}$  quarks as well as the energies of the neutrino and anti-neutrino. These constraints arise from the presence of BW propagators in the ME  $\mathcal{M}_i(\hat{\mathbf{y}})$  for the signal ( $i = 0$ ) and for the background ( $i = 1$ ) hypothesis. The effect of the BW propagators is that only those points  $\hat{\mathbf{p}}$  in the 6-particle phase space contribute to the value of the integral in Eq. (1) for which certain systems of final state particles satisfy certain mass conditions.

### 4.1 Relations for transfer functions

Following Ref. [6], we make the assumption that the directions of electrons, muons, and b-jets as well as the energies of electrons and muons are measured with negligible experimental resolution. The assumption implies that the TF for electrons and muons is given by:

$$W_\ell(\mathbf{p}|\hat{\mathbf{p}}) = f(E, \theta, \phi) \delta(E - \hat{E}) \cdot \delta(\theta - \hat{\theta}) \cdot \delta(\phi - \hat{\phi}), \quad (14)$$

while the TF for b-jets is given by:

$$W(\mathbf{p}|\hat{\mathbf{p}}) = f(E, \theta, \phi) W(E|\hat{E}) \delta(\theta - \hat{\theta}) \cdot \delta(\phi - \hat{\phi}), \quad (15)$$

where  $E$  denotes the energy,  $\theta$  the polar angle, and  $\phi$  the azimuthal angle of the electron, muon, or b-jet. The function  $W(E|\hat{E})$  quantifies the resolution for measuring the energy

of b-jets. We choose the function  $W(E|\hat{E})$  such that it satisfies the following normalization condition:

$$\int dE W(E|\hat{E}) \equiv 1.$$

The function  $f(E, \theta, \phi)$  ensures that the TF satisfy the normalization condition  $\int d^3\mathbf{p} \Omega(\mathbf{p}) W(\mathbf{p}|\hat{\mathbf{p}}) = 1$ . We only consider events that pass the event selection criteria, i.e. for which  $\Omega(\mathbf{p})$  is equal to one. With  $d^3\mathbf{p} = \beta E^2 \sin \theta dE d\theta d\phi$ , it follows that:

$$1 \equiv \int dE d\theta d\phi \beta E^2 \sin \theta f(E, \theta, \phi) W(E|\hat{E}) \delta(\theta - \hat{\theta}) \cdot \delta(\phi - \hat{\phi}),$$

which implies:

$$f(E, \theta, \phi) = \frac{1}{\beta E^2 \sin \theta}. \quad (16)$$

Eq. (16) holds for electrons and muons as well as for b-jets.

## 4.2 Relations for mass constraints

As explained in Section 1, the presence of BW propagators in the ME  $\mathcal{M}_i(\hat{\mathbf{p}})$  renders the numeric integration inefficient, unless the numeric integration is restricted to those narrow slices in the 6-particle PS where the mass constraints are satisfied. We achieve the desired restriction by inserting suitable  $\delta$ -functions into the integrand on the RHS of Eq. (7). In order to avoid that the insertion of the  $\delta$ -functions changes the value of the integral, we formally insert a factor of 1, which we write as:

$$\begin{aligned} 1 \equiv \text{BW} \cdot \text{BW}^{-1} &= \frac{\pi}{m_X \Gamma_X} \delta(E_X^2 - |\mathbf{p}_X|^2 - m_X^2) \cdot \left( (E_X^2 - |\mathbf{p}_X|^2 - m_X^2)^2 + (m_X \Gamma_X)^2 \right) \\ &= \pi m_X \Gamma_X \delta(E_X^2 - |\mathbf{p}_X|^2 - m_X^2), \end{aligned} \quad (17)$$

where we have used the narrow-width approximation to replace the first BW propagator by a  $\delta$ -function. The symbol X in Eq. (17) refers to the on-shell particle, of mass  $m_X$  and width  $\Gamma_X$ , which imposes the mass constraint.

We insert Eq. (17) into the integrand on the RHS of Eq. (7) and then use the  $\delta$ -function  $\delta(E_X^2 - |\mathbf{p}_X|^2 - m_X^2)$  to eliminate the integration over  $\hat{E}$  for one of the daughter particles that the particle X decays into. The  $\delta$ -function rule:

$$\delta(g(x)) = \frac{1}{|g'(x_0)|} \delta(x - x_0) \quad (18)$$

yields a factor of  $|g'(x_0)|^{-1} ::= |\frac{\partial g}{\partial x}|_{x=x_0}$ , which we account for when eliminating the integration over  $\hat{E}$ . The symbol  $x_0$  denotes the root of  $g(x)$ .

### 4.2.1 Energy of $\bar{b}$ produced in $H \rightarrow b\bar{b}$ decay

The condition that the mass of the (2-particle) system of b plus  $\bar{b}$  quark equals  $m_H$  implies that:

$$\begin{aligned} m_H^2 \equiv m_{b\bar{b}}^2 &= (\hat{E}_b + \hat{E}_{\bar{b}})^2 - (\hat{\mathbf{p}}_b + \hat{\mathbf{p}}_{\bar{b}})^2 \\ &= \hat{E}_b^2 + \hat{E}_{\bar{b}}^2 + 2 \hat{E}_b \hat{E}_{\bar{b}} - |\hat{\mathbf{p}}_b|^2 - |\hat{\mathbf{p}}_{\bar{b}}|^2 - 2 \hat{\mathbf{p}}_b \cdot \hat{\mathbf{p}}_{\bar{b}} \\ &= \underbrace{\hat{E}_b^2 - |\hat{\mathbf{p}}_b|^2}_{=m_b^2} + \underbrace{\hat{E}_{\bar{b}}^2 - |\hat{\mathbf{p}}_{\bar{b}}|^2}_{=m_{\bar{b}}^2} + 2 \underbrace{\hat{E}_b}_{::=a} \hat{E}_{\bar{b}} - 2 \underbrace{\sqrt{\hat{E}_b^2 - m_b^2} \hat{\mathbf{e}}_b \cdot \hat{\mathbf{e}}_{\bar{b}} \sqrt{\hat{E}_{\bar{b}}^2 - m_{\bar{b}}^2}}_{::=b} \\ \implies 0 &= \underbrace{\frac{m_H^2}{2} - m_b^2}_{::=\Delta_{m_H}} - a \hat{E}_{\bar{b}} + b \sqrt{\hat{E}_{\bar{b}}^2 - m_{\bar{b}}^2} ::= g(\hat{E}_{\bar{b}}), \end{aligned} \quad (19)$$

where the symbol  $\hat{\mathbf{e}}_b$  ( $\hat{\mathbf{e}}_{\bar{b}}$ ) denotes a unit vector in direction of the  $b$  ( $\bar{b}$ ) quark. Eq. (19) has two solutions:

$$\hat{E}_{\bar{b}} = \frac{a \Delta_{m_H} \pm |b| \sqrt{\Delta_{m_H}^2 - (a^2 - b^2) m_b^2}}{a^2 - b^2}. \quad (20)$$

The result for  $\hat{E}_{\bar{b}}$  is in agreement with the result given in Eq. (56) of Ref. [6]. We believe there is a typo in the definition of  $\Delta_{m_H}$  in Section A.2.3 of Ref. [6]. Following Ref. [6], we discard the solution of lower energy and consider the solution of higher energy only, i.e. we take the solution corresponding to the + sign in Eq. (20).

The derivative of the RHS of Eq. (19) with respect to  $\hat{E}_{\bar{b}}$  amounts to:

$$\begin{aligned} \frac{1}{|g'(\hat{E}_{\bar{b}})|} &= \frac{1}{\left| a - \frac{b \hat{E}_{\bar{b}}}{\underbrace{\sqrt{\hat{E}_{\bar{b}}^2 - m_b^2}}_{=\beta_{\bar{b}} \hat{E}_{\bar{b}}}} \right|} = \frac{1}{\left| a - \frac{1}{\beta_{\bar{b}}} b \right|} \\ &= \frac{1}{\left| \hat{E}_{\bar{b}} - \frac{1}{\beta_{\bar{b}}} \underbrace{\sqrt{\hat{E}_{\bar{b}}^2 - m_b^2}}_{=\hat{\beta}_{\bar{b}} \hat{E}_{\bar{b}}} \underbrace{\hat{\mathbf{e}}_b \cdot \hat{\mathbf{e}}_{\bar{b}}}_{::=\cos \angle(\hat{\mathbf{e}}_b, \hat{\mathbf{e}}_{\bar{b}})} \right|} = \frac{1}{\left| \hat{E}_{\bar{b}} \left( 1 - \frac{\hat{\beta}_{\bar{b}}}{\beta_{\bar{b}}} \cos \angle(\hat{\mathbf{e}}_b, \hat{\mathbf{e}}_{\bar{b}}) \right) \right|}. \end{aligned} \quad (21)$$

#### 4.2.2 Energy of $\nu$ produced in $W \rightarrow \ell \nu$ decay

The condition that the mass of the (2-particle) system of  $\ell$  plus  $\nu$  equals  $m_W$  implies that:

$$\begin{aligned} m_W^2 \equiv m_{\ell\nu}^2 &= (\hat{E}_\ell + \hat{E}_\nu)^2 - (\hat{\mathbf{p}}_\ell + \hat{\mathbf{p}}_\nu)^2 \\ &= \hat{E}_\ell^2 + \hat{E}_\nu^2 + 2 \hat{E}_\ell \hat{E}_\nu - |\hat{\mathbf{p}}_\ell|^2 - |\hat{\mathbf{p}}_\nu|^2 - 2 \hat{\mathbf{p}}_\ell \cdot \hat{\mathbf{p}}_\nu \\ &= \underbrace{\hat{E}_\ell^2 - |\hat{\mathbf{p}}_\ell|^2}_{=m_\ell^2 \approx 0} + \underbrace{\hat{E}_\nu^2 - |\hat{\mathbf{p}}_\nu|^2}_{=m_\nu^2 \approx 0} + 2 \hat{E}_\ell \hat{E}_\nu - 2 \underbrace{|\hat{\mathbf{p}}_\ell|}_{\approx \hat{E}_\ell} \underbrace{|\hat{\mathbf{p}}_\nu|}_{\approx \hat{E}_\nu} \underbrace{\hat{\mathbf{e}}_\ell \cdot \hat{\mathbf{e}}_\nu}_{::=\cos \angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)} \\ &= 2 \hat{E}_\ell \hat{E}_\nu \underbrace{\left( 1 - \cos \angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu) \right)}_{=2 \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2} \right)} = 4 \hat{E}_\ell \hat{E}_\nu \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2} \right) \\ \implies 0 &= m_W^2 - 4 \hat{E}_\ell \hat{E}_\nu \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2} \right) =: g(\hat{E}_\nu), \end{aligned} \quad (22)$$

which has the solution:

$$\hat{E}_\nu = \frac{m_W^2}{4 \hat{E}_\ell \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2} \right)}. \quad (23)$$

The symbol  $\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)$  refers to the angle between the directions of charged lepton and neutrino. The result for  $\hat{E}_\nu$  in Eq. (23) is in agreement with the result given in Eq. (54) of Ref. [6].

The derivative of the RHS of Eq. (22) with respect to  $\hat{E}_\nu$  yields:

$$\frac{1}{|g'(\hat{E}_\nu)|} = \frac{1}{4 \hat{E}_\ell \sin^2 \left( \frac{\angle(\hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\nu)}{2} \right)}. \quad (24)$$

### 4.2.3 Energy of $\nu^*$ produced in $H \rightarrow WW^* \rightarrow \ell\nu\ell^*\nu^*$ decay

As mentioned previously, we denote by  $\ell^*$  and  $\nu^*$  the charged lepton and the neutrino that originate from the decay of the off-shell W boson. The condition that the mass of the (4-  
165 particle) system of  $\ell$ ,  $\nu$ ,  $\ell^*$ , plus  $\nu^*$  equals  $m_H$  implies that:

$$\begin{aligned}
m_H^2 \equiv m_{\ell\nu\ell^*\nu^*}^2 &= \underbrace{(\hat{E}_\ell + \hat{E}_\nu + \hat{E}_{\ell^*} + \hat{E}_{\nu^*})^2}_{::=\hat{E}_{\ell\nu\ell^*}^2} - \underbrace{(\hat{\mathbf{p}}_\ell + \hat{\mathbf{p}}_\nu + \hat{\mathbf{p}}_{\ell^*} + \hat{\mathbf{p}}_{\nu^*})^2}_{::=\hat{\mathbf{p}}_{\ell\nu\ell^*}^2} \\
&= \hat{E}_{\ell\nu\ell^*}^2 + \hat{E}_{\nu^*}^2 + 2\hat{E}_{\ell\nu\ell^*}\hat{E}_{\nu^*} - |\hat{\mathbf{p}}_{\ell\nu\ell^*}|^2 - |\hat{\mathbf{p}}_{\nu^*}|^2 - 2\hat{\mathbf{p}}_{\ell\nu\ell^*} \cdot \hat{\mathbf{p}}_{\nu^*} \\
&= \underbrace{\hat{E}_{\ell\nu\ell^*}^2 - |\hat{\mathbf{p}}_{\ell\nu\ell^*}|^2}_{::=m_{\ell\nu\ell^*}^2} + \underbrace{\hat{E}_{\nu^*}^2 - |\hat{\mathbf{p}}_{\nu^*}|^2}_{=m_{\nu^*}^2 \approx 0} + 2 \underbrace{\hat{E}_{\ell\nu\ell^*}\hat{E}_{\nu^*}}_{::=a} \\
&\quad - 2 \underbrace{\sqrt{\hat{E}_{\ell\nu\ell^*}^2 - m_{\ell\nu\ell^*}^2} \hat{\mathbf{e}}_{\ell\nu\ell^*} \cdot \hat{\mathbf{e}}_{\nu^*}}_{::=b} \underbrace{|\hat{\mathbf{p}}_{\nu^*}|}_{\approx \hat{E}_{\nu^*}} \\
&= m_{\ell\nu\ell^*}^2 + 2a\hat{E}_{\nu^*} - 2b\hat{E}_{\nu^*} \\
\Rightarrow 0 &= \underbrace{\frac{m_H^2 - m_{\ell\nu\ell^*}^2}{2}}_{::=\Delta_{m_H}} - a\hat{E}_{\nu^*} + b\hat{E}_{\nu^*} =: g(\hat{E}_{\nu^*}), \tag{25}
\end{aligned}$$

which has the solution:

$$\begin{aligned}
\hat{E}_{\nu^*} &= \frac{\Delta_{m_H}}{a-b} = \frac{m_H^2 - m_{\ell\nu\ell^*}^2}{2 \left( \hat{E}_{\ell\nu\ell^*} - \underbrace{\sqrt{\hat{E}_{\ell\nu\ell^*}^2 - m_{\ell\nu\ell^*}^2}}_{=\hat{\beta}_{\ell\nu\ell^*} \hat{E}_{\ell\nu\ell^*}} \underbrace{\hat{\mathbf{e}}_{\ell\nu\ell^*} \cdot \hat{\mathbf{e}}_{\nu^*}}_{::=\cos \angle(\hat{\mathbf{e}}_{\ell\nu\ell^*}, \hat{\mathbf{e}}_{\nu^*})} \right)} \\
&= \frac{m_H^2 - m_{\ell\nu\ell^*}^2}{2 \hat{E}_{\ell\nu\ell^*} (1 - \hat{\beta}_{\ell\nu\ell^*} \cos \angle(\hat{\mathbf{e}}_{\ell\nu\ell^*}, \hat{\mathbf{e}}_{\nu^*}))}, \tag{26}
\end{aligned}$$

where, for the purpose of shortening the nomenclature, we denote by the symbol  $\hat{E}_{\ell\nu\ell^*}$  ( $\hat{\mathbf{p}}_{\ell\nu\ell^*}$ ) the energy (momentum) of the three-particle system comprised of the two charged leptons and of the neutrino originating from the decay of the on-shell W boson, and by  $\hat{\mathbf{e}}_{\ell\nu\ell^*}$  ( $m_{\ell\nu\ell^*}$ )  
170 a unit vector in direction of  $\hat{\mathbf{p}}_{\ell\nu\ell^*}$  (the mass of this system).

The derivative of the RHS of Eq. (25) with respect to  $\hat{E}_{\nu^*}$  amounts to:

$$\begin{aligned}
\frac{1}{|g'(\hat{E}_{\nu^*})|} &= \frac{1}{|a-b|} = \frac{1}{\hat{E}_{\ell\nu\ell^*} - \underbrace{\sqrt{\hat{E}_{\ell\nu\ell^*}^2 - m_{\ell\nu\ell^*}^2}}_{=\hat{\beta}_{\ell\nu\ell^*} \hat{E}_{\ell\nu\ell^*}} \cos \angle(\hat{\mathbf{e}}_{\ell\nu\ell^*}, \hat{\mathbf{e}}_{\nu^*})} \\
&= \frac{1}{\hat{E}_{\ell\nu\ell^*} (1 - \hat{\beta}_{\ell\nu\ell^*} \cos \angle(\hat{\mathbf{e}}_{\ell\nu\ell^*}, \hat{\mathbf{e}}_{\nu^*}))}. \tag{27}
\end{aligned}$$

#### 4.2.4 Energy of b ( $\bar{b}$ ) produced in $t \rightarrow bW^+ \rightarrow b\ell^+\nu$ ( $\bar{t} \rightarrow \bar{b}W^- \rightarrow \bar{b}\ell^-\bar{\nu}$ ) decay

The condition that the mass of the (3-particle) system of b quark, charged anti-lepton, plus neutrino equals  $m_t$  implies that:

$$\begin{aligned}
m_t^2 \equiv m_{b\ell^+\nu}^2 &= (\hat{E}_b + \underbrace{\hat{E}_{\ell^+\nu}}_{::=\hat{E}_{\ell^+\nu}})^2 - (\underbrace{\hat{\mathbf{p}}_b + \hat{\mathbf{p}}_{\ell^+\nu}}_{::=\hat{\mathbf{p}}_{\ell^+\nu}})^2 \\
&= \hat{E}_b^2 + \hat{E}_{\ell^+\nu}^2 + 2\hat{E}_b\hat{E}_{\ell^+\nu} - |\hat{\mathbf{p}}_b|^2 - |\hat{\mathbf{p}}_{\ell^+\nu}|^2 - 2\hat{\mathbf{p}}_b \cdot \hat{\mathbf{p}}_{\ell^+\nu} \\
&= \underbrace{\hat{E}_b^2 - |\hat{\mathbf{p}}_b|^2}_{=m_b^2} + \underbrace{\hat{E}_{\ell^+\nu}^2 - |\hat{\mathbf{p}}_{\ell^+\nu}|^2}_{=m_W^2} + 2\underbrace{\hat{E}_{\ell^+\nu}\hat{E}_b}_{::=a} \\
&\quad - 2\underbrace{\sqrt{\hat{E}_{\ell^+\nu}^2 - m_W^2}\hat{\mathbf{e}}_{\ell^+\nu} \cdot \hat{\mathbf{e}}_b}_{::=b} \sqrt{\hat{E}_b^2 - m_b^2} \\
\Rightarrow 0 &= \underbrace{\frac{m_t^2 - m_b^2 - m_W^2}{2}}_{::=\Delta_{m_t}} - a\hat{E}_b + b\sqrt{\hat{E}_b^2 - m_b^2} =: g(\hat{E}_b), \tag{28}
\end{aligned}$$

where we denote the energy (momentum) of the system of charged anti-lepton and neutrino by the symbol  $\hat{E}_{\ell^+\nu}$  ( $\hat{\mathbf{p}}_{\ell^+\nu}$ ), and use the symbol  $\hat{\mathbf{e}}_{\ell^+\nu}$  to denote a unit vector in direction of its momentum. The mass of the system of charged anti-lepton and neutrino equals  $m_W$ , as the W boson produced in the decay  $t \rightarrow bW^+$  is on-shell. Eq. (28) has two solutions:

$$\hat{E}_b = \frac{a\Delta_{m_t} \pm |b|\sqrt{\Delta_{m_t}^2 - (a^2 - b^2)m_b^2}}{a^2 - b^2}. \tag{29}$$

The result for  $\hat{E}_b$  is in agreement with the result given in Eq. (55) of Ref. [6]. Following Ref. [6], we discard the solution of lower energy and consider the solution of higher energy only, i.e. we take the solution corresponding to the + sign in Eq. (29).

The derivative of the RHS of Eq. (28) with respect to  $\hat{E}_b$  yields:

$$\begin{aligned}
\frac{1}{|g'(\hat{E}_b)|} &= \frac{1}{|a - \frac{b\hat{E}_b}{\underbrace{\sqrt{\hat{E}_b^2 - m_b^2}}_{=\beta_b\hat{E}_b}}|} = \frac{1}{|a - \frac{1}{\beta_b}b|} = \frac{1}{|\hat{E}_{\ell^+\nu} - \frac{1}{\beta_b}\underbrace{\sqrt{\hat{E}_{\ell^+\nu}^2 - m_W^2}}_{=\hat{\beta}_{\ell^+\nu}\hat{E}_{\ell^+\nu}}\underbrace{\hat{\mathbf{e}}_{\ell^+\nu} \cdot \hat{\mathbf{e}}_b}_{::=\cos\angle(\hat{\mathbf{e}}_{\ell^+\nu}, \hat{\mathbf{e}}_b)}|} \\
&= \frac{1}{|\hat{E}_{\ell^+\nu} \left(1 - \frac{\hat{\beta}_{\ell^+\nu}}{\beta_b} \cos\angle(\hat{\mathbf{e}}_{\ell^+\nu}, \hat{\mathbf{e}}_b)\right)|}. \tag{30}
\end{aligned}$$

The corresponding expressions for the case of  $\bar{b}$  quark, charged lepton, and anti-neutrino are identical to Eqs. (29) and (30), with b simply replaced by  $\bar{b}$  and  $\ell^+\nu$  by  $\ell^-\bar{\nu}$ .

## References

- [1] J. Neyman, Outline of a theory of statistical estimation based on the classical theory of probability, Phil. Trans. Roy. Soc. Lond. A 236 (767) (1937) 333. doi:10.1098/rsta.1937.0005.

- [2] I. Volobouev, Matrix element method in HEP: transfer functions, efficiencies, and likelihood normalization [arXiv:1101.2259](#).
- [3] J. D. Bjorken, Asymptotic sum rules at infinite momentum, *Phys. Rev.* 179 (1969) 1547. doi:10.1103/PhysRev.179.1547.
- 190 [4] A. D. Martin, W. J. Stirling, R. S. Thorne, G. Watt, Parton distributions for the LHC, *Eur. Phys. J. C* 63 (2009) 189. [arXiv:0901.0002](#), doi:10.1140/epjc/s10052-009-1072-5.
- [5] F. Fiedler, A. Grohsjean, P. Haefner, P. Schieferdecker, The matrix element method and its application in measurements of the top quark mass, *Nucl. Instrum. Meth. A* 624 (2010) 203. [arXiv:1003.1316](#), doi:10.1016/j.nima.2010.09.024.
- 195 [6] Bianchini, L. and others, Search for the associated  $t\bar{t}H$  ( $\rightarrow b\bar{b}$ ) production using the matrix element method, CMS Analysis Note 2013/313.
- [7] J. Alwall, A. Freitas, O. Mattelaer, The matrix element method and QCD radiation, *Phys. Rev. D* 83 (2011) 074010. [arXiv:1010.2263](#), doi:10.1103/PhysRevD.83.074010.
- 200 [8] Alexander, J. and others, Towards a robust  $p_T^{\text{miss}}$  significance with particle flow, CMS Analysis Note 2010/400.
- [9] D. Berdine, N. Kauer, D. Rainwater, Breakdown of the narrow width approximation for new physics, *Phys. Rev. Lett.* 99 (2007) 111601. [arXiv:hep-ph/0703058](#), <http://dx.doi.org/10.1103/PhysRevLett.99.111601> doi:10.1103/PhysRevLett.99.111601.
- 205 [10] T. Ohl, Vegas revisited: Adaptive Monte Carlo integration beyond factorization, *Comput. Phys. Commun.* 120 (1999) 13. [arXiv:hep-ph/9806432](#), [http://dx.doi.org/10.1016/S0010-4655\(99\)00209-X](http://dx.doi.org/10.1016/S0010-4655(99)00209-X) doi:10.1016/S0010-4655(99)00209-X.
- [11] G. P. Lepage, A new algorithm for adaptive multidimensional integration, *J. Comput. Phys.* 27 (1978) 192. doi:10.1016/0021-9991(78)90004-9.
- 210 [12] G. Aad, et al., Combined measurement of the Higgs boson mass in pp collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS experiments, *Phys. Rev. Lett.* 114 (2015) 191803. [arXiv:1503.07589](#), doi:10.1103/PhysRevLett.114.191803.
- 215 [13] K. A. Olive, et al., Review of particle physics, *Chin. Phys. C* 38 (2014) 090001. doi:10.1088/1674-1137/38/9/090001.