UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

Master's Thesis

A Study of Top Anomalous Couplings at a Future e^+e^- Collider

Julie Munch Torndal

Supervisor: Jørgen Beck Hansen

August 16th, 2021

Abstract

Particle Physics has arrived at an important moment in its history. The Standard Model has been probed up to several TeV and no new physics has been observed. There are no indications whether new physics is located at larger masses or smaller couplings. The only way to find out is to go look. The FCC-ee experiment is one of the proposed future e^+e^- colliders which would push the limits of the precision frontier, offering indirect sensitivity to new physics. The precise values of the couplings matter and any deviations would be a major discovery. The couplings of the top quark are of particular interest as it is the heaviest particle in the entire Standard Model suggesting it might couple more strongly to BSM fields.

In this thesis, the expected sensitivity to top anomalous couplings at the FCC-ee is determined. The work thereby addresses an aspect of the Snowmass 2021 Letter of Interest on top quark physics at FCC-ee, which specifically mentions sentitivity to top quark EFT couplings as a proposal for an analysis at the FCC-ee. The sensitivity is gauged in the semileptonic channel for top pair produced events in simulated datasets generated with FCCSW in the experimental environment of the IDEA detector at $\sqrt{s} = 365\,\mathrm{GeV}$. Jet performance studies consider various jet definitions with work flowing back into FCCAnalyses to reevaluate jet tools in future works. Event reconstruction is performed, an event selection developed, and a kinematic fit is applied using a software package written in connection with the analysis.

The 1σ confidence intervals are determined for a minimal set of top anomalous couplings. The intervals are found from the observables of angular distributions and total cross sections for each coupling separately.

Acknowledgements

First and foremost I would like to thank my supervisor Jørgen Beck Hansen for his expert guidance, his patience, and our working sessions, as well as for his mentorship these past years helping me achieving my academic goals. Further I would like to thank Mogens Dam for his introduction and insight into the FCC-ee project.

A special thank you to the conveners of the FCC-ee Physics Performance group, Patrizia Azzi and Emmanuel Perez, for their encouragement and their suggestions. They provided me with great insight into the inner works of research and have been a source of inspiration. Furthermore, a special thank you to Clement Helsens for his invaluable help and expertise regarding the software aspects of FCCSW. I would also like to acknowledge the people I have encountered as part of FCC-ee, for their friendly introduction into the scientific community.

Finally, I would like to thank a very special person for his constant support throughout this process of writing a master's thesis during a pandemic.

Contents

1. Introduction

Particle Physics studies the properties of the elementary particles and the interactions between them. Its manifestation is a consequence of humans striving for a more complete understanding of the Universe in its fundamental elements. The pièce de résistance of Particle Physics is the Standard Model which is the grand theory of *almost* everything. In 2012 the Standard Model was completed with the discovery of the Higgs boson at the Large Hadron Collider (LHC) [1, 2]. In addition to the discovery of the Higgs boson the LHC so far leaves a legacy where the Standard Model has been confirmed up to several TeV in the absence of new physics. The remaining decade of operation time has potential to change this picture. In the meantime, the future of Particle Physics is looking to return to lepton colliders. One of the proposed lepton colliders is the Future Circular Collider (FCC) hosted at the European Organization for Nuclear Research (CERN). The first stage will host electron-positron collisions at the FCC-ee, and the second stage will host proton-proton collisions at the FCC-hh. The history of particle physics have long shifted between lepton and hadron colliders with discoveries of equal importance made at both. They are both essential in the toolbox of modern high energy physics and the alternation between the two has lead to significant progress bringing us to where we are today. Hadron colliders are excellent for searching for new particles taking a model-dependent approach and breaking through the energy frontier. But as they splinter the proton to pieces, lots of uninteresting backgrounds are created adding up to the minimum bias and complicating the picture. Some examples of discoveries are the top quark that was discovered at the Tevatron (CERN) in 1995 and most recently the Higgs boson that was discovered at the LHC (CERN). Lepton colliders are excellent for probing the Standard Model and testing it in its limits taking the model-independent approach and breaking through the precision frontier. There is no underlying structure with cleaner final states but they do not reach the same energies as at hadron colliders. Some examples of important discoveries are the gluon discovered at Petra (DESY) in 1979 and the three neutrino generations discovered at LEP (CERN) in 1989. Precision measurements have also been precursors for the discovery of new phenomena and new particles playing an important role for the road-map of higher-energy machines.

With the Higgs boson discovery, the Standard Model has no flexibility to predict new phenomena such as non-baryonic dark matter, cosmological baryon-antibaryon asymmetry, the finite neutrino masses, etc. which are all experimental facts. So the question is where to look for new phenomena beyond the Standard Model. There are no indications whether new physics is at larger masses or smaller couplings. The only way to find out is to go look. More precision gives indirect sensitivity to new physics. With high luminosity, extreme statistical precision and experimental accuracy, the FCC-ee holds the potential to observe tiny violations of established symmetries. The precise values of the couplings matter and controls the structure of the universe and non-trivial deviations from the Standard Model predictions would indicate physics beyond the Standard Model and be a major discovery for unveiling a new theoretical framework. One particle of particular interest is the top quark. It is the heaviest particle in the entire Standard Model with a mass of around 173 GeV [3] comparable to the Higgs field expectation value [4], which suggest that the top quark might couple more strongly to fields beyond the Standard Model.

This thesis outlines a way to determine the sensitivity to top anomalous couplings at the FCC-ee. As the name might suggest, top anomalous couplings are absent in the framework of the Standard Model but they can be introduced in the framework of Effective Field Theory in which they are suppressed by a characteristic length scale. At length scales beyond those probed by today's colliders, anomalous couplings become relevant depending on the characteristic length scale. Here a minimal set of top anomalous couplings are studied. Section 2 summarises concepts of the Standard model formulated within Quantum Field Theory. The Effective Field Theory framework is presented in Section 3 where the minimal set of anomalous couplings are introduced as an extension to the Standard Model. Section 4 conceptualises the experimental environment at the FCC-ee. The planning of the detector design is still at a preliminary stage. The analysis is performed within the semileptonic channel of top pair produced events at a center-of-mass energy $\sqrt{s} = 365\,\mathrm{GeV}$. The Monte Carlo generation of signal and background samples for the event selection is described in Section 5 in addition to the Monte Carlo generation of top pair production with the individual anomalous couplings. Jets are crucial to the signature of semileptonic $t\bar{t}$ events. The performance of a set of jet algorithms has been been studied in Section 6. Section 7 presents the event selection where a significance-optimised selection cut strategy has been applied for deciding the signal region. Following the event selection, the event reconstruction is improved by applying kinematic fitting to the measured events. The method is described in Section 8 along with a description of the software package that was written in connection with this analysis. Finally, the analysis is outlined in Section 9 and the results are presented.

2. The Standard Model

The Standard Model of Particle Physics (SM) is a theoretical framework that embodies our current knowledge of the *fundamental constituents* of the Universe: the elementary particles and the interactions between them, the forces. The definition of *fundamental constituents* has been a long journey since their existence was first postulated by Democritus in 400 BCE. He held that everything is composed of the "atomi", the smallest indivisible bodies, and that they exist in the vacuum. Anything else is merely thought to exist.

Since Democritus' time we have evolved from understanding the world in terms of four classical elements to understanding the world in terms of four fundamental forces and a "zoo" of particles. We have moved through different length scales surrounding the

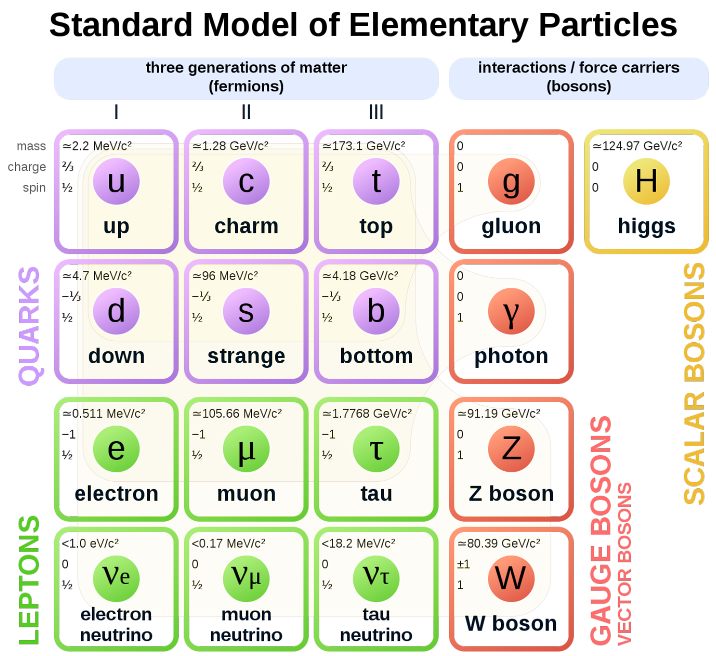


Figure 2.1.: The Standard Model Elementary Particles: The fermions are divided into three generations (column 1-3). Each generation contains an up-like quark (up, charm, top) with electric charge $-1/3$, a down-like quark (down, strange, bottom) with charge $+2/3$, a negatively charged lepton (electron, muon, tau) and a corresponding neutrino with zero charge. The three generations only differ in masses which increases with each the generation. The fermions are spin–$1/2$ particles, the gauge bosons are spin–1 particles and the Higgs is a spin–0 particle. Figure taken from [5].

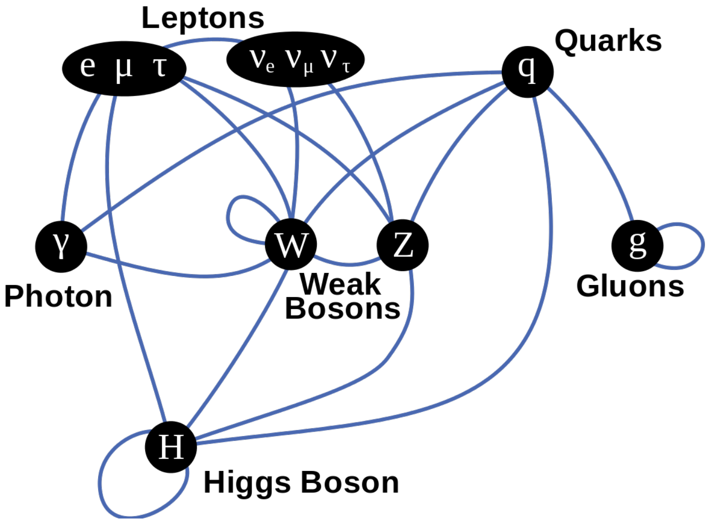


Figure 2.2.: Standard Model Interactions

building blocks of the Universe from studying the chemical elements, to the atoms, to the nucleus and, at present, we have arrived at the elementary particles.

The elementary particles are classified by fermions that make up matter and by the force carriers that mediate the interactions between the particles. They are summarised in Figure 2.1 along with some of their properties. The fermions are spin–$1/2$ particles divided into quarks and leptons. For each fermion there exists an anti-particle with the same properties but quantum numbers of opposite sign. The gauge bosons are spin–1 particles. They are responsible for three of the four fundamental forces which are described by the exchange of a gauge boson. The electromagnetic force is mediated by the photons, and only fermions with electric charge experience this force. The weak force is mediated by the W and Z bosons and is experienced by all fermions since all fermions carry weak isospin. The strong force is mediated by the gluons, and only the quarks that carry colour charge experience this force. Extending the picture to the gravitational force, the existence of a graviton can be hypothesised, however since it is too weak on the mass and length scales characteristic for particle physics experiments, the graviton remains hypothetical. The Higgs boson is a scalar boson with spin–0 and particles acquire mass through the Higgs mechanism as they interact with the Higgs field. The Standard Model was completed with the discovery of the Higgs boson in 2012 [1, 2]. A visualised summary of the interactions can be seen in Figure 2.2.

The Standard Model of Particle Physics (SM) as we know it today was primarily developed in the 1970's by S. Glashow, S. Weinberg, and A. Salam [6, 7]. Its mathematical description is formulated within Quantum Field Theory (QFT) frameworks. For the SM, the QFT frameworks include quantum chromodynamics (QCD) and the Glasgow-Weinberg-Salam model for electroweak unification, which also contains quantum electrodynamics (QED). The particles are described by excitations of a quantum field and are expressed in terms of Lagrangians.

The forces in the SM are associated with underlying symmetries. The SM is based on the gauge group

$$SU(3)_C \times SU(2)_L \times U(1)_Y \tag{2.1}$$

where the indices represent the charges of the groups.

2.1. Quantum Electrodynamics

Quantum electrodynamics (QED) is the quantum field theory of the electromagnetic interaction. QED is gauge invariant under the local phase transformation U(1). The theory of QED can be described in an easy and compact way in the Lagrangian formalism where particles are excitations of a quantum field that satisfies the appropriate QM field equations. Classically, the Lagrangian is defined as the difference between kinematic energy and potential energy

$$L = T - V \tag{2.2}$$

and the Lagrangian is specified by a discrete set of particles and their time derivatives $L(q_i, \dot{q}_i)$. Substituting the Lagrangian into the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \tag{2.3}$$

gives the equations of motion for the system. Instead of the discrete system of particles above, the physical system for fields are described by continuous systems. The Lagrangian formalism can be extended by introducing a Lagrangian density, $\mathcal{L}$, such that

$$L\left(q_i, \frac{\mathrm{d}q_i}{\mathrm{d}t}\right) \to \mathcal{L}(\phi_i, \partial_\mu \phi_i). \tag{2.4}$$

In the Lagrangian density the set of particles is replaced by the fields $\phi_i(t, x, y, z)$ and the time derivatives are replaced by the derivatives of the fields with respect to each of the four space-time coordinates. It can be shown that the equivalent of the Euler-Lagrange equation is [4]

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\right) - \frac{\partial \mathcal{L}}{\partial \phi_i} = 0. \tag{2.5}$$

The Lagrangian itself is the spatial integral of the Lagrangian density. Classically, a symmetry of the Lagrangian is related to a conservation law. In field theory, a symmetry of the Lagrangian is related to a conserved current by Noether's theorem [4]. Henceforward the Lagrangian density will simply be denoted as the Lagrangian. The Lagrangian of QED describes the electron field, the photon field, and the interactions between them and it can be written as

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu \partial_\mu - m_e)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.6}$$

The first term describes the spin-half spinor field, ψ, of an electron with mass m_e. Substituting the first term into the Euler-Lagrange equation, it can be shown that the field satisfies the Dirac equation given by [4]

$$i\gamma^\mu(\partial_\mu \phi) - m\phi = 0. \tag{2.7}$$

A Lagrangian that satisfies the Dirac equation is not sufficient to describe the whole of electromagnetism as it does not satisfy a local $U(1)$ gauge symmetry. The symmetry is restored by introducing the second term that describes the interactions between the electron and the photon field A_μ, and require that the spinor field and photon field transform as

$$\psi(x) \to e^{i\alpha(x)}\psi(x) \tag{2.8}$$

$$A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x) \tag{2.9}$$

which ensures invariance of the Lagrangian under a transformation of a local phase, $\alpha(x)$. The last term describes the photon field itself from which Maxwell's equations can be recovered. The Lagrangian for QED sums of the whole of electromagnetism. It can be rewritten on the form

$$\mathcal{L}_{QED} = -\bar{\psi}(i\not{D} + m_e)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.10}$$

where $\not{D}$ is slash notation defined as $\not{a} \equiv \gamma^\mu a_\mu$ and D_μ is the covariant derivative of electromagnetic interactions

$$D_\mu = \partial_\mu - ieA_\mu(x) \tag{2.11}$$

Similarly to QED, Lagrangians can be derived for charged-current weak and strong interactions by requiring a local $SU(2)_L$ gauge symmetry and $SU(3)$ gauge symmetry respectively for a Lagrangian that satisfies the Dirac equation. The generator for a $U(1)$ transformation simply corresponds to a multiplication by a complex phase as in Eq. 2.8. The $SU(2)$ symmetry group has three generators that define a non-Abelian Lie algebra. The generators can be expressed as $\mathbf{T} = \boldsymbol{\sigma}/2$ where $\boldsymbol{\sigma}$ are the three Pauli spin matrices that do not commute. The covariant derivative of charged current weak interactions is

$$D_\mu = \partial_\mu + ig_W\boldsymbol{T}\cdot\boldsymbol{W}_\mu(x) \tag{2.12}$$

where g_W is the coupling of the W boson, T are the three generators of the group and $W_\mu(x)$ are three fields associated with the group. The $SU(3)$ symmetry group has eight generators and is also non-Abelian. For a non-Abelian gauge theory, the associated gauge fields are not independent. In order to conserve the gauge invariance additional gauge boson self-interaction terms have to be added to the Lagrangians for weak interactions. The principles of the Lagrangian formalism introduced here shows how compact field theories can be summarised. In fact with the right substitutions, the Lagrangian for the whole SM can fit on a coffee mug.

2.2. Quantum Chromodynamics

Quantum chromodynamics (QCD) is the quantum field theory of the strong interaction and it is gauge invariant under the local phase transformation $SU(3)$. The strong interaction bears many similarities to the electromagnetic interaction. Similarly to QED, there is a charge associated with QCD which is denoted colour charge. It is mediated by the massless gluons that couple to particles that carry colour charge similar to how the massless photon couples to particles that carry electric charge in QED. However the gauge group symmetry leads to a distinct behaviour of QCD. Non-Abelian $SU(3)$ has eight generators, $\mathbf{T}$, where the covariant derivative can be written as

$$D_\mu = \partial_m u + ig_S\mathbf{G}_\mu \cdot \mathbf{T}, \tag{2.13}$$

where $\mathbf{G}$ are the eight fields for each of the generators. The Dirac Lagrangian with the covariant derivative above is invariant under the local phase transformation $SU(3)$ provided that the fields transform as

$$G^k_\mu(x) \to G^k_\mu(x) - \partial_\mu\alpha_k(x) - g_S f_{ijk}\alpha_i(x)G^j_\mu(x) \tag{2.14}$$

where f_{ijk} are the structure constants of the $SU(3)$ group defined by the commutation relations $[T_i, T_j] = if_{ijk}T_k$. Comparing to the transformation of the photon field in Eq. 2.9, the last terms for the gluon fields arise from the non-Abelian behaviour and give rise to gluon self-interactions. Since gluons couple to particles that carry colour charge, they themselves must also carry colour. There are 3 colour states labeled *red*, *blue*, and *green* and they represent 3 degrees of freedom introduced from SU(3). Quarks carry colour charge, their anti-particles carry anti-colour charge and gluons carry a combination of the two. The quantum numbers representing these states are colour isospin I_3^c and colour hypercharge Y^c. The colour of quarks and anti-quarks can be represented in a Y^c vs. I_3^c diagram as seen in Figure 2.3.

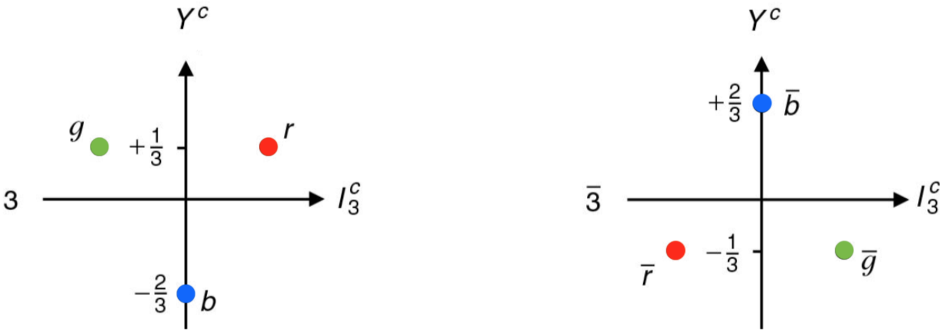


Figure 2.3.: Y^c vs. I_3^c diagram representing the three colour states for quarks and the three anti-colour states for anti-quarks.

Colour is an always conserved quantity which follows from the colour symmetry of QCD. It is an exact symmetry which means the strength of the interaction is independent of colour charge. The strength of an interaction is determined by the coupling in the vertex. In strong interactions, the coupling strength changes rapidly with the momentum transfer. In electromagnetic interactions depends the coupling strength on the momentum transfer as well, but the behaviour is much less pronounced. The running of the couplings are due to higher order diagrams which contribute to the effective coupling. The total amplitude, $\mathcal{M}_{fi}$, for any process has to be summed over all individual Feynman diagrams that have the same final state

$$\mathcal{M}_{fi} = \mathcal{M}_{LO} + \sum \mathcal{M}_{HO}. \tag{2.15}$$

The sum is formed over an infinite set of higher order diagrams that will contribute. The infinities can be renormalised away by knowing the effective coupling at some scale $q^2 = \mu^2$, and the running coupling of α in QED is [4]

$$\alpha(q^2) = \frac{\alpha(\mu^2)}{1 - \alpha(\mu^2)\frac{1}{3\pi}\ln\left(\frac{q^2}{\mu^2}\right)}. \tag{2.16}$$

It increases slowly with q^2 and $\alpha \approx 1/137$ is a good approximation. For leading order (LO), the matrix element is directly proportional to α, next-to-leading order (NLO) is proportional to α^2, next-to-next-to-leading order (NNLO) is proportional to α^3 and so on. With $\alpha \approx 1/137$, the series of matrix elements converges rapidly and is dominated by the LO term which is necessary for applying perturbation theory to QED interactions. The running of α_S decreases with increasing q^2

$$\alpha_S(q^2) = \frac{\alpha_S(\mu^2)}{1 + B\alpha_S(\mu^2)\ln\left(\frac{q^2}{\mu^2}\right)} \quad , \quad B = \frac{11N_c - 2N_f}{12\pi} \tag{2.17}$$

where N_c is the number of colour states and N_f is the number of quark flavours. The distinct behaviour is a consequence of gluon self-interactions that also contribute to the effective coupling which leads to the asymptotic freedom of QCD. QCD can be divided into two regimes. In the high-energy regime at $|q| > 100$ GeV, $\alpha_S \sim 0.1$ which is small enough that perturbation theory can be applied. However perturbation expansion does not converge as fast as for QED and higher order corrections cannot be neglected. In the low-energy regime, perturbation theory cannot be applied and no calculations from first principles exists. Instead we have to rely on phenomenological models that can be investigated in elastic scattering experiments.

Free quarks have never been observed. In collider experiments quarks are initially produced as free particles at very high energies in the process

$$ee \rightarrow q\bar{q} \rightarrow \text{hadrons}. \tag{2.18}$$

As the quarks travel away from each other their colour charges give rise to a colour field where virtual gluons can interact with themselves. This behaviour squeezes the colour field such that the energy stored in the field is proportional to the distance. At some point the field contains enough energy to produce a new pair of quarks which breaks up the field. This process will repeat itself until the energy is sufficiently low to form hadrons. The process is known as hadronisation. It belongs in the low-energy regime and can only described qualitatively. The gluon self-interactions give rise to colour confinement where only colour singlet states can exist as free particles. In high energy experiments quarks and gluons are always observed as jets of hadrons.

2.3. Electroweak unification

QED interactions are invariant under local U(1) phase transformations and QCD interactions are invariant under local SU(3) phase transformations. The underlying symmetry of charged current (CC) weak interactions mediated by W^+ and W^- is SU(2), however the neutral current (NC) weak interactions mediated by the Z do not share the same symmetry of CC weak interaction

Starting from SU(2), there are three generators of the symmetry that are written in terms of the three 2×2 Pauli matrices. The local gauge invariance is satisfied by introducing three gauge fields with three associated gauge bosons, $W^{(1)}$, $W^{(2)}$, and $W^{(3)}$. The wave functions are represented by weak isospin, I_W. Since CC weak interactions only couple to left-handed (LH) chiral state particles and right-handed (RH) chiral state anti-particles, the symmetry is referred to as SU$(2)_L$. These particles are placed in isospin doublets with $I_W = \frac{1}{2}$ because of the dimensions of the Pauli matrices

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L , \quad \begin{pmatrix} u \\ d' \end{pmatrix}_L ,$$

where d' is the weak Eigenstate associated with the mass Eigenstate through the CKM matrix. Weak and mass Eigenstates are different. The RH chiral state particles and LH chiral state anti-particles are placed in isospin singlets with $I_W = 0$

$$\left(\nu_e\right)_R, \quad \left(e^-\right)_R, \quad \left(u\right)_R, \quad \left(d'\right)_R.$$

The physical $W^\pm$-bosons can be identified as the linear combinations of the SU(2) bosons $W^{(1)}$ and $W^{(2)}$

$$W^\pm = \frac{1}{\sqrt{2}}(W^{(1)} \pm iW^{(2)}).$$

We are left with the $W^{(3)}$-boson but we still have two physical, neutral spin–1 gauge bosons, γ and Z. $W^{(3)}$ can be associated with both of these by introducing an additional neutral spin–1 boson field B. The photon and Z-boson can be written as the linear combinations of B and $W^{(3)}$

$$\begin{aligned} A &= +B\cos\theta_W + W^{(3)}\sin\theta_W \\ Z &= -B\sin\theta_W + W^{(3)}\cos\theta_W \end{aligned}$$

The B field has the underlying symmetry $\mathrm{U}(1)_Y$ which is similar to the U(1) symmetry of QED. Y is the weak hypercharge that the fields couples to which is different from electric charge coupling in QED. It is given by

$$Y = 2Q - 2I_W^3,$$

where Q is the electric charge and I_W^3 is the third component of the weak isospin - the projected component. The neutrinos and the up-like quarks have $I_W^3 = +\frac{1}{2}$ while the leptons and the down-like quarks have $I_W^3 = -\frac{1}{2}$. Thus weak hypercharge relates EM and weak interactions. The mixing angle relates the coupling strengths of the weak interaction and the EM interaction to the $\mathrm{U}(1)_Y$ symmetry by

$$e = g_W\sin\theta_W = g'\cos\theta_W,$$

where e is electric charge of an electron, g_W is the coupling strength in interactions involving $W^\pm$, and g' is the coupling strength of the B-field.

The $\mathrm{SU}(2)_L$ symmetry couples the weak interaction bosons, $W^\pm$ and Z, to LH particles and RH anti-particles only. But the Z-boson is a combination of two different underlying symmetries which means that it also couples to RH particles and LH anti-particles, however unevenly. The left- and right-handed couplings are given by

$$c_L = I_W^3 - Q\sin^2\theta_W \quad , \quad c_R = -Q\sin^2\theta_W.$$

Both couplings depends on electric charge and the mixing angle, but only the left-handed coupling also depends on weak isospin. The couplings reflect the underlying symmetries for the Z-boson. The weak interaction has a V-A structure and the couplings can be re-expressed in terms of vector and axial-vector coupling where $c_V = c_L + c_R$ and $c_A = c_L - c_R$. The electroweak unification works for describing how the Z-boson couples differently from the $W^\pm$-bosons. But it comes at the cost of introducing a new free parameter, and it does not provide any explanation from higher theoretical principles.

2.4. Higgs Mechanism

The Higgs mechanism gives mass to particles and it is key for the success of the SM. Without the Higgs mechanism the SM would be an inconsistent theory. The QFT formalism describing the boson fields only works for massless bosons. Forcefully introducing mass terms into the Lagrangian breaks the gauge invariance and the SM would not be renormalisable according to the proof by t'Hooft [4]. Introducing masses via the Higgs mechanism ensures that the SM remains renormalisable, otherwise it would loose its predictive power. This is not a problem in QED and QCD since the photon and the gluons are massless, however the weak sector bosons do carry mass. In the Salam-Weinberg theory, the simplest Higgs model is derived from

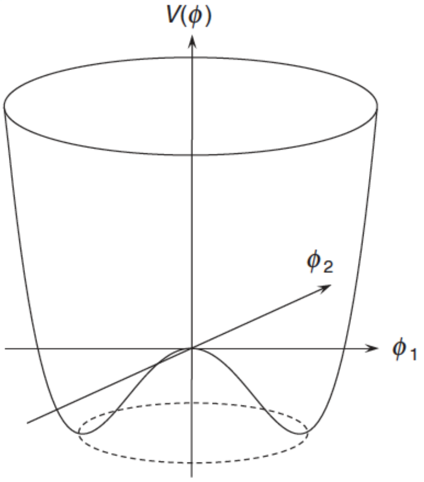


Figure 2.4.: Higgs potential

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.19}$$

which consists of two complex scalar fields placed in a weak isospin doublet. Scalar particles have spin–0. The Lagrangian for complex scalar fields is

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \tag{2.20}$$

where D_μ is the covariant derivative and $V(\phi)$ is the Higgs potential. The covariant derivative is defined by

$$D_\mu = \partial_\mu + i g_W \boldsymbol{T} \cdot \boldsymbol{W}_\mu(x) + i g' \frac{Y}{2} B_\mu \tag{2.21}$$

and it ensures that the Lagrangian for the Higgs field transforms under $U(1)_Y \times SU(2)_L$ corresponding to the symmetry of the electroweak sector in the SM. The Higgs potential is defined by

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{2.22}$$

and the Higgs potential for one of the complex scalar fields can be seen in Figure 2.4 for the case of $\mu^2 < 0$. The potential has an infinite set of degenerate minima which lie on the ring. They correspond to the vacuum state i.e. the lowest energy state of the field. The vacuum state obtains non-zero minima and gives rise to the relation of the Vacuum Expectation Value (VEV)

$$v = \sqrt{\frac{-\mu^2}{\lambda}}. \tag{2.23}$$

Via spontaneous symmetry breaking, a choice is made for the vacuum state corresponding to the physical Higgs field. In unitary gauge the complex scalar fields can be chosen so that they are entirely real. Expanding the scalar field around the vacuum state, the Higgs doublet can be written as

$$\phi = \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.24}$$

where v is the VEV and $h(x)$ is the physical Higgs field. In unitary gauge, the physical fields correspond to physical particles. The combined Lagrangian for the Higgs doublet and gauge boson fields describes the massive Higgs scalar, the massive and massless gauge bosons, as well as the interactions between the Higgs scalar and the gauge bosons, and Higgs self-interactions. The gauge bosons masses are generated from the $(D_\mu\phi)^\dagger(D^\mu\phi)$ term, and the Higgs mass and self-interactions are generated from the potential. The mass of the Higgs boson scales by the VEV:

$$m_H = \sqrt{2\lambda}v \tag{2.25}$$

as does the mass of the W boson

$$m_W = \frac{1}{2}g_W v, \tag{2.26}$$

and the mass of the Z boson

$$m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2} = \frac{1}{2}\frac{gWv}{\cos\theta_W}, \tag{2.27}$$

as well as the fermion masses

$$m_f = \frac{1}{\sqrt{2}}g_f v. \tag{2.28}$$

The fermions acquire mass through Yukawa coupling, g_f, where the associated coupling strength is scaled by the VEV. The Higgs mechanism does not predict this coupling. For all Dirac fermions, gauge invariant mass terms are constructed from Lagrangians [4]

$$\mathcal{L} = -g_f[\bar{L}\phi R + (\bar{L}\phi R)^\dagger] \quad \text{or} \quad \mathcal{L} = g_f[\bar{L}\phi_c R + (\bar{L}\phi_c R)^\dagger], \tag{2.29}$$

where the conjugate doublet is constructed from $\phi_c = -i\sigma_2\phi*$, L is the doublet for left-handed fermion fields with $\bar{L} = L^\dagger\gamma^0$, and R is the singlet for right-handed fermion fields. The Lagrangians above also give rise to interactions between the fermion and the Higgs which leads to both bosonic and fermionic decay modes of the Higgs boson. Since the Higgs vacuum state obtains a non-zero value, the particles of the SM are always trapped in the Higgs field. As particles travel through the Higgs field, they will be slowed down as they interact with the field. This translates to acquiring an “intrinsic” mass, where the mass is proportional to the strength of the interaction. All while preserving local gauge invariance of the SM. The discovery of the Higgs boson gives experimental evidence for the Higgs field as its existence is implied by the observation of the Higgs boson. In 2012 a boson consistent with the Higgs boson of mass near 125 GeV was discovered [1, 2]. Further studies in Run-1 and Run-2 of The Large Hadron Collider (LHC) have been conducted [8–10] in order to shed light into its role in electroweak symmetry breaking and the mechanism of mass generation. Fully investigating all of its properties and confirming theoretical predictions requires still further studies as the success of the SM relies strongly on the interplay between theory and experiment.

3. Anomalous couplings

At every given length scale there is an appropriate Hamiltonian to describe the physics at said length scale. It should include all the possible operators, however, the operators of high dimension are irrelevant so most of the operators of the Hamiltonian will be irrelevant. The relevant operators are consistent with the degrees of freedom i.e. the fields, and the symmetries, and they have the lowest dimension. The SM is the Hamiltonian constructed from the symmetries and degrees of freedom that are relevant at the length scales probed by modern colliders. In Section 2.1, the Hamiltonian was presented in the Lagrangian formalism. The degrees of freedom are the fields of the photon, the gluons, the W-bosons, the Z-boson, the Higgs, and the fermions. This is an exhaustive list of the fields. The symmetries are $SU(3)_C \times SU(2)_L \times U(1)_Y$, and the most general Hamiltonian is then constructed from fields compatible with the symmetries to the lowest possible dimension. The lowest possible dimension is dimension 4 (from now on denoted dim-4) [11]. The SM is a consistent theory but it is incomplete. It should be considered as a low energy Effective Field Theory (EFT): Standard Model Effective Field Theory (SMEFT).

As limits of the energy frontier of particle physics are being pushed, the set of relevant operators will change. In EFT, the Lagrangian includes terms for all relevant and irrelevant operators and therefore includes operators beyond dim-4. It is given by an infinite series of terms of increasing operator dimension

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{SM} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} O_i^{(6)} + \sum_j \frac{c_j^{(8)}}{\Lambda^4} O_j^{(8)} + \ldots, \tag{3.1}$$

where $O_i^{(D)}$ are the operators with their dimension denoted by the superscript, c_i is an associated coefficient, and Λ is the characteristic length scale which is not known. It ensures that the coefficients are dimensionless. Only operators that are compatible with the SM symmetries are included. It can be shown that the operators with dim-5 cannot conserve both baryon and lepton number therefor the lowest dimension operators with $D > 4$ are the dim-6 operators [12]. Assuming baryon conservation there are 59 independent operators for dim-6 terms in the SM Lagrangian [13]. The higher dimension operators are suppressed by a factor $\frac{1}{\Lambda^{d-4}}$ where d is the dimension of the operator. The values of the new higher dimensional operators/interactions will reveal the "selection rules" of the SM intimately linked to new structure/symmetries. In this thesis only operators of dimension 6 relevant for top anomalous couplings will be investigated as higher dimension operators have larger suppression factors. There are 7 independent operators that contribute to the Wtb, $Zt\bar{t}$, and $\gamma t\bar{t}$ vertices and in the notation of [14] they are

$$\begin{aligned}
O_{\phi q}^{(3)} &= i(\phi^\dagger \tau^I D_\mu \phi)(\bar{q}_L \gamma^\mu \tau^I q_L), \\
O_{\phi q}^{(1)} &= i(\phi^\dagger D_\mu \phi)(\bar{q}_L \gamma^\mu q_L), \\
O_{\phi\phi} &= i(\tilde{\phi}^\dagger D_\mu \phi)(\bar{t}_R \gamma^\mu b_R), \\
O_{\phi t} &= i(\phi^\dagger D_\mu \phi)(\bar{t}_R \gamma^\mu t_R), \\
O_{tW} &= (\bar{q}_L \sigma^{\mu\nu} \tau^I t_R)\tilde{\phi} W_{\mu\nu}^I, \\
O_{bW} &= (\bar{q}_L \sigma^{\mu\nu} \tau^I b_R)\phi W_{\mu\nu}^I, \\
O_{tB\phi} &= (\bar{q}_L \sigma^{\mu\nu} t_R)\tilde{\phi} B_{\mu\nu}
\end{aligned} \tag{3.2}$$

with the covariant derivative

$$D_\mu = \partial_\mu + ig_s \frac{\lambda^a}{2} G_\mu^a + ig \frac{\tau^I}{2} W_\mu^I + ig' Y B_\mu \tag{3.3}$$

where G_μ^a, W_μ^I and B_μ are the gauge fields for $SU(3)$, $SU(2)_L$ and $U(1)_Y$, λ^a are the Gell-Mann matrices with $a = 1, \ldots 8$, τ^I are the Pauli matrices for $I = 1, 2, 3$, and Y is the hypercharge of the field. ϕ is the SM Higgs doublet with $\tilde{\phi} = \epsilon\phi^*, \epsilon = i\tau^2$. The quark weak interaction Eigenstates are the doublet and singlets

$$g_L = \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \quad t_R, \quad b_R. \tag{3.4}$$

The Lagrangian for the effective Wtb vertex including SM contributions is [14]

$$\begin{aligned}
\mathcal{L}_{Wtb} = &- \frac{g}{\sqrt{2}} \bar{b} \gamma^\mu (V_L P_L + V_R P_R) t W_\mu^- \\
&- \frac{g}{\sqrt{2}} \bar{b} \frac{i\sigma^{\mu\nu} q\nu}{M_W} (g_L P_L + g_R P_R) t W_\mu^- + H.c.
\end{aligned} \tag{3.5}$$

where $q = p_i - p_j$ is the outgoing boson momentum. P_R and P_L are the right- and left-handed chiral projection operators

$$P_R = \frac{1}{2}(1 + \gamma^5) \text{ and } P_L = \frac{1}{2}(1 - \gamma^5). \tag{3.6}$$

and V_L, V_R, g_L and g_R are the couplings. In the SM only the V_L remains while the rest of the couplings vanish at tree level. The dim-6 operators in Eq. 3.2 present new physics on the Wtb vertex where contributions to the couplings are related to the operator coefficients by

$$\begin{aligned}
\delta V_L &= C_{\phi q}^{(3)*} \frac{v^2}{\Lambda^2}, & \delta g_L &= \sqrt{2} C_{bW}^* \frac{v^2}{\Lambda^2}, \\
\delta V_R &= \frac{1}{2} C_{\phi\phi} \frac{v^2}{\Lambda^2}, & \delta g_R &= \sqrt{2} C_{tW} \frac{v^2}{\Lambda^2}.
\end{aligned} \tag{3.7}$$

The Lagrangian for the effective Ztt vertex including SM contributions is [14]

$$\begin{aligned}
\mathcal{L}_{Ztt} = &- \frac{g}{2c_W} \bar{t} \gamma^\mu (X_{tt}^L P_L + X_{tt}^R P_R - 2s_W^2 Q_t) t Z_\mu \\
&- \frac{g}{2c_W} \bar{t} \frac{i\sigma^{\mu\nu} q\nu}{M_Z} (d_V^Z + i d_A^Z \gamma_5) t Z_\mu
\end{aligned} \tag{3.8}$$

where s_W and c_W are the sine and cosine of the weak angle θ_W, respectively, $Q_t = 2/3$ is the top quark electric charge, X^L_{tt} and X^R_{tt} are couplings in terms of the chiral parts, and d^Z_V and d^Z_A are couplings parameterised in terms of the vector and axial parts. Within the SM, only X^L_{tt} is non-zero at the tree level. The new physics contributions from the dim-6 operators in Eq. 3.2 are

$$\begin{aligned} \delta X^L_{tt} &= \mathrm{Re}\left[C^{(3)}_{\phi q} - C^{(1)}_{\phi q}\right]\frac{v^2}{\Lambda^2}, & \delta d^Z_V &= \sqrt{2}\,\mathrm{Re}\left[c_W C_{tW} - s_W C_{tB\phi}\right]\frac{v^2}{\Lambda^2}, \\ \delta X^R_{tt} &= -\,\mathrm{Re}\, C_{\phi t}\frac{v^2}{\Lambda^2}, & \delta d^Z_A &= \sqrt{2}\,\mathrm{Im}\left[c_W C_{tW} - s_W C_{tB\phi}\right]\frac{v^2}{\Lambda^2}. \end{aligned} \tag{3.9}$$

Lastly, the Lagrangian for the effective γtt vertex including SM contributions is [14]

$$\mathcal{L}_{\gamma tt} = -\, eQ_t\bar{t}\gamma^\mu A_\mu - e\bar{t}\frac{i\sigma^{\mu\nu}q\nu}{m_t}(d^\gamma_V + id^\gamma_A\gamma_5)tA_\mu \tag{3.10}$$

where the dim-6 operators in Eq. 3.2 contribute to the couplings d^γ_V and d^γ_V by

$$\begin{aligned} \delta d^\gamma_V &= \frac{\sqrt{2}}{e}\,\mathrm{Re}\left[c_W C_{tB\phi} + s_W C_{tW}\right]\frac{vm_t}{\Lambda^2}, \\ \delta d^\gamma_A &= \frac{\sqrt{2}}{e}\,\mathrm{Im}\left[c_W C_{tB\phi} + s_W C_{tW}\right]\frac{vm_t}{\Lambda^2}. \end{aligned} \tag{3.11}$$

There are 10 anomalous couplings which depend on 7 operator coefficients relevant for lepton collider physics.[1] More details can be found in [14]. The couplings can be inferred from the observed cross section. Physical observables, such as the cross section, depend on the matrix element squared. Including the higher order operators modifies the matrix element such that

$$\mathcal{M} = \mathcal{M}_{SM} + \mathcal{M}_{D=6}. \tag{3.12}$$

Since the Lagrangian for the sum of dim-6 operators is proportional to the operator coefficients, they can be factored out. The matrix element squared has the following dependence on a single coupling parameter

$$\begin{aligned} |\mathcal{M}|^2 &= |\mathcal{M}_{SM} + c_i\mathcal{M}'_{D=6}|^2 \\ &= |\mathcal{M}_{SM}|^2 + c_i(\mathcal{M}^*_{SM}\mathcal{M}'_{D=6} + \mathcal{M}_{SM}\mathcal{M}'^*_{D=6}) + c_i^2|\mathcal{M}'_{D=6}|^2. \end{aligned} \tag{3.13}$$

hence the cross section is expected to have a quadratic polynomial dependence on a single non-zero coupling parameter. For multiple non-zero couplings the matrix element will have a collinear dependence on the couplings, and the picture will be less simple. The observation of top anomalous couplings is a fast way to find beyond the Standard Model (BSM) physics.

---

[1] In hadron collider physics there is an additional, independent operator coefficient and two anomalous couplings for the gtt vertex.

4. FCC-ee Experiment

Particle physics is at an important point in its history. Following the discovery of the Higgs boson in 2012 at the LHC completing the SM, there has been an increased interest to return to e^+e^- collider physics in order to study its properties at a dedicated Higgs factory and verify whether it is consistent with the SM. FCC-ee is one of the proposed future e^+e^- colliders that complies with the ESPP 2013 guidelines aiming to "study the properties of the Higgs boson and other particles with unprecedented precision and whose energy can be upgraded". It is a large and ambitious project with a timeline spanning over multiple decades. The FCC-ee experiment is the first of two stages for the FCC project where the focus is on the precision frontier. The second stage seeks to push the limits on the energy frontier with the FCC-hh experiment. The two-stage concept bears resemblance with the successful LEP-LHC programme. The Conceptual Design Report (CDR) [15] for FCC-ee covers the physics opportunities, technical designs for the accelerator, detector concepts, civil engineering, cost and schedule. Figure 4.1 summarises the roadmap of the project which has already been underway since the Higgs discovery. The FCC programme is proposed as the next flagship project at CERN to be ready for the LHC shut down. With this ambitious project, CERN stands to maintain its leading role at the forefront of particle physics. The project is for obvious reasons driven by the physics and spans a 70 year time scale from feasibility studies to the end of operation and will therefore cover most of the 21st century. Extensive planning and proof of feasibility are crucial. Currently feasibility studies are ongoing in the 8 year long preparatory phase. Following that will be a 10 year construction phase, and 15 year operation time for the FCC-ee stage.

Precision requires luminosity. In Figure 4.2, the expected baseline luminosities as a function of the centre-of-mass energy is shown for four e^+e^- collider projects. The projects

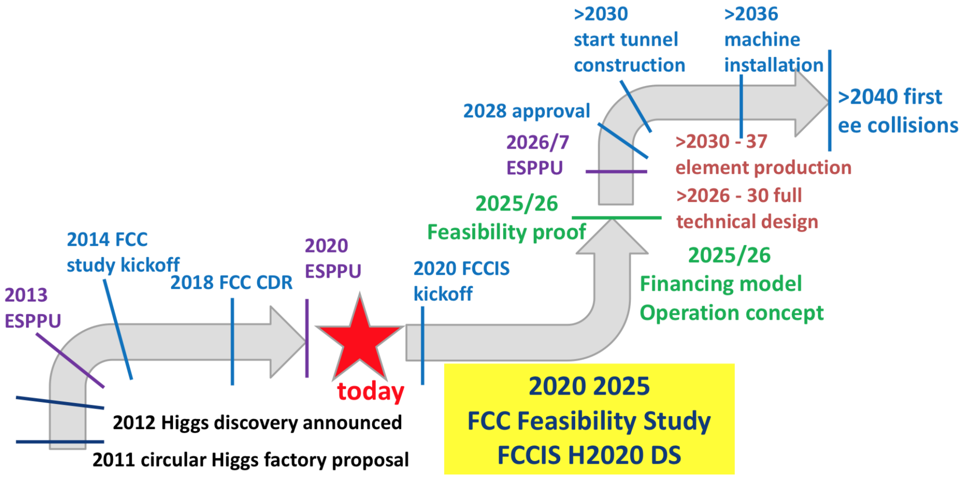


Figure 4.1.: FCC roadmap

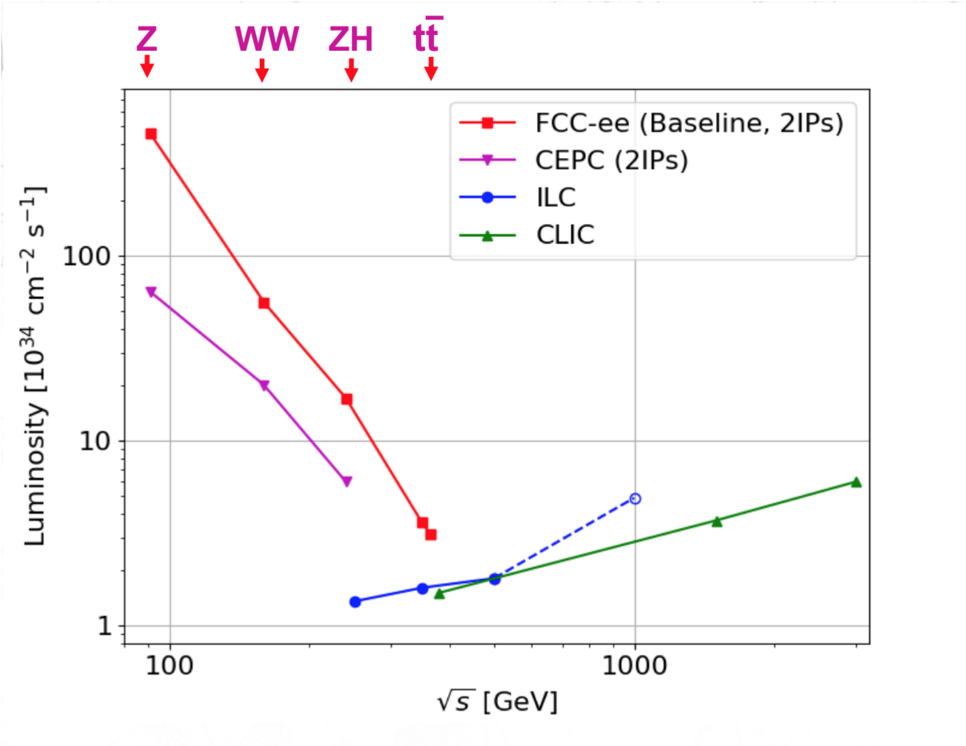


Figure 4.2.: Expected baseline luminosities as a function of the centre-of-mass energy $\sqrt{s}$, at each of the four worldwide e^+e^- collider projects [16].

are the FCC-ee, The Circular Electron-positron Collider (CEPC), The International Linear Collider (ILC) and The Compact Linear Collider (CLIC). FCC-ee and CEPC are both proposed circular collider project with plans to run with e^+e^- collisions at threshold energies between 90 GeV to 350 GeV and 90 GeV to 240 GeV respectively followed by a second stage with proton-proton collisions. ILC and CLIC are linear colliders with plans to run with e^+e^- collisions at threshold energies between 250 GeV to 500 GeV and 380 GeV to 3 TeV respectively. Below 400 GeV, circular colliders trump linear colliders since higher luminosities can be reached, however the energy and therefore the luminosity of circular colliders are limited by synchrotron radiation

$$\Delta E \propto \frac{1}{R}\left(\frac{E}{m}\right)^4 \tag{4.1}$$

which is defined by the ability to keep particles on a circular orbit. Since electrons have a mass that is a factor 2000 lighter than the proton energy, the maximum energy between the two types of experiments behaves very differently. The synchrotron radiation can be compensated by the radius to some extend. Above 400 GeV there is an advantage to using linear colliders for lepton physics since they can reach higher luminosities. The choice of project depends on what physics processes are prioritised in future searches. In an ideal world, the more independent experiments the merrier. Since there are currently no hints of where to look for BSM physics, the aim of the FCC-ee project is to have as broad a scope for research as possible. The FCC-ee is planned to operate at the Z-pole (91 GeV), at the WW threshold (161 GeV), as a Higgs factory (240 GeV), and around the $t\bar{t}$ threshold (340 GeV to 365 GeV). The precision frontier of FCC-ee allows access to much smaller couplings with a focus on electroweak measurements. Up to the $t\bar{t}$ threshold,

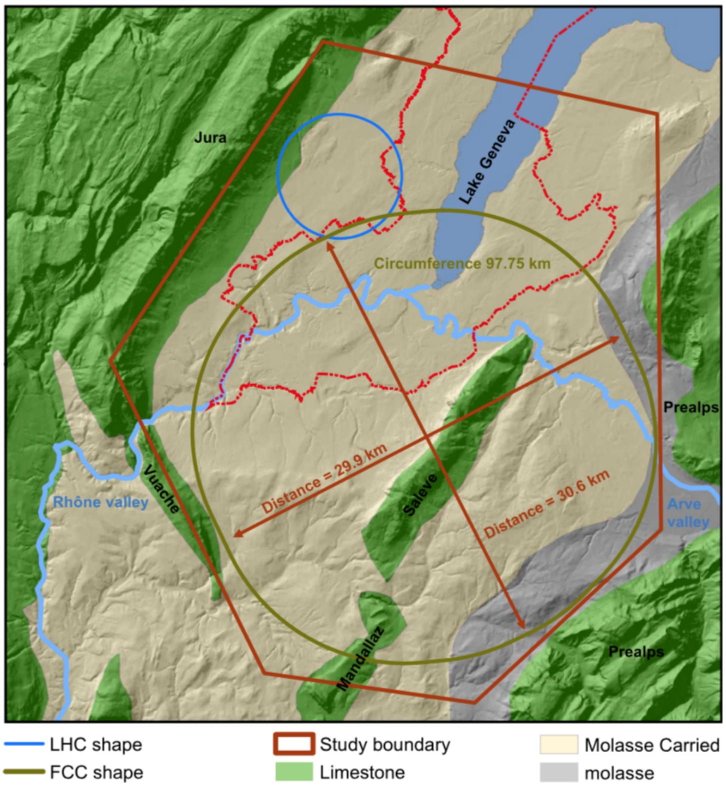


Figure 4.3.: Placement of the collider ring with main topographical and geological structures [15].

the FCC-ee collider provides the highest luminosity at thresholds below 400 GeV where there is a lot of interesting physics to be explored. At the $t\bar{t}$ threshold a scan over the energies around $2m_{\text{top}}$ would be performed followed by 4 years at $\sqrt{s} = 365\,\text{GeV}$ for which the optimum precision on the couplings is reached [17].

The construction of the tunnel is planned to start around 2030. Deciding on the optimal placement of a 100 km collider tunnel requires extensive feasibility studies. The construction has been deemed feasible at this point. The tunnel will be build such that the amount of molasse is maximised while the amount of limestone is minimised. The placement of the tunnel w.r.t. the Large Hadron collider can be seen in Figure 4.3. The double-ring collider will have a circumference of 97.756 km. There will be two Interaction Points (IPs) with the possibility of multiple detector designs. Each collider ring will accelerate bunches of electrons or protons using two Radiofrequency (RF) cavities by producing a “radio frequency” oscillating electric field in the beam direction [18]. At the $t\bar{t}$ threshold there will be 48 bunches in total with an average bunch spacing of 3396 ns and a bunch population of $2.3 \cdot 10^4$. The energy loss from synchrotron radiation at this threshold is the highest with a loss of 9.2 GeV per turn. The energy loss is accounted for in the scaling of the magnets apart from the solenoids which are kept constant at 2 T. The high luminosity reach of the FCC-ee is boosted from a top-up injection scheme where the beams are refilled during operation. Without the top-up injection, the integrated luminosity would be expected at one magnitude lower than shown in the figure for the baseline luminosity.

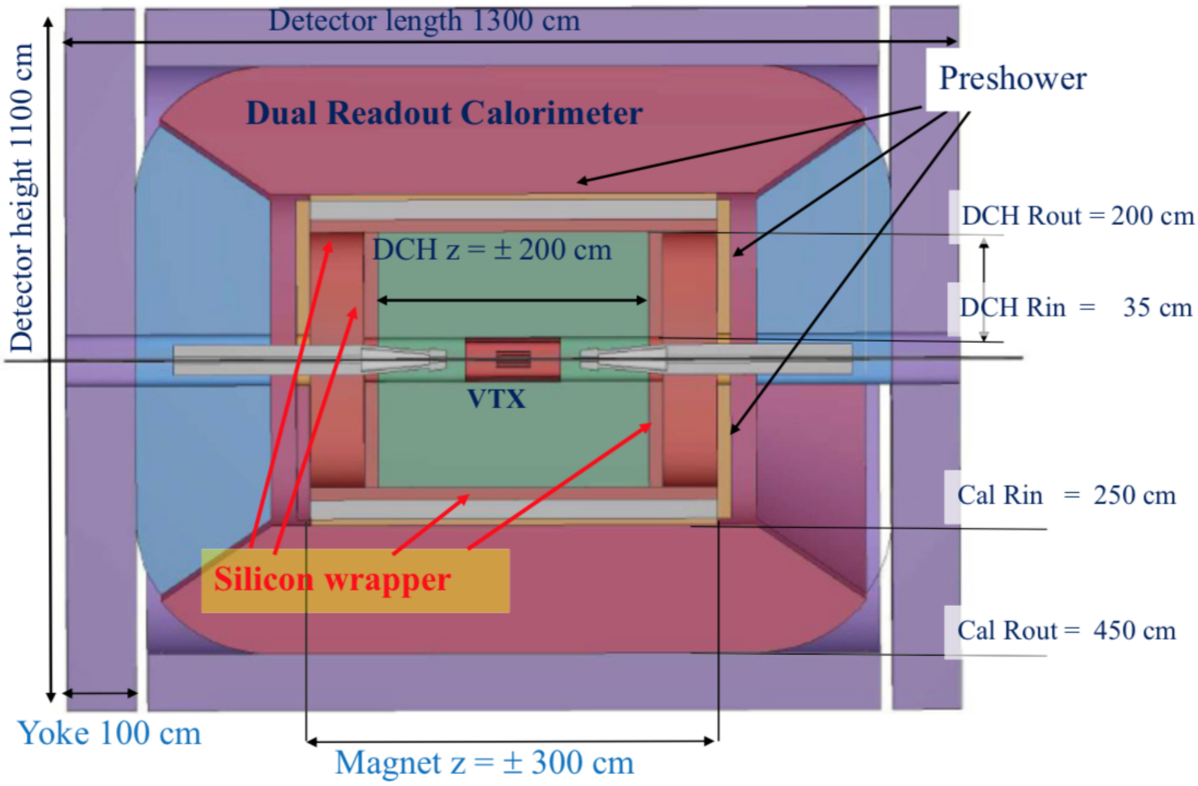


Figure 4.4.: Schematic layout of the IDEA detector [15]

4.1. IDEA Detector

The International Detector for Electron-positron Accelerator (IDEA) is one of the proposed detector concepts and it is developed specifically for the FCC-ee programme. The feasibility of the detector performance has been demonstrated in the CDR [15], albeit the specifics of the design are still being explored for optimisation. The International Detector for Electron-positron Accelerator (IDEA) detector is a multi-purpose detector and complies with the usual layout consisting of a tracking, calorimeter and muon detector system with a solenoid to produce a magnetic field. The innermost layer of the IDEA tracking system is the silicon pixel vertex detector with excellent resolution to provide excellent b- and c-tagging performance. Surrounding it is a large-volume, extremely-light, short-drift wire chamber which covers a cylindrical region from inner radius $R_{\text{in}} = 0.35\,\text{m}$ to outer radius $R_{\text{out}} = 2.0\,\text{m}$ with length $L = 4\,\text{m}$ and consists of a total of 56448 drift cells. The material used is a light gas mixture consisting of 90 % He and 10 % iC_4H_{10} (isobutane) which ensures a high transparency and minimises multiple scattering. It is designed for good tracking, high-precision momentum measurements and excellent particle identification by cluster counting instead of the $\text{d}E/\text{d}x$ technique. The last of the IDEA tracking system is a layer of silicon micro-strip detectors which provides an additional space point and is used for precisely defining the tracker acceptance. The spatial resolution of the tracking system is expected to be better than 100 µm with transverse momentum resolution of $\sigma(1/p_T) \simeq a \oplus b/p_T$ with $a \simeq 3 \cdot 10^{-5}\ \text{GeV}^{-1}$ and $b \simeq 0.6 \cdot 10^{-3}$. The calorimeter system consists of a lead-fibre dual-readout calorimeter. The number of fibres is of order 10^8 and the region is 2 meters deep. It is sensitive to scintillation light and Čerenkov light which can be used to separate electromagnetic showers from hadron showers. Isolated electrons will have a resolution of $10\%\sqrt{E}$ and isolated pions will have a resolution of $30\%\sqrt{E}$. The magnetic field is provided by a low-mass superconducting solenoid coil which sits after the tracking detector in the barrel region. It works as an absorber of $1X_0$. It is followed by a layer of Micro Pattern Gas Detector (MPGD) chambers, another $1X_0$ absorber layer of lead and a second layer of chambers. In the endcap region, the setup is similar with am $1X_0$ absorber layer of lead instead of the solenoid coil. The MGPD works as a preshower detector which provides tracks to be matched to calorimeters showers in case they are close. The muon system is placed in the outermost layer. It consists of layers of chambers embedded in a magnet return yoke. The yoke protects the beams and limits the impact on the luminosity from the 2 T magnetic field of the solenoid. There are still solutions needed for both the engineering of the solenoid system, and for chamber technology that is more cost-effective. The clean experimental environment of lepton colliders means that the readout can be trigger-less. The clean experimental environment at lepton colliders can be attributed to the fact that leptons are elementary particles that have no (measurable) substructure. There is no underlying event, the initial state is known in all directions since there are no parton distribution functions, and the final state is cleaner. The limiting factor will be data storage and therefore the amount of data needs to be reduced. One option is a traditional hardware trigger system where the minimum information needed is transferred to be analysed in real time by a fast readout algorithm which filters out noise and background hits, however the trigger efficiency would have to be known with a precision of 10^{-5}. Another option is to use software algorithms to perform the event selection on the full detector readout as in LHCb. Currently, the trigger versus “trigger-less” schemes are under investigation.

5. Monte Carlo Samples

This analysis is performed with simulated data that are generated within the Future Circular Collider Software (FCC-SW) [19] The framework is made up of modular components. The important components here are `DELPHES`, `Pythia8`, and EDM4HEP which lies under the `Key4HEP` umbrella. `Key4HEP` is the turnkey software for future colliders and is not limited to be used only for FCC. Alongside the planning of future experiments in particle physics, great efforts are being put into developing new software where projects like `Key4HEP` are gathering software components and tools into just one common framework that can be tailored to the specific experiment at hand such as the FCC, ILC or CLIC. The intentions of centralising software are to have easy access and easy deployment of the many functionalities that exist, to avoid duplicating efforts, and to have it be well-maintained which is vital.

The Monte Carlo samples are generated by calling the `DelphesPythia8_EDM4HEP` command with a `DELPHES` card and a `Pythia` command card. The `Pythia` card specifies the hard process that is generated and `Pythia` subsequently takes care of the fragmentation where parton showers and their hadronisation process is generated. `DELPHES` reconstructs the particles using fast detector simulation within the detector environment specified in the `DELPHES` card. In this analysis the IDEA detector card has been used. The `DELPHES` card is the only component where the experimental conditions specific to the FCC-ee are set. Everything else is generic. This shows how easily `Key4HEP` can be adapted to many different experiments. From the `DELPHES` process an EDM4HEP output file is created.

Process	σ [pb^{-1}]	N_{expected}
$t\bar{t} \to b\bar{b}q\bar{q}\ell\nu_\ell$	0.1933 ± 0.00019	289,950
$\mu^+\mu^-$	0.7942 ± 0.0007	1,191,300
$\tau^+\tau^-$	0.7937 ± 0.00022	1,190,500
$\sum_{q=u,d,s,c} q\bar{q}$	4.143 ± 0.006	6,214,500
$b\bar{b}$	0.7448 ± 0.0007	1,117,200
γZ	3.386 ± 0.0021	5,079,000
W^+W^-	10.72 ± 0.04	16,080,000
ZZ	0.6428 ± 0.0021	964,200
ZH	$0.1173 \pm < 0.0001$	175,950
ZW^+W^-	$(15.91 \pm 0.03) \cdot 10^{-3}$	23,865
ZZZ	$(0.7633 \pm 0.0010) \cdot 10^{-3}$	1,145
Single top	$(3.337 \pm 0.003) \cdot 10^{-3}$	5,006

Table 5.1.: Expected statistics at $\sqrt{s} = 365\,\text{GeV}$ with $\mathcal{L} = 1.5\,\text{ab}^{-1}$

EDM4HEP is a generic event data model for future HEP collider experiments wherein the structure of the output file is given. In Table 5.1 the signal and all the considered background processes are summarised with their respective cross sections and number of expected events given a total luminosity of $1.5\,\mathrm{ab}^{-1}$ during a four year runtime. At the time of writing, the expected luminosity is quoted to be $0.34\,\mathrm{ab}^{-1}$ per year in the CDR [15] which over a run time of 4 years adds up to $1.36\,\mathrm{ab}^{-1}$. However a luminosity target of $1.5\,\mathrm{ab}^{-1}$ at $\sqrt{s} = 365\,\mathrm{GeV}$ is also quoted. With the first e^+e^- collisions to only start around 2040, it could be assumed that there is ample time for technical advancements in order to reach the desired goal on the total luminosity. The target luminosity was used to find the expected number of events, but in the analysis the expected target luminosity was used reducing the number of events slightly. It should be noted that the luminosity is quoted for 2 IPs. The goal of this project is to analyse the reach of the FCC-ee, not the reach of the IDEA detector. Hence the use of IDEA detector should be considered as a detector example. The number of expected events can be calculated from

$$N_{\text{expected}} = \mathcal{L} \cdot \sigma \tag{5.1}$$

where $\mathcal{L}$ is the total luminosity and σ is the cross section. For the event selection the number of simulated events is rescaled to the number of expected events with $\mathcal{L} = 1.36\,\mathrm{ab}^{-1}$.

5.1. Signal

The Monte Carlo $t\bar{t}$ samples used in this analysis are generated within the semileptonic channel only. The top-quark predominantly decays by $t \to bW^+$ hence the top quark pair production mostly results in

$$e^+e^- \to t\bar{t} \to bW^+\bar{b}W^-. \tag{5.2}$$

Each W-boson can decay either hadronically ($W \to q\bar{q}'$) or leptonically ($W \to \ell\nu_{ell}$) where the hadronic decay mode has a branching fraction of 67.5% [4]. Thus $t\bar{t}$ decays semi-leptonically in 44% of the cases where one W-boson decays hadronically and one

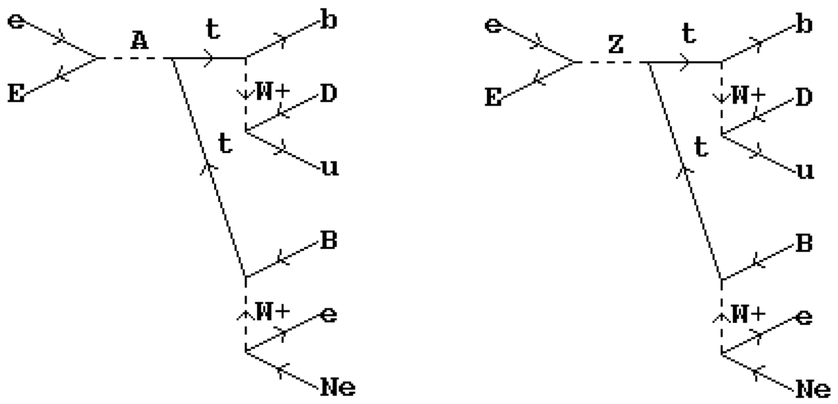


Figure 5.1.: Feynman diagrams for semileptonic top pair production. The diagrams are made with `CompHep`.

W-boson decays leptonically

$$t\bar{t} \to bW^+\bar{b}W^- \to b\bar{b}q\bar{q}'\ell\nu_\ell. \tag{5.3}$$

This leads to a signature in the detector with 1 lepton, some missing energy, and 4 jets where ideally two jets are b-tagged. The signal process is generated for all three lepton flavours separately with $\ell = \{e, \mu, \tau\}$. The processes are generated with `Pythia`. Since the W boson decay can only be set for them separately, the half of the events are generated for $W^+ \to q\bar{q}$ and $W^- \to \ell^- n\bar{u}_\ell$ and the other half for $W^+ \to \ell^+ nu_\ell$ and $W^- \to q\bar{q}$. Lepton universality is assumed. The cross section in the table is calculated for the semileptonic channel only using `MadGraph`.

5.2. Backgrounds

The backgrounds considered in this analysis can all lead to a final state resembling the expected signature in the detector. The list includes all two-fermion processes, the diboson processes for γZ, W^+W^-, ZZ, and ZH, the tri-boson processes and single top. The largest contribution comes from W^+W^-. The hard processes for the last three signatures are generated with `MadGraph`. The Les Houches Event (LHE) files are imported with a `Pythia` command card and `Pythia` takes care of the fragmentation. The cross sections in the table are all calculated using `MadGraph` and cross checked using `CompHep`. The single top production from SM processes such as $e^+e^- \to e^-\bar{\nu}_e t\bar{b}$ have contributions from numerous Feynman processes some of which are top pair production, bottom pair production and WW production. Excluding top and/or bottom quarks as internal particles and exclude the top and/or bottom pair production contribution to the single top production. In the generated single top production only top quarks are excluded as internal particles. Removing some of the Feynman diagrams that contribute to the same final state breaks Lorentz invariance. It also introduces an error on the cross section from missing interference amplitudes. However it is necessary for this analysis in order to separate signal from background. In the worse case scenario it introduces an error estimated at 20 %.

5.3. Anomalous couplings in top pair production

For the analysis, data samples with anomalous coupling in top pair production have been simulated with `Whizard`. Similar to `MadGraph`, `Whizard` produces LHE files that are imported with a `Pythia` command card. `Whizard` takes care of the hard process, `Pythia` takes care of the fragmentation, and `DELPHES` takes care of the detector reconstruction. The anomalous couplings are introduced via Feynman rules that already exists in a predefined model denoted “SM_top_anom”. There are 10 coupling parameters corresponding to the 10 couplings in Eqs. 3.7, 3.9 and 3.11. Since the couplings are linked with there dependence on the 7 operator coefficients, three of the parameters in the model are fixed in order to retain gauge invariance. The choice of for which parameters to keep fixed is based on the work in [20]. The parameters and their corresponding couplings are summarised in Table 5.2.

Parameter	Coupling	Gauge freedom
ta_ttA	δd_A^γ	Free
tv_ttA	δd_V^γ	Free
ta_ttZ	δd_A^Z	Fixed
tv_ttZ	δd_V^Z	Fixed
vl_ttZ	δX_{tt}^L	Fixed
vr_ttZ	δX_{tt}^R	Free
tl_tbW_Re	δg_L	Free
tr_tbW_Re	δg_R	Free
vl_tbW_Re	δV_L	Free
vr_tbW_Re	δV_R	Free

Table 5.2.: Model parameters and the corresponding couplings

6. Jet Performance Studies

Quarks cannot exist as free particles due to colour confinement. Instead they will go through the process of hadronisation. As the hadrons travel through the detector they will decay producing a "jet" of particles in each event. The jets create structures that are visually obvious in an event display providing footprints of the original parton configuration before hadronisation. It is important to note that at *no point* do we actually measure the original partons. The jets should be considered as event properties that reflect the energy and direction of the original partons as the hadrons are formed roughly collinear with the original partons [21]. Translating the qualitative footprints of the jets into quantitative event properties requires a jet definition that maps the set of final state particles on to a set of jets

$$\{p_i\} \rightarrow \{j_k\}. \tag{6.1}$$

In the words of M.H. Seymour, "What You See Is What You Get" [22]. Jet definitions are not unique and there are many possible ways of defining jets. The most suitable choice is specific to the analysis at hand. In this section the performance of different jet definitions is studied.

6.1. Jet Definition

The history of jets in collider physics dates back to the 1980's where cone algorithms were first used at hadron colliders such as the Tevatron [23]. The structures of the hadrons resemble cones around an axis so the algorithm is simple to define as it looks for the direction that maximises the energy flowing through a cone drawn around it. However problems arise when cones start to overlap and the same particle is found in multiple jets which means that the jet definition is collinear unsafe in addition to being infrared unsafe. The later is caused as the algorithm tries to maximise its energy. It will pull in as much neighbouring energy as possible - including soft particles. In perturbative calculations this will lead to soft divergence and in experiments it creates a bias on triggers as well as background noise [22]. The differences in cone algorithms arise from how they handle these problems. Although simple to define they are complicated in practice.

In the 1990's The Large Electron-Positron Collider (LEP) brought on the golden era for jets where sequential recombination algorithms were introduced. These schemes are more complicated to define but simple in practice. A sequential recombination algorithm finds the pair of particles that are closest in some distance measure specific to the algorithm. Then the particle pair is recombined into a pseudoparticle, and the procedure is iterated until some stopping criteria is reached. The jet definition is simple in practice because it solves the problem of overlapping cones hence the algorithms are trivially collinear and infrared safe [23]. Only neighbouring particles are merged if they are close enough to do so. Traditionally, invariant mass was used for the distance measure e.g. the scaled pair mass is defined as

$$y_{ij} = M_{ij}^2 / E_{\mathrm{CM}}^2. \tag{6.2}$$

The issue with using the above measure is that it is extremely non-local in angle, which means that two particles with sufficiently low energy can be close in phase space and merged regardless of the geometrical distance in angle [22]. The $\mathrm{k}_\perp$ algorithm solves this issue by incorporating the transverse momentum such that only soft particles that are additionally close in angle will be merged. The $\mathrm{k}_\perp$, also known as the Durham algorithm, is still one of the best-known algorithms today.

With the next collider experiment of the LHC came a new and very different collider environment and new jet definitions soon followed e.g the Cambridge/Aachen algorithm and the anti-$\mathrm{k}_\perp$ algorithm. In the anti-$\mathrm{k}_\perp$ algorithm, hard particles are merged first and if no other hard particle is close by, the algorithm will give perfect cones and the evolution of jets has come full circle [23]. The anti-$\mathrm{k}_\perp$ algorithm has the advantage that it is resilient to background which there is plenty of at hadron colliders. The future of particle physics is looking to return to lepton collider physics. Lepton collider experiments are notorious for having clean environments with very little backgrounds hence the common standard for jet definitions is likely to change once again.

In the following sections, only sequential recombination algorithms are studied. The jet definition depends on *which* particles are chosen to be combined into the jet and *how* they are combined into the jet. The choice of jet algorithm takes care of the *which*, where two particles are selected to be combined into a new pseudoparticle. As the jet algorithms are iterative in nature, this choice is made in each iteration. The *how* is taken care of by the choice of recombination scheme. The considered jet algoritms and recombinations schemes are summarized in Table 6.1. As exactly 4 jets are expected, the iteration is stopped once 4 jets have been reonstructed. If there are not enough particles in the event, no jets will be reconstructed.

Table 6.1.: Jet definitions included in performance studies. Possible combinations are formed between jet algorithms and recombination schemes.

Jet Algorithms
$k_\perp$
Durham (e^+e^- $k_\perp$)
e^+e^- Anti-$k_\perp$
e^+e^- Cambridge/Aachen
Valencia
Jade

$\otimes$

Recombination Schemes
E-scheme
$E0$-scheme
p-scheme

The jets are reconstructed using the Jet Clustering Interface in the `FCCAnalyses` software [24]. The interface uses `FastJet` [25] and was written such that it could be used for later stage process adaptive jet clustering. After the MonteCarlo (MC) samples are generated, the interface makes it possible to run jet reconstructions for multiple jet definitions at once, it allows the user to specify the input particles, and it provides access to the jet constituents.

6.2. Jet Algorithms

The jet algorithms are available through `FastJet` which provides native implementations of all widely used sequential recombination jet algorithms for both pp and e^+e^- collisions as well as the possibility to add external algorithms via plugins. The `FastJet` Package is written in the current hadron collider era and therefore targets mainly hadron colliders, but it also supplies algorithms suitable for lepton collider experiments.

Jade For the Jade algorithm, the distance measure is defined as

$$d_{ij} = 2E_iE_j(1 - \cos\theta_{ij}). \tag{6.3}$$

The Jade algorithm is the oldest algorithm used in this analysis. It is an invariant mass type algorithm as it is an extension of the scaled pair mass defined in Eq. 6.2. The issue with invariant mass type algorithms is that it is not guaranteed to handle soft particles correctly.

Durham (e^+e^- $k_\perp$) For the Durham algorithm, the distance is defined as

$$d_{ij} = \min(E_i^2, E_j^2)(1 - \cos\theta_{ij}). \tag{6.4}$$

The Durham algorithm came after the Jade algorithm and the two bear some resemblance. The Durham algorithm is based on the invariant mass type metric but it also incorporates the transverse momentum, such that only particles that are both close in invariant mass phase space and angular space are combined.

$k_\perp$ The $k_\perp$ algorithm in `FastJet` is focused towards hadron colliders. The distance measure is defined as

$$d_{ij} = \min(p_{ti}^2, p_{tj}^2)\frac{\Delta R_{ij}^2}{R^2} \tag{6.5}$$

It is an extension of the Durham algorithm which was the original $k_\perp$ algorithm, but it has been altered towards hadron colliders where the initial state is only known in the plane transverse to the beam axis (z-direction). The algorithm is defined with parameters that are more suitable at hadron colliders with the squared cone radius $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ where $y_i = \frac{1}{2}\ln\frac{E_i+p_{zi}}{E_i-p_{zi}}$ is the rapidity and ϕ_i is the azimuthal angle. R is a parameter known as the jet radius which determines the angular reach of the algorithm.

e^+e^- Cambridge/Aachen For the Cambridge/Aachen algorithm, the distance measure is defined as

$$d_{ij} = \min(E_i, E_j)\frac{1 - \cos\theta_{ij}}{1 - \cos R} \tag{6.6}$$

The algorithm resembles the Durham algorithm except for the power of the energy which is not squared here. Additionally it is normalised with a jet radius parameter. In this analysis $R = 1$ so that the only difference lies in the power of E. The distance measure is written in parameters suitable at lepton colliders.

e^+e^- Anti-$k_\perp$ For the anti-$k_\perp$ algorithm, the distance measure is defined as

$$d_{ij} = \min(E_i^{-2}, E_j^{-2})\frac{1 - \cos\theta_{ij}}{1 - \cos R} \tag{6.7}$$

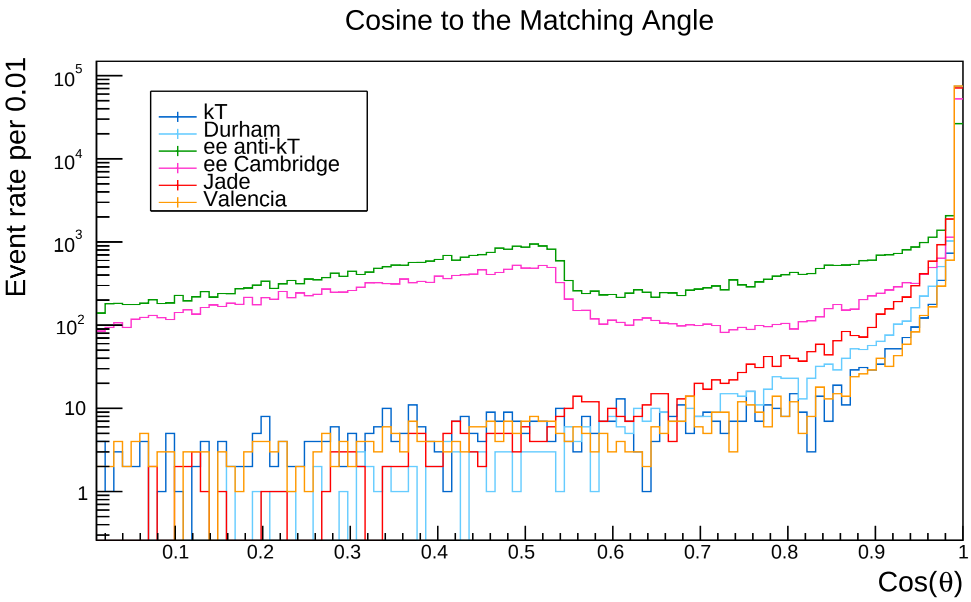


Figure 6.1.: Matching angle between reco jets and particle jets

Here the $k_\perp$-algorithm has been modified such that it depends on the inverse of the energy squared instead. The distance measure is also here written in parameters suitable at lepton colliders. It is more widely used at hadron colliders than lepton colliders. This stems from the fact that the anti-$k_\perp$ algorithm is significantly more resilient to backgrounds than the $k_\perp$ algorithm, which is of high importance at hadron colliders.

Valencia The distance measure for the Valencia algorithm is defined as

$$d_{ij} = \min(E_i^{2\beta}, E_j^{2\beta})(1 - \cos\theta_{ij})/R^2 \tag{6.8}$$

The Valencia algorithm is the newest algorithm used in this analysis. Compared to the Durham algorithm, the power of the energy depends on a parameter β which can tune the dependence of the energy. The algorithm has been shown to be particularly robust at other lepton colliders ILC and CLIC where there is a large background contribution from $\gamma\gamma \to$ hadrons due to the Initial State Radiation (ISR) [26, 27].

A comparative study of the six jet algorithms has been performed to choose a jet algorithm suitable for this analysis of semileptonic top pair produced events. The jets we reconstruct from the final state particles at the true MC level after the hadronisation are called particle jets while the jets reconstructed from final state particles at the detector level (reconstructed by `DELPHES`) are referred to as a reco jet. In Figure 6.1, the cosine to the matching angle between reco jets and particle jets is shown. Each reco jet in an event is matched to the particle jet closest in angle. The vast majority of the reco jets are matched to particle jets with $cos(\theta)$ close to 1. Durham, $k_\perp$, Valencia and Jade algorithms display similar behaviour whereas the anti-$k_\perp$ and Cambridge algorithms have larger fraction of jets with larger matching angles. They both have a bump around 0.5 which might be explained by the definitions of the different algorithms. The Durham algorithm uses a bottom-up approach where it combines pseudoparticles that are close in phase space i.e. both smaller energies and close in angular space. When using exclusive clustering, the recombination stops when it has exactly 4 jets, disregarding of how close they are in phase space at the end.

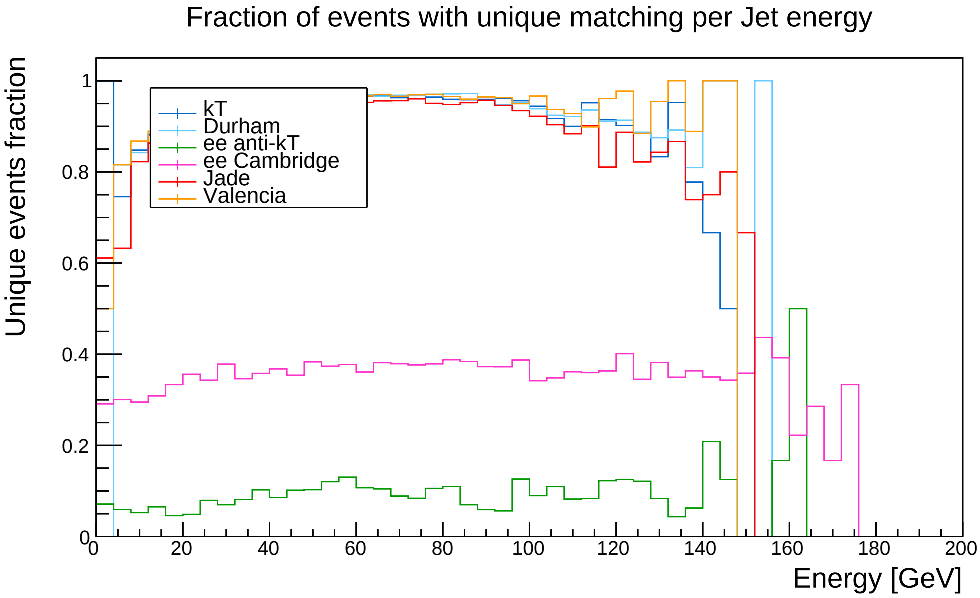


Figure 6.2.: Unique matching between reco jets and particle jets

For the anti-$k_\perp$ algorithm, the priority of combining the pseudoparticles does not have the same physical meaning with merging particles that are close in energy and angle. Instead it combines the hard particles first, and with exclusive clustering no cut-off value is provided for how close those hard particles are allowed to be. So it might combine two pseudo-jets that are far apart in angle, before it has found exactly 4 jets. The peak at 0.5 could indicate that assuming a particle jet is combined somewhat correctly from constituents close in angle, the reco jet might contain two hard particles from opposite directions and the jet will lie in the middle of those and hence be orthogonal to the particle jet. Durham is better at preventing particles to be combined if they are far apart in angle because of how its distance measure is defined, which is very different for the physical meaning of the anti-$k_\perp$ distance measure.

As a cross check for the matching, the number of unique events across reco jet energies has been considered, which can be seen in Figure 6.2. Events are considered uniquely matched if all four reco jets are matched to four different particles jets. If more than one reco jet is matched to the same particle jet, the event is not considered uniquely matched. The figure shows that the matching is unique in the majority of the events across the energies for the Durham, $k_\perp$, Valencia and Jade algorithms, while there are fewer uniquely matched events for the Cambridge and especially the anti-$k_\perp$ algorithms. The Durham algorithm has 89.5 % uniquely matched events while the fraction for $k_\perp$ is a little lower with 81.5 %. The Cambridge algorithm has in total 45.1 % uniquely matched events while anti-$k_\perp$ only has 6.0 % uniquely matched events and they do not indicate a close correspondence between particle level and detector level within the setup of the jet reconstruction.

The reco jet constituents are reconstructed by `DELPHES` from the true final state MC particles, so they can also be associated with the particle jet constituents. To further investigate the correspondence between the reco jets and particle jets, the overlapping of jet contituents has been considered. For each reco jet, the fraction of associated jet constituents that were found in the matched particle jet is plotted in Figures 6.3 and 6.4.

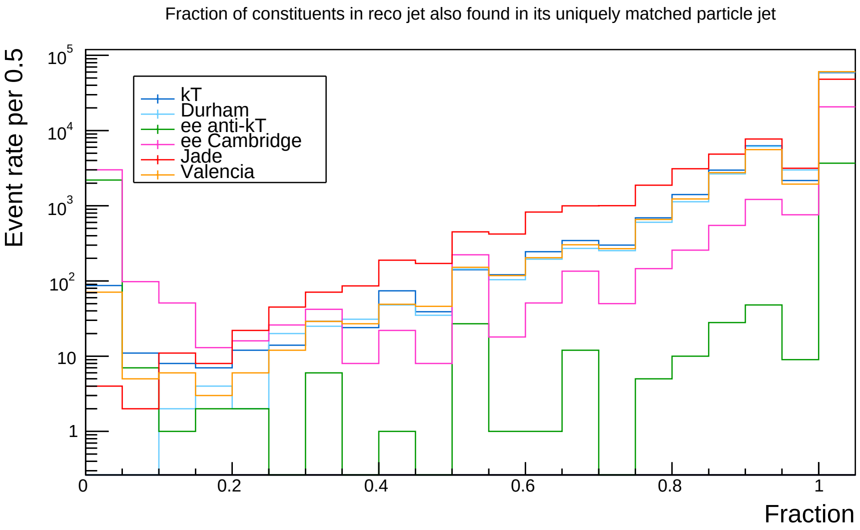


Figure 6.3.: Matching constituents between reco jets and particle jets for events with unique matching

Figure 6.3 shows the fraction of associated constituents between reco jets and particle jets in uniquely matched events and Figure 6.4 shows the fraction in non-uniquely matched events. For most of the reco jets in uniquely matched events, all of the MC final state particles associated to the reco jet constituents are found in the matched particle jet. Since not all final-state particles are reconstructed, there can be more constituents in particle jets than reco jets. The plot also shows that there are jets where none of the associated MC particles are actually found in the matched particle jet. This can be due to detector coverage or lack thereof. If a quark flies through an area not covered by the detector, none of the final-state particles will be reconstructed, but the jet definition still requires 4 jets.

Detector coverage does not explain the amount of jets with 0 % overlap since it occurs more frequently for jets reconstructed with anti-$k_\perp$ and Cambridge, so for some of the events the lack of overlap must be accounted for by reconstruction effects. Again, the Durham, $k_\perp$, Valencia and Jade algorithms display similar behaviour where the number of events increases with an increasing fraction of overlap. The Jade algorithm stands out a bit from the other indicating that it catches a smaller overlap than the three other algorithms. For non-unique events, the distributions looks more flat and there are less jets with 100 % overlap. The unique matching is not a guarantee for a good correspondence between detector level and reco level but it does show a bias towards good correspondence.

With a good correspondence between the reco jets and particle jets, the resolution is found by creating a profile plot of the relative energy difference

$$\frac{E_{\text{reco}} - E_{\text{particle}}}{E_{\text{particle}}} \tag{6.9}$$

For a binned set of jet energies, the profile plot calculates the mean and the standard deviation on the relative energy difference assuming the distribution is Gaussian. The standard deviation corresponds to the resolution which can be seen for all six jet algorithms in Figure 6.5. For Gaussian distributed measurements, the standard deviation of $\hat{\sigma}$ (or the error on the error) is $\sigma/\sqrt{2n}$ for large n [28]. The figure shows that the

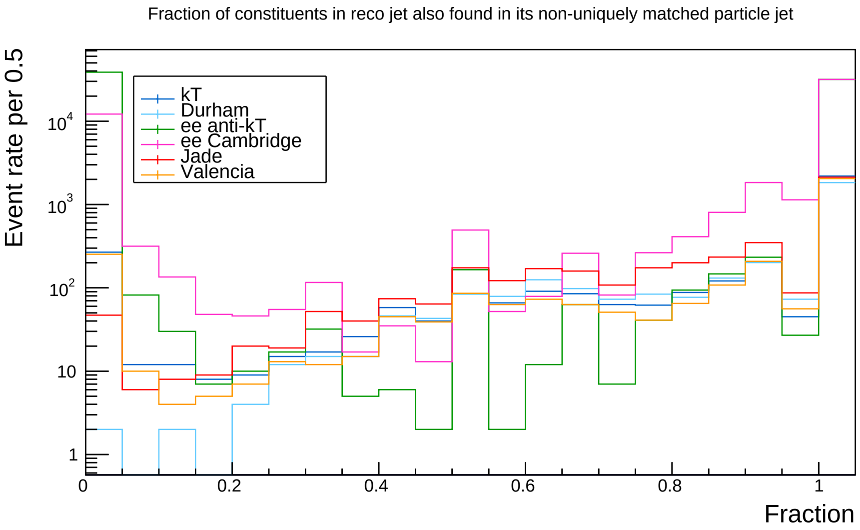


Figure 6.4.: Matching constituents between reco jets and particle jets for events with non-unique matching

Cambridge and anti-$k_\perp$ algorithms have poorer resolutions compared to the four other algorithms and that the resolution worsens with the jet energy. Zooming in on the other four algorithms, their resolutions with the error on the error versus jet energy are shown in Figure 6.6. The four algorithms have very similar resolutions except for a single point at the W mass, where the resolution for the $k_\perp$ algorithm increases.

Two of the jets are b-jets where the jet constituents should be the b-hadron decay products. To investigate how well the jet algorithms separate the b-hadron decay product into just one, all jet constituents parent history has been used to differentiate which constituents belongs to which b-hadrons.

In Figure 6.7 the largest fraction of b-hadrons decay products that ended up in the same jet is shown. The Durham algorithm performs the best at separating all b-hadron decay products into the same jet where full separation is achieved in 79.2 % of the cases. Valencia is second best with 76.9 %, followed by $k_\perp$ with 76.3 % and Jade with 73.8 %. Cambridge achieved full separation 61 % of the times and for anti-$k_\perp$ it was only 12.3 %.

In Figure 6.8, the fraction of jets with full separation across jet energy is shown. The fraction falls off for higher energy jets. For particle jets, the Valencia algorithm performed better at separating the jets than the Durham algorithm with 64.1 % versus 62.8 %. For particle jets there are more low energy particles to apply the jet algorithm to, which complicates the event. These particles are not necessarily reconstructed and available to the reco jets. The performance of jet algorithms depends on the physics at hand and the most suitable choice for a given analysis is not known a priori.

The presented studies showed that while the Valencia, Durham and $k_\perp$ all exhibit good performance, the Durham algorithm is favoured due to its ability for separating the b-hadron decay products into the jets. The Jade algorithm also showed good performance however, the jets are slightly more polluted with wrong constituents. The Cambridge and especially the anti-$k_\perp$ algorithms showed poor performance in comparison. In these studies, the algorithms were set up to have as equal conditions as possible, which put the Cambridge and anti-$k_\perp$ algorithms at a disadvantage. Requiring exclusive clustering on

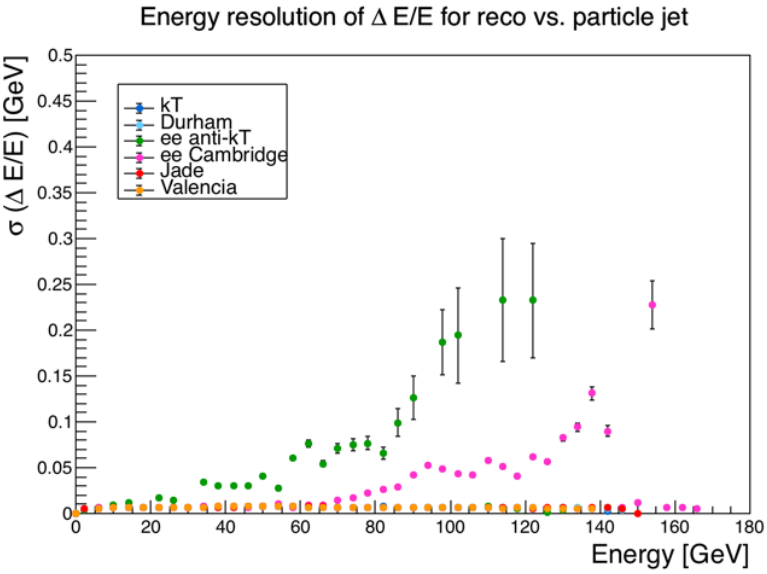


Figure 6.5.: Energy resolution of jets (only events with unique matching) for $k_\perp$, Durham, e^+e^- anti-$k_\perp$, e^+e^- Cambridge/Aachen, Valencia, and Jade.

these algorithms has no physical meaning for the anti-$k_\perp$ since it will merge hard particles first disregarding the phase space. Without a cut-off value on the distance measure, which is provided in inclusive clustering, hard particles that are far apart can still be merged. In this thesis, the Durham algorithm has been used for the jet reconstruction.

6.3. Recombination Schemes

Once the jet algorithm has selected which two partons to combine into a new pseudo jet, the recombination scheme defines *how* to combine them. In the E-scheme, parton i and j are replaced by a pseudojet k with four-momentum

$$\mathbf{p}_k = \mathbf{p}_i + \mathbf{p}_j$$

This scheme is Lorentz invariant and it conserves the energy and momentum. However it also introduces a non-zero mass for the pseudojet k, which cannot consistently be accounted for in the QCD calculations [29]. Alternative recombination schemes such as the $E0$-scheme and the p-scheme account for the invariant mass of the jet such that it remains zero. In the $E0$-scheme, the four-momentum of pseudojet k is rescaled to have zero invariant mass

$$E_k = E_i + E_j \quad , \quad \vec{p}_k = \frac{E_k}{|\vec{p}_i + \vec{p}_j|} \cdot (\vec{p}_i + \vec{p}_j)$$

However this scheme is not Lorentz invariant and only the energy is conserved while the momentum is not. Additionally, it can only be applied in the laboratory frame. In the p-scheme, the four-momentum is constructed to have zero invariant mass.

$$\vec{p}_k = \vec{p}_i + \vec{p}_j \quad , \quad E_k = |\vec{p}_k|$$

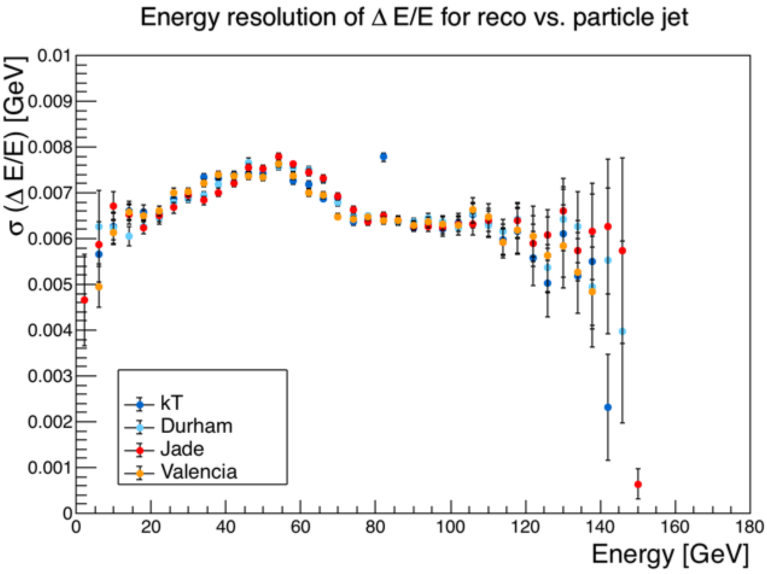


Figure 6.6.: Energy resolution of jets (only events with unique matching) for $k_\perp$, Durham, Valencia, and Jade.

This scheme is also not Lorentz invariant and it only conserves momentum not energy. The total energy sum gradually decreases with each recombination.

In order to compare the three recombination schemes, the same studies as for the jet algorithms have been performed. The Durham algorithm, as the selected algorithm, was used with each of the recombination schemes in the jet reconstruction. The cosine to the matching angle is shown in Figure 6.9,the relative number of unique events across the energies is shown in Figure 6.10, and the resolution is shown in Figure 6.11. All three recombination schemes exhibit similar performance and no single scheme seems to outperform the others. The figures for the matching of the constituents and the separation of the b-hadron decay products can be found in Appendix A, where the performance is comparable for all of them. In this analysis, Lorentz invariance has been prioritised in the jet reconstruction and the E-scheme with the Durham algorithm has been applied.

6.4. Jet Clustering & Jet Tagging in `FCCAnalyses`

As part of this thesis the Jet Clustering Interface and Jet Tagging Utilities were written. `DELPHES` has the option to reconstruct jets by setting up a jet definition in the `DELPHES` card. It is only possible to set up one jet definition at a time and particles cannot be deselected from the clustering. The clustering and tagging tools have been written to be used as a flexible "after burner" following the `DELPHES` generation of the MC samples. It's architecture and design has been laid out to be compatible with `FCCAnalyses`, which is a package that produces flat `ROOT` trees from FCC-SW Event Data Model (EDM) files. In this analysis the FCC-SW EDM files where generated with `DELPHES` and EDM4HEP as mentioned in section 5. The `FCCAnalyses` software uses `ROOT` dataframes to produce the files which allows for an efficient analysis using a `Python` framework. Behind the dataframe, there is a library of `C++` analysers which now also include the Jet Clustering Interface and the Jet Tagging Utilities. They are available as part of the `FCCAnalyses`

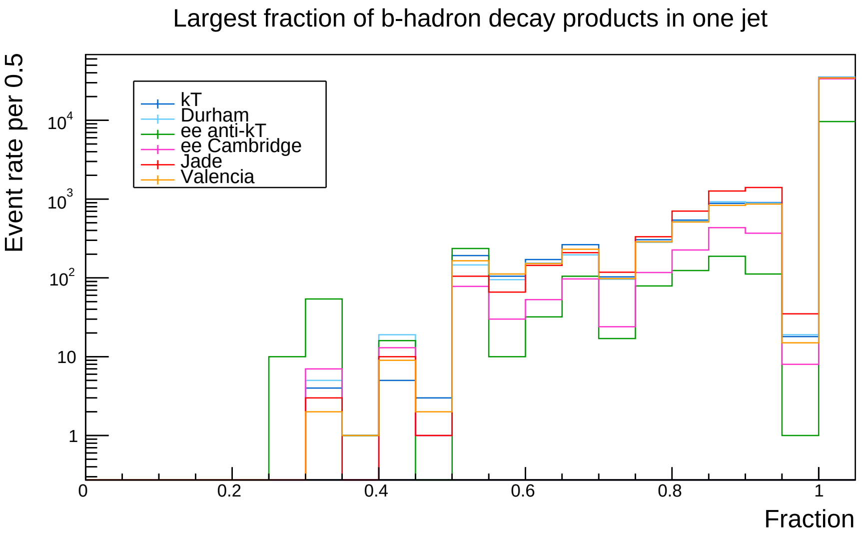


Figure 6.7.: Largest separation of b-hadron decay products in reco jets

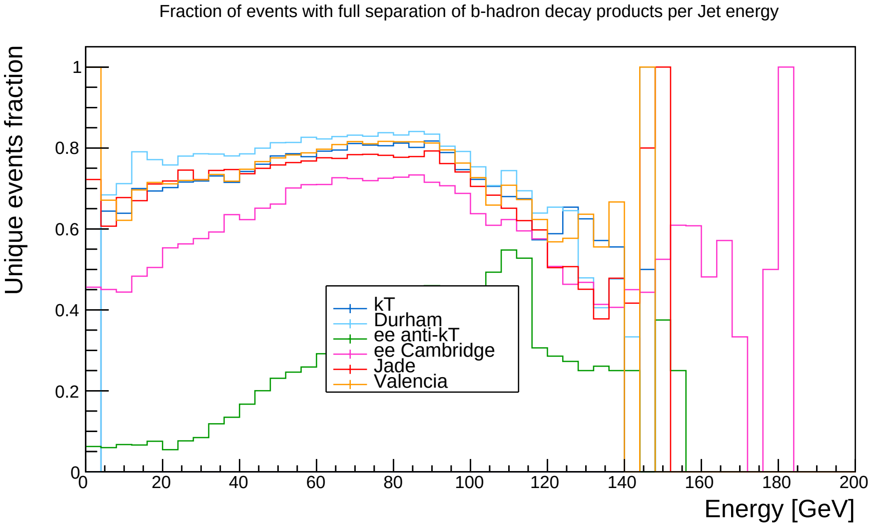


Figure 6.8.: Full separation of b-hadron decay products in reco jets.

software at [24].

The Jet Clustering Interface is an interface to `FastJet` and has been made available for later stage process adaptive jet clustering. Since `FastJet` is written in the "hadron collider era", the native jet algorithms and recombination schemes are also focused towards these detector environments. The interface only contains a select number of the jet algorithms and recombination schemes implemented in `FastJet`, which are summarised in Table 6.2. The Jade and Valencia algorithms are available as plugins which is an option that is made possible through `FastJets` design. The list can be expanded to include further jet algorithms whether or not they are already available in `FastJet`. For the recombination schemes already in `FastJet`, only the E-scheme is suitable at lepton colliders. External recombination schemes for the $E0$-scheme and the p-scheme have been written as well as part of this thesis. The contribution has already reached upstream repositories[1] and the schemes are available as part of `FCCAnalyses`.

---

[1] `https://github.com/HEP-FCC/FCCAnalyses/pull/83`

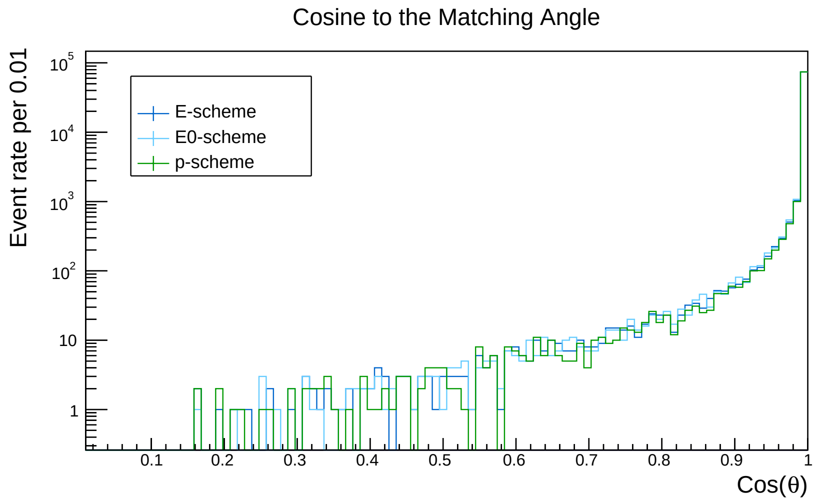


Figure 6.9.: Matching angle between reco jets and particle jets

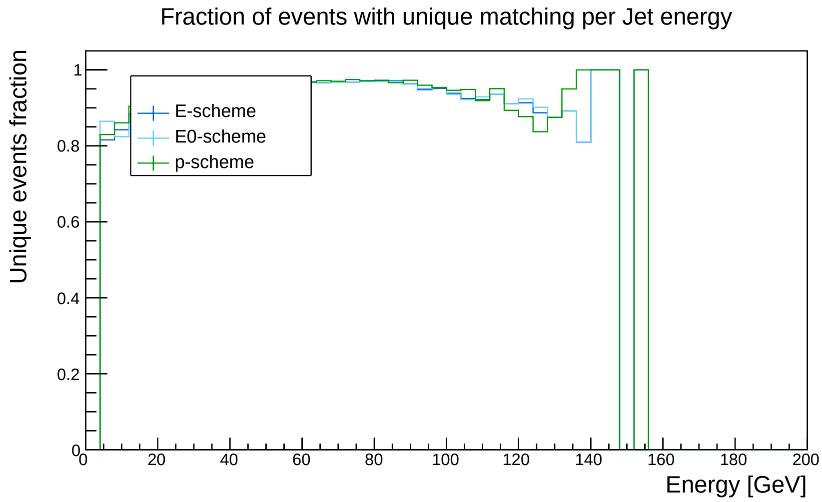


Figure 6.10.: Unique matching between reco jets and particle jets

Listing 6.1: Jet Clustering Interface arguments shown without default values.

```
struct clustering {
  clustering (
    float arg_radius ,
    int   arg_exclusive ,
    float arg_cut ,
    int   arg_sorted ,
    int   arg_recombination
  );
}
```

The interface works by first building pseudo jets from momentum and energy components of a set of particles. This set can be the entire event or a true subset. In this analysis e.g. the highest energy lepton was excluded from the clustering. The jet clustering is run by choosing one of the algorithms with a set of arguments. The arguments are in the order as stated in Listing 6.1 and its possible values are listed below:

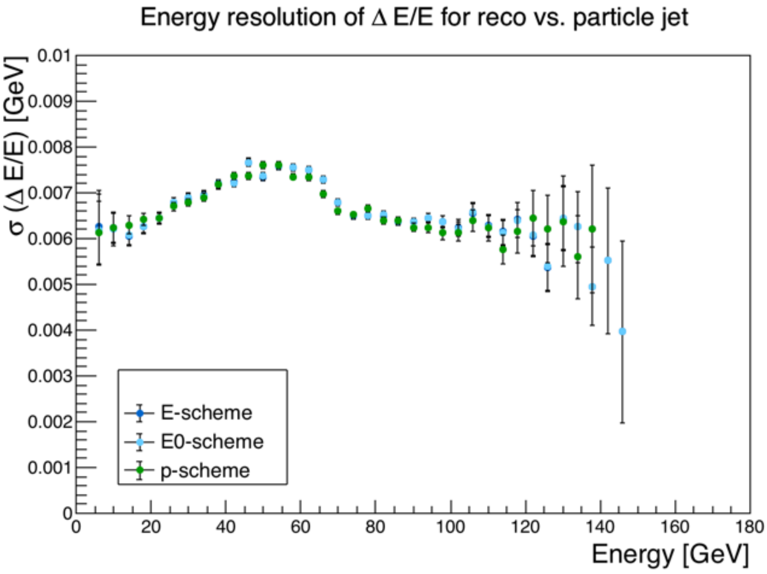


Figure 6.11.: Energy resolution of jets (only events with unique matching) for E-scheme, $E0$-scheme, and p-scheme.

`arg_radius` Jet cone radius

`arg_exclusive` Clustering

0 → inclusive clustering,
1 → exclusive clustering with dcut,
2 → exclusive clustering to exactly njets,
3 → exclusive clustering up to exactly njets,
4 → exclusive clustering with ycut.

`arg_cut` Cut-value depending on clustering

`arg_sorted` Ordering of returned jets

0 → sorted by p_t,
1 → sorted by E.

`arg_recombination` Recombination scheme

0 → E-scheme
1 → p_t-scheme
2 → p_t^2-scheme
3 → E_t-scheme
4 → E_t^2-scheme
5 → Boost-invariant p_t-scheme
6 → Boost-invariant p_t^2-scheme

others Additional input parameters specific to jet algorithm

The jet clustering returns a `struct` that contains a vector of the jets and vector of vectors with the indices of the constituents for each jet. Accessing the jet constituent indices like

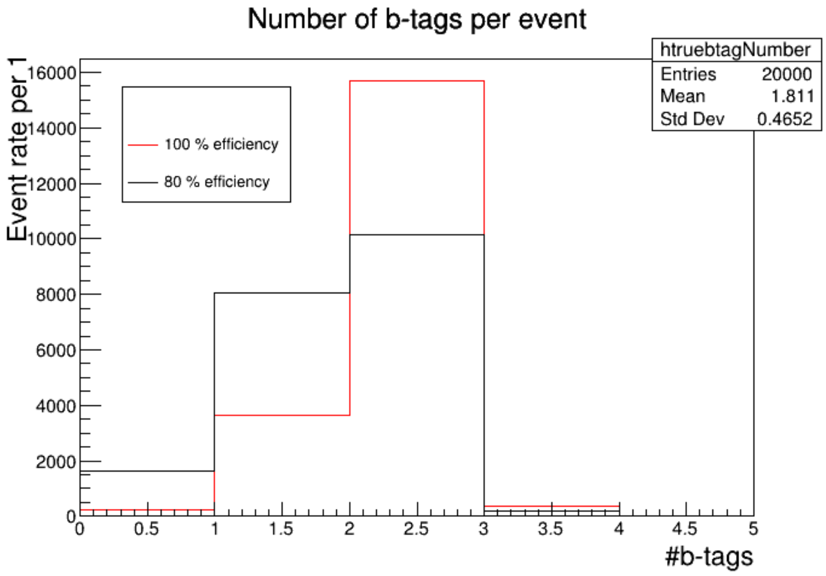


Figure 6.12.: Number of b-tags with 80 % efficiency (black) and 100 % efficiency (red).

this is also different to `DELPHES`.

Table 6.2.: Jet definitions included in performance studies

Jet Algorithms
$k_\perp$
Anti-$k_\perp$
Cambridge/Aachen
Generalised $k_\perp$
Durham (e^+e^- $k_\perp$)
Generalised $k_\perp$ for e^+e^-
Valencia
Jade

$\otimes$

Recombination Schemes
E-scheme
p_t-scheme
p_t^2-scheme
E_t-scheme
E_t^2-scheme
Boost-invariant p_t-scheme
Boost-invariant p_t^2-scheme
$E0$-scheme
p-scheme

The Jet Tagging Utilities written for and used in this thesis, are intended to be used as an extension of the Jet Clustering Interface. The method is in line with tools used by `DELPHES` where a jet is matched to the true MC partons if it fulfils the requirement $\Delta R(\text{jet,parton}) < 0.5$ i.e. the parton and jet lie within a cone radius of 0.5. In `FCCAnalyses` the angular distance is calculated in terms of the vector angle instead, such that a jet and parton are matched if $\alpha(\text{jet,parton}) < 0.3$.

The hierarchy of the parton flavours is the following. If a jet is matched to a b-parton, the jet is b-tagged. If a jet is matched to a c-parton but no b-partons, the jet is c-tagged. And finally if a jet is matched to neither b- nor c-partons, it is considered a light flavour jet ($\{u, d, s\}$). Realistically, we would not have access to the true MC partons. and hence the Jet Tagging Utilities perform too well in comparison to what could be achieved with real data. To account for this discrepancy, an efficiency is applied for each jet flavour separately. In the `DELPHES` card for the IDEA detector, the efficiency for b-tagging is 80 %

which will be used in the following analysis. The number of b-tags with 100 % efficiency and 80 % efficiency is shown in Figure 6.12. In an ideal scenario with 100 % efficiency we could most frequently see events with two b-tagged jets. However some events only have one b-tagged jet, which is most likely due to the second b-jet not being reconstructed properly. If the b-jet flies into a direction not covered by the detector area, it is simply not possible to reconstruct it.

7. Event Selection

The number of events for the backgrounds considered in Section 5 outweighs the number of events for the signal. The signal can be recovered from the background by determining a signal region that optimises the signal to background ratio. In the first section, the identification of the objects is discussed. In the second section, the signal region is determined from the event selection.

7.1. Object Identification

From the signature of the semi-leptonic channel for top pair production, there are 6 reconstructed objects expected to be found in the detector. The objects corresponds to one lepton, one neutrino, and four quarks where two of them are b-quarks. This analysis searches for leptons of the flavour $\ell = \{e, \mu\}$ assuming perfect Particle IDentification (PID). In collider experiments, electrons and muons give clear signals and are easily identified. Tau-leptons are much heavier compared to electrons and muons, and since its lifetime is proportional to its inverse mass to the fifth power ($\tau_\tau \propto 1/m_\tau^5$), the tau-lepton will decay much faster. It decays within the detector and can be identified from its decay products. A dedicated search for τ-leptons has been omitted to be revisited in future works. The main tau-lepton decay modes are $\tau^- \to e^- \bar{\nu}_e \nu_\tau$ (17.8 %), $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$ (17.4 %), $\tau^- \to \pi^-(n\pi^0)\nu_\tau$ (48 %) and $\tau^- \to \pi^-\pi^+\pi^-(n\pi^0)\nu_\tau$ (15 %) [4]. Hence $\sim 35\,\%$ of the tau-leptons do decay into a lepton flavour searched for in this analysis which increases the number of signal events. In an event there can be more than one observed lepton from the hadronisation process. The signature lepton from the $W \to \ell\nu$ decay process is assumed to carry the highest energy since the signature lepton is created at an earlier stage in the decay chain. The assumption can be used as a selector which has an acceptance of $(96.6 \pm 0.7)\,\%$. This number was found by using the parent history of the true MC particle associated to the reconstructed lepton with the highest energy. The lepton is accepted as a signature lepton when a particle in the parent history has PDG and status code that matches an intermediate W boson in the hardest sub-process. For 0.02 % of the events in the test sample, the highest energy lepton is matched to a W-boson of opposite charge which indicates that the W-boson has decayed hadronically and the hadronic d or s decay products have decayed to a lepton. The remaining highest energy leptons are all accounted for as they originate from a b-quark in the hardest sub–process. The objects corresponding to the quarks are the jets. In this analysis, the jets are reconstructed with the interface to FastJet discussed in the previous section. Exactly 4 jets are reconstructed using exclusive clustering. The highest energy lepton has been excluded from the clustering. The jet definition consists of the Durham algorithm with the E-scheme. b-jets are identified using the jet tagging tools also discussed in the previous section. The efficiency has been set to 80 % which is the efficiency set in the Delphes card for the IDEA detector. The object corresponding to the neutrino is the missing energy. The cut-flow assumes a 100 % trigger efficiency.

7.2. Signal Selection

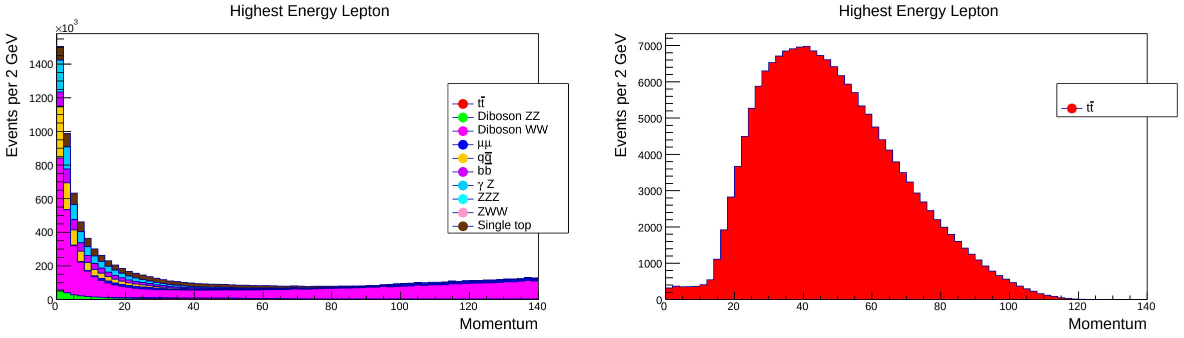
The signal region is determined by using a significance-optimised selection cut strategy. The strategy is elaborated as follows. In Figure 7.1, the distribution for the momentum of the highest energy lepton is shown for both the signal and the backgrounds (left) and for the signal only (right). The signal is barely visible with respect to the background. Comparing the signal distribution to the sum of distributions, a large amount of the background events can be removed while preserving a large amount of the signal events by applying a lower cut. This optimises the significance. The significance along with efficiency and purity are defined as

$$\text{Significance} = \frac{\text{sig}}{\sqrt{\text{sig} + \text{bkg}}} \tag{7.1}$$

$$\text{Efficiency} = \frac{\text{sig}}{\text{sig}_{\text{tot}}} \tag{7.2}$$

$$\text{Purity} = \frac{\text{sig}}{\text{sig} + \text{bkg}} \tag{7.3}$$

where sig is the number of selected signal events, bkg is the number of selected background events, and sig_{tot} is the total number of signal events before selection. In the limit of large numbers of events, the significance indicates the number of standard deviations by which the signal exceeds statistical fluctuations of the background [30]. The efficiency indicates the ability to retain signal events and the purity indicates the proportion of signal events to the total of selected events. All three measures should be optimised by the selection simultaneously. Optimising the significance will take care of the other two as well. For the event selection a set of observables have been considered. The list is summarised in Table 7.1. The strategy is an iterative procedure. For each observable, the significance is scanned based on where a cut would be made for both upper and lower cuts. The observable that has the largest maximum significance is chosen and a cut is applied if



(a) Signal and background

(b) Signal only

Figure 7.1.: Highest energy lepton distributions

List of observables

- Highest energy lepton
- 2nd highest energy lepton
- Lepton momentum
- Lepton momentum excluding highest energy lepton
- Momentum difference between highest and second highest energy lepton
- Missing momentum
- Invariant mass of lepton-neutrino pair
- Invariant mass of 1st and 2nd highest energy leptons
- Invariant mass of event excluding highest energy lepton
- Invariant mass of entire event
- Thrust of event excluding highest energy lepton
- Thrust of entire event
- Mass of jet
- Energy of jet
- Number of b-tagged jets
- Significance distribution
- Minimum of distance measure, d_{min}
- Invariant mass of one jet, $\Delta(m_i - m_W/2)$
- Invariant mass of di-jet system, $\Delta(m_ij - m_W)$
- Invariant mass of tri-jet system, $\Delta(m_ijk - m_t)$
- Invariant mass of lepton-neutrino-jet system, $\Delta(m_{l,nu,i} - m_t)$

Table 7.1.: Jet definitions included in performance studies

possible. The procedure is reiterated on the remaining set until there are no observables left in the set for which a cut is suitable.

The event selection is chosen in two steps. The strategy was first applied to the set of observables that do not require jets, and afterwards applied to the jet relevant observables. The observables in these two sets are assumed uncorrelated. Applying a cut selection for one set first will not affect the cut selection on the other set afterwards. The observables that do not require jets include momentum of leptons for either specific ones or sets of leptons, invariant masses of systems to target expected W or Z bosons and more generally for the event, and the thrust of the event. Thrust is a measure used for quantifying the event shape by tracking the quark and lepton directions. It is defined as [31]

$$T = \max \frac{\sum |\vec{p}_i| \cdot \hat{n}_T}{|\vec{p}_i|} \tag{7.4}$$

where $\hat{n}_T$ defines the thrust axis that maximises the energy flow. The value of thrust lies between 0.5 and 1 where $T = 0.5$ for an isotropic event shape and $T = 1$ for a fully aligned ff' system. With a non-isotropic event shape, the semileptonic decay products of the lepton, the neutrino and one jet is expected on one side of the semi-major axis with the three other jets on the other side.

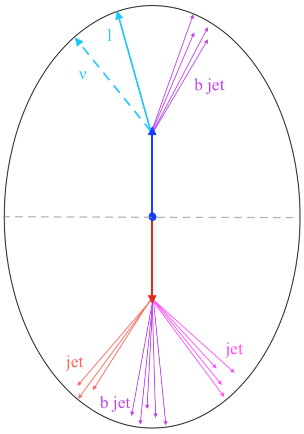
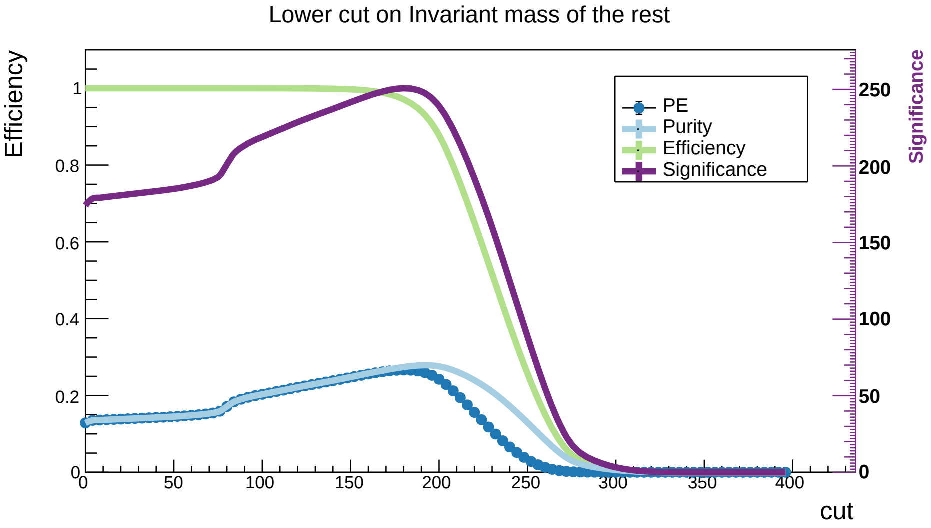
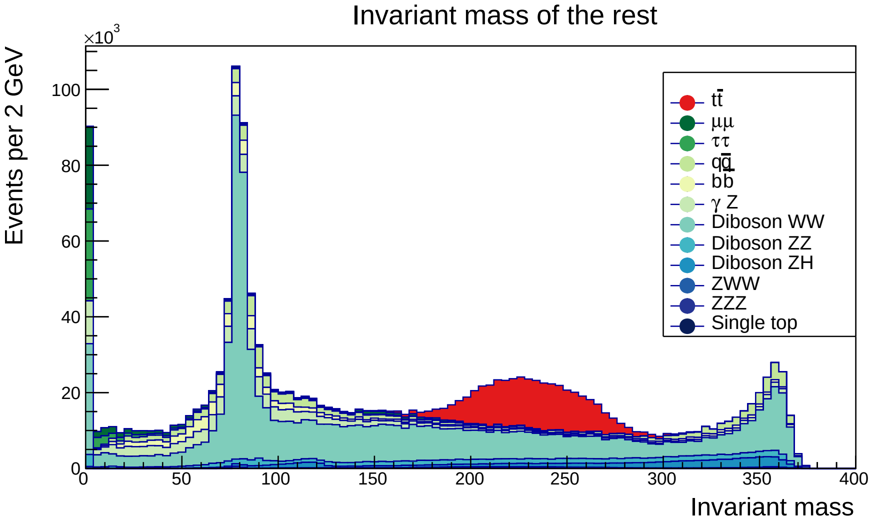


Figure 7.2.: Event shape

The first cut applied removes any events with zero leptons since at least on lepton is required for the signature. This can be considered as a sanity check. The second cut is applied on the thrust on the entire event since its significance scan returned the highest maximum in the first iteration. The cut is an upper cut of Thrust < 0.85. In the second iteration, the observable that returns the highest maximum for the significance scan is for the invariant mass of the event excluding the highest energy lepton. Figure 7.3 shows the two plots relevant for choosing a cut, where Figure 7.3a shows the significance scan for a lower cut and Figure 7.3b shows the sum of the backgrounds and the signal distributions. The significance has a peak around 200 GeV which corresponds well with the distributions. The invariant mass for the signal lies primarily in a region between 160 GeV and 300 GeV. The background can easily be removed below this region since the efficiency stays intact. A lower cut is chosen at $M(\text{rest}) > 160\,\text{GeV}$ which is slightly below the peak since the improvement of the significance for a tighter cut is relatively



(a) Significance as well as efficiency, purity and their product



(b) Sum of distributions for the background and signal

Figure 7.3.: Significance plot and sum of distributions for invariant mass of event excluding highest energy lepton prior to selecting a lower cut of $M(\text{rest}) > 160\,\text{GeV}$

Selection	**Signal**$/10^3$	**Background**$/10^3$	$\boldsymbol{\epsilon}_{\textbf{sig}}$	$\boldsymbol{\epsilon}_{\textbf{bkg}}$	**Signif**
Initial	262.89 ± 0.24	$29\,052 \pm 22$	1.0	1.0	48.6
At least 1 lepton	246.01 ± 0.23	$13\,565 \pm 14$	0.94	0.47	66.2
Thrust < 0 .85	236.07 ± 0.23	1599.1 ± 4.7	0.90	0.055	174.3
$M(rest) > 160$ GeV	234.61 ± 0.23	676.4 ± 2.9	0.89	0.023	245.8
$M(\ell_{HE},\not{E}) > 50$ GeV	181.21 ± 0.20	178.3 ± 1.5	0.69	0.0061	302.2
$p_{\ell_{HE}} < 100$ GeV	179.72 ± 0.20	133.6 ± 1.3	0.68	0.0046	321.0
$p_{\ell_{HE}} > 15$ GeV	176.01 ± 0.20	117.8 ± 1.2	0.67	0.0041	324.7
$p_{\ell_{2nd_{HE}}} < 40$ GeV	175.07 ± 0.20	103.5 ± 1.1	0.67	0.0036	331.7
Exactly 4 jets	175.07 ± 0.20	93.4 ± 1.1	0.67	0.0032	337.9
At least 1 b-tag	160.62 ± 0.19	11.14 ± 0.25	0.61	0.000 38	387.6

Table 7.2.: Simple cut-flow for signal events per 10^3, background events per 10^3, their efficiencies ϵ_{sig} and ϵ_{bkg}, and the significance (signif).

small. This practice is general for choosing the cuts. Signal events are rather preserved if the gain in significance is too small. The distributions would also indicate that an upper cut could be applied. However the observables are correlated and as the following iterations bring additional cuts, the majority of the background events above 300 GeV are removed in the M(rest) distribution. The additional cuts are a lower cut on the lepton-neutrino pair of $M(\ell_{HE},\not{E}) > 50$ GeV, an upper and a lower on the momentum of the highest energy lepton of $15\,\text{GeV} < p_{\ell_{HE}} < 100\,\text{GeV}$, and an upper cut on the momentum of the second highest energy lepton of $p_{\ell_{2nd_{HE}}} < 40$ GeV. In the second step of the event selection, the first cut applied to the jet-relevant observables requires exactly 4 jets. This cut can also be considered as a sanity check. If there are not excactly 4 reconstructed jets, it means that there were too few reconstructed particles available in the jet reconstruction. The final cut is a lower cut on the number of b-tagged jets where at least one b-tag is required. Any other observables does not improve the significance enough to warrant a cut. Prior to each cut the two relevant plots for the significance scan and the distributions can be found in Appendix B.

Table 7.2 summarises the cut-flow for the signal and background events where the number of events for signal and background, their efficiencies, and the significance are listed after each cut. The efficiency for the signal is 61 % while the efficiency for the backgrounds is 0.038 %. The event selection improves the significance on the signal to background ratio by a factor of ~ 8 from 48.6 to 387.6. The event selection is optimised to look for semileptonic top pair produced events with $\ell = \{e, \mu\}$. Since $\sim 35\,\%$ the tau-leptons decay into one of the other lepton flavours, the cut-flow numbers of the signal are given for the process $t\bar{t} \to \ell\nu_\ell b\bar{b}$ with $\ell = \{e, \mu, \tau\}$. The cut-flow for the number of events on the signal for each of the lepton flavours is summarised in Table 7.3. Taking the branching ratio of the tau-lepton into account, the process $t\bar{t} \to \tau\nu_\tau b\bar{b}q\bar{q}$ only contribute half as often as the processes for the other two flavours. Table 7.4 summarises the cut-flow for the number of events on each background. The largest contributions to the signal region comes from ZH followed by ZZ and single top.

Selection	$\boldsymbol{t\bar{t} \rightarrow e\nu_e b\bar{b}q\bar{q}}$	$\boldsymbol{t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}q\bar{q}}$	$\boldsymbol{t\bar{t} \rightarrow \tau\nu_\tau b\bar{b}q\bar{q}}$
Initial	87629±139	87629±139	87629±139
At least 1 lepton	87495±138	87281±138	71236±125
Thrust < 0 .85	84121±136	83932±136	68015±122
$M(rest) > 160$ GeV	83597±135	83381±135	67630±122
$M(\ell_{HE},\not{E}) > 50$ GeV	81727±134	81667±134	17815±62
$p_{\ell_{HE}} < 100$ GeV	81000±133	80903±133	17813±62
$p_{\ell_{HE}} > 15$ GeV	80557±133	80556±133	14896±57
$p_{\ell_{2nd_{HE}}} < 40$ GeV	80101±132	80090±132	14876±57
Exactly 4 jets	80101±132	80090±132	14876±57
At least 1 b-tag	73633±127	73743±127	13246±54

Table 7.3.: Cut-flow for the number of signal events for each lepton flavour, $\ell = \{e, \mu, \tau\}$.

Selection	$\boldsymbol{\mu\mu}$	$\boldsymbol{\tau\tau}$	$\boldsymbol{q\bar{q}}$	$\boldsymbol{b\bar{b}}$	$\boldsymbol{\gamma Z}$
Initial	1080110±2415	1079430±2414	5634480±12599	1012930±2265	4604960±10297
At least 1 lepton	1039600±2369	588822±1783	995185±5295	651416±1816	1239460±5342
Thrust < 0 .85	71498±621	50377±521	130269±1916	84300±653	154749±1888
$M(rest) > 160$ GeV	25285±370	6898±193	56373±1260	20659±323	37991±935
$M(\ell_{HE},\not{E}) > 50$ GeV	17136±304	3039±128	2986±290	2603±115	9762±474
$p_{\ell_{HE}} < 100$ GeV	4067±148	2958±126	2986±290	2598±115	6539±388
$p_{\ell_{HE}} > 15$ GeV	4067±148	2936±126	1775±224	1752±94	4996±339
$p_{\ell_{2nd_{HE}}} < 40$ GeV	1188±80	2704±121	1662±216	1722±93	2947±260
Exactly 4 jets	0±0	173±31	1662±216	1722±93	2003±215
At least 1 b-tag	0±0	0±0	56±40	1611±90	1381±178

Selection	$\boldsymbol{WW}$	$\boldsymbol{ZZ}$	$\boldsymbol{ZH}$	$\boldsymbol{ZWW}$	$\boldsymbol{ZZZ}$	**single top**
Initial	14579200±14579	874208±1955	159528±357	21638±48	1038±2	4538±10
At least 1 lepton	8511610±11140	406907±1334	111476±298	15118±40	627±2	4332±10
Thrust < 0 .85	898808±3620	103479±673	87473±264	13623±38	591±2	3907±9
$M(rest) > 160$ GeV	378695±2350	67493±543	68751±234	10696±34	500±2	3065±8
$M(\ell_{HE},\not{E}) > 50$ GeV	112989±1283	12641±235	9515±87	5174±24	66.3±0.6	2373±7
$p_{\ell_{HE}} < 100$ GeV	91484±1155	8545±193	7471±77	4685±23	60.9±0.6	2243±7
$p_{\ell_{HE}} > 15$ GeV	81498±1090	7015±175	6930±74	4548±22	52.2±0.5	2217±7
$p_{\ell_{2nd_{HE}}} < 40$ GeV	77401±1062	4218±136	5458±66	4126±21	30.1±0.4	2079±7
Exactly 4 jets	72138±1026	4078±134	5401±66	4110±21	29.9±0.4	2079±7
At least 1 b-tag	423±79	2352±101	2627±46	764±9	15.7±0.3	1910±7

Table 7.4.: Cut-flow for the number of background events for each background process.

8. Kinematic Fitting

Measurements are biased. They reflect the physical world - “das Ding an sich” [32] - but they are biased from the experimental lens that we perceive the world through. In a statistical sense measurements are random variables drawn from an underlying distribution. A single measurement on its own has little to no meaning, but combining it with other measurements and statistical tools, measurements are key to understanding the physical world. They are never fully exact, but they can be used to make estimates of the true world. Estimates themselves from their very nature are also random and therefore include imprecision. They can and will differ from the *true value* because of statistical fluctuations. The aim is to find the “best” value for the estimate that most accurately describes the *true value* while being acutely aware of how reliable it is i.e. the extend of the imprecision. The method of least squares is a simple and powerful tool for the analysis of experimental data and it can be used as a method of estimation. Generally, a model can be expressed in the form of conditions or constraints

$$f(a_1, a_2, \dots, a_p, y_1, y_2, \dots, y_n) = 0, \tag{8.1}$$

where y_i is a set of measurements consisting of n measured values, and a_i is a set of parameters consisting of p values, that the model depends on. Measurements will lie within some confidence region of the true underlying distribution, and they are unlikely to fulfil the conditions a priori, but corrections Δy_i can be added such that $y_i + \Delta y_i$ exactly fulfil the conditions. The principle of least squares requires that the sum of the corrections squared form a minimum:

$$S = \sum_{i=1}^{n} \Delta y_i^2 = \text{min}. \tag{8.2}$$

As a simple example, the mean of a set of measured values is found from

$$\hat{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{8.3}$$

which satisfies the principle of least squares. The principle has many applications, where the most common one would be estimating the parameters for a model based on a set of measurements. Another common use is found in testing the hypothesis that the measured data are consistent with the assumed model parametrisation.

In this analysis, the principle of least squares is applied in order to improve the accuracy of the event reconstruction by determining the corrections of the measurements. As an example an important set of constraints will be the conservation law of energy and momentum which defines four conditions. The energy and momentum of the final state must equal the initial state $(\sqrt{s}, 0, 0, 0)$ in order to be conserved. Kinematic fitting works particularly well at an e^+e^- collider because the initial state (E, p_x, p_y, p_z) is already known in all directions unlike at hadron colliders, where the initial state is only known in the plane transverse to the beam axis (z-direction).

8.1. Constrained Fit

The derivations in this section are based on [33]. In a constrained fit, the underlying model for a system is expressed in the form of m constraints

$$f_k(\bar{a}, \bar{y}) = 0 \quad , \quad k = 1, \ldots, m \tag{8.4}$$

with n measured parameters for $\bar{y}_i$, $i = 1, \ldots, n$ and p unmeasured parameters for a_j, $j = 1, \ldots, p$. The model is exactly fulfilled by a set of measurements whose expectation values are the true values $\bar{y}_i$ without bias

$$E(y_i) = \bar{y}_i. \tag{8.5}$$

For an uncorrelated set of measurements, the covariance matrix is a diagonal matrix

$$\mathbf{V} = \delta_{ij}\sigma_i\sigma_j \tag{8.6}$$

where σ_i is the resolution of the i-th measurement. In the case of no free parameters, the sum of squares can be rewritten as

$$S(y) = \Delta y^T \mathbf{V}(y)^{-1} \Delta y, \tag{8.7}$$

where Δy is a vector of the corrections to the measurements such that the sum $S(y)$ obtains a minimum and $\mathbf{V}(y)$ is the covariance matrix. For measurements that follow a Gaussian distribution, this form corresponds to the general form for calculating the χ^2. Lagrangian multipliers can be used to introduce constraints to the function we wish to minimise. The sum of squares can be extended by

$$L(y) = S(y) + 2\sum_{k=1}^{m} \lambda_k f_k(a, y). \tag{8.8}$$

Under the minimisation, the multipliers must vanish such that $f_k(a, y) = 0$. Only solutions that exactly fulfil the constraints are searched for, hence minimising $L(y)$ is equivalent to minimising $S(y)$ while fulfilling the constraints. The solution simultaneous fulfils

$$\frac{\partial L}{\partial y} = 0 \quad , \quad \frac{\partial L}{\partial a} = 0 \quad , \quad \frac{\partial L}{\partial \lambda} = 0. \tag{8.9}$$

Generally, least squares problems are solved by iteration. In the special case where the problem is linear, it can be solved analytically. Nonlinear problems can be reduced to a set of linear problems by linearisation techniques and solved numerically. Numerical methods require some attentiveness in order for the solution to (hopefully) converge and initial starting values should be given some thought. The initial values for the measured parameters should be the measurements themselves, y_i. The initial values for the unmeasured parameters depend on the problem.

In the case of semi-leptonic top pair produced events there is only one neutrino that has free parameters. Naively, one would expect that its four-momentum can be reconstructed from the missing four-momentum, but that assumes that the energy-momentum conservation can be satisfied a priori. While neutrinos do carry away some of the missing four-momentum, it also strongly depends on measurement effects for the reconstruction of the remaining event and missing coverage in the detector where some particles can escape detection. Nonetheless, the missing four-momentum is a reasonable initial value for the neutrino as it solves the energy and momentum conditions. In fully-leptonic events there are two undetected neutrinos and hence unmeasured parameters for them both. The directions of the two neutrinos contribute both to the missing four-momentum. Reasonable initial values could be found by solving selected conditions such as invariant mass of particle systems. In fully-hadronic events there are no unmeasured parameters which simplifies the problem of initial values, instead the challenge is figuring out the combinatorics of particle systems from the 6-jet final state.

Non-linear conditions can be linearised by Taylor-expansion

$$f_k(a^n, y^n) + \sum_j \frac{\partial f_k}{\partial a_j^{n+1}} (\Delta a_j^{n+1} - \Delta a_j^n) + \sum_i \frac{\partial f_k}{\partial y_i^{n+1}} (\Delta y_i^{n+1} - \Delta y_i^n) \approx 0 \tag{8.10}$$

and the function L for the $(n+1)$-th iteration can be rewritten as

$$L = \Delta y^T \mathbf{V}(y)^{-1} \Delta y + 2\lambda^T (A\Delta a + B\Delta y - c) \tag{8.11}$$

with

$$c = A\Delta a^n + B\Delta y^n - f \tag{8.12}$$

and

$$A = \begin{pmatrix} \partial f_1/\partial a_1 & \partial f_2/\partial a_2 & \dots & \partial f_1/\partial a_p \\ \partial f_2/\partial a_1 & \partial f_2/\partial a_2 & \dots & \partial f_2/\partial a_p \\ \dots & & & \\ \partial f_m/\partial a_1 & \partial f_m/\partial a_2 & \dots & \partial f_m/\partial a_p \end{pmatrix}, \quad f = \begin{pmatrix} f_1(a^n, y^n) \\ f_2(a^n, y^n) \\ \dots \\ f_m(a^n, y^n) \end{pmatrix},$$
$$B = \begin{pmatrix} \partial f_1/\partial y_1 & \partial f_1/\partial y_2 & \dots & \partial f_1/\partial y_n \\ \partial f_2/\partial y_1 & \partial f_2/\partial y_2 & \dots & \partial f_2/\partial y_n \\ \dots & & & \\ \partial f_m/\partial y_1 & \partial f_m/\partial y_2 & \dots & \partial f_m/\partial y_n \end{pmatrix}. \tag{8.13}$$

The system of equations to be solved can be written in the form of one matrix equation that is solved by

$$\begin{pmatrix} V^{-1} & 0 & B^T \\ 0 & 0 & A^T \\ B & A & 0 \end{pmatrix} \begin{pmatrix} \Delta y \\ \Delta a \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} \Delta y \\ \Delta a \\ \lambda \end{pmatrix} = \begin{pmatrix} C_{11} & C_{21}^T & C_{31}^T \\ C_{21} & C_{22} & C_{32}^T \\ C_{31} & C_{32} & C_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} = \begin{pmatrix} C_{31}^T \\ C_{32}^T \\ C_{33} \end{pmatrix} c \tag{8.14}$$

where the elements C_{11} through C_{33} are

$$\begin{aligned} C_{11} &= V - VB^T W_B BV + VB^T W_B A W_A^{-1} A^T W_B BV \\ C_{21} &= -W_A^{-1} A^T W_B BV \\ C_{22} &= W_A^{-1} \\ C_{31} &= W_B BV - W_B A W_A^{-1} A^T W_B BV \\ C_{32} &= W_B A W_A^{-1} \\ C_{33} &= -W_B + W_B A W_A^{-1} A^T W_B \end{aligned} \tag{8.15}$$

with $W_B = (BVB^T)^{-1}$ and $W_A^{-1} = (A^T W_B A)^{-1}$. Since the covariance matrix $\mathbf{V}$ for the measurements is symmetric, the elements C_{11}, C_{22}, C_{31}, and C_{33} are also symmetric. The symmetric covariance matrix for the solution is of the form

$$V \begin{pmatrix} \hat{y} \\ \hat{a} \\ \hat{\lambda} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{21}^T & 0 \\ C_{21} & C_{22} & 0 \\ 0 & 0 & -C_{33} \end{pmatrix}. \tag{8.16}$$

The above covariance matrix shows that the estimates for both measured and unmeasured particles are independent of the estimates of the Lagrange multipliers. They are an aid in treating the minimisation problem of least squares without an explicit parametrisation in terms of constraints, but they do not enter the solution.

Inclusion of Probability Distribution Functions

Above, the energy and momentum conservation were given as examples of constraints on the form of Eq. 8.4. Other examples of constraints are the mass constraints in the jet systems. On the hadronic side of a semileptonic top pair produced event, the top decays into a W-boson and a b-quark and the W decays further into a quark pair. Therefore there is a system of 3 jets whose mass should equal the top mass and a system of 2 light-flavour jets whose system should equal the W mass. On the leptonic side of the event, there is a jet, a lepton, and a neutrino where the system of all three particles should also equal the top mass and the system of the lepton and neutrino should equal the W mass. The masses of the jet system are not expected to be exact since the particles are unstable. Unstable particles have a total decay rate $\Gamma = 1/\tau$ hence their masses have a Breit-Wigner resonance where the Full Width at Half Maximum (FWHM) corresponds to Γ. Since the masses are not expected to be exactly on resonance, a constraint can be loosened by introducing a Probability Distribution Function (PDF) which further extends the function L by

$$L(y) = S(y) + g(x) + 2\sum_{k=1}^{m} \lambda_k f_k(a, y, x). \tag{8.17}$$

The additional function $g(x)$ only depends on a scalar variable, and it represents a penalty function. Since the parameters of the distributions are not free in the fit, the PDFs can be handled by extending the B-matrix such that the total set of measured parameters is $\{\bar{y}_i, x\}$

$$B = \frac{\partial f(a, y, x)}{\partial (y, x)}. \tag{8.18}$$

and the extended covariance matrix

$$\tilde{\mathbf{V}} = \begin{pmatrix} \mathbf{V} & 0 \\ 0 & \left(\frac{1}{2}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\Big|_{x=x^n}\right)^{-1} \end{pmatrix} \tag{8.19}$$

The corrected measured parameters in the $(n+1)$-th iteration are calculated from

$$\begin{aligned} \begin{pmatrix} y^{n+1} \\ x^{n+1} \end{pmatrix} &= \begin{pmatrix} y_0 \\ x^n \end{pmatrix} - \tilde{\mathbf{V}} \begin{pmatrix} 0 \\ \frac{1}{2}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\Big|_{x=x^n}^{-1} \end{pmatrix} + \tilde{\mathbf{V}} B^T (B\tilde{\mathbf{V}}B^T)^{-1} \\ &\times \left[A(a^n - a_0) + B \begin{pmatrix} y^n - y_0 \\ \frac{\mathrm{d}g}{\mathrm{d}x}\Big|_{x=x^n}^{-1} \Big/ \frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\Big|_{x=x^n} \end{pmatrix} - f(a^n, y^n, x^n) \right] \end{aligned} \tag{8.20}$$

and the corrected free parameters in the $(n+1)$-th iteration are calculated from

$$\begin{aligned} a^{n+1} = & \, a_0 + W_A^{-1} A^T W_B \\ & \times \left[A(a^n - a_0) + B \left(\frac{y^n - y_0}{\frac{\mathrm{d}g}{\mathrm{d}x}\big|^{-1}_{x=x^n} / \frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\big|_{x=x^n}} \right) - f(a^n, y^n, x^n). \right] \end{aligned} \tag{8.21}$$

8.2. ABC-parametrisation

In the constrained fit, the parameters of the underlying quarks and leptons should be Gaussian distributed. Otherwise the interpretation of the χ^2 would have little meaning. Additionally, Gaussian distributed parameters are vital for the convergence of the fit [21]. Using the measured energies and angles directly does not ensure Gaussian distributed parameters. The 3-vector for a reconstructed particle can be written as

$$\vec{p}_j^{\,r} = a_j |\vec{p}_j^{\,m}| \vec{p}_j^{\,a} + b_j \vec{p}_j^{\,b} + c_j \vec{p}_j^{\,c} \tag{8.22}$$

where a_j, b_j and c_j are the parameters used in the fit and $\vec{p}_j^{\,a}$, $\vec{p}_j^{\,b}$ and $\vec{p}_j^{\,c}$ are the unit vectors which form a Cartesian system. They are determined from the measured jet momentum $\vec{p}_j^{\,m}$,

$$\begin{aligned} \vec{p}_j^{\,a} &= \frac{\vec{p}_j^{\,m}}{|\vec{p}_j^{\,m}|}, \\ \vec{p}_j^{\,b} &= \frac{1}{\sqrt{p_{x,m}^2 + p_{y,m}^2}} (p_y^m, -p_x^m, 0), \\ \vec{p}_j^{\,c} &= \frac{1}{\sqrt{|\vec{p}_j^{\,m}|^2 (p_{x,m}^2 + p_{y,m}^2)}} (-p_x^m p_z^m, -p_y^m p_z^m, p_{x,m}^2 + p_{y,m}^2). \end{aligned} \tag{8.23}$$

$\vec{p}_j^{\,a}$ is defined to follow the direction of the measured jet, $\vec{p}_j^{\,b}$ is defined so that the dot product between itself and $\vec{p}_j^{\,a}$ equals zero, and $\vec{p}_j^{\,c}$ is defined as the cross product of $\vec{p}_j^{\,a}$ and $\vec{p}_j^{\,b}$. The initial values of the parameters are set to $\{a_j, b_j, c_j\} = \{1, 0, 0\}$ so that the reconstructed particle vector initially overlaps with the measured particle vector.

8.3. `ABCfit++` software package

The `ABCfit++` software package was written in connection with this analysis. It is based on ABCfit developed by Oliver Buchmuller and Jørgen Beck Hansen. It contains a set of classes described in the following:

Coordinate Representation The first base class contains the coordinate representation for the particles where `CoordRepr.h` defines the base class. There are four derived classes defined in `PxPyPzM.h`, `PxPyPzE.h`, `PtEtaPhiM.h`, and `ABCD.h`. Each derived class defines how the coordinates transform between itself (this) and the internal representation, the derivative of itself w.r.t the internal representation, the derivative of the internal representation w.r.t. itself, and default expectation parameter values and expectation covariance matrix. The internal representation is PxPyPzM which defines the coordinates in p_x, p_y, p_z, and mass. PxPyPzE defines the coordinates in p_x, p_y, p_z, and energy. As an example, the transformations between PxPyPzM and PxPyPzE are straightforward since only the fourth parameter is transformed by

$$E = \sqrt{m^2 + p_x^2 + p_y^2 + p_z^2} \quad \longleftrightarrow \quad m = \sqrt{E^2 - p_x^2 - p_y^2 - p_z^2} \tag{8.24}$$

PtEtaPhiM defines the coordinates in terms of transverse momentum p_T, pseudorapidity η, and the polar angle in the transverse plane ϕ. The derivatives are 4×4 matrices and convenient for applying the chain rule when calculating the A and B matrices in Eq. 8.13. The final coordinates representation is the ABCD representation which defines the A, B, and C coordinates as described in the previous section. The fourth parameter denoted D is set to the mass of the particle. The ABCD representation is currently the only suitable choice in the software package for representing the parameters in Eq. 8.17 since the parameters will be Gaussian distributed.

In the other three representations, the default expectation values are simply set to coordinates of the particle itself. In ABCD, the default expectation values are $\{1, 0, 0, m\}$ as discussed above. Since the parameters are defined in terms of the measured jet, the unit vectors must be saved. In addition to the four parameters for ABCD, the representation also includes the auxiliary coordinates for the unit vectors bringing the total number of coordinates to twelve.

The base class representation contains an option to overwrite the default expectation values and covariance matrices for individual particles. This is useful since different particles have different expectations as will be seen in the following section.

Particle Object The base class for particle objects is defined in `ParticleObject.h`. A particle object contains a set of coordinates and an input coordinate representation in which the coordinates for a particle are defined. Only measured particles contribute to the χ^2, so their objects contain an additional χ^2 coordinate representation for how the parameters of the χ^2 are defined. By default, all parameters of a particle object are subject to change in a fit. The particle object class contains the option to fix certain parameters in the constrained fit.

Constraint Setting up constraints is defined by the base class `constraint.h`. There are five derived classes defined in `SumP{x,y,z}Constraint.h`[1], `SumEConstraint.h`, and `InvMassConstraint.h`. The first four set up the 4-momentum constraints and the last one sets up the constraint for the invariant mass of a particle system. Each constraint takes a list of particles which is a vector of pointers to particle objects. The list defines the particle system. There are two types of constraints.

For the first type, the constraint has to be exactly fulfilled and a constraint value is given as input. The class contains a constraint function which in the case of the SumEConstraint calculates the sum of energies. The constraint is fulfilled when the sum exactly matches the constraint value.

For the other type, the constraint is allowed to vary, and instead of a constraint value a probability distribution function is given as input. These types of constraints

---

[1] `SumPXConstraint.h`, `SumPyConstraint.h`, `SumPzConstraint.h`

contributes to the χ^2 via the penalty function $g(x)$ as seen in equation 8.17. The more the constraint function value varies, the higher the penalty.

Composite constraint The base class of composite constraints is used to create linear combinations of constraints. The class is defined in `CompositeConstraint.h`. It takes a list of constraints as input where the list is a vector of pairs. Each pair contains a constraint and a constant for how the constraints is added in the linear combination. It also takes either a constraint value or a PDF as input, depending on how tight the composit constraint should be defined. Examples of linear combinations are equal energy constraints or equal mass constraints for two particle systems.

Probability distribution functions PDFs are defined in `ProbDistFunc.h`. Currently there is only one derived class in the software package which sets up a Gaussian PDF in `GaussianPDF.h`. The option to expand the software package and add derived classes for other PDFs such as the Breit-Wigner and ISR exists. The penalty function is calculated from the PDF by

$$g(x) = -2\ln(\mathrm{pdf}(x)) \tag{8.25}$$

Since both first and second order derivatives of the penalty function are needed in the fit, the class contains functions for calculating these. The class also contains an expectation value for initially setting up the fit.

Matrix Algebra The base class of `MatrixAlgebra.h` contains helper functions for the tedious matrix calculations. The helper functions include functions for vector and matrix algebra such as addition, multiplication, and calculating inverse and transposed matrices. These functions are called upon many times in the fitting procedure and for optimising performance important features have been efficient allocation of memory and pass-by-pointer.

ABC Fit The fitting procedure is defined in the base class of `ABCFit.h`. It takes a list of composite constraints and a maximum number of iterations. In each iteration, corrections to the parameters are calculated as seen by Eqs. 8.20 and 8.21. The fit stops either when it has converged i.e. the difference in the χ^2-values between two iterations is small, or when the maximum number of iterations has been reached. The fit returns a statuscode indicating whether the fit converged. It also returns the number of iterations it took before the fit converged, a χ^2-value on how well the fitted particles fulfil the constraints, number of degrees of freedom, and a list of the fitted particles. Since the particle objects are passed as pointers into the constraints, a constrained fit can be applied in stages where the fit is first applied for one set of constraints followed by another set of constraint instead of requiring all constraints to be fulfilled at once. This could be relevant in cases of convergence issues.

The full `ABCfit++` software package is published at [34].

9. Analysis

9.1. Observables

In this analysis the observables investigated are the cross section and angular distributions.

In total there are five angles of interest. The first is the polar angle between the electron-beam and the reconstructed top particle which is given by

$$\cos\theta_{et} = \frac{\mathbf{p}_e \cdot \mathbf{p}_t}{|\mathbf{p}_e||\mathbf{p}_t|} \tag{9.1}$$

where $\mathbf{p}_e$ is the unit vector for the electron beam that lies along the z-axis (0,0,1), and $\mathbf{p}_t$ is the momentum of the reconstructed top particle. The angles between the top and its decay product can be calculated by boosting the reference frame from the laboratory frame to that of the top particle. The component parallel to the flight direction of the top particle transforms under the Lorentz transformation, and it is calculated from

$$\mathbf{p}_{t_\parallel} = (\mathbf{p}_t \cdot \mathbf{p}_d)\frac{\mathbf{p}_t}{|\mathbf{p}_t|^2} \tag{9.2}$$

and the perpendicular component is calculated from

$$\mathbf{p}_{t_\perp} = \mathbf{p}_t - \mathbf{p}_{t_\parallel} \tag{9.3}$$

The parallel component is transformed by

$$\mathbf{p}^*_{t_\parallel} = \gamma(\mathbf{p}_{t_\parallel} - \beta E) \tag{9.4}$$

where γ is the Lorentz factor. The polar angle between the top quarks and its decay product can now be calculated from

$$\cos\theta^*_{td} = \frac{\mathbf{p}^*_{t_\parallel}}{\sqrt{(\mathbf{p}^*_{t_\parallel})^2 + (\mathbf{p}_{t_\perp})^2}} \tag{9.5}$$

In the reference frame of the top particle, the z-axis lies parallel to its flight direction hence $\hat{z'} = \frac{\mathbf{p}_t}{\mathbf{p}_t}$. In this reference frame, the x'-axis is found from the cross product between the z'-axis and the z-axis hence $\mathbf{x'} = \hat{\mathbf{z'}} \times \mathbf{p}_e$ and the y'-axis perpendicular to x' and z' is found their cross product $\mathbf{y'} = \mathbf{x'} \times \hat{z'}$. The normalisation or lack thereof cancels, since the azimuthal angle is calculated by

$$\phi_{td} = \arctan\frac{\mathbf{x'} \cdot \mathbf{p}_d}{\mathbf{y'} \cdot \mathbf{p}_d} \tag{9.6}$$

$t\bar{t}$ production	$t \rightarrow Wb$	$W \rightarrow \ell\nu$
θ_{et}	θ^*_{tb}, ϕ^*_{tb}	$\theta^*_{W\ell}$, $\phi^*_{W\ell}$

Table 9.1.: Summary of the angles of interest

Similarly, the reference frame can be boosted in that of the W particle, where the angle between the W boson and one of its decay products can be calculated by substituting the momentum of the top particle with that of the W. None of the considered operators contain CP violating terms hence the physics of the tW^+b and the $\bar{t}W^-\bar{b}$ are expected to be alike and hence only the angles of one top is of interest. For the W decays, the hadronic side is of no use since light flavour jets cannot be distinguished at detector level. In total there are five angles of interest summarised in Table 9.1.

9.2. Selection on Angular Phase Space & Total Cross Section

The event selection improved the significance of the signal to background. In the process, some of the signal was also removed where the event selection showed to have an efficiency of 61 % for the signal. By comparing the angular distributions on quark level before and after the event selection, effects on angular phase space can be investigated. The angular distributions are shown in Figure 9.1 for the angles $\cos(\theta_{et})$, θ^*_{tb} and ϕ^*_{tb}. The distributions have been normalised in order to compare the shapes, since the event selection also changes the cross section. Qualitatively, the two distributions look comparable even though they do not match exactly. The ratio plots indicate that for $\cos(\theta_{et})$ and θ^*_{tb} the differences in the distributions are not due to any systematic effects from the event selection, while there seems to be a systematic effect on θ^*_{tb}. The comparison of two distribution can be quantified by the Kolmogorov-Smirnov test (KS test) which is a goodness-of-fit test statistic. In this case though, the two distributions are highly correlated which is not accounted for in the absolute value of the test statistics. Here only the relative KS test statistics are evaluated with respect to each other. The KS-test is a general test for comparing whether two distributions in 1D are the same, and it is sensitive to differences in both location and shape. In the two-sample case, EDFs (Empirical Distribution Functions) are constructed from the data. The test measures the maximal distance between the EDFs and gives a probability of being from the same distribution. The KS-test should be unbinned but can also be applied to binned data as long as the bin width is small. The KS-test was applied to the two histograms with smaller binning than what is shown in the figure. The distributions are rebinned to match the plots in the next section. The KS-test on $\cos(\theta_{et})$ yields a p-value of 0.066, on θ^*_{tb} it yields a p-value of $5 \cdot 10^{-35}$, and on ϕ^*_{tb} it yields a p-value of 0.40. ϕ^*_{tb} shows the highest compatibility between the distributions before and after the event selection for the three angles. $\cos(\theta_{et})$ has relative large p-value which in combination with the ratio plot shows that the distribution is mostly unaffected by the event selection. There might be a systematic effect to the left. θ^*_{tb} has by far the smallest p-value indicating that the distribution is in fact affected by the event selection which is confirmed by the ratio plot.

Since the event selection removes part of the signal it influences the total cross section. In Figure 9.2, the efficiency for all the couplings as well as for the SM are shown. The

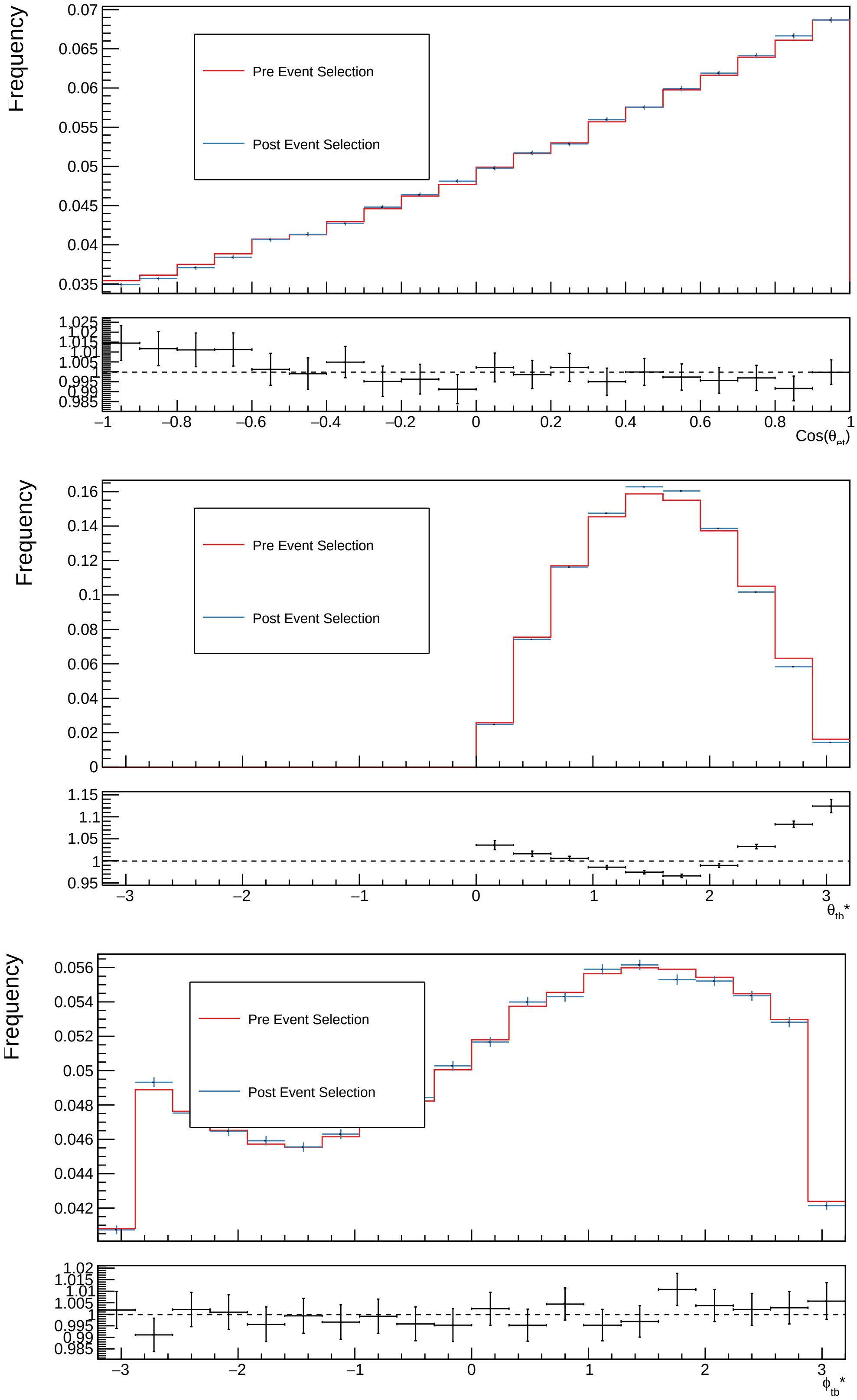


Figure 9.1.: Normalised angular distributions before and after event selection with ratio plots. The angles are in the order $\cos(\theta_{et})$, θ^*_{tb} and ϕ^*_{tb}.

efficiency is lower for the vr_tbW_re and especially vl_tbW_re coupling where only 56.5 % of the events are kept compared to 61.4 % for the SM. For the rest of the couplings, the efficiency is comparable with the efficiency of the SM. This shows that the event selection affects the total cross section differently for some of the couplings than it affect the total cross section of the SM.

9.2.1. Event Reconstruction of Semileptonic $t\bar{t}$ with Top Systems

At quark level, the mother-daughter history is known for the six decay particles making it trivial to match top quarks and W-bosons to their respective decay products when calculating the angles. At detector level, the set of six measured particles consists of a lepton, a neutrino reconstructed from the missing energy, and 4 jets with at least one of them being b-tagged. Jet systems can be combined into reconstructed W-particles and top particles. The leptonic top-system consists of the lepton, the neutrino and one (b-)jet where the leptonic W-system consists of the lepton and the neutrino, and the hadronic top-system consists of the remaining three jets where the hadronic W-system consists of two light flavour jets. Which jets belong to which systems is not known. For 2 b-tags there are two possible combinations for the reconstruction since the two light flavour jets combine into a W-boson leaving a choice for dividing the two b-jets into each of the two top systems. In the following section only the top-systems are of interest, so for 1 b-tag there are four possible combinations since any of the four jets can belong to the leptonic

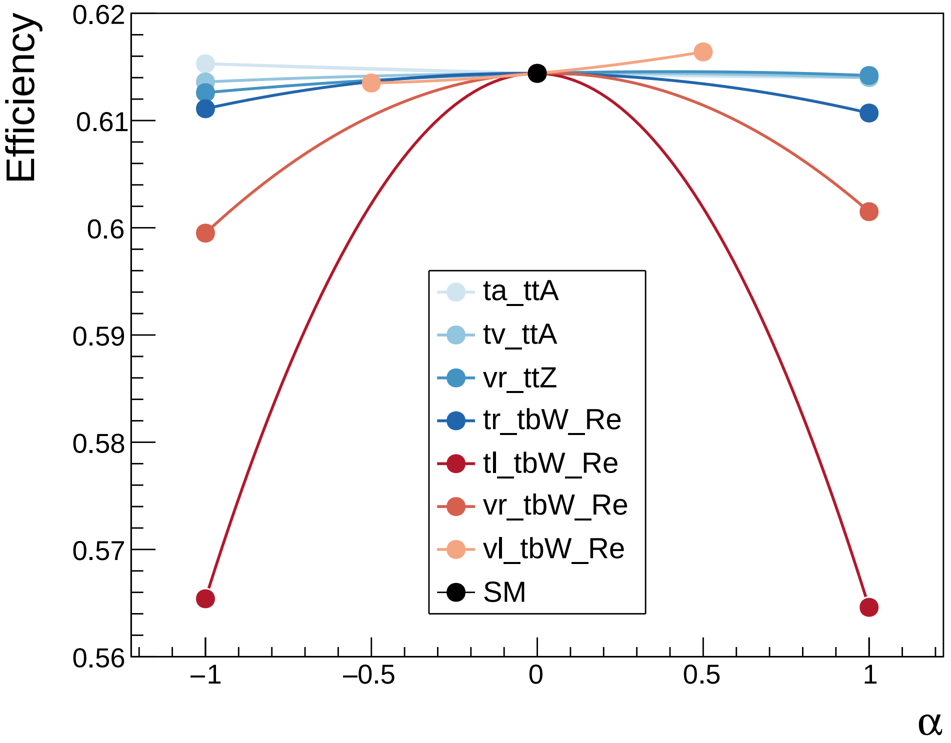


Figure 9.2.: Efficiency of event selection for all couplings and the Standard Model.

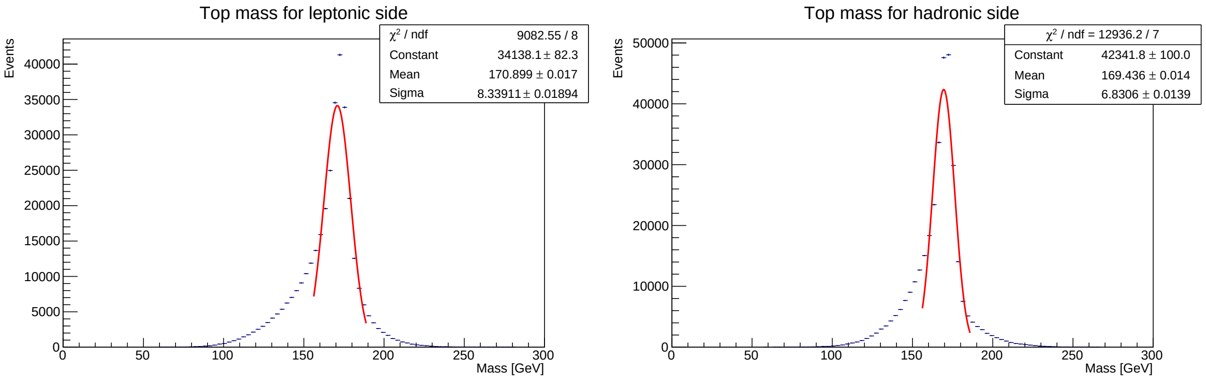
top system with the remaining three jets belonging to the hadronic top system. The task is to choose the combination that most likely corresponds to the combination at quark level. As a first test, the top mass χ^2 for the leptonic side could be used

$$(m_{\ell\nu j} - m_{\text{top}})^2 = min \tag{9.7}$$

where the combination is chosen such that the leptonic top system has an invariant mass closes to the top mass. Naively, one could expect that the combination that has the lowest χ^2 for the leptonic top system will also correspond to the lowest χ^2 for the hadronic top system. This assumption is biased towards the leptonic side. In order to include the hadronic side, the χ^2 can be expanded by

$$\frac{(m_{\ell\nu j} - m_{\text{top}})^2}{\sigma_{\text{lep}}^2} + \frac{(m_{jjj} - m_{\text{top}})^2}{\sigma_{\text{had}}^2} = min \tag{9.8}$$

where each of the contributions to the χ^2 has to be weighted by the variance σ^2. The variances are estimated from the top mass distributions in Figure 9.3 where a Gaussian is fitted to the distribution around the central value. The distributions are found by matching the measured jets to the true quarks in the Monte Carlo data where each jet is matched to the quark closest in angle. As a cross check for the matching, the invariant top masses are only kept for events where the matching was unique i.e. each of the four jets was matched to the four different quarks. If two or more jets are matched to the same quark, the matching is not unique. The Gaussian distribution was chosen as an approximation for the mass distribution even though it does not pick up the full behaviour which take after a Breit-Wigner distribution convoluted with a Gaussian distribution. The Breit-Wigner behaviour follows from the behaviour of unstable particles and the Gaussian behaviour follows from the fact that the particles are (simulated) measurements. The Gaussian fit is a good approximation and it was chosen partly because it is consistent with the interpretation for the χ^2 and partly because in the later section the same approximation with a Gaussian mass distribution is used for the constrained fit. Both fit results have a central value just below the top mass, which is slightly lower on the hadronic side and both fits estimate a similar standard deviation of the same order of magnitude around 7-8 GeV in the fitted region. The poor χ^2/ndf values of the fits are



(a) Leptonic top mass at reco level.

(b) Hadronic top mass at reco level.

Figure 9.3.: Gaussian fits of invariant top mass distributions in a region around the central value for the leptonic side (left) and the hadronic side (right) of the event.

due to the tails from Breit-Wigner part of the distributions. The standard deviation for the leptonic top mass is slightly higher than the standard deviation from the hadronic top mass which is expected since the standard deviation on the neutrino picks up the measurement effect for the entire event as well as detector effects such as missing coverage. The central values slightly below the top mass are a consequence of the jet energy measurements. In Figure 9.4, the W mass distributions are shown. For the leptonic W mass distribution, the central value fits with the W mass whereas for the hadronic W mass distribution, the central value is slightly lower than the W mass. This confirms that it must be the measured jet energies that are responsible for the shift in central value of the invariant mass distributions. The jet energies can be corrected with a constrained fit which is elaborated on in the following section.

For each event, a mass χ^2 value from Eq. 9.8 can be calculated for each combination, where the σ-values are conservatively chosen to be slightly above the fit results with $\sigma_{\text{lep}} = 10$ GeV and $\sigma_{\text{had}} = 9$ GeV. This choice has been made to compensate for the tails. The true combination is expected to have the lowest χ^2 value however for some events it is not possible to clearly distinguish which combination is most likely to match the underlying quark systems. In order to reduce bias, probability weights are introduced for the event reconstruction. For each combination a probability is calculated from the probability density function of the χ^2 distribution with 2 degrees of freedom – one for each of the top systems. The probability density function is given by

$$p_k(x) = \frac{1}{\Gamma(k/2)2^{k/2}} x^{k/2-1} e^{-x/2)} \tag{9.9}$$

where k is the number of degrees of freedom and $\Gamma(k/2)$ denotes the gamma function. For each combination, the five angles of interest are calculated and each of them are weighted their probability value weighted by the sum of probabilities for each event. In Figure 9.5 the five angular distributions are plotted for different choices of the event reconstruction. The effects of the different choices are best seen by the first plot for cosine to the angle between the beam and the top particle. The two blue lines shows the angular distributions where all 2 or 4 possible combinations are weighted equally. The darker blue line shows the angular distributions for the measured parameters and the dashed blue line shows the angular distributions for the parameters reconstructed with a constrained fit. Reconstructing the events based on how well the combinations fit with the top systems makes shifts the angular distribution making it more pronounced. For

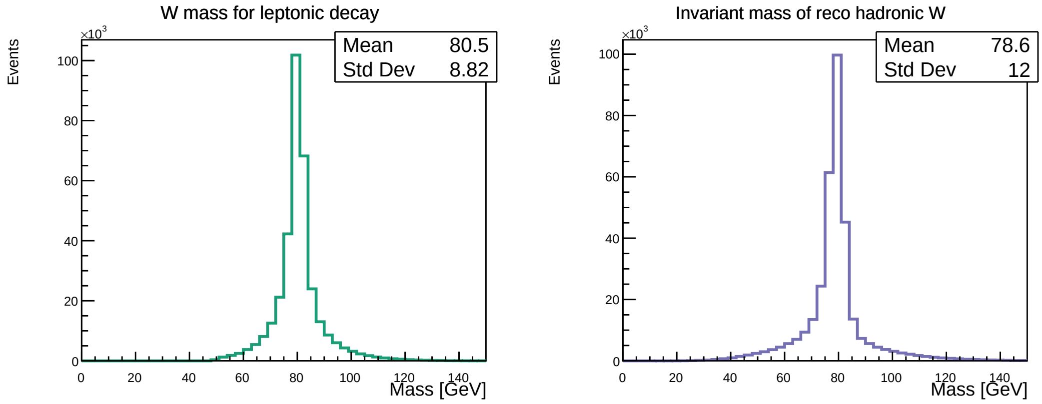


Figure 9.4.: W mass distribution for leptonic decay (left) and hadronic decay (right).

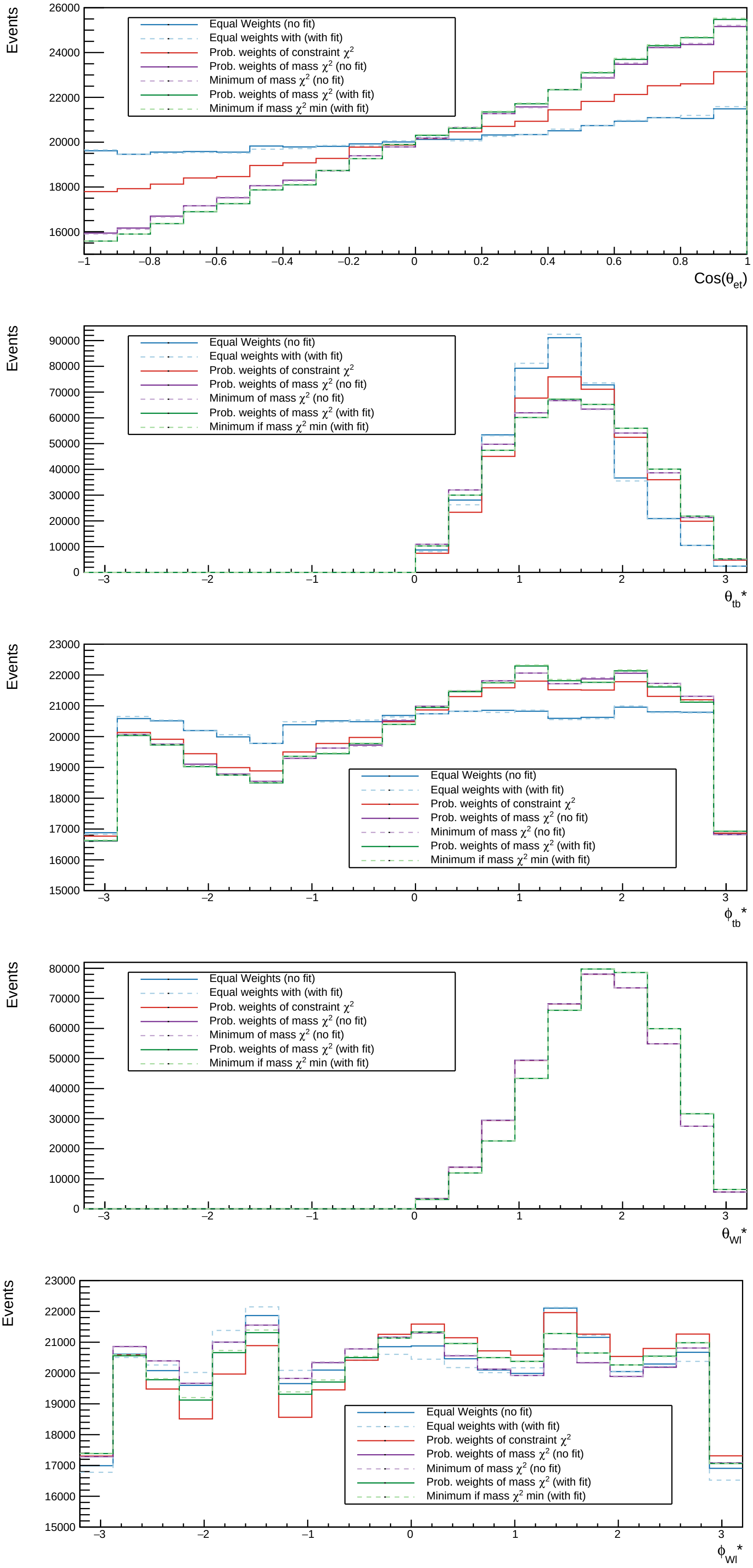


Figure 9.5.: Angular distributions with different choices for the event reconstruction. The angles are in the order $cos(\theta_{et})$, θ_{tb}, ϕ_{tb}, $\theta_{W\ell}$, and $\phi_{W\ell}$.

the two purple lines, the angles are reconstructed for the measurements directly. For the dashed line, only the combination that obtained the minimum for the mass χ^2 is kept, and for the solid line, all combinations are weighted based on the probability. The two purple lines are overlapping which shows that the combination with the lowest χ^2 has the largest weight. The angular distribution can be shifted further upwards by adding corrections to the measurements with a kinematic fit. Only the combination that achieved the lowest constraint χ^2 defined in Eq. 8.17 is kept for the green lines and is plotted as the dashed line. The solid line cycles through the combinations for the reconstructed parameters, calculates their mass χ^2 and weighs them by their probability. Again the two lines are overlapping which shows that the combination with the lowest χ^2 still has the largest weight. In fact they are overlapping more than for the purple lines which means that the constrained fit pushes the “good” combination further away from the “bad” one(s). The red line shows the distribution in the case where all combinations are reconstructed with a constraint fit and all reconstructed combinations contribute weighted by their probability. In this case “bad” events are made to fit better with the top systems and they will therefore weigh more than they should, pulling down the distribution.

9.2.2. Event Reconstruction of Semileptonic $t\bar{t}$ with a Constrained Fit

In this section, the constrained fitting technique is applied to the semileptonic $t\bar{t}$ events in the signal region. The fit contains six constraints in total. The first four constraints are the energy and momentum conservation constraints which are required to be exactly fulfilled such that

$$\sum \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} 365 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{GeV} \tag{9.10}$$

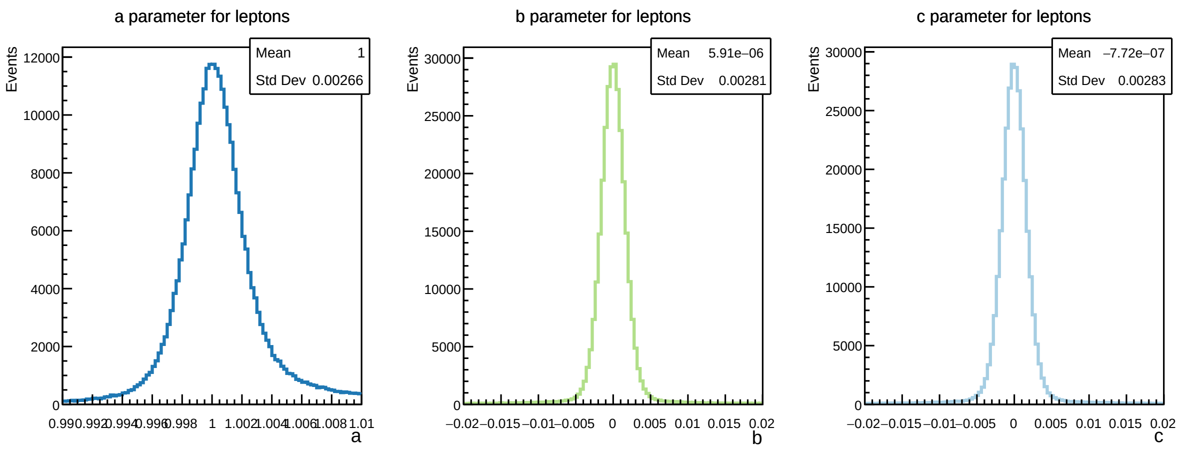
and the last two constraints are the mass constraints for the two top systems where a Gaussian term for $g(x)$ is included in Eq. 8.17 for each of the top system masses. The Gaussian terms have $\mu = 173$ GeV, $\sigma_{\text{lep}} = 10$ GeV and $\sigma_{\text{had}} = 9$ GeV. The order of magnitude for the σs were found in the previous section from the Gaussian fit of the top masses which can seen in Figure 9.3. The σ-values are kept slightly above the fit results to compensate for the tails and stay consistent with the previous section. The event reconstruction is performed by cycling through each possible combination and performing a constrained fit. The combination that returns the lowest constraint χ^2 calculated from Eq. 8.17 is kept with the corrections returned by the fit. The fit uses the ABCD-parametrisation detailed in the previous section. Assuming that the measurements are uncorrelated, the covariance matrix is a diagonal matrix as in Eq. 8.6. The resolution for the a, b, and c parameters can be found in Figure 9.6. Here the a, b, and c parameter distributions are plotted for leptons, b-jets and light flavour jets respectively. The distributions are found by matching the measured lepton and jets to the true lepton and quarks in the Monte Carlo data where each jet is matched to the quark closest in angle. The same cross check as before is applied where the parameters are only kept for events that had unique matching. For leptons the mean of the a parameter is 1 which means that on average the quark level lepton and the measured lepton have equal energies. For both b-jets and light flavour jets however the mean of the a parameter is 1.03 which means that on average the jets have measured energies lower than the energies of their corresponding quarks. This confirms the observation from the previous section where the hadronic W mass and the hadronic top was were observed to have lower central values than the leptonic W mass and the leptonic top mass. The leptons are well measured with resolutions 0.003 for all three parameters. For both types of jets the resolution of a is 0.07, for the b and c the resolutions are 1.2 for the b-jets and 1.4 for the light flavour jets. The covariance matrices for leptons, b-jets and light flavour jets are

$$\begin{gathered} V_{\text{lepton}} = \begin{pmatrix} 0.003 & 0 & 0 & 0 \\ 0 & 0.003 & 0 & 0 \\ 0 & 0 & 0.003 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}, \\ V_{\text{b-jet}} = \begin{pmatrix} 0.07 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 \\ 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}, \quad V_{\text{lf-jet}} = \begin{pmatrix} 0.07 & 0 & 0 & 0 \\ 0 & 1.4 & 0 & 0 \\ 0 & 0 & 1.4 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix} \end{gathered} \tag{9.11}$$

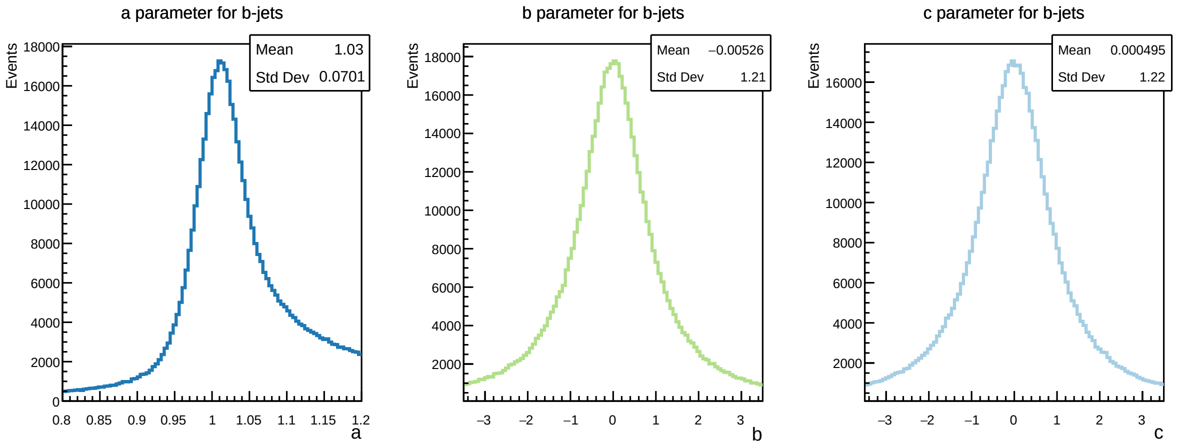
The resolution on d or the mass resolution is here set to 10. It is not used in the fit because the mass parameter is kept fixed after rescaling the energy and momentum of the input particles to have zero mass. For the jet reconstruction, the recombination scheme was chosen such that the sum of the pseudo particles 4-momentum was Lorentz invariant. The fit procedure deals with the non-zero mass by rescaling the jets. In total there are 15 parameters for the five measured particle objects; the lepton, the two b-jets and the two light flavour jet, there are 3 free parameters for the unmeasured neutrino object, and there are two constraint parameters for the two Gaussian terms. The parametrisation function for the jets has been changed from the default (1.0,0.0,0.0,m) to (1.3,0.0,0.0,m) to match the mean values of the a distributions. Since the neutrino is unmeasured, there is no associated covariance matrix. The constrained fit has 6 constraints and 3 free parameters leaving 3 degrees of freedom. The final choice on the event reconstruction corresponds to the solid green line in Figure 9.5. Following the event selection, the constrained fit is performed for each of the 2 to 4 combinations of the jet systems. The corrections to the measurements are kept for the combination that gave the smallest constraint χ^2, and the particles are reconstructed with these corrections. The 2 to 4 combinations are again cycled through where the mass χ^2 in Eq. 9.8 is calculated with the reconstructed particles instead of the measured particles. The angles are weighted with the probability associated with the mass χ^2 for each combination. Each weight is normalised by the sum of weights for the event.

9.3. Results

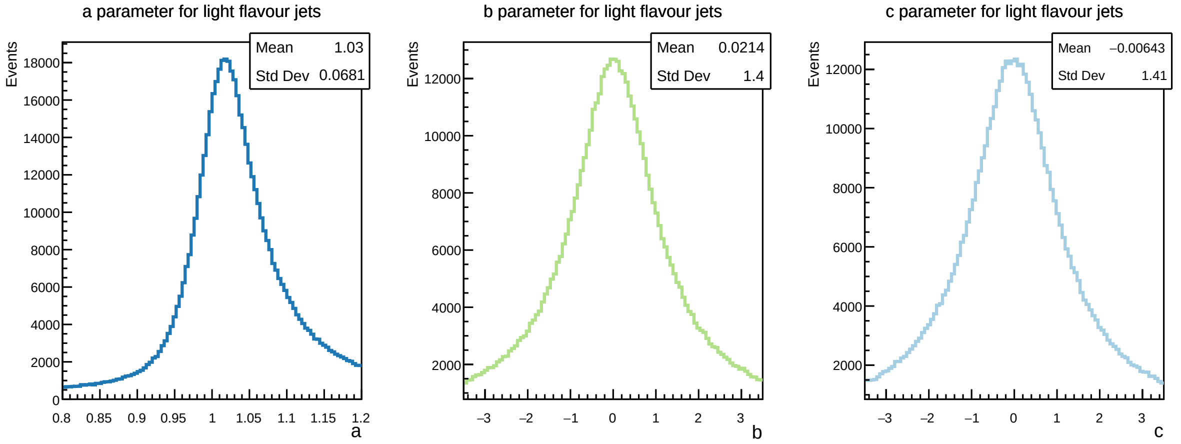
In this section confidence intervals for each coupling is determined. These confidence intervals provide insight on the sensitivity to anomalous top couplings at the FCC-ee. The confidence intervals are determined by performing a fit of the signal region to a parabola. The signal regions constitutes the top pair produced events and all the backgrounds post the event selection. In Section 3, it was shown that the matrix element has a parabolic dependence for a single non-zero coupling parameter. The parameters of the parabola can be determined from three of its points. For this analysis, events have been simulated



(a) a, b, and c parameter distribution for leptons



(b) a, b, and c parameter distribution for b-jets



(c) a, b, and c parameter distribution for light flavour jets

Figure 9.6.: a, b, and c parameter distributions for the three types of reconstucted objects. The parameters are defined in Section 8.2.

for one coupling at a time on either side of the SM where the model parameter is set to $\pm\alpha$. The SM events correspond to a model parameter of $\alpha = 0$. The model parameter is set to $\alpha = 1.0$ for all model parameters except vl_tbW_Re where the model parameter is set to $\alpha = 0.5$. Having observables corresponding to three points of the parabola for

$\{\alpha^+, \alpha^0, \alpha^-\}$

$$f(\alpha^0) = C \tag{9.12}$$

$$f(\alpha^+) = A\alpha^2 + B\alpha + C \tag{9.13}$$

$$f(\alpha^-) = A\alpha^2 - B\alpha + C \tag{9.14}$$

where $f(\alpha)$ is one of the angular distributions. It is trivial to determine the parameters

$$A = \frac{f(\alpha^+) + f(\alpha^-)}{2\alpha^2} - \frac{f(SM)}{\alpha^2} \tag{9.15}$$

$$B = \frac{f(\alpha^+) - f(\alpha^-)}{2\alpha} \tag{9.16}$$

$$C = f(SM) \tag{9.17}$$

Including the background the fit model is

$$f(x) = A \cdot x^2 + B \cdot x + f(SM) + f(Bkg). \tag{9.18}$$

The confidence intervals are determined by using a least-squares method where the histograms of the signal region are fitted to the model above. The χ^2 is calculated from the sum over each bin

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - f(x_i;\alpha))^2}{\sigma_{y_i}^2 + \sigma_{f(x_i;\alpha)}^2} \tag{9.19}$$

where the numerator is the square of difference in the bin content between the signal region histogram and the fit model for the ith bin and the denominator is the square of the error of the difference found from error propagation. The fit model histogram and the signal region histogram are constructed from the Monte Carlo samples where the number of simulated events far exceeds the number of expected events in order to reduce the relative statistical error. The angular distributions are rescaled from their respective cross section to match the number of expected events while the relative statistical error is preserved. Therefore the errors on the fit model histogram is small but not zero. In the fit, the signal region histograms acts as data so the errors have to also match the expected errors for real data. The errors for the rescaled histograms cannot be used. One option is to simulate the signal region so that the number of simulated events match the number of expected events. In this case the errors resemble the expected errors. The statistical fluctuations of drawing random variables from the underlying distribution for the signal region will be carried over into the fit. The fit returns an estimate for the confidence intervals and they are therefore also prone to statistical fluctuations. The statistical fluctuations could be reduced by repeating the fit for multiple randomly drawn distributions and taking the average on the estimates for the confidence intervals. Another option is to use the scaled signal region histogram as an Asimov dataset and then simply rescaling the errors. The error on each bin is the square root of the bin content according to Poisson statistics. Asimov datasets are "perfect datasets" where all expectation values are exactly equal to the true values i.e. [35]

$$\hat{X} = \bar{X} \text{ for all parameters X.} \tag{9.20}$$

Applying the fit to the Asimov dataset gives the median result immediately. The dataset is constructed to always fit perfectly with the SM which can be seen from Figure 9.7 for

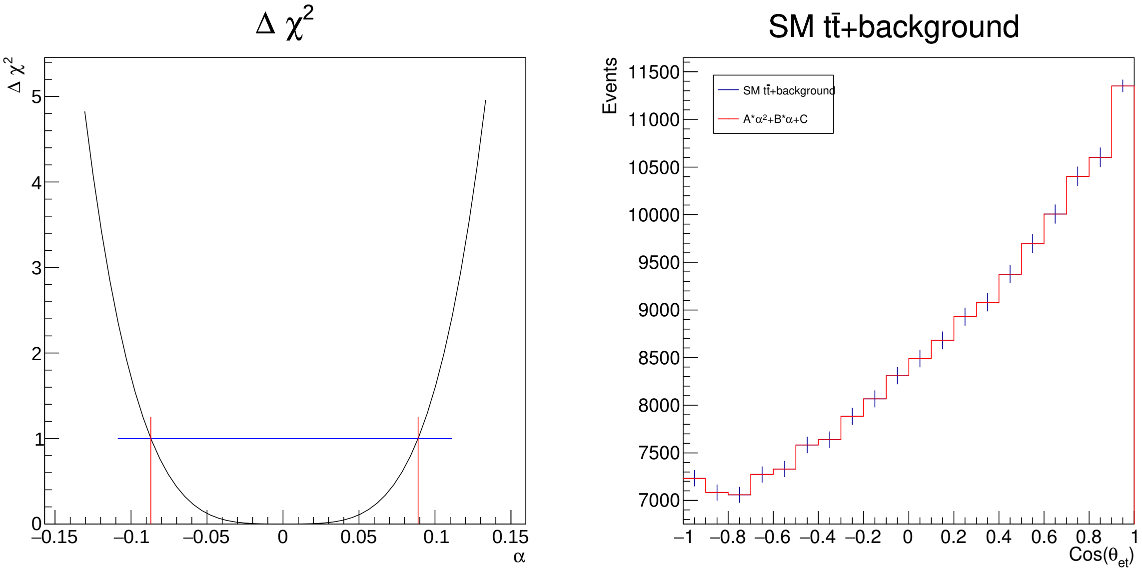


Figure 9.7.: θ_{et} for ta_ttA

the case of θ_{et} for ta_ttA. The corresponding figures for the remaining set of angles and couplings can be found in Appendix C. The left side shows that the minimum of $\Delta\chi^2$ coincides with the SM where α=0 and the right side shows how the signal region data coincides with the fit model. The red lines on the left side encloses the 1σ confidence intervals. The 1σ confidence intervals are found from the upper and lower values of α that yield $\Delta\chi^2 = 1$, since there is only one degree of freedom in the fit.

Observables of both the cross section and the angular distributions provide information that contributes to the sensitivity to anomalous top couplings. Excluding the information about the cross section, the fit can be applied to the angular distributions by normalising the fit model to the data by a factor $\int data / \int f(x)$, and including the information about the cross section the $f(\alpha)$ are normalised by their respective cross sections and used in calculating the fit model. The total cross sections obtained by Whizard are summarised in Table 9.2. Whizard provides the total cross section for top pair production. They are normalised by the branching fraction of 0.4 for the semileptonic channel such that they are normalised to the expected number of events found in the event selection.

In Table 9.3, all 70 confidence intervals achieved from the fit are summarised. For each coupling the confidence intervals are obtained for all five angular distributions twice where the cross section has been either included or excluded. Including the cross section normalisation improves the sensitivity for all couplings across all angles. $\theta^*_{W\ell}$ seems to have the least dependence on the cross section where for the couplings ta_ttA, tl_tbW_Re and vr_tbW_Re, the sensitivity improves very little. However for these confidence intervals, the $\theta^*_{W\ell}$ already has the best sensitivity of all the angles for these three couplings respectively. That is not the case for the other couplings where the sensitivity is comparable across the angles. In other cases the sensitivity improves by a factor of up to 50 (see $\cos(\theta_{et})$ for vr_ttZ). The best sensitivity expected at the FCC-ee is the sensitivity to tv_ttA followed by tr_tbW_Re, vl_tbW_Re and vr_ttZ. The experiment has the least sensitivity to vr_tbW_Re followed by tl_tbW_Re. As seen in Table 9.2, the two couplings tl_tbW_Re and vr_tbW_Re do change the total cross section for top pair produced events. Therefore including the cross section information in the fit should not affect the estimates

coupling	α	σ_{tot} [pb]
SM	0.0	0.4834
ta_ttA	+1.0	0.6505
	−1.0	0.6505
tv_ttA	+1.0	2.874
	−1.0	7.416
vr_ttZ	+1.0	0.7907
	−1.0	0.4854
tl_tbW_Re	+1.0	0.4834
	−1.0	0.4834
tr_tbW_Re	+1.0	2.591
	−1.0	4.129
vl_tbW_Re	+0.5	0.8308
	−0.5	0.4453
vr_tbW_Re	+1.0	0.4834
	−1.0	0.4834

Table 9.2.: Cross sections calculated by Whizard. Not normalised by the semileptonic branching ratio.

of the confidence intervals which it does. The explanation lies in Figure 9.2, which shows that the efficiency of the event selection differs the most from the efficiency of the SM for these two couplings.

coupling	σ_{tot}	θ_{et}	θ^*_{tb}	ϕ^*_{tb}	$\theta^*_{W\ell}$	$\phi^*_{W\ell}$
ta_ttA	Yes	-0.08698 to 0.0889	-0.09064 to 0.09425	-0.07231 to 0.07241	-0.05336 to 0.05394	-0.09048 to 0.0941
	No	-0.1599 to 0.1529	-0.7906 to 0.6069	-0.09011 to 0.09155	-0.0552 to 0.05534	-0.2485 to 0.2338
tl_tbW_Re	Yes	-0.1954 to 0.188	-0.1665 to 0.1698	-0.1888 to 0.1781	-0.07797 to 0.07842	-0.1434 to 0.1439
	No	-0.5892 to 0.6208	-0.1969 to 0.215	-0.3068 to 0.2918	-0.07843 to 0.07913	-0.155 to 0.1602
tr_tbW_Re	Yes	-0.001894 to 0.001922	-0.001858 to 0.001884	-0.001732 to 0.001752	-0.001808 to 0.001833	-0.001882 to 0.001909
	No	-0.04373 to 0.05687	-0.01955 to 0.01966	-0.00446 to 0.004435	-0.006526 to 0.006775	-0.02864 to 0.03579
tv_ttA	Yes	-0.0006407 to 0.0006424	-0.0006316 to 0.0006333	-0.0006368 to 0.0006385	-0.0005212 to 0.0005224	-0.0006395 to 0.0006411
	No	-0.01293 to 0.01298	-0.0264 to 0.02269	-0.006161 to 0.006088	-0.0009246 to 0.0009211	-0.008812 to 0.008416
vl_tbW_Re	Yes	-0.003745 to 0.003703	-0.003674 to 0.003631	-0.003576 to 0.003539	-0.002955 to 0.002898	-0.003738 to 0.003694
	No	-0.01547 to 0.01611	-0.0402 to 0.0388	-0.01129 to 0.01152	-0.004909 to 0.004655	-0.03073 to 0.03162
vr_tbW_Re	Yes	-0.3157 to 0.35	-0.3139 to 0.365	-0.1892 to 0.2101	-0.1185 to 0.1258	-0.2361 to 0.204
	No	-0.5007 to 0.4515	-0.4884 to 0.4586	-0.1964 to 0.2127	-0.1193 to 0.1258	-0.2546 to 0.2073
vr_ttZ	Yes	-0.007987 to 0.007884	-0.009419 to 0.009245	-0.008588 to 0.008472	-0.007714 to 0.007477	-0.009588 to 0.009407
	No	-0.01598 to 0.01597	-0.2185 to 0.344	-0.01901 to 0.01925	-0.01362 to 0.01235	-0.1732 to 0.1563

Table 9.3.: Confidence regions for each of the coupling. The first column denotes the name of the coupling, the second column denotes where the information of the cross section was included in the fit and the last columns denotes the confidence interval for the five angles of interest.

10. Conclusion

The results shown in Section 9.3 conclude the determined sensitivity to top anomalous couplings expected at the FCC-ee. The FCC-ee holds the potential to observe small but significant deviations with respect to the Standard Model predictions for the couplings of the top quark. The sensitivity was determined in the semileptonic channel for top pair produced events in simulated datasets generated in the experimental environment of the IDEA detector at the FCC-ee at $\sqrt{s} = 365\,\text{GeV}$. The method of fitting compares the angular distributions of the signal region to a parabolic fitting function whose parameters are determined from three angular distributions corresponding to three values of a single coupling. The 1σ confidence interval is determined by a least squares fit for each coupling separately. The sensitivity improves by including the information of the cross section in the fit. With real data acquired at the future e^+e^- collider FCC-ee, a study of the top anomalous couplings offers a fast way to observe BSM physics provided the sensitivity has the required reach. Even if no discovery would be made, it would not prove that the anomalous couplings are absent. It might just be that the effect is too small to be revealed by the experiment. With this analysis, we now have some knowledge about the required coupling size.

Within this analysis an event selection and event reconstruction has been performed. The reconstruction of jets is crucial for the event reconstruction and the jet studies shed light on additional work to be done for the FCC-ee. The performance of the different jet definitions depends on the experimental environment, so additional studies for optimising the choice of jet definition might be required at a later point. As part of this thesis, software tools for jet clustering and jet tagging have been developed and are fed into `FCCAnalyses` to be used in future works, where also more realistic efficiencies for the flavour tagging could be investigated. The sensitivity to top anomalous couplings has been improved by performing a kinematic fit in which the event reconstruction is made more accurate. In connection with the kinematic fitting, the `ABCfit++` software package was written with the intention to also contribute to future works. It is developed in a modern programming language which makes it easy to maintain and options to extend it with additional features remain open. It could also potentially be included as an external software package into `FCCAnalyses`.

This work paves the way for determining the sensitivity to top anomalous couplings from event simulation to statistical evaluation. However, as science is an iterative process the individual steps leave room for optimisation. As the feasibility studies for the FCC-ee project progress, further advancements will be made. Software tools will be developed, a technical design report will more accurately describe the experimental setup, fitting methods taking the full multidimensional kinematic information into account can be applied, and other channels can be investigated. All of which could potentially further improve on the sensitivity reach found in this study.

A. Figures for Jet Studies

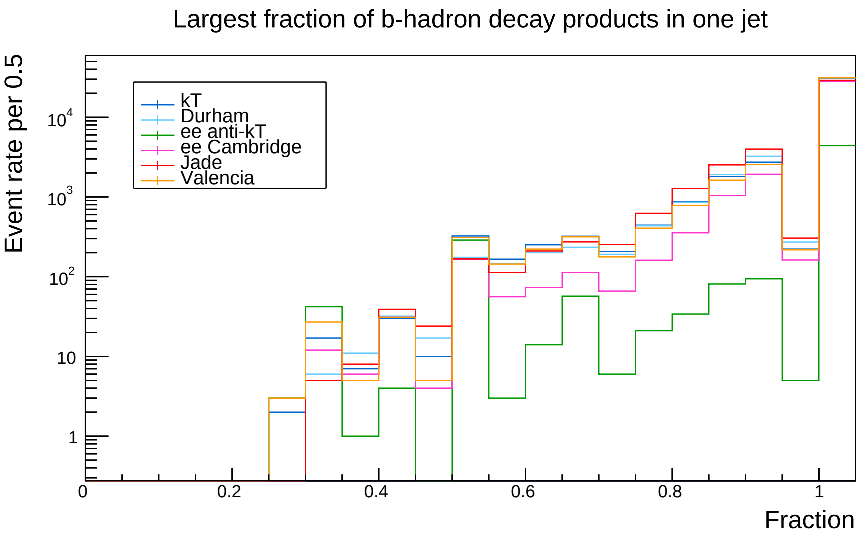


Figure A.1.: Distribution of b-hadron decay products in particle jets

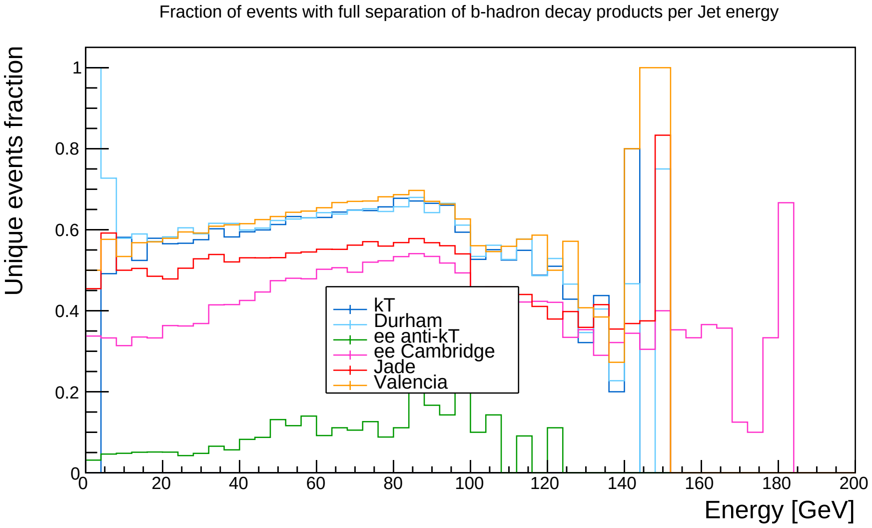
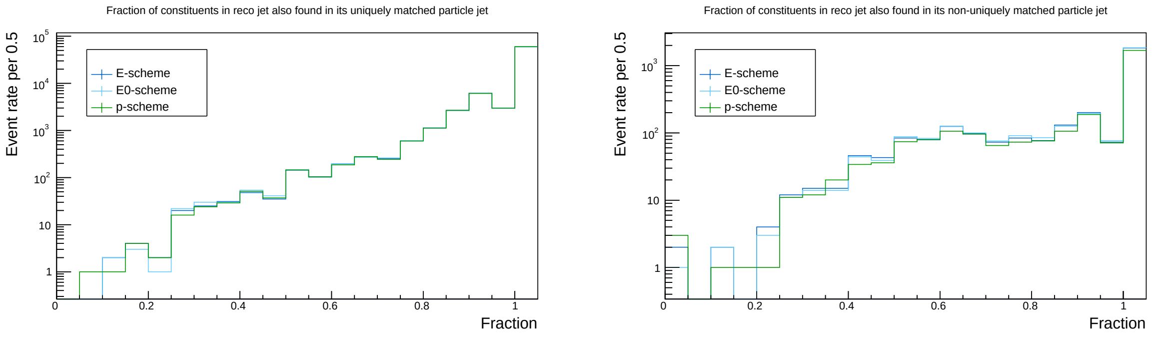


Figure A.2.: Full separation of b-hadron decay products in particle jets.

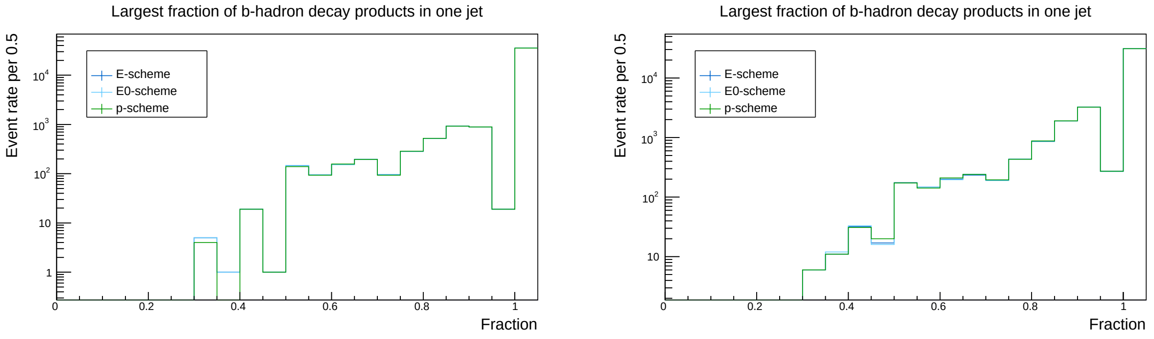
Figure A.3.: Matching constituents between reco and particle jets



(a) Events with unique matching

(b) Events with non-unique matching

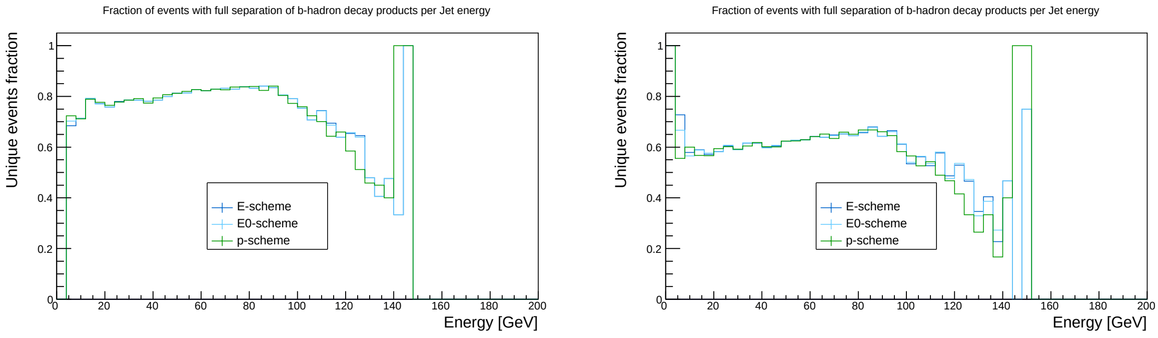
Figure A.5.: Distribution of b-hadron decay products



(a) Reco jets

(b) Particle jets

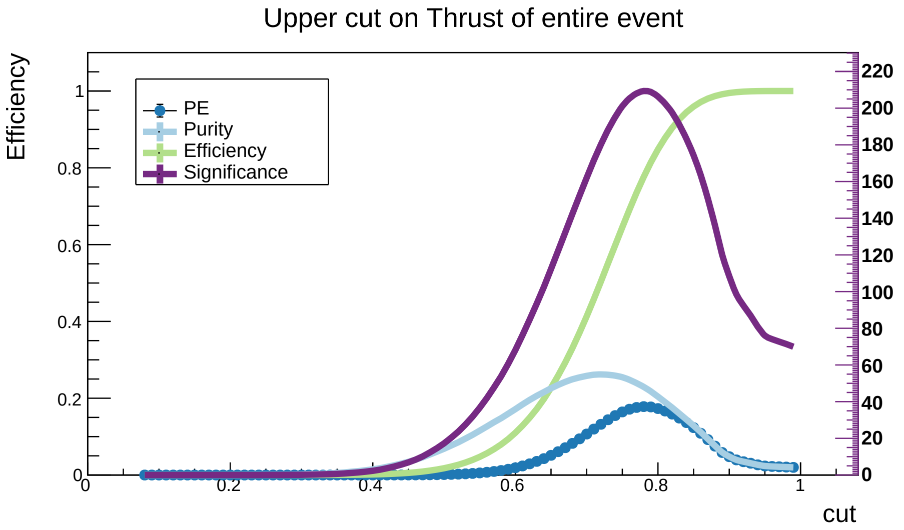
Figure A.7.: Full separation of b-hadron decay products



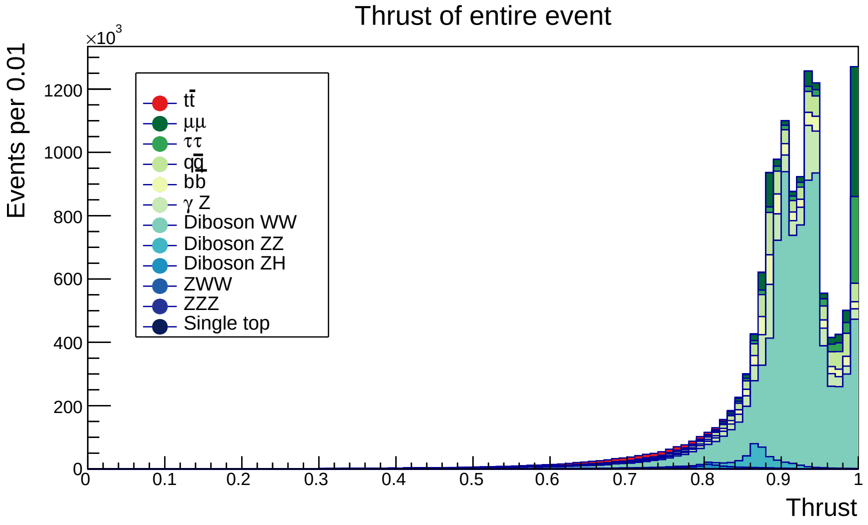
(a) Reco jets

(b) Particle jets

B. Figures for Event Selection

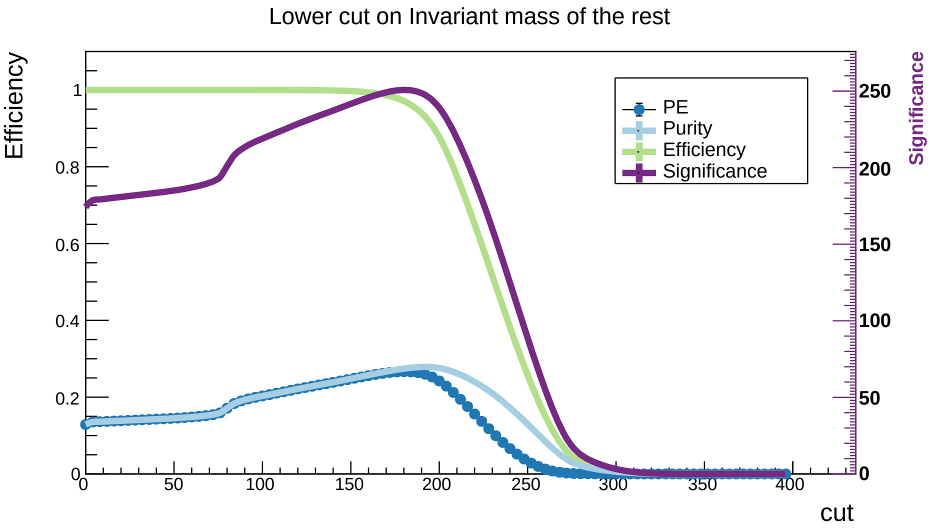


(a) Significance as well as efficiency, purity and their product

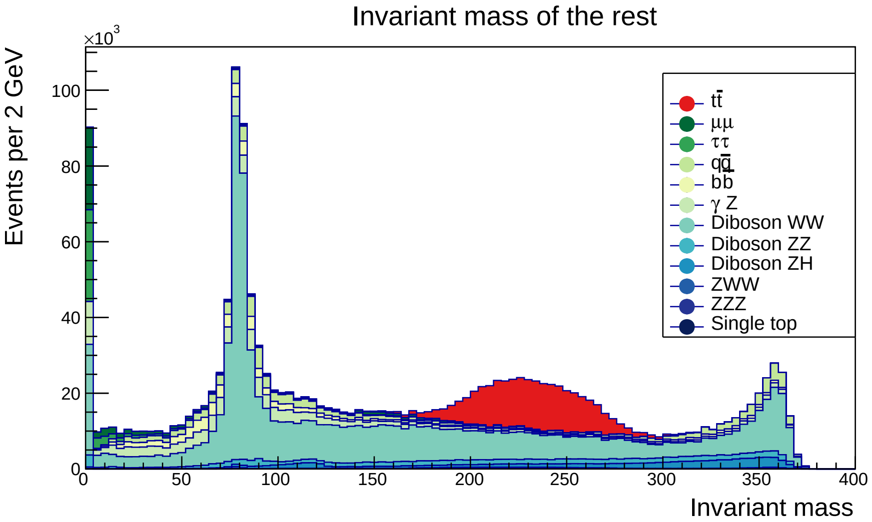


(b) Sum of distributions for the background and signal

Figure B.1.: Significance plot and sum of distributions for thrust on entire event prior to selecting an upper cut of Thrust < 0.85

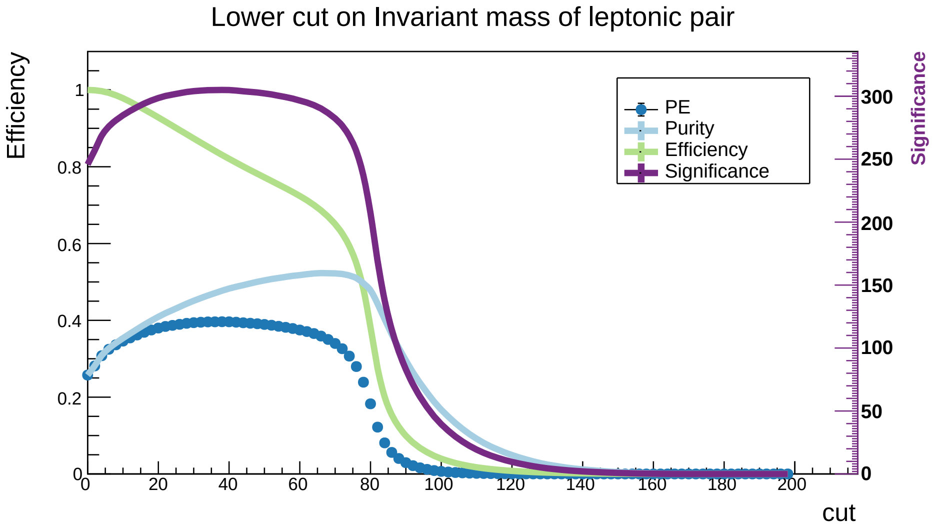


(a) Significance as well as efficiency, purity and their product

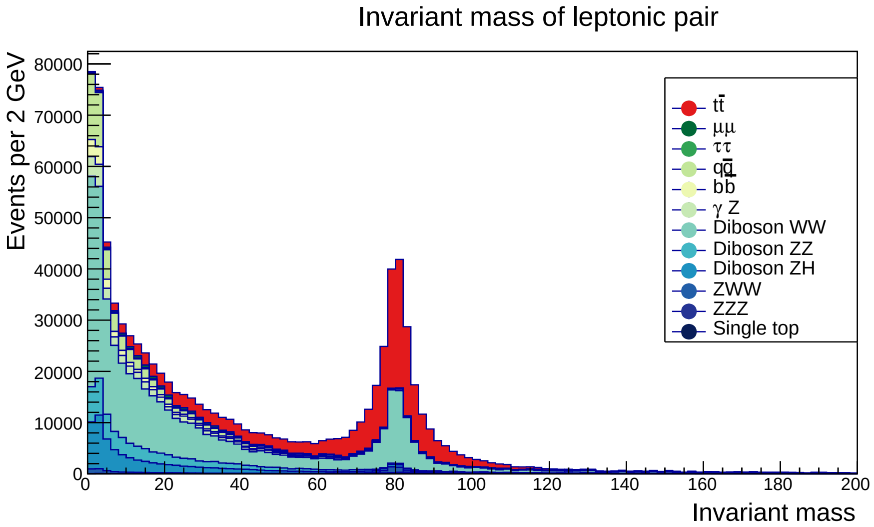


(b) Sum of distributions for the background and signal

Figure B.2.: Significance plot and sum of distributions for invariant mass of event excluding highest energy lepton prior to selecting a lower cut of $M(\text{rest}) > 160$ GeV

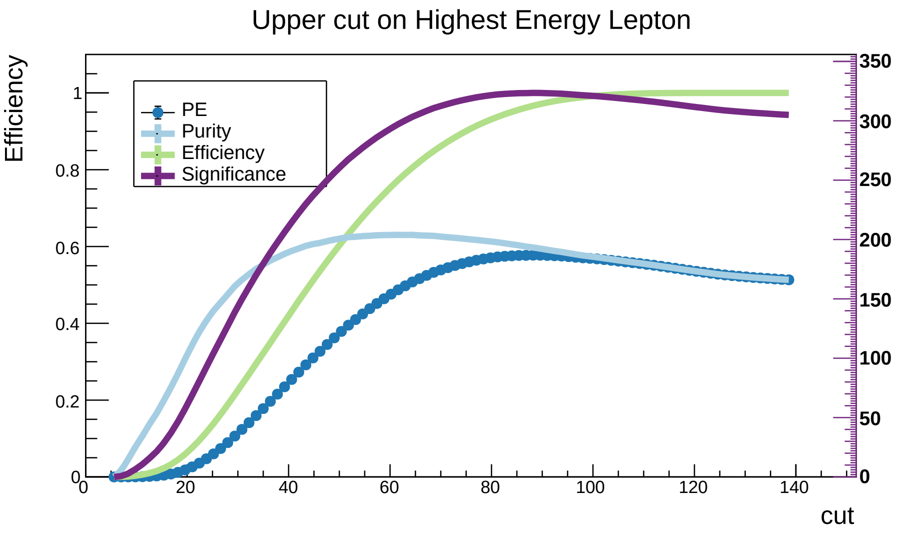


(a) Significance as well as efficiency, purity and their product

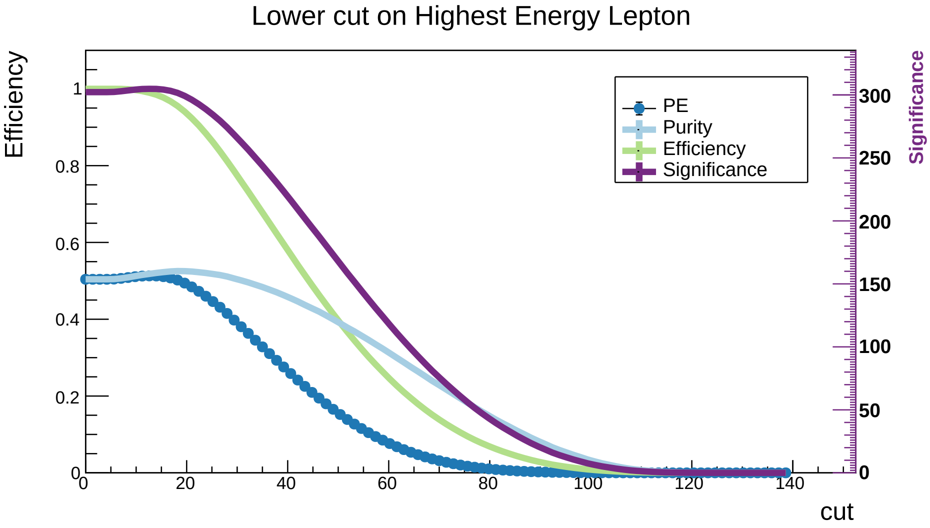


(b) Sum of distributions for the background and signal

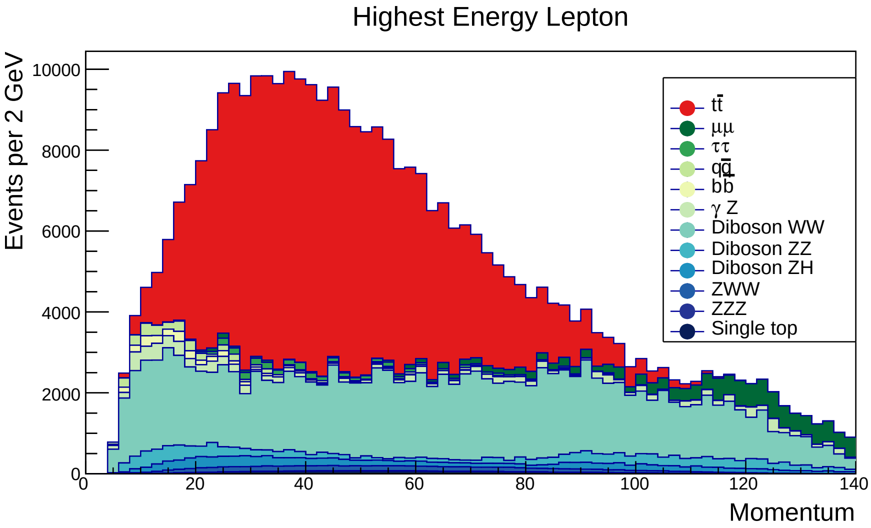
Figure B.3.: Significance plot and sum of distributions for invariant mass of lepton-neutrino pair prior to selecting a lower cut of $M(\ell_{HE}, cancelE) > 50$ GeV



(a) Significance as well as efficiency, purity and their product

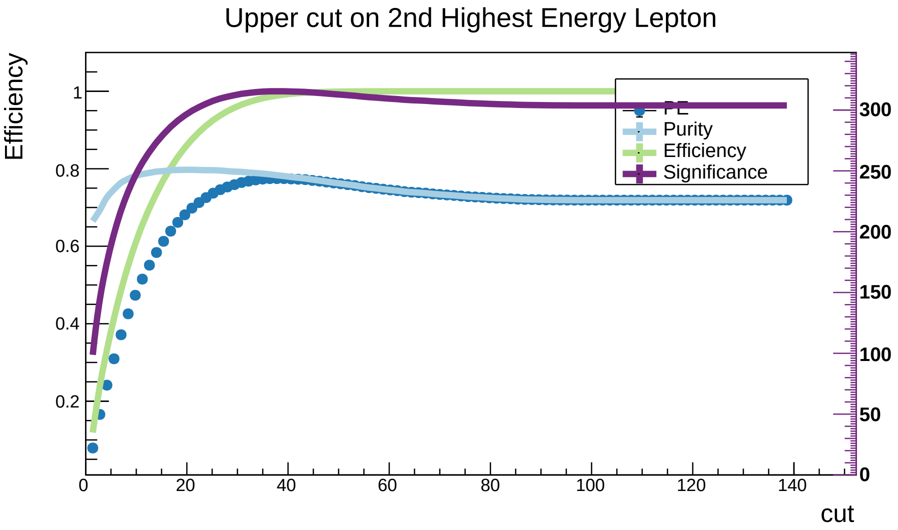


(b) Significance as well as efficiency, purity and their product

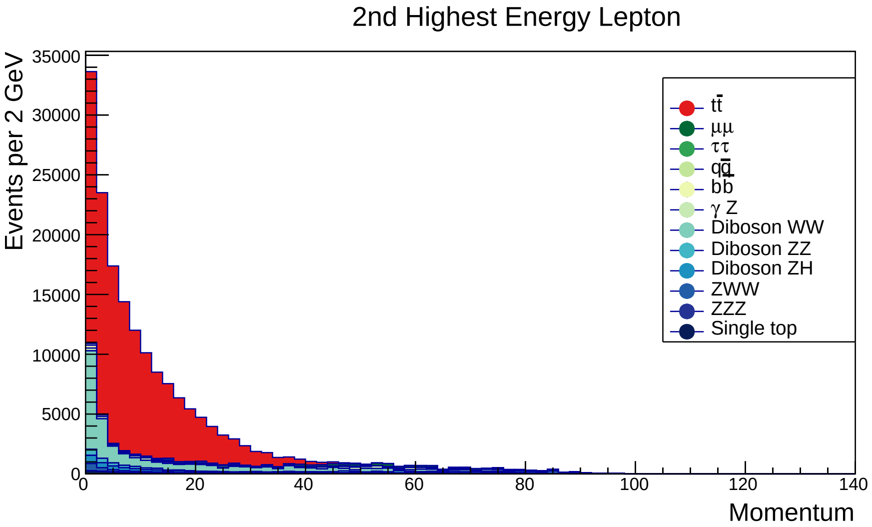


(c) Sum of distributions for the background and signal

Figure B.4.: Significance plot and sum of distributions for momentum of highest energy lepton prior to selecting an upper cut of $p_{\ell_{HE}} < 100$ GeV and a lower cut of $p_{\ell_{HE}} > 15$ GeV

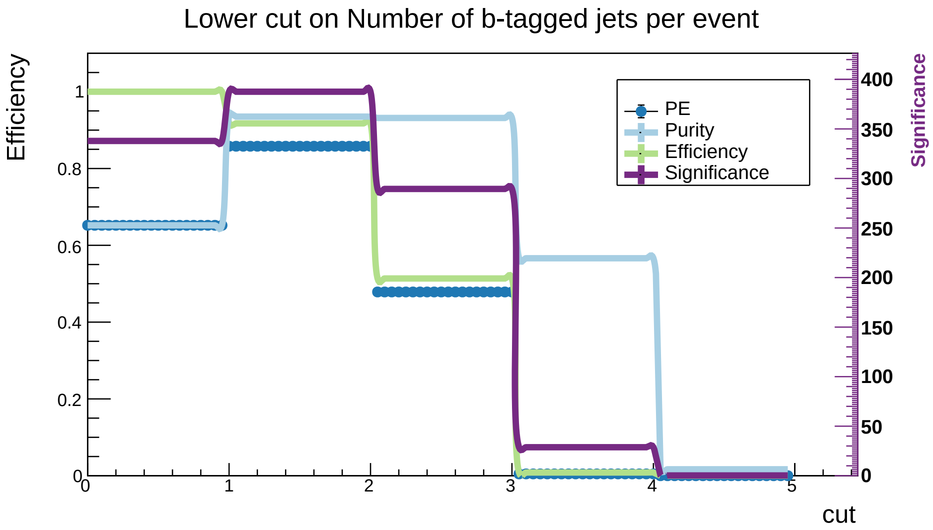


(a) Significance as well as efficiency, purity and their product

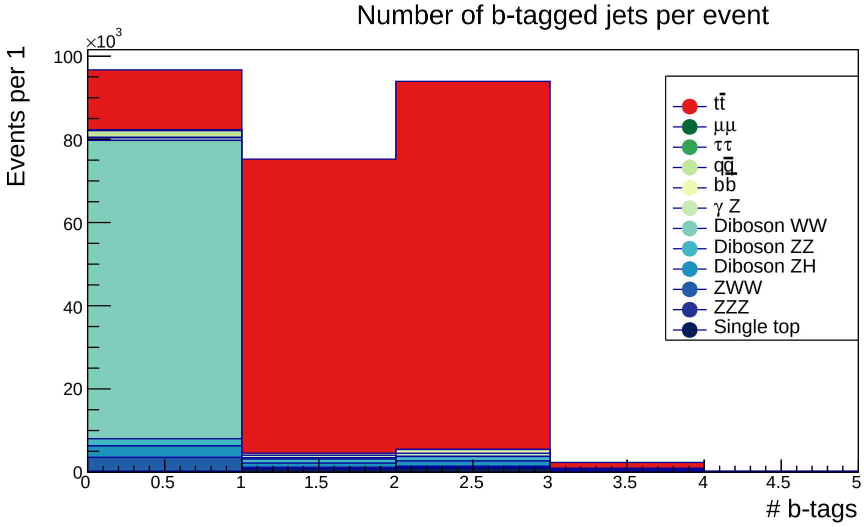


(b) Sum of distributions for the background and signal

Figure B.5.: Significance plot and sum of distributions for momentum of 2nd highest energy lepton prior to selecting an upper cut of $p_{\ell_{2nd_{HE}}} < 40$ GeV



(a) Significance as well as efficiency, purity and their product



(b) Sum of distributions for the background and signal

Figure B.6.: Significance plot and sum of distributions for number of b-tagged jets prior to selecting a lower cut of #b-tags > 0

C. Figures for Analysis

ta_ttA

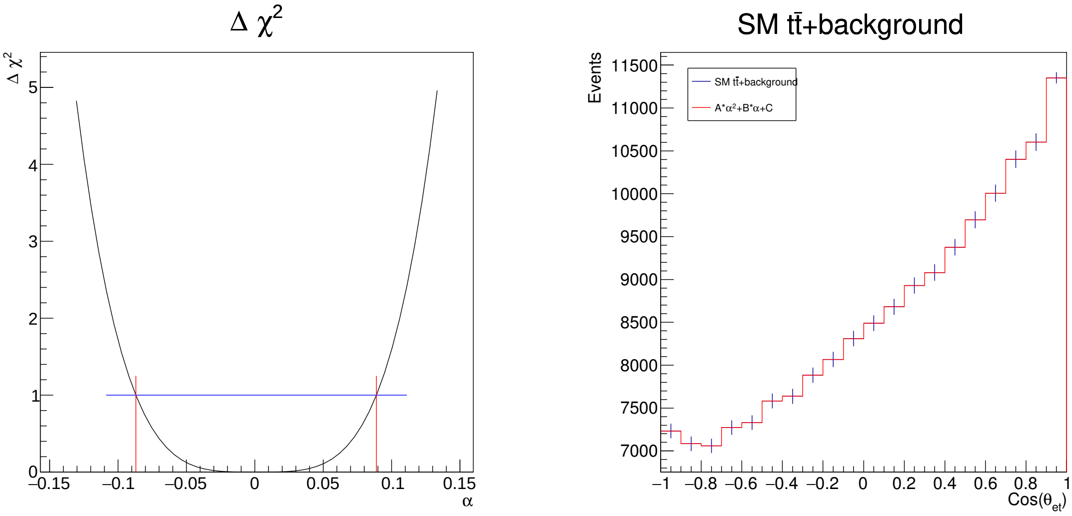


Figure C.1.: θ_{et} for ta_ttA

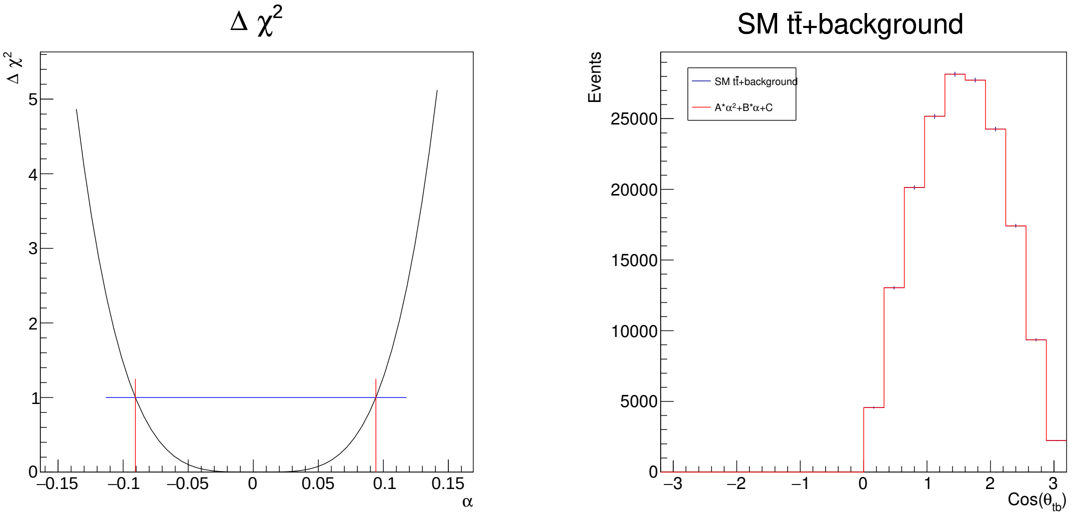


Figure C.2.: θ_{tb} for ta_ttA

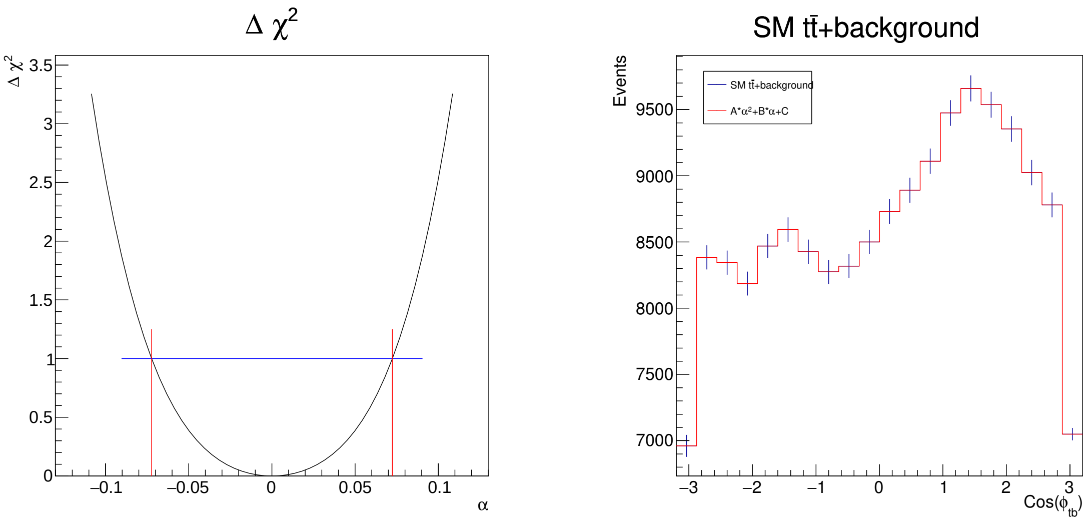


Figure C.3.: ϕ_{tb} for ta_ttA

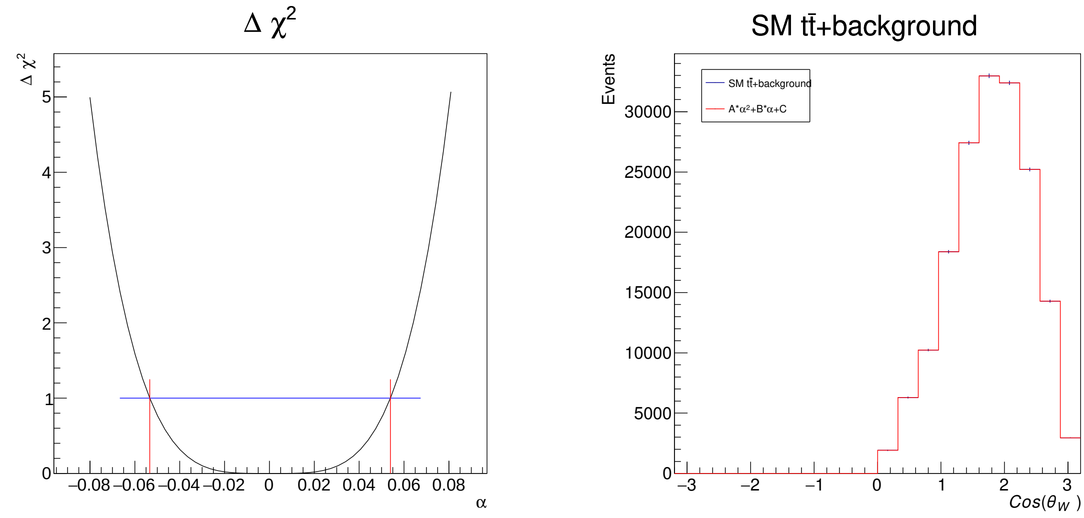


Figure C.4.: $\theta_{W\ell}$ for ta_ttA

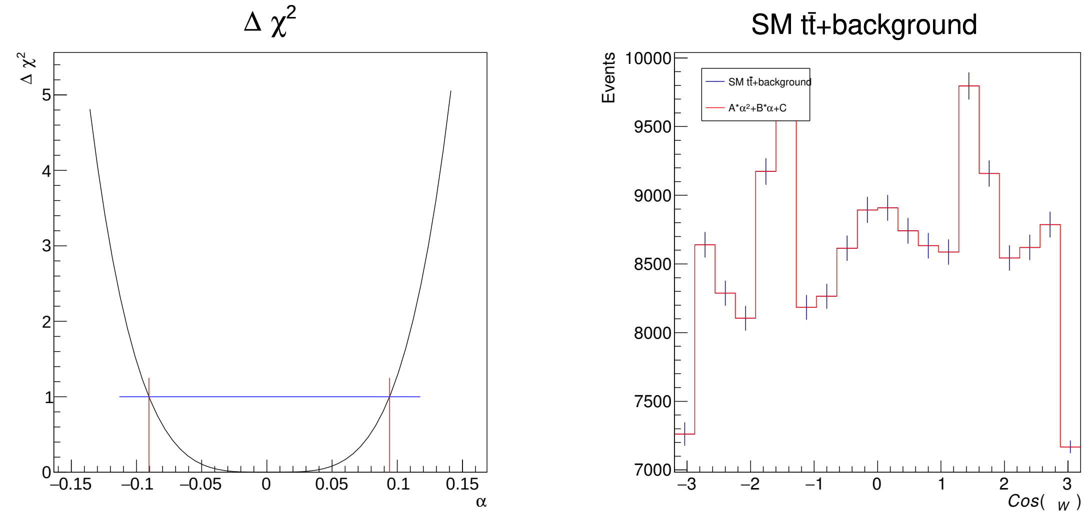


Figure C.5.: $\phi_{W\ell}$ for ta_ttA

tv_ttA

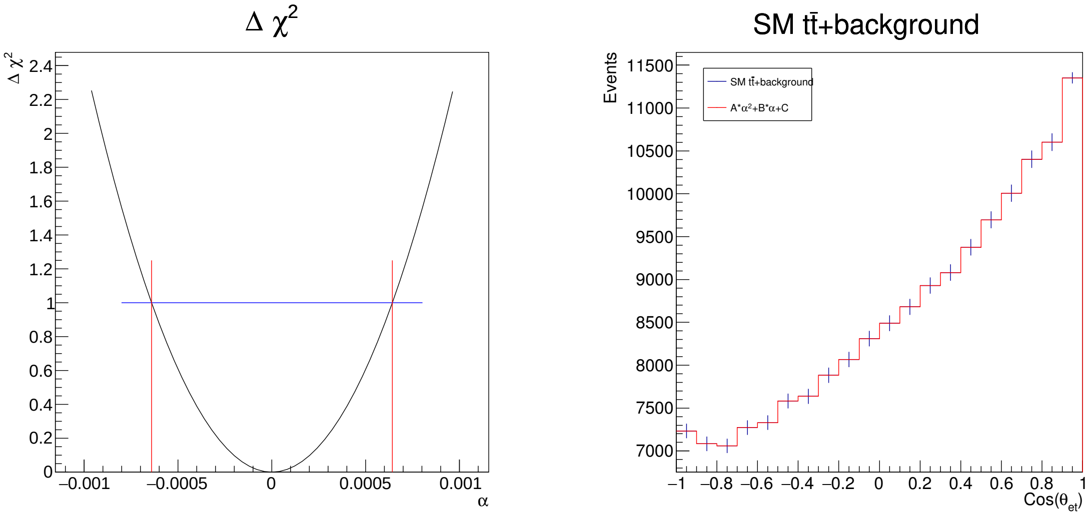


Figure C.6.: θ_{et} for tv_ttA

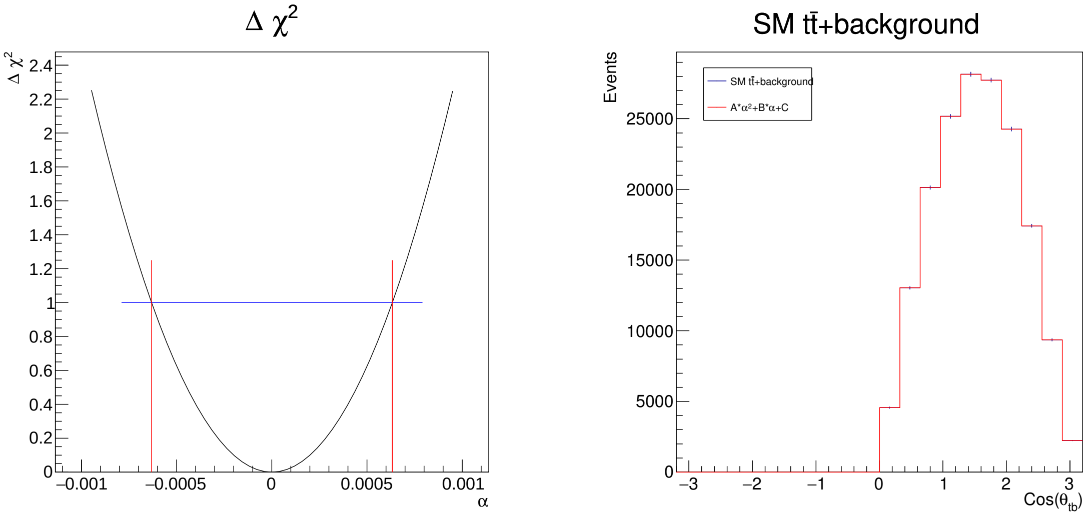


Figure C.7.: θ_{tb} for tv_ttA

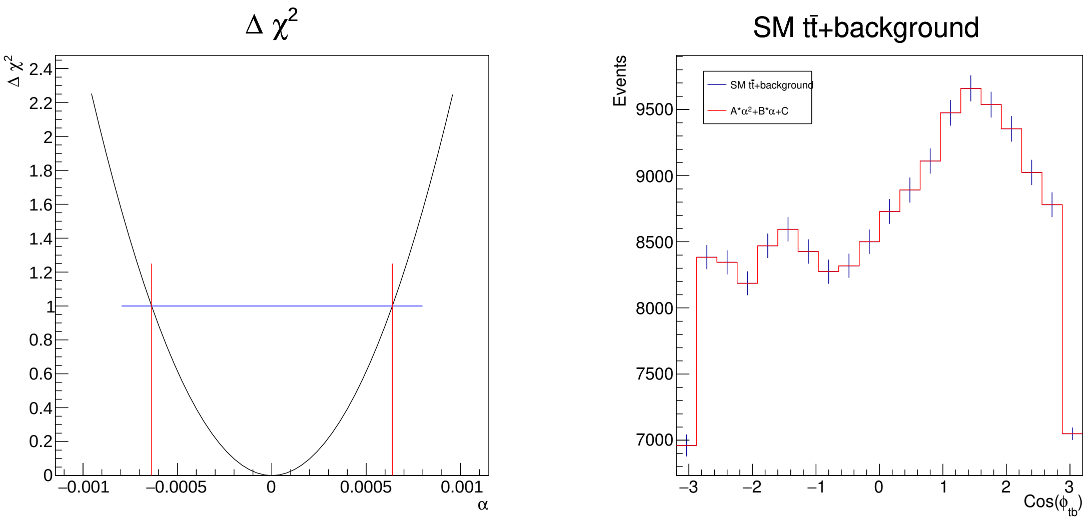


Figure C.8.: ϕ_{tb} for tv_ttA

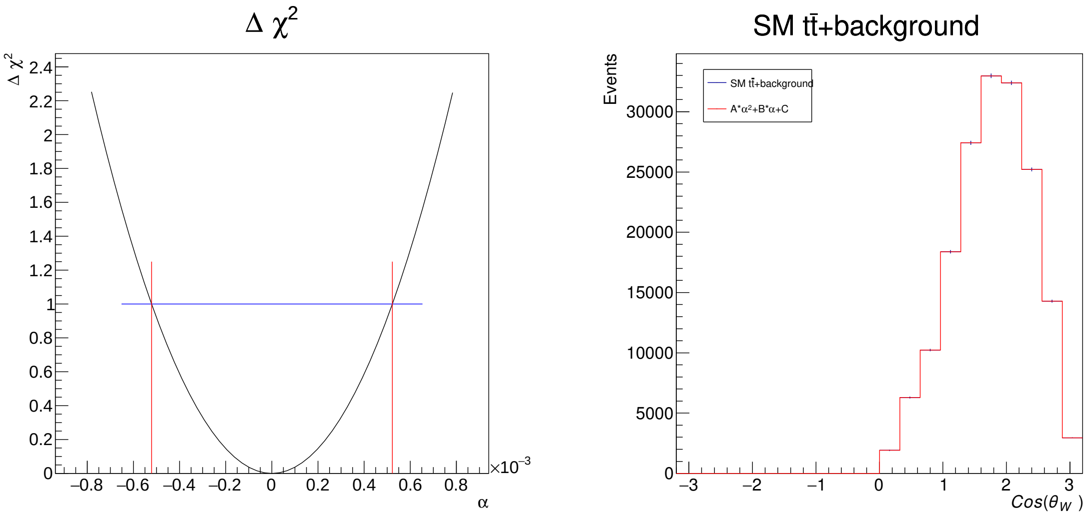


Figure C.9.: $\theta_{W\ell}$ for tv_ttA

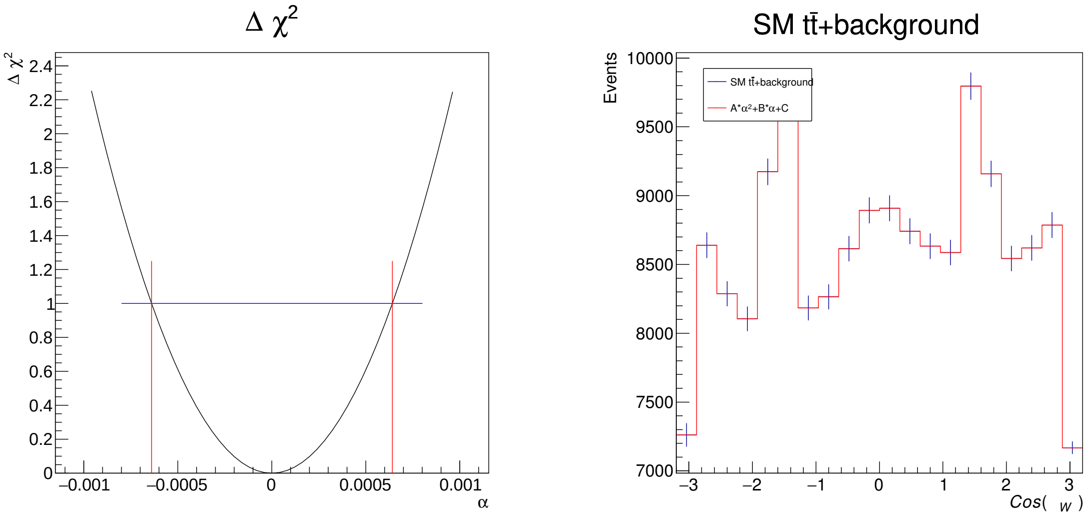


Figure C.10.: $\phi_{W\ell}$ for tv_ttA

vr_ttZ

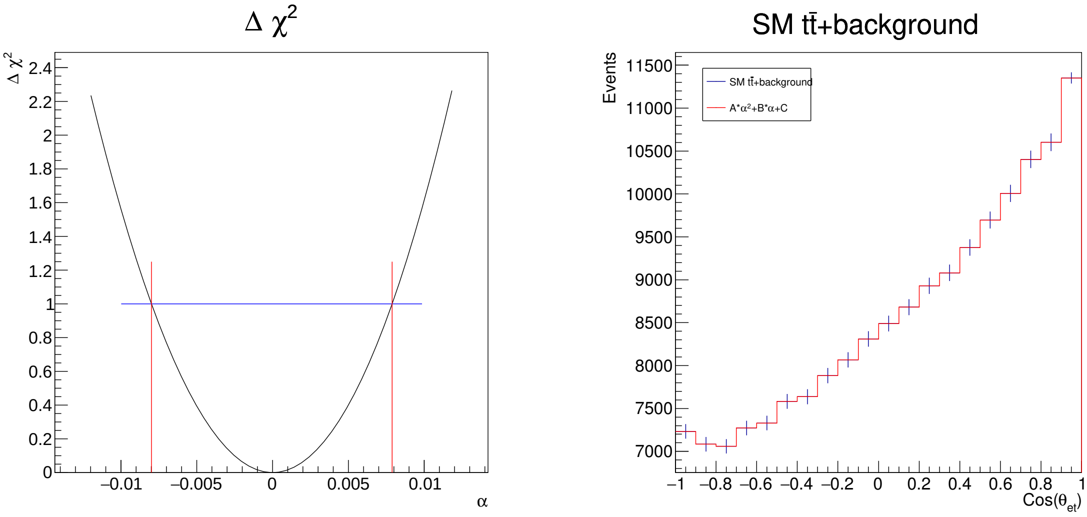


Figure C.11.: θ_{et} for vr_ttZ

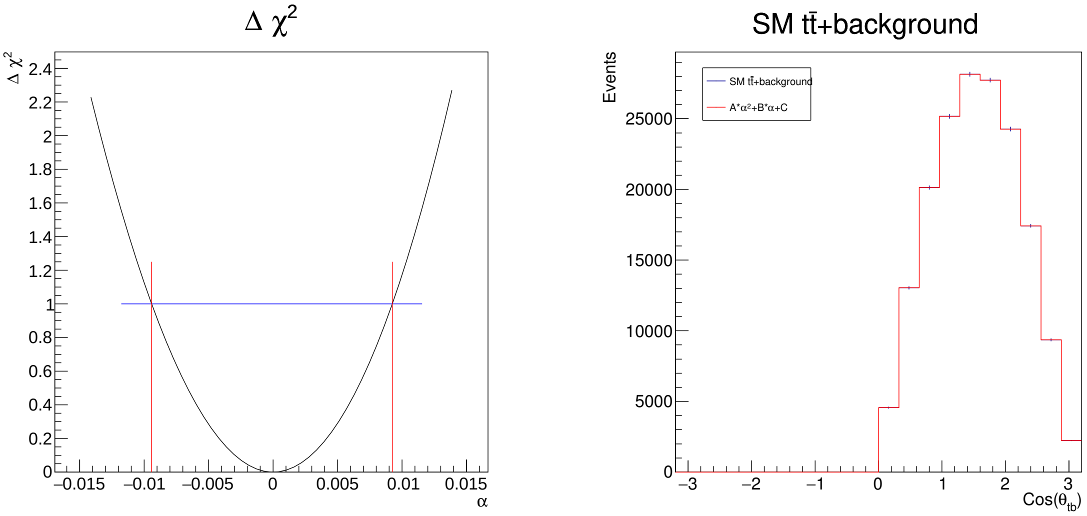


Figure C.12.: θ_{tb} for vr_ttZ

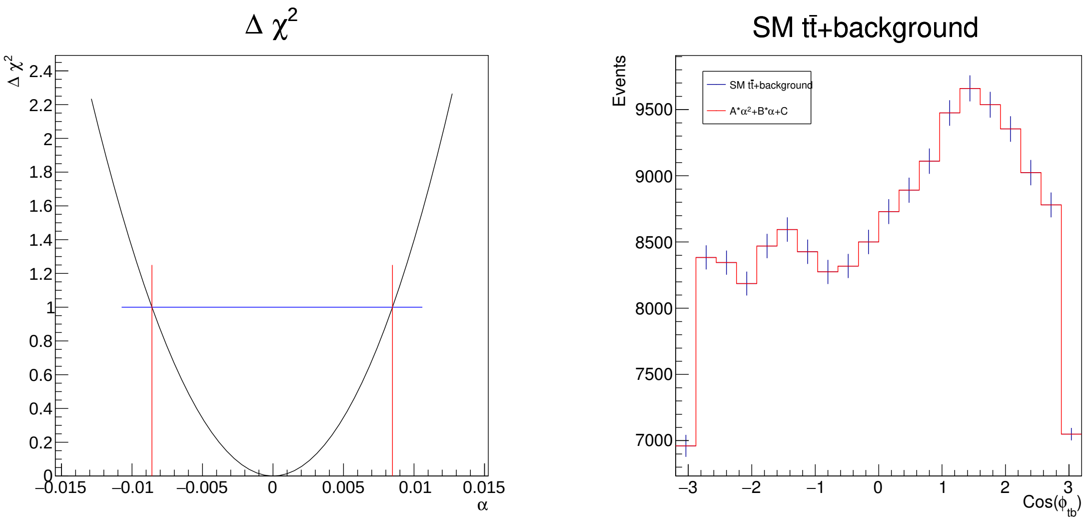


Figure C.13.: ϕ_{tb} for vr_ttZ

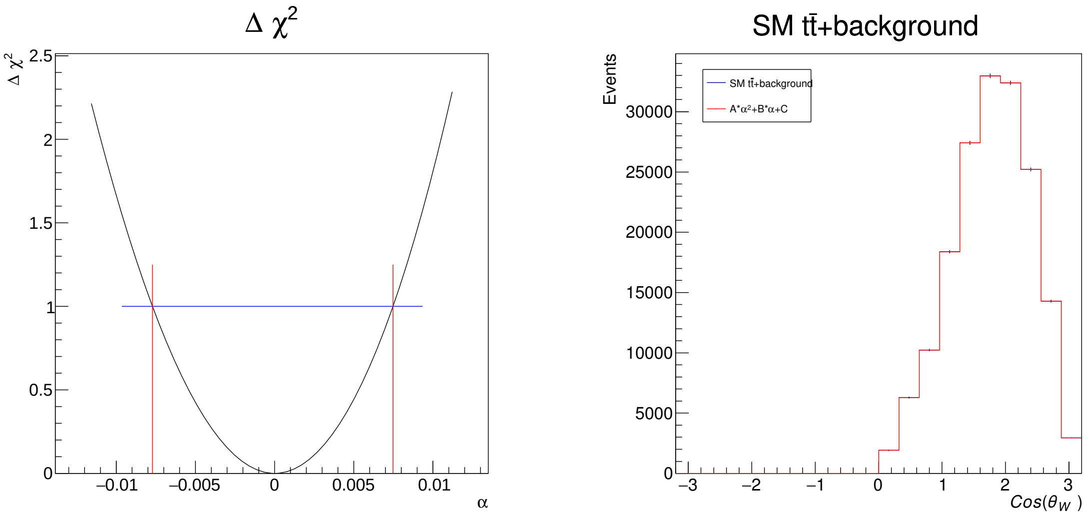


Figure C.14.: $\theta_{W\ell}$ for vr_ttZ

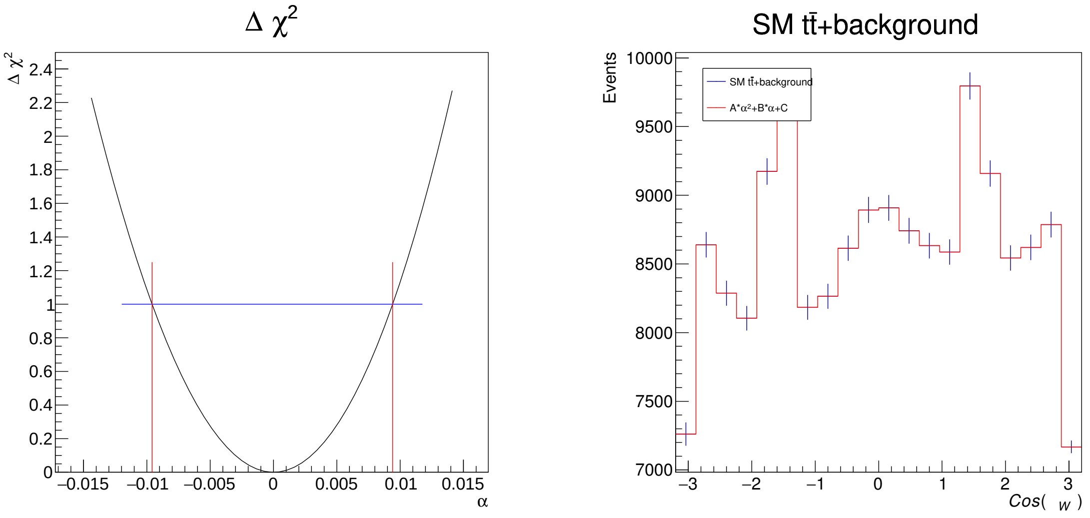


Figure C.15.: $\phi_{W\ell}$ for vr_ttZ

tl_tbW_Re

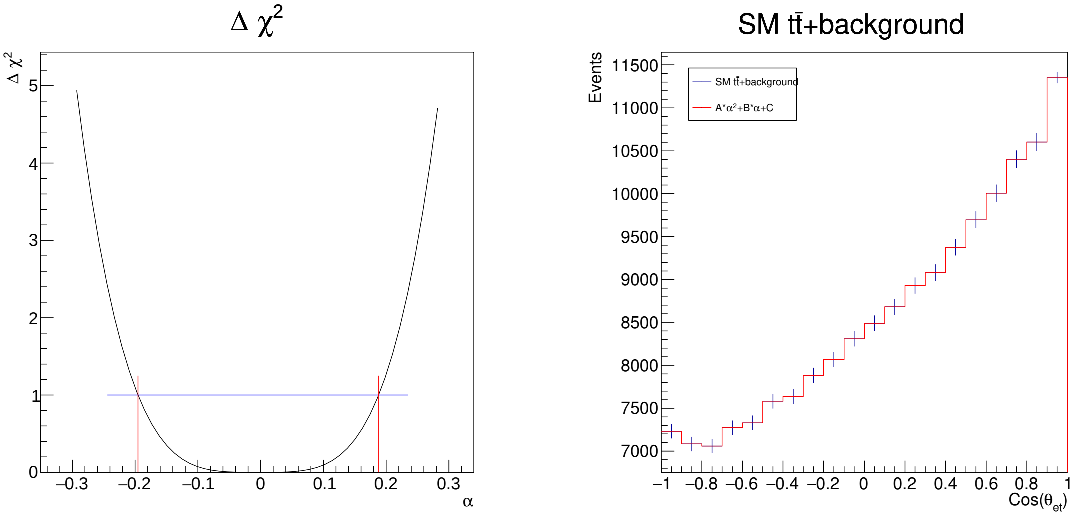


Figure C.16.: θ_{et} for tl_tbW_Re

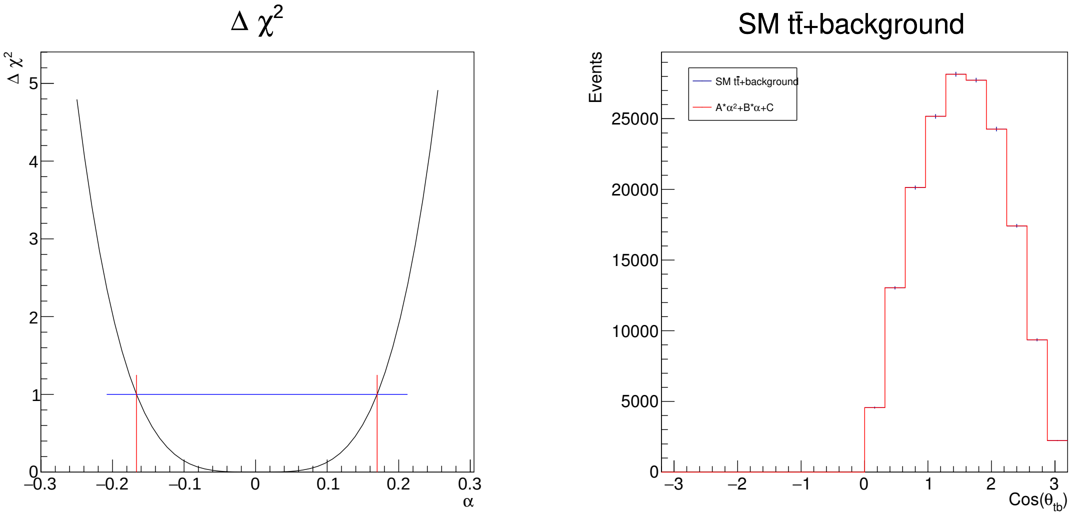


Figure C.17.: θ_{tb} for tl_tbW_Re

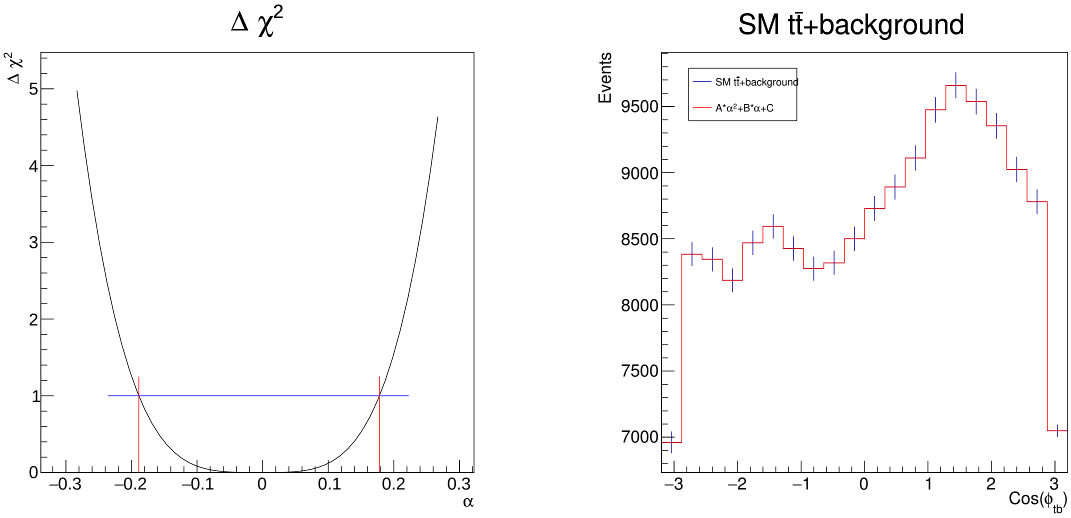


Figure C.18.: ϕ_{tb} for tl_tbW_Re

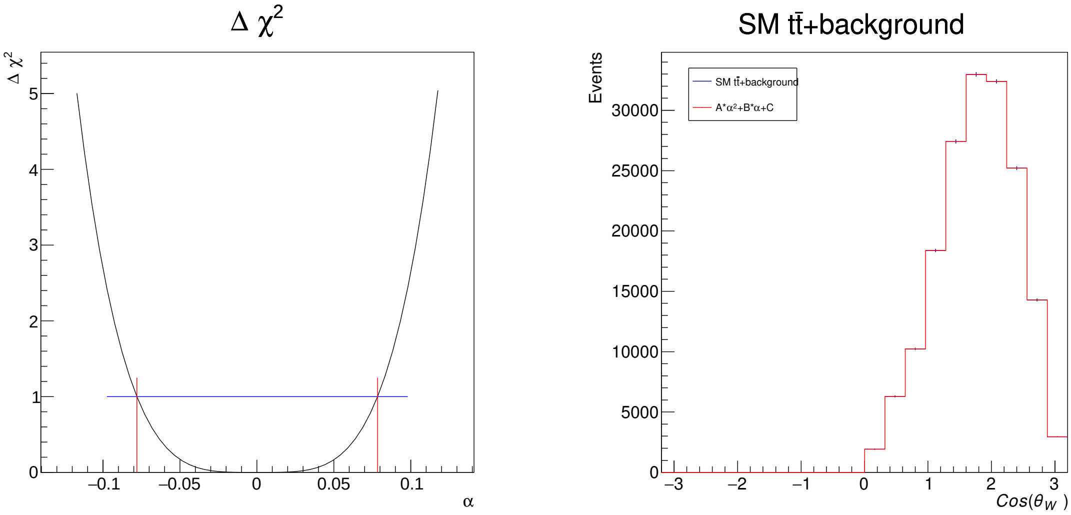


Figure C.19.: $\theta_{W\ell}$ for tl_tbW_Re

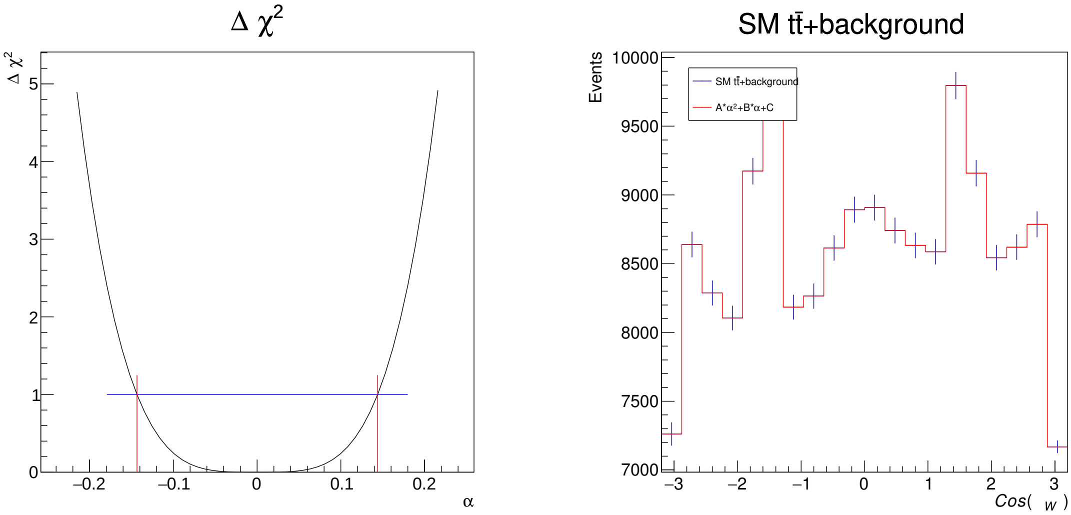


Figure C.20.: $\phi_{W\ell}$ for tl_tbW_Re

tr_tbW_Re

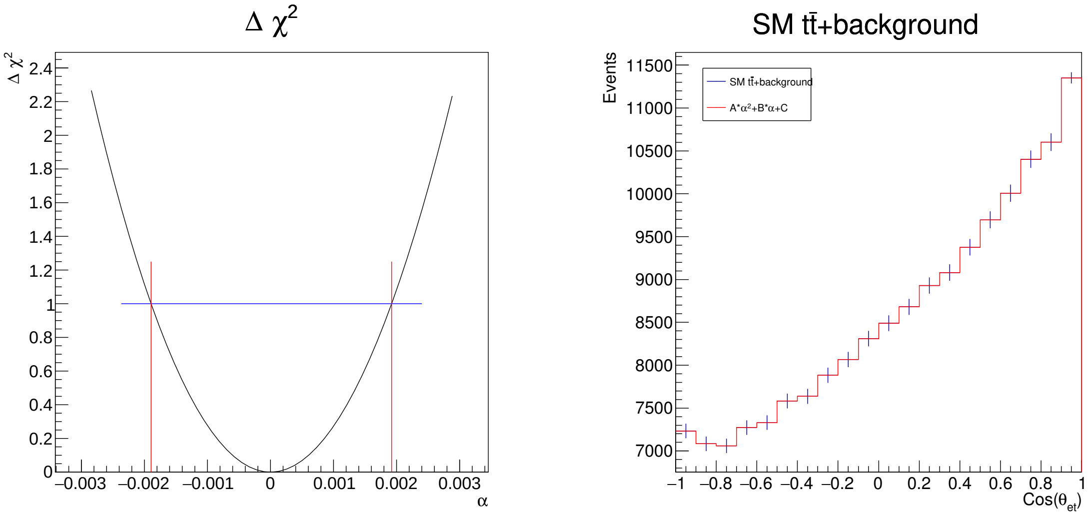


Figure C.21.: θ_{et} for tr_tbW_Re

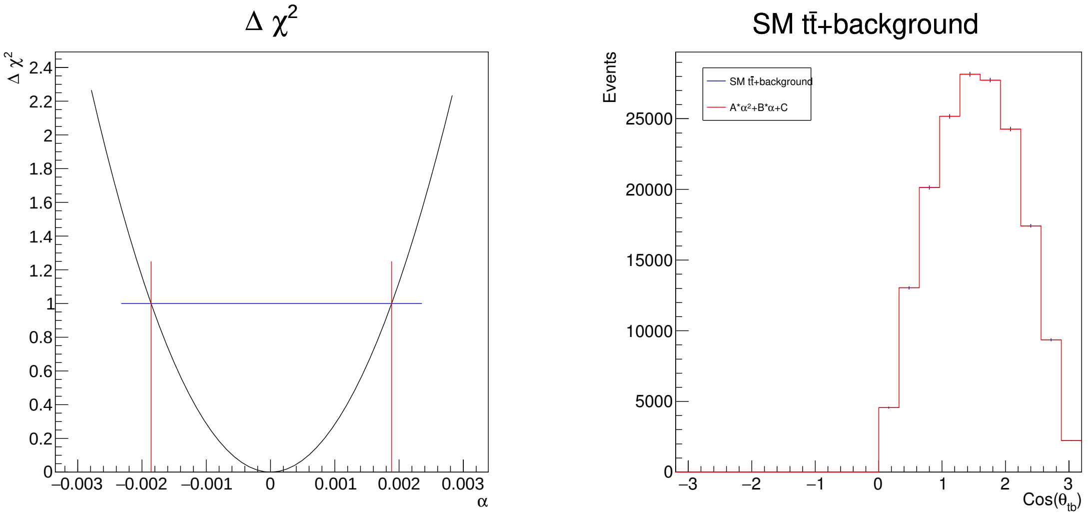


Figure C.22.: θ_{tb} for tr_tbW_Re

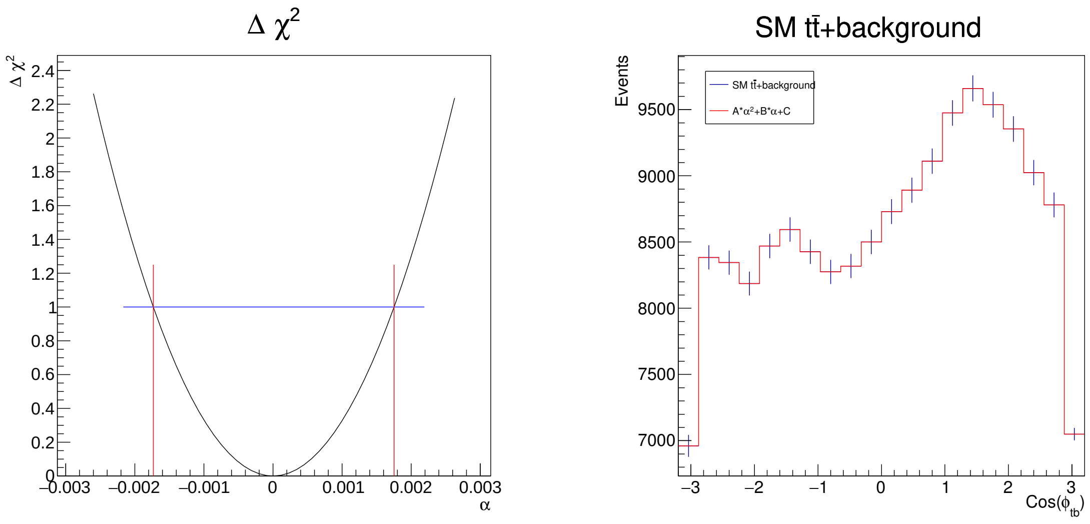


Figure C.23.: ϕ_{tb} for tr_tbW_Re

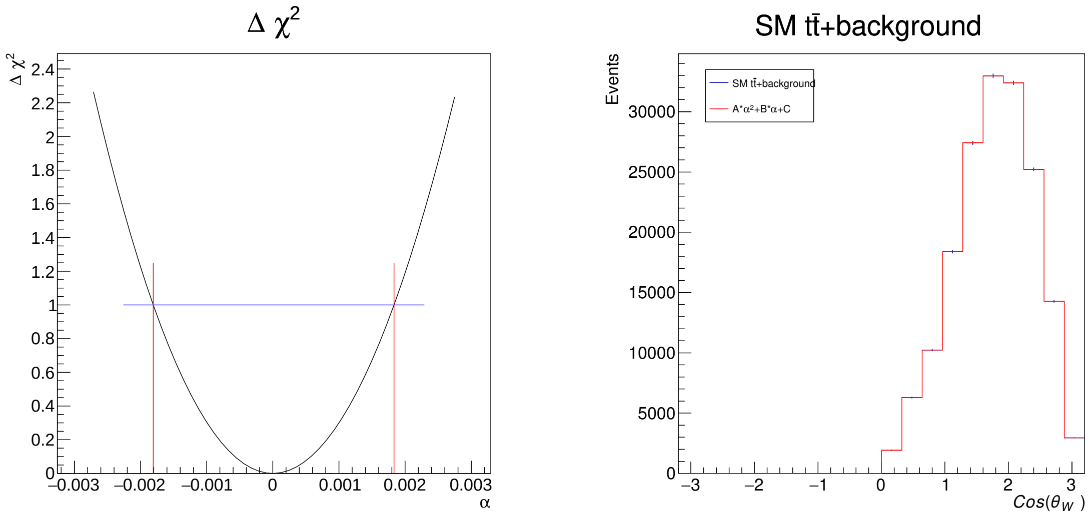


Figure C.24.: $\theta_{W\ell}$ for tr_tbW_Re

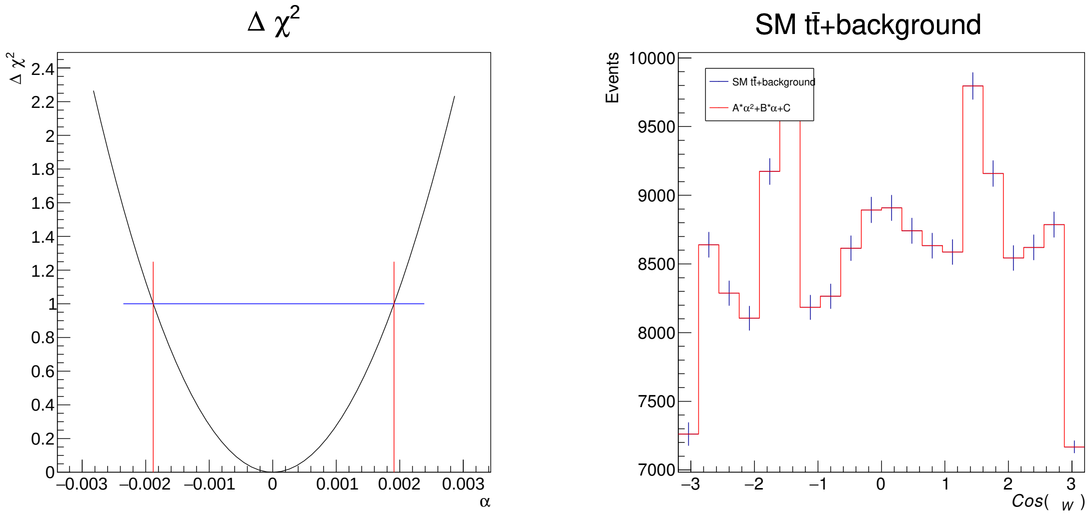


Figure C.25.: $\phi_{W\ell}$ for tr_tbW_Re

vl_tbW_Re

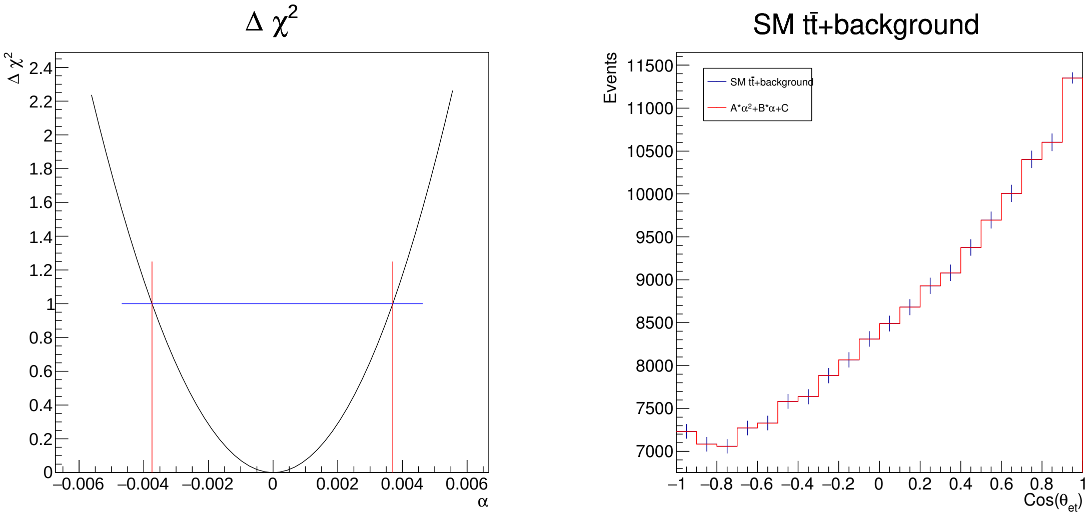


Figure C.26.: θ_{et} for vl_tbW_Re

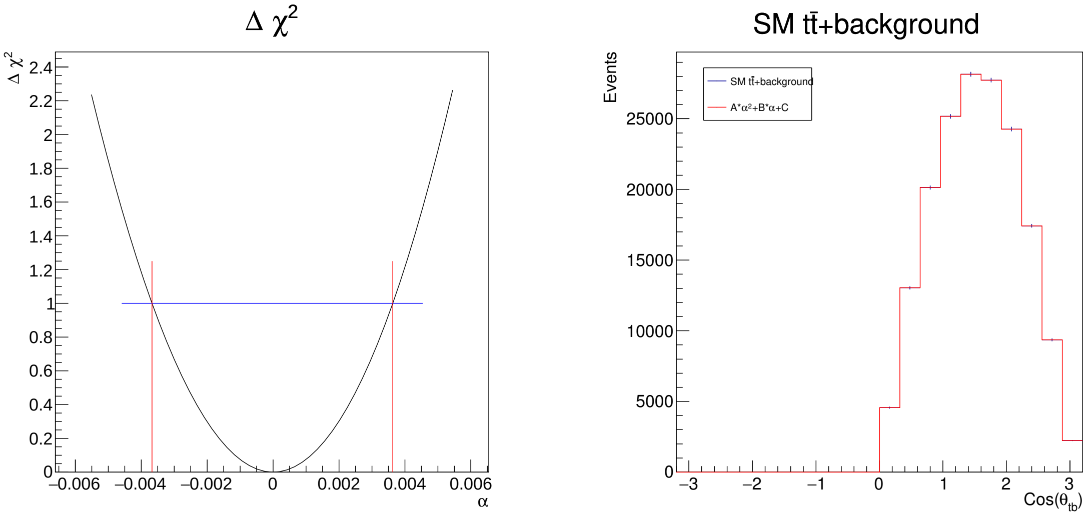


Figure C.27.: θ_{tb} for vl_tbW_Re

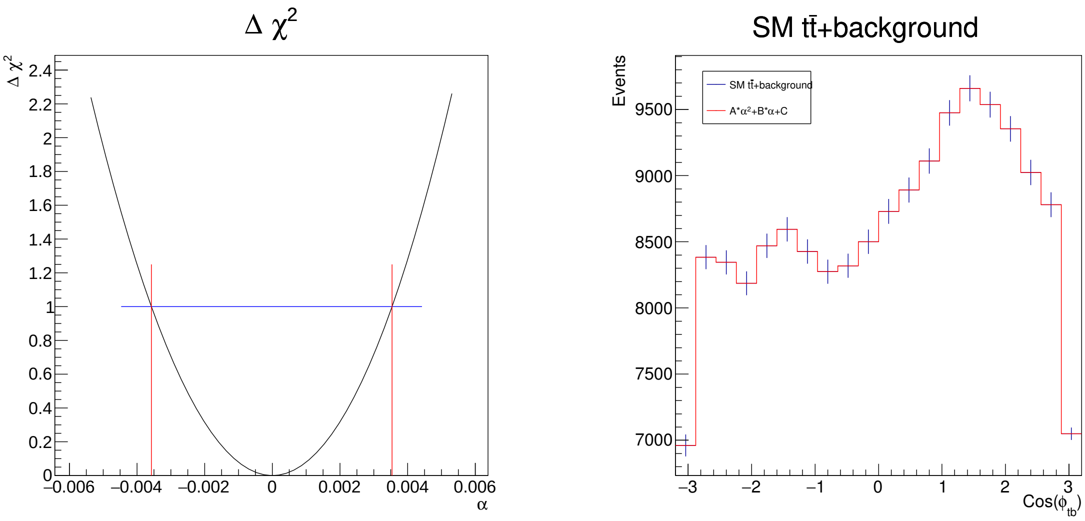


Figure C.28.: ϕ_{tb} for vl_tbW_Re

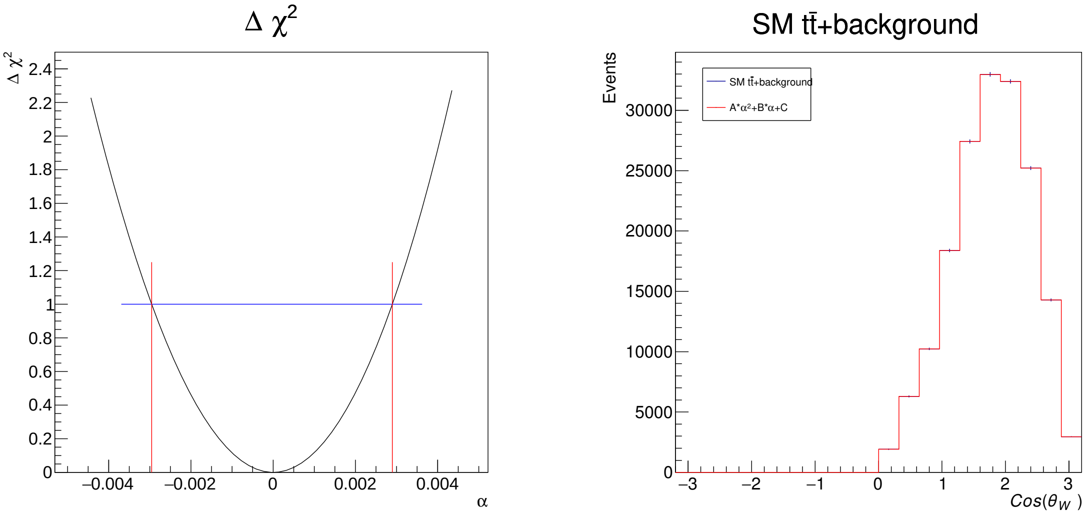


Figure C.29.: $\theta_{W\ell}$ for vl_tbW_Re

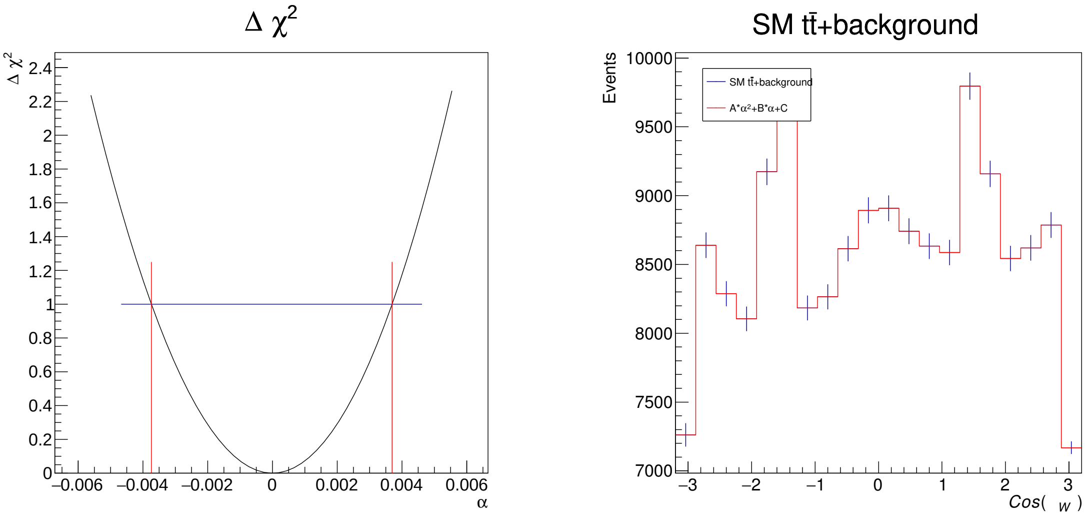


Figure C.30.: $\phi_{W\ell}$ for vl_tbW_Re

vr_tbW_Re

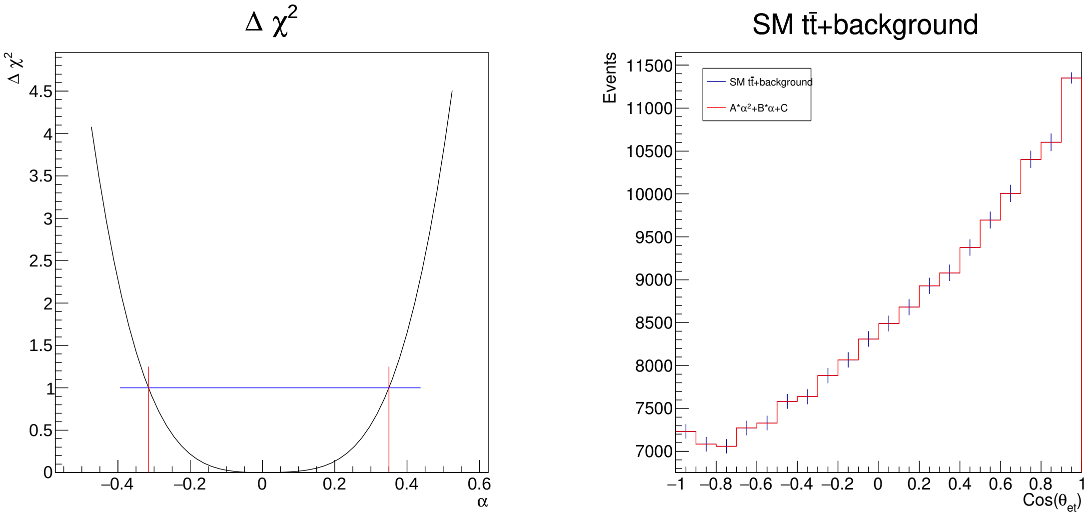


Figure C.31.: θ_{et} for vr_tbW_Re

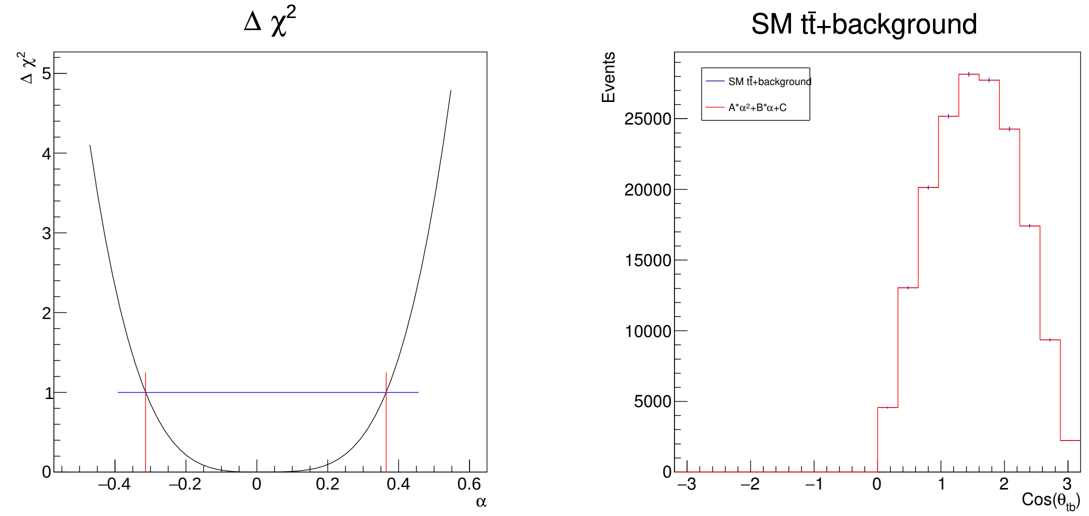


Figure C.32.: θ_{tb} for vr_tbW_Re

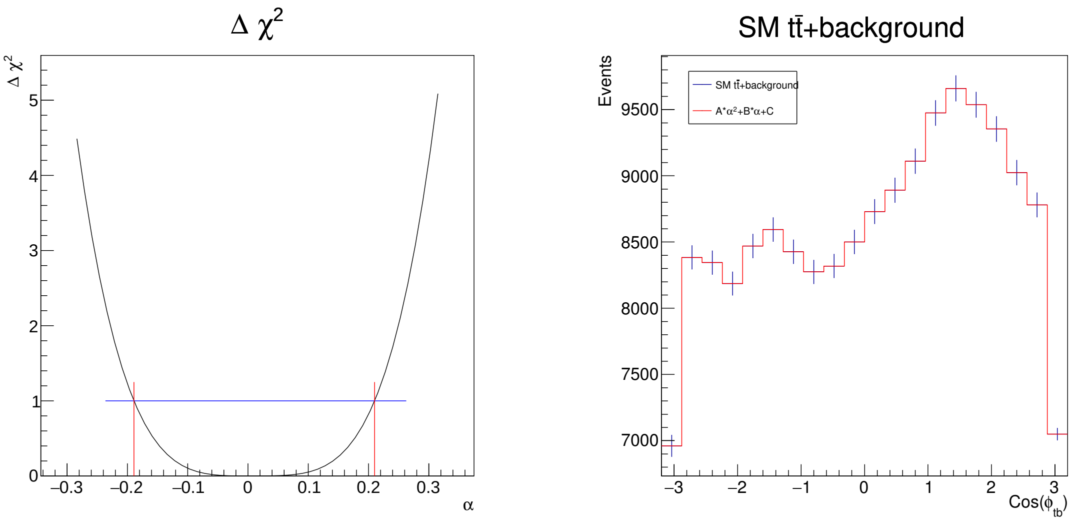


Figure C.33.: ϕ_{tb} for vr_tbW_Re

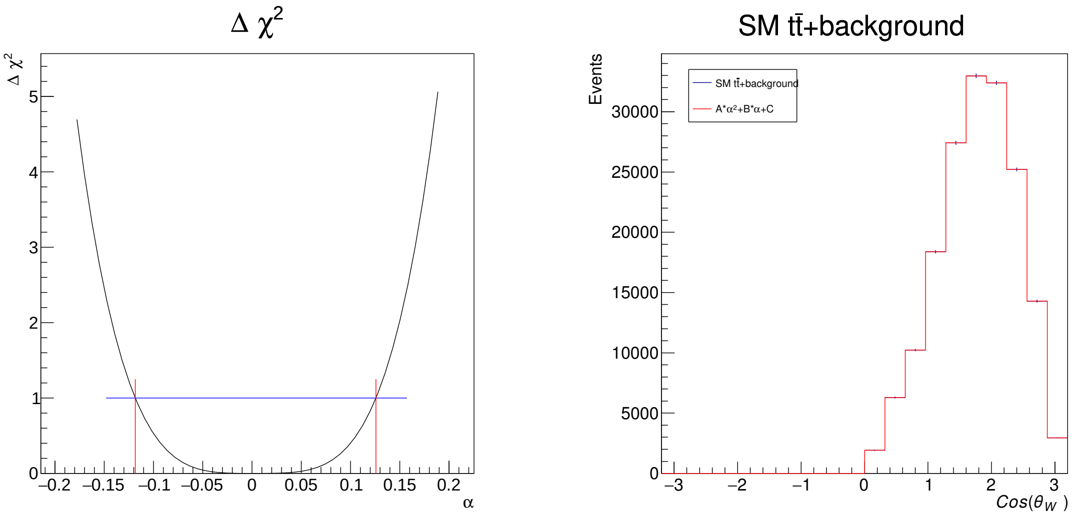


Figure C.34.: $\theta_{W\ell}$ for vr_tbW_Re

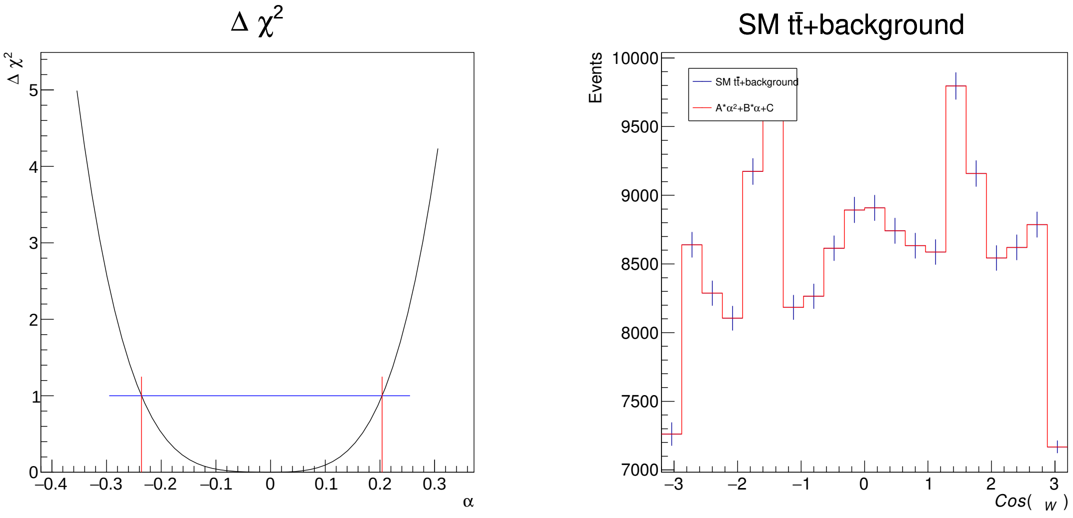


Figure C.35.: $\phi_{W\ell}$ for vr_tbW_Re

List of Figures

List of Tables

References

[1] G. Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2012.08.020`. URL: `https://www.sciencedirect.com/science/article/pii/S037026931200857X` (cit. on pp. 1, 4, 11).

[2] S. Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2012.08.021`. URL: `https://www.sciencedirect.com/science/article/pii/S0370269312008581` (cit. on pp. 1, 4, 11).

[3] *2021 Review of Particle Physics - Top Quark*. URL: `https://pdg.lbl.gov/2021/reviews/rpp2020-rev-top-quark.pdf` (visited on 06/16/2021) (cit. on p. 2).

[4] Mark Thomson. *Modern Particle Physics*. en (cit. on pp. 2, 5, 7, 10, 11, 24, 41).

[5] MissMJ and Cush. *Standard Model of Elementary Particles.svg - Wikipedia, the free encyclopedia*. URL: `https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg` (visited on 05/09/2021) (cit. on p. 3).

[6] Sheldon L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: `https://doi.org/10.1016/0029-5582(61)90469-2`. URL: `https://www.sciencedirect.com/science/article/pii/0029558261904692` (cit. on p. 4).

[7] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: `10.1103/PhysRevLett.19.1264`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.19.1264` (cit. on p. 4).

[8] A. M. Sirunyan et al. “Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Journal of High Energy Physics* 2018.11 (Nov. 2018). ISSN: 1029-8479. DOI: `10.1007/jhep11(2018)185`. URL: `http://dx.doi.org/10.1007/JHEP11(2018)185` (cit. on p. 11).

[9] G. Aad et al. “Measurement of Higgs boson production in the diphoton decay channel in *pp* collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”. In: *Physical Review D* 90.11 (Dec. 2014). ISSN: 1550-2368. DOI: `10.1103/physrevd.90.112015`. URL: `http://dx.doi.org/10.1103/PhysRevD.90.112015` (cit. on p. 11).

[10] V. Khachatryan et al. “Observation of the diphoton decay of the Higgs boson and measurement of its properties”. In: *The European Physical Journal C* 74.10 (Oct. 2014). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-014-3076-z`. URL: `http://dx.doi.org/10.1140/epjc/s10052-014-3076-z` (cit. on p. 11).

[11] Andrew Cohen. *Theoretical Concepts in Particle Physics - 5/5. Theoretical Concepts in Particle Physics - 5/5.* July 2019. URL: `https://cds.cern.ch/record/2682610` (cit. on p. 13).

[12] Aneesh V. Manohar. *Introduction to Effective Field Theories.* 2018. arXiv: `1804.05863 [hep-ph]` (cit. on p. 13).

[13] B. Grzadkowski et al. "Dimension-six terms in the Standard Model Lagrangian". In: *Journal of High Energy Physics* 2010.10 (Oct. 2010). ISSN: 1029-8479. DOI: `10.1007/jhep10(2010)085`. URL: `http://dx.doi.org/10.1007/JHEP10(2010)085` (cit. on p. 13).

[14] J.A. Aguilar-Saavedra. "A minimal set of top anomalous couplings". In: *Nuclear Physics B* 812.1-2 (May 2009), pp. 181–204. ISSN: 0550-3213. DOI: `10.1016/j.nuclphysb.2008.12.012`. URL: `http://dx.doi.org/10.1016/j.nuclphysb.2008.12.012` (cit. on pp. 13–15).

[15] the FCC Collaboration et al. "FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2". en. In: *The European Physical Journal Special Topics* 228.2 (June 2019), pp. 261–623. ISSN: 1951-6355, 1951-6401. DOI: `10.1140/epjst/e2019-900045-4`. URL: `http://link.springer.com/10.1140/epjst/e2019-900045-4` (cit. on pp. 17, 19, 20, 24).

[16] Mogens Dam. *Experimental Physics at Lepton Colliders.* CERN Summer Student Lecture. July 2019. URL: `https://indico.cern.ch/event/817612/attachments/1865560/3114919/LeptonColliders2019_2.pdf` (cit. on p. 18).

[17] Patrick Janot. "Top-quark electroweak couplings at the FCC-ee". en. In: *Journal of High Energy Physics* 2015.4 (Apr. 2015), p. 182. ISSN: 1029-8479. DOI: `10.1007/JHEP04(2015)182`. URL: `http://link.springer.com/10.1007/JHEP04(2015)182` (visited on 09/04/2020) (cit. on p. 19).

[18] Peter Hansen. *Particle Detectors and Accelerators Lecture Notes.* 3rd ed. University of Copenhagen, Oct. 2016 (cit. on p. 19).

[19] Volkl, Valentin et al. *FCCSW.* `https://github.com/HEP-FCC/FCCSW/`. 2021 (cit. on p. 23).

[20] G. Roland. "Phenomenological study of anomalous $t - \bar{t}$ production at FCC-ee". Bachelor's Thesis. University of Copenhagen, June 2021 (cit. on p. 25).

[21] J. B. Hansen. "Triple Gauge-boson Couplings in W pair Production via e^+e^- Annihilation". PhD thesis. Copenhagen U., 1996 (cit. on pp. 27, 53).

[22] Seymour, Michael H. "Jets in QCD". In: *AIP Conf. Proc* 357 (June 1995), 568–587. 20 p. DOI: `10.1063/1.49625`. URL: `https://cds.cern.ch/record/283896` (cit. on pp. 27, 28).

[23] M. Cacciari. *Jet algorithms and jet substructure.* Hard Probes 2016 Student Lectures. Université Paris Diderot, Oct. 2016. URL: `https://indico.cern.ch/event/502239/contributions/2279188/attachments/1341921/2021242/jets.pdf` (cit. on pp. 27, 28).

[24] Helsens, Clement et al. *FCCAnalyses.* `https://github.com/HEP-FCC/FCCAnalyses/`. 2021 (cit. on pp. 28, 36).

[25] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *The European Physical Journal C* 72.3 (Mar. 2012). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-012-1896-2`. URL: `http://dx.doi.org/10.1140/epjc/s10052-012-1896-2` (cit. on p. 28).

[26] M. Boronat et al. "A robust jet reconstruction algorithm for high-energy lepton colliders". In: *Physics Letters B* 750 (Nov. 2015), pp. 95–99. ISSN: 0370-2693. DOI: `10.1016/j.physletb.2015.08.055`. URL: `http://dx.doi.org/10.1016/j.physletb.2015.08.055` (cit. on p. 30).

[27] M. Boronat et al. *A new jet reconstruction algorithm for e+e- colliders*. 37th International Conference on High Energy Physics (ICHEP 2014), Valencia (Spain), 2 Jul 2014 - 9 Jul 2014, July 2014. URL: `https://indico.ific.uv.es/event/2025/contributions/1475/attachments/1385/1533/Valencia_algorithm_ICHEP.pdf` (cit. on p. 30).

[28] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01, Statistics chapter. ISSN: 2050-3911. DOI: `10.1093/ptep/ptaa104`. eprint: `https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf`. URL: `https://doi.org/10.1093/ptep/ptaa104` (cit. on p. 32).

[29] The OPAL Collaboration. "A study of the recombination scheme dependence of jet production rates and of $\alpha_s(M_{Z_0})$ in hadronic Z_0 decays". In: (Oct. 1990). DOI: `10.1007/BF01549689` (cit. on p. 34).

[30] Dan Green. *High P_T Physics at Hadron Colliders*. en. Cambridge monographs on particle physics, nuclear physics, and cosmology 22. Cambridge, UK ; New York: Cambridge University Press, 2005. ISBN: 978-0-521-83509-1 (cit. on p. 42).

[31] M. Shapiro. *Lecture 12: QCD and $e^+e^- \rightarrow$ Hadrons Continued*. Berkeley University, Oct. 2016. URL: `https://www-atlas.lbl.gov/~shapiro/Physics226/lecture12.pdf` (cit. on p. 43).

[32] Immanuel Kant. *Critik der reinen Vernunft*. de. Riga: Hartknoch, 1781, p. 856 (cit. on p. 49).

[33] Volker Blobel. "Least Squares Methods". In: *Formulae and methods in experimental data evaluation, with emphasis on high energy physics v.3 : Articles on statistical and numerical methods*. Vol. 3. CERN, Geneva: The European Physical Society, 1983 (cit. on p. 50).

[34] Hansen, Jørgen Beck and Torndal, Julie Munch. *ABCfit++*. `https://github.com/Torndal/ABCfitplusplus/`. 2021 (cit. on pp. 55, 106).

[35] N. Berger. *Statistical methods for particle physics – Lecture 3*. Lecture Slides. LAPP, Feb. 2021. URL: `https://indico.in2p3.fr/event/23254/contributions/90869/attachments/62128/84941/LAPP-Stats-2021_lecture3.pdf` (cit. on p. 67).

[36] E. Boos et al. "CompHEP 4.4—automatic computations from Lagrangians to events". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 534.1-2 (Nov. 2004), pp. 250–259. ISSN: 0168-9002. DOI: `10.1016/j.nima.2004.07.096`. URL: `http://dx.doi.org/10.1016/j.nima.2004.07.096` (cit. on p. 106).

[37] A. Pukhov et al. "CompHEP - a package for evaluation of Feynman diagrams and integration over multi-particle phase space. User's manual for version 33". In: (Aug. 1999) (cit. on p. 106).

[38] J. de Favereau et al. "DELPHES 3: a modular framework for fast simulation of a generic collider experiment". In: *Journal of High Energy Physics* 2014.2 (Feb. 2014). ISSN: 1029-8479. DOI: `10.1007/jhep02(2014)057`. URL: `http://dx.doi.org/10.1007/JHEP02(2014)057` (cit. on p. 106).

[39] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: `10.1007/jhep07(2014)079`. URL: `http://dx.doi.org/10.1007/JHEP07(2014)079` (cit. on p. 106).

[40] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. ISSN: 0010-4655. DOI: `10.1016/j.cpc.2015.01.024`. URL: `http://dx.doi.org/10.1016/j.cpc.2015.01.024` (cit. on p. 106).

[41] Wolfgang Kilian, Thorsten Ohl, and Jürgen Reuter. "WHIZARD—simulating multi-particle processes at LHC and ILC". In: *The European Physical Journal C* 71.9 (Sept. 2011). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-011-1742-y`. URL: `http://dx.doi.org/10.1140/epjc/s10052-011-1742-y` (cit. on p. 106).

[42] Mauro Moretti, Thorsten Ohl, and Jurgen Reuter. "O'Mega: An Optimizing matrix element generator". In: (Feb. 2001). Ed. by Ties Behnke et al. arXiv: `hep-ph/0102195` (cit. on p. 106).

Acronyms and Glossary

Acronyms

BSM beyond the Standard Model. 15, 18

CC charged current. 8

CDR Conceptual Design Report. 17

CEPC The Circular Electron-positron Collider. 18

CLIC The Compact Linear Collider. 18, 23, 30

EDM Event Data Model. 35

EFT Effective Field Theory. 13

FCC-SW Future Circular Collider Software. 23, 35

FWHM Full Width at Half Maximum. 52

IDEA International Detector for Electron-positron Accelerator. 20, 23, 39

ILC The International Linear Collider. 18, 23, 30

IPs Interaction Points. 19

ISR Initial State Radiation. 30, 55

LEP The Large Electron-Positron Collider. 27

LH left-handed. 8, 9

LHC The Large Hadron Collider. 11, 17, 28

LHE Les Houches Event. 25

LO leading order. 7, 8

MC MonteCarlo. 28, 30–32, 35, 39, 107

MPGD Micro Pattern Gas Detector. 21

NC neutral current. 8

NLO next-to-leading order. 7

NNLO next-to-next-to-leading order. 7

PDF Probability Distribution Function. 52, 55

QCD quantum chromodynamics. 4, 6–8, 10

QED quantum electrodynamics. 4–10

QFT Quantum Field Theory. 4, 10

reco jet reconstructed jet. 30–33, 36, 37, 95, 107, *Glossary:* reconstructed jet

RF Radiofrequency. 19

RH right-handed. 8, 9

SM Standard Model of Particle Physics. 3, 4, 6, 10, 11, 13–15, 17, 106

SMEFT Standard Model Effective Field Theory. 13

VEV Vacuum Expectation Value. 10, 11

Glossary

`ABCfit++` Library for applying a general constrained kinematic fit using ABC-parametrisation [34]. 53, 55, 71

`CompHep` package for evaluation of Feynman diagrams[36, 37]. 24, 25, 95

`DELPHES` Framework for fast simulation of a generic collider experiment [38]. 23, 25, 30, 31, 35, 39, 107

`Key4HEP` turnkey software for future colliders: Key4HEP docs. 23

`MadGraph` framework for SM and BSM phenomenology: cross sections, generation of hard events, matrix elements etc. [39]. 25

`Pythia` Monte Carlo generator for high-energy physics collider events [40]. 23, 25

`Whizard` Monte Carlo generator that calculates multi-particle scattering cross sections [41, 42]. 25

EDM4HEP a generic event data model for future HEP collider experiments. 23, 24, 35

Higgs mechanism mechanism by which the particles of the SM acquire mass. 4, 10, 11

Jet Clustering Interface An Interface to `FastJet` for later stage process adaptive jet clustering. 28, 35–37, 39

Jet Tagging Utilities A tool used to assign a flavour to the jet based on the MC truth. 35, 39

particle jet reconstructed from the final state particles at the true MC level after the hadronisation. 30–33, 37, 95

reconstructed jet reconstructed from the final state particles at the detector level (reconstructed by `DELPHES`), often shortened to reco jet. 30

Index