Novel Unfolding Methods and Measurements of $t\bar{t}$ Differential Cross Sections with SMEFT interpretation using the ATLAS detector at the LHC

Pim Jordi Verschuuren

Supervisor: Professor Glen Cowan

Submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Particle Physics

Department of Physics

Royal Holloway University

September 2021

I hereby declare that I am the sole author of this thesis. In any other case, it is indicated in the text.

_____

Pim Jordi Verschuuren

Abstract

Studies on unfolding methods and measurements of $t\bar{t}$ differential cross sections in the l+jets decay channel at the Large Hadron Collider are presented in this thesis. Available and novel unfolding algorithms are compared using known and new figures of merit and a response matrix estimation method that is less dependent on the size of the training data. Lastly, a novel covariance matrix estimation method is presented that includes both statistical and systematic uncertainties from various sources. Measurements of $t\bar{t}$ production cross sections in the l+jets decay channel at the Large Hadron Collider are presented. A total integrated luminosity of 139 fb^{-1} of proton-proton collision data at a centre-of-mass energy of $\sqrt{s} = 13$ TeV has been collected at the ATLAS detector. The production cross section as a function of several kinematic variables are measured, unfolded and compared to theory. The unfolded differential cross sections are used to estimate and constrain the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients of the Standard Model Effective Field Theory with individual and two-parameter fits using LO and NLO QCD SMEFT predictions. The results are found to be in good agreement with the SM.

Acknowledgements

My first and foremost thanks go to Glen Cowan. He has been an amazing teacher on statistics, physics, collaboration within the HEP community and research in general. He has definitely shown me how exciting and enjoyable research can be.

Secondly, I would like to thank Francesco Spanó. He has been an infinite well of knowledge regarding anything related to top quark physics and the LHC Top community. I don't remember a single question that he did not have an answer for.

More thanks goes out to Lydia Brenner, Carsten Burgard and Vince Croft. I admire their way of working and enjoyed our collaboration during the unfolding side-projects. I see them all as great examples of a physicist.

Thanks goes out to the Royal Holloway Particle Physics group. They supplied a good discussion platform and gave me a fresh point of view whenever I needed it throughout my PhD. Special thanks goes out to James, Sean and Callum for the many breaks and much needed distractions in the office.

Many thanks to all the fellows and supervisors of the Marie Sklodowska-Curie Innovative Training Network INSIGHTS. All the fellows were great intellectual sparring partners in physics, statistics, machine learning and beyond. They've made every shared training event and conference incredibly enjoyable.

I would also like to thank Sophie. For putting up with the stress-induced grumpiness, the overworking in the weekends and supplying of a listening ear to any of my PhD-related complaints. It is your unconditional emotional support and our moments together that kept me sane.

Lastly, I want to thank my parents. Not just for their support during my PhD but also for everything before that. I would not have made it this far without all their support, guidance and teachings on many things in life. I am forever grateful for having them.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Any contemporary particle physicist will acknowledge that they are standing on the shoulders of giants. The development of the Standard Model (SM) has been a theoretical success story and is one of the most accurate quantum field theories (QFT) to date. Also, the massive collaborative effort, technological innovation and experimental creativity that produced the evidence to support the Standard Model is impressive to say the least. The Large Hadron Collider (LHC) at the Conseil Européen pour la Recherche Nucléaire (CERN) has been probing the fundamental building blocks of nature for 13 years and affirmed our view on nature with the discovery of the Higgs boson in 2012 by both the ATLAS and CMS experiments. However, many unanswered questions about our universe remain.

Measurements on the top quark, the most massive particle of the Standard Model, can be used to both improve measurements of Standard Model parameters and test the validity of new QFTs. The Large Hadron Collider provides unique conditions to study the top quark and produce high precision data. This thesis presents top quark pair differential cross sections in the lepton+jets decay channel using pp collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ as recorded by the ATLAS detector.

Many SM extensions, i.e. Beyond the Standard Models (BSM), have emerged that posed solutions for unanswered questions. However, none have been supported with sufficient evidence to become the new status quo. We need to take a step back, widen our scope and redefine our direction in our theoretical landscape. Effective field theory (EFT) is the ideal framework to provide this wide model-agnostic scope. This thesis therefore presents estimates and confidence bounds on parameters of the Standard Model Effective Field Theory using the top quark pair differential cross section measurements.

The proper statistical treatment of measurements and their comparison with theory is crucial. A vital part of many particle physics statistical frameworks is *unfolding*: the process of correcting measured differential distributions for detector effects. This thesis therefore also presents research on novel deconvolution methods that can be applied to differential cross section measurements such as the one presented in this thesis.

The thesis is structured as follows: Chapter 2 provides an introduction to quantum field theory in the context of particle physics and how its principles are used to define the Standard Model. It also summarizes some of the phenomena the Standard Model fails to describe, an introduction to effective field theory and the specific EFT that is used in this thesis known as the Standard Model Effective Field Theory (SMEFT). Lastly, it describes the phenomenological models used to simulate top quark pair production at the LHC. Chapter 3 provides an overview of the experimental setup used to collect the data, i.e. the Large Hadron Collider and the ATLAS detector. Chapter 4 describes the data analysis methods used to reconstruct meaningful physics analysis objects from the data collected at the ATLAS detector. Chapter 5 presents

the research on unfolding methods. Chapter 6 presents the measurement of the $t\bar{t}$ differential cross sections in the l+jets decay channel. Chapter 7 presents how these measurements are used to constrain SMEFT Wilson coefficients and Ch. 8 closes with some conclusions.

Personal contribution

The ATLAS Collaboration consists of 5500 members spanning 41 countries that contribute as physicists, engineers, technicians, students and support staff. This thesis has been made possible by the collaborative effort of these members. Any scientific results that has been produced in part or completely by someone else is referenced accordingly. Here I would like to outline my personal contributions.

Chapter 5 presents unfolding studies that emerged whilst maintaining, developing and supporting the RooUnfold software package. I was one of the two first authors of the comparison study and the main investigator of the response matrix and covariance estimation studies.

Chapter 6 presents a differential cross section measurement that has been done in collaboration with an ATLAS analysis team. The event and object selection has been developed and applied by the analysis team and previous analysis teams that performed a $t\bar{t}$ differential cross section measurement. The data and Monte Carlo samples were produced by members of ATLAS dedicated to these specific tasks. I contributed by validating the produced data and Monte Carlo samples, studying data-MC agreement, setting up the unfolding framework and producing the final unfolded result.

Chapter 7 presents measurements of SMEFT Wilson coefficients produced by myself. This includes defining the strategy and statistical framework, producing the SMEFT Monte Carlo samples, applying the Rivet object and event selection and writing the code for the fitting procedures.

Chapter 2

Theoretical Framework

This chapter presents the theoretical framework needed to make predictions for *pp* collisions at the LHC. Over the years many theorists have shown great creativity and effort in the development of theories that describe the character of matter on the smallest possible scales. The goal of these theories is to model nature in a mathematical framework resulting in predictions for observable quantities that can be compared with experiment.

Quantum Field Theory (QFT) was born in the necessity to improve the concepts of relativistic quantum mechanics. In the 1920s Paul Dirac took the first steps by quantizing classical electromagnetic fields in an attempt to describe the creation and annihilation of particles. However, it was only in the 1949 that Richard Feynman reformulated Dirac's theory of Quantum Electrodynamics (QED) into one with calculable quantities. This resulted in the first QFT which was followed by many others. The Standard Model of Particle Physics is a group of quantum field theories that have been amassed over the years. Together they supply one of the most accurate theories in history and is the current status quo of the subatomic world.

Section 2.1 will start with a brief superficial introduction on the Standard Model i.e. a general description of all the subatomic particles and fundamental forces it contains. Section 2.2 continues with some basic concepts of a quantum field theory, how they are used to describe particles and how they manifest themselves into measurable observables such as cross sections. Section 2.3 continues by introducing *gauge symmetries* and how they are used to define the quantum field theories that make up the Standard Model. Section 2.4 describes the *Higgs mechanism* that is needed to explain particle masses. Section 2.5 discusses some open issues of the Standard Model that motivate the experiments at the LHC. Section 2.6 discusses a special type of quantum field theory known as an *effective field theory* (EFT) that is used to explore new physics in a BSM theory agnostic way. Section 2.7 discusses the theoretical principles used to simulate top quark pair production at the LHC.

2.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics is a renormalizable quantum field theory [1, 2] that describes the most fundamental building blocks of the universe and the interactions between them. This section will give a general summary of the particles it contains, some of their characteristics and how they interact.

In Fig. 2.1 an overview is shown of all the currently known elementary particles along with their mass, colour charge, electric charge and spin. The latter divides the elementary particles into two groups: the fermions with half integer spin and the bosons with integer spin. The fermions consist of the quarks and leptons which make up all matter around us whilst the gluons, photons, Z and W bosons mediate the strong, electromagnetic and weak force, respectively. The Higgs boson is an excitation of the Higgs field but does not introduce a new force like the before mentioned

bosons. Instead the Higgs field "gives" particles mass through the Higgs mechanism, discussed in Sec. 2.4. The top quark is the most massive fermion of the Standard Model and will be the focus of this thesis.

At the bottom right of Fig. 2.1 a space is left blank for the hypothesized graviton boson [3] which should mediate the gravitational force. No particle with fitting characteristics has yet been found to support a theory of quantum gravity [4, 5]. However, in the upcoming sections the gravitational force can be neglected as its strength is negligible on subatomic scales [6].
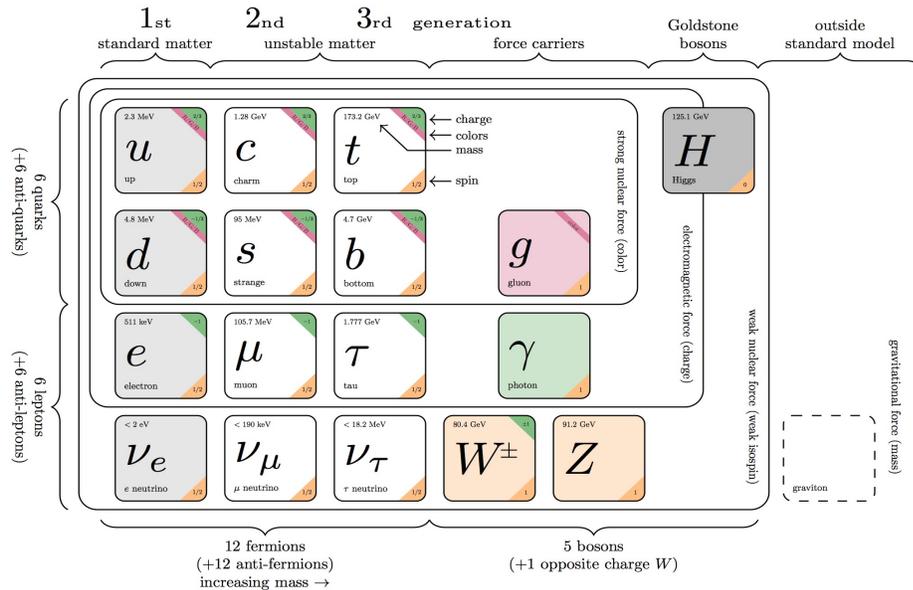
Figure 2.1: The particles of the Standard Model. From [7].

This zoo of particles has a very intricate mathematical framework behind them. Treating the particles as point- or sphere-like objects, as many envision subatomic particles, can be a good classical approximation in certain scenarios such as in Rutherford scattering [8]. However, a more complete description can be given with quantum field theories that consider particles as excitations of quantum fields. The Lagrangian formalism is one of the narratives used within quantum field theories to build models and make predictions.

2.2 Lagrangian Formalism

The Lagrangian formalism [9, 10] initially was a reformulation of the Newtonian classical mechanics. However, in this section we will show how the concepts of this mathematical framework can be used as a natural starting point for a quantum field theory. There are two different ways to generalize the classical Lagrangian formalism to a quantum field theory: the path integral description and the canonical quantization. We will not discuss the canonical quantization here. However, additional material on the canonical quantization can be found in [11, 12] and Appendix A.1. The rest of this chapter will follow the Lagrangian formalism with the path integral narrative.

2.2.1 Principle of Least Action

The path integral approach starts with the principle of least action. This is a variational principle that, when applied to the action, gives the equations of motion of a system [9]. The action is defined as

$$\mathcal{S}[\mathbf{q}(t)] = \int_{t_1}^{t_2} L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \, dt. \tag{2.1}$$

The action is a time integral over the Lagrangian $L(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ which is a function of a set of generalized coordinates $\mathbf{q}(t) = (q_1(t), q_2(t), ..., q_N(t))$ and their time derivates $\dot{\mathbf{q}}(t) = (\frac{dq_1(t)}{dt}, \frac{dq_2(t)}{dt}, ..., \frac{dq_N(t)}{dt})$. The evolution from one state at time t_1 to another at t_2 can go along different paths in the space spanned by the generalized coordinates and its time derivatives. The *principle of least action* states that the path taken by the system will be the one for which the action is stationary to the first order, i. e. $\delta \mathcal{S} = 0$ [13], as depicted in Fig. 2.2. From this the Euler-Lagrange equations can be derived

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial L}{\partial \mathbf{q}} = 0. \tag{2.2}$$

The equations of motion of a system can be derived from the Euler-Lagrange equations by defining an appropriate Lagrangian.
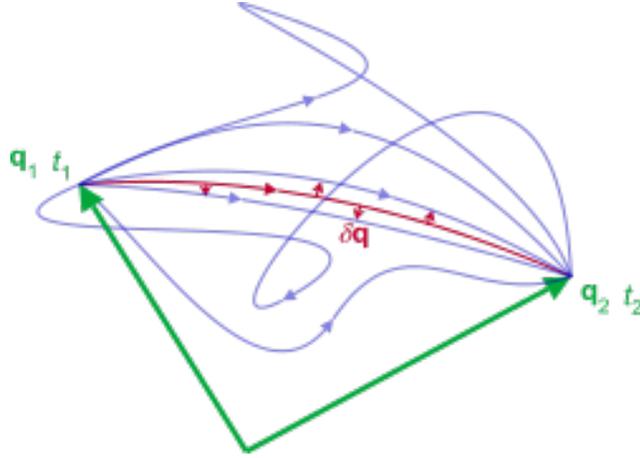
Figure 2.2: Different paths through configuration space. The red line indicates the path for which $\delta S = 0$. From [14].

2.2.2 Path Integrals in Quantum Mechanics

In 1948 Feynman showed that there is an analogous approach to the classical path integral approach for non-relativistic quantum mechanical systems [15]. He argued that the evolution of a quantum system is not deterministically determined by the principle of least action but probabilistic of nature. Each possible path in $\boldsymbol{q}$-space between an initial and a final quantum state comes with a probabilistic amplitude $e^{i\mathcal{S}}$, where $\mathcal{S}$ is the previously introduced action. The squared sum of these amplitudes is the probability that the transition from an initial state $|\mathbf{q}_i, t_i\rangle$ to a final quantum state $|\mathbf{q}_f, t_f\rangle$ will occur. The probability is defined as

$$P(\mathbf{q}_f, t_f; \mathbf{q}_i, t_i) = |\langle \mathbf{q}_f, t_f | \mathbf{q}_i, t_i \rangle|^2 \tag{2.3}$$

with the probability amplitude defined as

$$\langle \mathbf{q}_f, t_f | \mathbf{q}_i, t_i \rangle = \int \mathcal{D}\mathbf{q}(t) \, e^{i/\hbar \int_{t_i}^{t_f} dt L(\mathbf{q}, \dot{\mathbf{q}})} \tag{2.4}$$

where $\mathcal{D}\mathbf{q}(t)$ denotes an integral taken over all possible paths between $\mathbf{q}_f$ and $\mathbf{q}_i$ [2].

2.2.3 Free Scalar Field Theory

In this section we will introduce the concept of fields and how they represent particles with a simple toy example. Particles are considered to be excited states, or quanta, of different fields that permeate all space. To include fields in the path integral formalism one can make the following replacements.

$$\mathbf{q}(t) \longrightarrow \phi(\mathbf{x}, t) \text{ (Field)}$$

$$L(\mathbf{q}(t)), \dot{\mathbf{q}}(t) \longrightarrow \mathcal{L}(\phi, \partial_\mu \phi) \text{ (Lagrangian Density)} \tag{2.5}$$

Here we introduce the four-gradient $\partial_\mu = \left(\frac{1}{c} \frac{\partial}{\partial t}, \boldsymbol{\nabla} \right)$ and the scalar field ϕ. A scalar field was chosen over vector and tensor fields to keep the example simple. Let us also introduce the space-time coordinates $x = (\mathbf{x}, t)$ for notational ease. The action and Euler-Lagrange equations are resp. redefined as

$$S[\phi(x)] = \int \mathcal{L}(\phi, \partial_\mu \phi) \, d^4 x \tag{2.6}$$

$$\partial_\mu \left(\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} \right) - \frac{\delta \mathcal{L}}{\delta \phi} = 0 \tag{2.7}$$

Now let us introduce the Lagrangian density corresponding to the free scalar field theory i.e. a non-interacting 0-spin particle.

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 \tag{2.8}$$

10

The probability amplitude of Eq. (2.4) changes into the functional

$$Z_0[J] \equiv \langle 0|0 \rangle_J = \int \mathcal{D}\phi \, e^{i/\hbar \int d^4x (\mathcal{L}_0 + J\phi)} \tag{2.9}$$

also known as the *functional integral*. $J(x)$ is known as the *source* and is added to the lagrangian density $\mathcal{L}$ as a mathematical trick to calculate probability amplitudes. $|0\rangle$ corresponds to the vacuum state i.e. a state with no particles present. For free field theory we can define the functional as [2],

$$Z_0[J] = \exp\left[\frac{i}{2} \int d^4x \, d^4x' \, J(x)\Delta(x - x')J(x') \right] \tag{2.10}$$

with the *Feynman propagator*,

$$\Delta(x - x') = \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik(x-x')}}{k^2 + m^2 - i\epsilon}. \tag{2.11}$$

We can use these definitions to calculate *n-point time-ordered correlation functions* which can be used to calculate transition probability amplitudes with the Lehmann–Symanzik–Zimmermann(LSZ) reduction formula [2, 16]. A n-point time-ordered correlation function [2, 1] can be defined as

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)...\phi(x_n)|0\rangle = \frac{\int \mathcal{D}\phi \, \phi(x_1)\phi(x_2)...\phi(x_n)e^{i/\hbar \int d^4x \, \mathcal{L}_0}}{\int \mathcal{D}\phi \, e^{i/\hbar \int d^4x \, \mathcal{L}_0}}$$
$$= \frac{1}{i} \frac{\delta}{\delta J(x_1)} \frac{\delta}{\delta J(x_2)}...\frac{\delta}{\delta J(x_n)} Z_0[J]\bigg|_{J=0}. \tag{2.12}$$

Here we introduce the functional derivative $\frac{\delta}{\delta J(x)}$ and the *time ordering symbol* $\mathcal{T}$. See Appendix A.2 for some example correlation functions for the free scalar field theory. According to *Wick's theorem* [17] we can redefine the correlation functions in terms of the Feynman propagators (Eq. (2.11)) as

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)...\phi(x_{2n})|0\rangle = \frac{1}{i^n}\sum_p \Delta(x_{p_1}-x_{p_2})...\Delta(x_{p_{2n-1}}-x_{p_{2n}}). \qquad (2.13)$$

with $\sum_p$ summing over all the permutations between the indices $p_1, p_2, ..., p_{2n}$. It is possible to pictorially represent the terms of the correlation functions with graphs known as *Feynman diagrams* [18]. These graphs can make computations for more complex correlation functions easier by following Feynman rules [2, 19]. These rules are different for each underlying theory but can make the calculation of probability amplitudes a lot simpler. Some example Feynman diagrams for the free scalar field theory can be found in Appendix A.2.

Now let us use Eq. (2.13) to calculate the probability amplitude of an initial state with n massive scalar particles with momenta $(k_1, k_2, ..., k_n)$ to a final state with m massive scalar particles with momenta $(p_1, p_2, ..., p_m)$. The LSZ reduction formula [2, 16] states that this probability amplitude can be defined as

$$\langle p_1, p_2, ..., p_m | k_1, k_2, ..., k_n\rangle =$$
$$\prod_{j=1}^{m}\left[\int d^4x_j \, ie^{p_j x_j}(\partial_{x_j}^2 + m^2)\right] \times \prod_{l=1}^{n}\left[\int d^4x_l \, ie^{k_l x_l}(\partial_{x_l}^2 + m^2)\right] \qquad (2.14)$$
$$\times \langle 0|\mathcal{T}\phi(x_1)\phi(x_2), ..., \phi(x_{2n})|0\rangle.$$

Here we see how the correlation functions determine the scattering amplitude. Now let us assume a simple example of 2 incoming and 2 outgoing massive scalar particles with momenta p_1, p_2, p_3 and p_4. For 2 incoming and 2 outgoing particles we need the free scalar field 4-point correlation of Eq. (A.5). The scattering amplitude becomes

$$\langle p_1, p_2 | p_3, p_4 \rangle_0 =$$
$$\int d^4x_1 \, d^4x_2 \, d^4x_3 \, d^4x_4 \, e^{i(p_1x_1 + p_2x_2 - p_3x_3 - p_4x_4)} (\partial_{x_1}^2 + m^2)(\partial_{x_2}^2 + m^2)(\partial_{x_3}^2 + m^2)(\partial_{x_4}^2 + m^2)$$
$$\times \left(\Delta(x_1 - x_2)\Delta(x_3 - x_4) + \Delta(x_1 - x_3)\Delta(x_2 - x_4) + \Delta(x_1 - x_4)\Delta(x_2 - x_3) \right)$$
$$(2.15)$$

If we fill in the definition of the propagator (Eq. (2.11)) and apply Fourier transformations [2] we get

$$
\begin{aligned}
\langle p_1, p_2 | p_3, p_4 \rangle_0 = {} & (2\pi)^4 \delta^4(p_1 - p_3)(2\pi)^4 \delta^4(p_2 - p_4) F(p_1, p_3) F(p_2, p_4) \\
& + (2\pi)^4 \delta^4(p_1 - p_4)(2\pi)^4 \delta^4(p_2 - p_3) F(p_1, p_4) F(p_2, p_3) \\
& + (2\pi)^4 \delta^4(p_1 - p_2)(2\pi)^4 \delta^4(p_3 - p_4) F(p_1, p_2) F(p_3, p_4)
\end{aligned}
\tag{2.16}
$$

with F denoting a Fourier transform term that does not need specifying for now. The most important part is to note the delta functions δ. These tell us that all the three terms contribute to "no scattering" events as they require the incoming and outgoing momenta to be the same. See Appendix A.2 for Feynman diagrams that represent these "no scattering events". These are also known as *vacuum diagrams* because they represent particles appearing and disappearing without any interaction. We are interested in pp collisions where particles interact i.e. in diagrams that are said to be *fully connected*.

2.2.4 Weak Coupling Perturbation Theory

One way of introducing an interaction is by including a perturbative contribution in the free scalar Lagrangian,

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\text{int}} \tag{2.17}$$

with $\mathcal{L}_0$ being the free scalar field Lagrangian of Eq. (2.8) and $\mathcal{L}_{\text{int}}$ some interaction term. The generating functional of Eq. (2.9) is redefined as

$$Z[J] = \int \mathcal{D}\phi \, \exp\left[i/\hbar \int d^4x \left(\mathcal{L}_0 + \mathcal{L}_{int} + J\phi\right)\right]. \qquad (2.18)$$

Now let us assume a quartic self-interaction between the scalar fields with coupling strength λ.

$$\mathcal{L}_{\text{int}} = -\frac{\lambda}{4!}\phi^4 \qquad (2.19)$$

In Appendix A.2 it is shown how generating functionals, Taylor expansions and functional derivatives can be used in a similar way to the previous section to calculate n-point time-ordered correlation functions. Again we can plug the 4-point correlation function of the weak coupling perturbation Lagrangian (See Appendix A.2) into the $2 \to 2$ LSZ formula and apply Fourier transformations as we did with Eq. (2.16).

$$\langle p_1, p_2 | p_3, p_4 \rangle = \langle p_1, p_2 | p_3, p_4 \rangle_0 - i\lambda(2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) + ... \qquad (2.20)$$

In Appendix A.2 it is shown how the commonly used terms *leading order* (LO) and *next-to-leading order* (NLO) accuracy correspond to the first and second order of the Taylor expansion of the correlation functions in λ, respectively. One can include higher order λ-contributions to improve the accuracy of the scattering amplitude. For now we define the interaction scattering amplitude at leading order as

$$\langle p_1, p_2 | p_3, p_4 \rangle_{\text{int}} = -i\lambda(2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \qquad (2.21)$$

2.2.5 Cross Sections

Now that we know how to calculate a scattering amplitude $\langle p_1, p_2 | p_3, p_4 \rangle_{\text{int}}$ we need to convert it into an observable that can be measured by particle collider experiment. One of the go-to observables of particle collider experiments is the cross section. One can define the total cross section as [2]

$$\sigma = \oint \frac{|\mathcal{M}|^2}{64\pi^2 s} \, d\Omega \tag{2.22}$$

with the $\sqrt{s}$ being the center-of-mass energy of the collision, the integral done over a whole sphere solid angle and $\mathcal{M}$ the matrix element. The matrix element $\mathcal{M}$ is connected to the scattering amplitude $\langle p_1, p_2 | p_3, p_4 \rangle_{\text{int}}$ by [2]

$$\langle p_1, ..., p_m | k_1, ..., p_n \rangle_{\text{int}} = i\mathcal{M}(2\pi)^4 \delta^4(k_1 + ... + k_n - p_1 - ... - p_m). \tag{2.23}$$

For a $2 \rightarrow 2$ scattering and leading order weak coupling perturbation interaction of Sec. 2.2.4 the total cross section becomes

$$\sigma = \frac{\lambda^2}{16\pi s} \tag{2.24}$$

Particle collider experiments like the ATLAS Experiment try to identify and count the number of events N with the outgoing particles from the scattering process of interest. This is coupled to the total cross section via the definition

$$N = \sigma \int L \, dt \tag{2.25}$$

with L being the luminosity, the number of interactions between the incoming particles per unit of time. Additionally, one can count the number of events with a certain kinematic variable value e.g. the number of events with a certain transverse

momentum for one of the outgoing particles. These type of observables are known as differential cross sections and are the focus of this thesis. An analysis approach would be to compare measured cross sections with the cross sections predicted by the weak coupling perturbation theory determined by the λ coupling parameter. In practice, the quantum field theories that we test are more complicated and contain many more parameters. However, the general approach of comparing measured cross sections with predictions for different parameter values remains the same. The Standard Model is the current status quo when it comes to predictions in particle physics and has a much larger set of fields, types of interactions and parameters. In the upcoming sections we will show how gauge symmetries are used to introduce the terms that define the Standard Model Lagrangian.

2.3   Gauge Symmetries

Noether's theorem [20] states that if the Lagrangian of a system is invariant under a transformation then such a symmetry corresponds to a conservation law. In this section it is shown how the invariance of a Lagrangian under certain groups of transformations results in the introduction of the fundamental interactions between various fields.

2.3.1   Quantum Electrodynamics

Quantum Electrodynamics [18, 21, 22] was the first gauge theory constructed for the description of particle physics. In this section it is used to show how a gauge symmetry results in a gauge field that mediates a fundamental interaction, the electromagnetic interaction. Quantum electrodynamics describes the interaction between charged particles mediated by the gauge boson known as the photon. Let us define the Lagrangian for a charged free spin-$\frac{1}{2}$ particle as

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi. \tag{2.26}$$

Here γ^μ are the gamma matrices [22] and ψ a four-component spinor field. Now let us apply the simplest possible gauge transformation, a $U(1)$ local gauge transformation, on the spinor field,

$$\psi \to \psi' = e^{iq\xi(x)}\psi. \tag{2.27}$$

Note that the dependence of the phase $\xi(x)$ on arbitrary space-time coordinates makes the gauge transformation local. This results in the transformed Lagrangian

$$\mathcal{L} \to \mathcal{L}' = ie^{iq\xi(x)}\bar{\psi}\gamma^\mu\left[e^{iq\xi(x)}\partial_\mu\psi + iq(\partial_\mu\xi(x))e^{-iq\xi(x)}\psi\right] - me^{-iq\xi(x)}\bar{\psi}e^{iq\xi(x)}\psi. \tag{2.28}$$

To restore the local gauge invariance the derivative needs to be replaced with the *covariant derivative*,

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu, \tag{2.29}$$

where a new four-component gauge field A_μ is introduced. One can impose a gauge field transformation

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu\xi(x). \tag{2.30}$$

Only by introducing a new gauge field with very specific transformation properties can the Lagrangian stay invariant under a local gauge transformation. The $U(1)$ local gauge invariant Lagrangian for a spin-$\frac{1}{2}$ fermion therefore reads,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - q\bar{\psi}\gamma^\mu A_\mu\psi \tag{2.31}$$

17

which shows that the result of imposing a local gauge invariance is an additional interaction term between the fermionic fields ψ, $\bar{\psi}$ and the gauge field A_μ. This can be interpreted as two fermions interacting via the electromagnetic interaction mediated by a photon. The final QED Lagrangian can be defined as

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - q\bar{\psi}\gamma^\mu A_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.32}$$

with the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.33}$$

The introduced gauge field "predicts" a gauge boson which we now know as the photon.

2.3.2 The Standard Model Gauge Theory

The Standard Model is described by a Lagrangian that is invariant under transformations from three different gauge groups.

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.34}$$

The $SU(3)_C$ and $SU(2)_L \otimes U(1)_Y$ group represents the strong and electroweak interaction, respectively. The subscripts refer to the colour charge (C), hyper charge (Y) and left-handed fermions (L). The Standard Model Lagrangian is defined as,

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} + \mathcal{L}_{\text{Higgs}}, \tag{2.35}$$

where each of the Lagrangians $\mathcal{L}_{EW}$, $\mathcal{L}_{QCD}$ and $\mathcal{L}_{\text{Higgs}}$ will be introduced in following sections.

2.3.3 Electroweak Interaction

The gauge group $SU(2)_L \otimes U(1)_Y$ is also known as the Glashow-Salam-Weinberg model [23, 24, 25] of electroweak interactions and is the unified theory of the weak and electromagnetic interaction. Again a local gauge transformation is applied, the appropriate gauge fields are added by introducing a covariant derivative and specific transformations are imposed for the corresponding gauge fields to acquire new interactions. A more general form for a local gauge transformation can be defined as

$$U(x) = e^{iA_j(x)B_j} \tag{2.36}$$

where $A_j(x)$ is the space-time coordinate dependent phase parameter and B_j the generator of the gauge group. A group is called *abelian* if the corresponding generators commute and *non-abelian* if otherwise. $SU(2)_L$ is a non-abelian group associated to the weak isospin (I_3) with the Pauli matrices $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ as generators. Fermions with left-handed chirality have weak isospin $I_3 = \pm\frac{1}{2}$ and are $SU(2)_L$ doublets while fermions with right-handed chirality have $I_3 = 0$ and are $SU(2)_L$ singlets. This means that the associated gauge fields $\boldsymbol{W_\mu^i} = (W_\mu^1, W_\mu^2, W_\mu^3)$ only interact with left-handed fermions. Hence the subscript L in the group notation.

The $U(1)_Y$ is an abelian group associated to the weak hypercharge Y_W with generator $B = 1$. It acts on fermions with all chiralities and generates a gauge field B_μ. The weak hypercharge is linked to the electric charge Q and weak isospin I_3 by the relation $Q = I_3 + \frac{Y_W}{2}$. In Table 2.1 gives an overview of all the fermions in the Standard Model in the $SU(2)_L$ representation. See Appendix A.3 for the gauge transformations, covariant derivatives, gauge fields and gauge field transformations of the $SU(2)_L \otimes U(1)_Y$-group. The electroweak Lagrangian for the $SU(2)_L \otimes U(1)_Y$ group is defined as

$$\mathcal{L}_{EW} = \sum_{\psi_L} \bar{\psi}_L i \gamma^\mu D_\mu \psi_L + \sum_{\psi_R} \bar{\psi}_R i \gamma^\mu D_\mu \psi_R - \frac{1}{4} W_i^{\mu\nu} W_{\mu\nu}^i - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \qquad (2.37)$$

		1st gen.	2nd gen.	3rd gen.	I_3	Y_W	Q
Quarks		$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$
		u_R	c_R	t_R	0	$\frac{4}{3}$	$\frac{2}{3}$
		d_R	s_R	b_R	0	$-\frac{2}{3}$	$-\frac{1}{3}$
Leptons		$\begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$
		e_R	μ_R	τ_R	0	-2	1

Table 2.1: Classification and properties of the Standard Model fermions in the $SU(2)_L$-representation

2.3.4 Strong Interaction

The $SU(3)_C$ group is the last part of the Standard Model gauge theory and describes the strong interaction. The gauge group is non-abelian and is associated with the colour charge (C) which can take either the value *green, blue* or *red.* The eight

generators of this $SU(3)_C$ symmetry are called the Gell-Mann matrices λ_i [26],

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad (2.38)$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ 0 & 0 & i \end{pmatrix}, \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

$$(2.39)$$

Quarks are the only fermions with colour charge and can be described in the $SU(3)_C$-representation with a triplet,

$$\psi = \begin{pmatrix} q_g \\ q_b \\ q_r \end{pmatrix}. \quad (2.40)$$

The gauge fields G_μ^i do not interact with leptons as they do not carry any colour charge. See Appendix A.4 for the gauge transformations, covariant derivatives, gauge fields and gauge field transformations of the $SU(3)_C$-group. The Quantum Chromodynamic (QCD) Lagrangian is defined as

$$\mathcal{L}_{QCD} = \sum_i \bar{q}_i^f (i\gamma_\mu D^\mu - m_f) q_i^f - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (2.41)$$

Again, the new gauge field "predicts" the existence of a gauge boson which we now know as the gluon.

2.4 Brout-Englert-Higgs Mechanism

The local gauge principle provides an elegant description of the interactions in the Standard Model and have many high-precision measurements that agree with it. But until recently this picture of the subatomic world was not yet complete and dealt with some serious problems.

2.4.1 Particle Masses

The fundamental interactions require that the Standard Model is invariant under transformations of any of the previously mentioned gauge groups. However, this symmetry is broken when any mass terms are added to the Standard Model Lagrangian. Let us consider a fermionic field mass term $-m_f \bar{\psi}\psi$ and decompose the expression in helicity states.

$$ -m_f \bar{\psi}\psi = -m_f(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R) \tag{2.42} $$

However, Sec. 2.3.3 states that the helicity states behave differently under a $SU(2)_L \otimes U(1)_Y$ transformation.

$$ -m_f(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R) \rightarrow -m_f(\bar{\psi}_R e^{i\alpha_i(x)\frac{\tau_i}{2}} \psi_L + \bar{\psi}_L e^{-i\alpha_i(x)\frac{\tau_i}{2}} \psi_R) \neq -m_f\bar{\psi}\psi \tag{2.43} $$

The mass terms show that an arbitrary fermionic field breaks the electroweak gauge invariance. This is a problem as it is experimentally proven that fermions such as electrons or quarks have mass. A similar argument is made for the bosons. Consider a boson field mass term $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$ for the photon. The theory of Quantum Electrodynamics required a local gauge invariance that imposed a specific gauge field transformation.

$$\frac{1}{2}m_\gamma^2 A_\mu A^\mu \rightarrow \frac{1}{2}m_\gamma^2 (A_\mu + \partial_\mu \xi(x))(A^\mu + \partial^\mu \xi(x)) \neq \frac{1}{2}m_\gamma^2 A_\mu A^\mu \tag{2.44}$$

We know that the photon and gluon are massless which means we can just omit these terms in the Lagrangian. However, this broken symmetry applies as well to the Z- and W-boson and we know from experiment that they are massive. We need a mechanism that introduces mass for the right fields.

2.4.2 Violating Unitarity

Violating unitarity refers to a scattering amplitude exceeding its unitary limit with increasing energy that will result in probabilities larger than 1. This is the case for longitudinal polarized weak bosons. When bosons are longitudinal polarized their spin z-component is perpendicular to their moving vector. A polarization four vector for longitudinal polarization can be defined as

$$\epsilon_L = \frac{1}{m_V}\begin{pmatrix} p_z \\ 0 \\ 0 \\ E \end{pmatrix}. \tag{2.45}$$

This can be used to describe the scattering amplitude of a $W_L^+ W_L^+ \rightarrow W_L^+ W_L^+$ process. In lowest order, this process will depend on three processes, the four-point interaction and the t-channel γ/Z exchanges. The Feynman diagrams of the leading order contributions to the scattering amplitude are shown in Fig. 2.3.

The scattering amplitude of the four-point interaction is proportional to

$$\mathcal{M} \sim \epsilon_L \epsilon_L \epsilon_L \epsilon_L \sim E^4. \tag{2.46}$$

The scattering amplitude diverges and would violate unitarity for increasing energy.
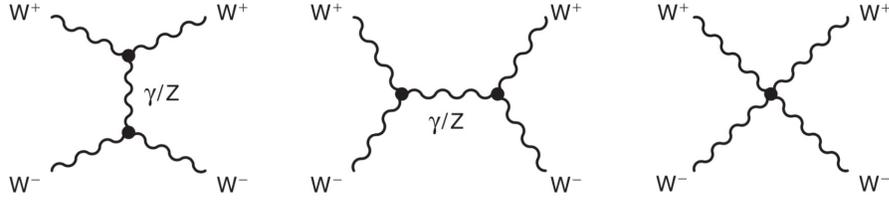
Figure 2.3: The leading order contributions to the amplitude of the $W^{\pm}W^{\mp} \rightarrow W^{\pm}W^{\mp}$. From [22].

2.4.3 Spontaneous Symmetry Breaking

A solution for the above mentioned problems is to spontaneously break the electroweak symmetry through the Brout-Englert-Higgs mechanism [27, 28]. To describe this procedure the Higgs Lagrangian is introduced.

$$\mathcal{L}_{\text{Higgs}} = \mathcal{L}_{HG} + \mathcal{L}_{HF} \qquad (2.47)$$

where $\mathcal{L}_{HG}$ and $\mathcal{L}_{HF}$ contain couplings of the Higgs boson to resp. the gauge bosons and fermions, that will generate their masses and remove the violation of unitarity. The derivation starts with an introduction of a complex scalar field in the $SU(2)_L$ representation

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_3 \end{pmatrix}. \qquad (2.48)$$

of which ϕ^0 represents a neutral scalar field and ϕ^+ a charged scalar field. The Lagrangian for this doublet of complex scalar fields can be written as,

$$\mathcal{L}_H = (\partial^\mu \Phi)^\dagger (\partial_\mu \Phi) - V(\Phi) \qquad (2.49)$$

with the Higgs potential defined as

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \qquad (2.50)$$

The λ-term describes the quartic self-interaction of the scalar field and is required to be positive $\lambda > 0$ for vacuum stability. By minimising the potential the Higgs ground state i.e. the vacuum expectation value (VEV) can be found. For $\mu^2 > 0$ the trivial VEV of $\langle \Phi \rangle_0 = 0$ is obtained and the symmetry will remain unbroken. However, for $\mu^2 < 0$ a non-zero VEV is obtained.

$$\langle \Phi \rangle_0 = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}. \tag{2.51}$$

The minimum at the origin is no longer stable but is accompanied by an infinite set of degenerate vacuum states. These states all satisfy

$$\Phi^\dagger \Phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \frac{v^2}{2}. \tag{2.52}$$

By choosing an arbitrary minimum, e.g. $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$, it is possible to expand the field Φ around this minimum for small deviations $h(x)$, i.e.

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \chi_1(x) + i\chi_2(x) \\ v + h(x) + i\chi_3(x) \end{pmatrix}. \tag{2.53}$$

The spontaneously symmetry breaking introduced three massless Goldstone bosons χ_i and one massive scalar boson $h(x)$. By taking the scalar field $\Phi(x)$ in the appropriate basis the Goldstone boson components χ_i can be set to zero. This is also known as the unitary gauge and results in the scalar field

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.54}$$

As a result the Higgs to gauge couplings become,

$$\mathcal{L}_{HG} = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{g^2}{4}(v+h)^2 W_\mu^+ W^{-\mu} + \frac{1}{8}\frac{g^2}{\cos^2(\theta_W)}(v+h)^2 Z_\mu Z^\mu$$
$$+ \frac{\mu^2}{2}(v+h)^2 - \frac{\lambda}{4}(v+h)^4. \tag{2.55}$$

The full procedure of "gauging-away" the Goldstone bosons and constructing the above terms for the gauge bosons can be found in Appendix A.5. The new terms introduce the $W^\pm$- and Z^0-bosons as superpositions of the W_μ^i and B_μ fields. The Higgs-Gauge Lagrangian generates the mass terms of the W and Z bosons with the v^2-factors but leaves the photon field (A_μ) massless. The new masses are defined as

$$m_W = \frac{vg}{2}, \qquad m_Z = \frac{m_W}{\cos(\theta_W)}, \qquad m_h = \sqrt{-\mu^2}. \tag{2.56}$$

Similar to the Higgs to gauge couplings one can use the concepts of spontaneous symmetry breaking to generate mass terms for the fermions. The Higgs to fermion couplings can be written as

$$\mathcal{L}_{HF} = -\sum_{\text{u type}} g_u \bar{Q}_L \tilde{\Phi} u_R - \sum_{\text{d type}} g_d \bar{Q}_L \Phi d_R - \sum_{\text{l type}} \left(g_\nu \bar{L}_L \tilde{\Phi} \nu_{l,R} + g_l \bar{L}_L \Phi l_R \right) \tag{2.57}$$

with g_u, g_d, g_ν and g_l being the Yukawa couplings. By inserting the scalar field definition of Eq. 2.54 we get both mass terms for the fermions and interaction terms with the Higgs field,

$$
\begin{aligned}
\mathcal{L}_{HF} = & \overbrace{- \sum_{\text{u type}} \frac{g_u v}{\sqrt{2}} \big(\bar{Q}_L u_R - \bar{u}_R Q_L \big) - \sum_{\text{d type}} \frac{g_d v}{\sqrt{2}} \big(\bar{Q}_L d_R - \bar{d}_R Q_L \big)}^{\text{mass terms}} \\
& - \sum_{\text{l type}} \frac{g_l v}{\sqrt{2}} \big(\bar{L}_L l_R - \bar{l}_R L_L \big) \\
& \overbrace{- \sum_{\text{u type}} \frac{g_u}{\sqrt{2}} \big(\bar{Q}_L h\, u_R - \bar{u}_R h\, Q_L \big) - \sum_{\text{d type}} \frac{g_d}{\sqrt{2}} \big(\bar{Q}_L h\, d_R - \bar{d}_R h\, Q_L \big)}^{\text{higgs interaction terms}} \\
& - \sum_{\text{l type}} \frac{g_l}{\sqrt{2}} \big(\bar{L}_L h\, l_R - \bar{l}_R h\, L_L \big),
\end{aligned}
\tag{2.58}
$$

with the fermion masses

$$
m_i = -\frac{g_i v}{\sqrt{2}}, \qquad i = u,\, d,\, l. \tag{2.59}
$$

The introduction of the Higgs boson also solves the unitarity violation. Recall that the scattering amplitude of longitudinally polarized W bosons is dependent on the contribution of the 4-point interaction diagram. However, the Higgs boson introduces a new scattering process with the Higgs as mediator which adds new terms to the scattering amplitude. At large energies beyond the Higgs boson mass the result is,

$$
\mathcal{M}_W + \mathcal{M}_H = g^2 \frac{m_H^2}{4 m_W^2}. \tag{2.60}
$$

The $\mathcal{M}_W$ is the amplitude contribution of the 4-point interaction and $\mathcal{M}_H$ of the Higgs mediated interaction. The energy dependence has been canceled out by the Higgs contribution and the unitarity requirement has been restored [29].

2.5 Beyond the Standard Model

As accurate as the Standard Model might be in many cases, it still leaves many questions open. Many adjustments or replacement QFTs have been proposed and are also known as Beyond the Standard Model (BSM) theories. In this section a few open issues of the Standard Model are given to illustrate the necessity of extensions and thus motivate our measurements at the ATLAS experiment.

2.5.1 Dark Matter

One of the most prominent and obscure problems in particle physics is *dark matter*. Dark matter is matter that has only been observed via gravitation effects but does not seem to interact via the electromagnetic interaction. Many astrophysical and cosmological observations support the idea of dark matter through observations in rotation curves of galaxies [30, 31, 32, 33], the expansion of the universe [34] or gravitational lensing [35]. Many physicists are therefore set on finding a candidate particle that fits the dark matter profile. Some contemporary candidates are supersymmetric particles [36] such as the *neutralino*, particles in the Little Higgs model [37], *pyrgons* introduced by extra spatial dimensions such as in the Kaluza-Klein model [38] or *axions* introduced via the Peccei–Quinn mechanism [39]. However, none of the searches have shown any significant signals thus far [40, 41]. Many of these models predict new top quark pair production processes which can be tested or constrained with top quark pair data [42, 43].

2.5.2 Naturalness

A natural theory is one where dimensionless ratios of the parameters of the theory are of the order 1. This idea of naturalness can be seen as subjective and does not give any mathematical inconsistencies. However, parameters differing by many orders

often indicate that there are more advanced theories that would be an improvement. Some examples of such naturalness problems are the differences in fermion masses

$$\frac{m_\nu}{m_t} \sim 10^{-11} \tag{2.61}$$

or between the gravitational and weak force constants

$$\frac{G_N}{G_F} \sim 10^{-34} \tag{2.62}$$

for a neutrino mass m_ν, the top quark mass m_t, the gravitational force constant G_N and weak force constant G_F constants. In more general terms, the Standard Model will only be valid up until some energy scale Λ, generally believed to be the Planck scale $\Lambda_P = 1.22 \times 10^{19}$ GeV. This means that the Standard Model can be considered as an *effective field theory*, a type of quantum field theory that will be further explained in Sec. 2.6. The top quark is the only particle that has a coupling to the Higgs boson of order one. This makes the top quark ideal for studying the issue of naturalness in the SM.

2.5.3 Matter-Antimatter Asymmetry

Wherever we look, either here on earth or at far away stars and planets, we see that everything is predominantly made out of protons, neutrons and electrons i.e. matter. Some direct searches have looked for bodies of antimatter [44, 45, 46] in several regions of space but no significant finds have been made. However, according to the model of the Big Bang, equal amounts of matter and anti-matter were created from energy at the beginning of the universe. In 1967, Sakharov proposed three conditions [47] that are needed for the asymmetry:

1. Violation of baryon number

2. Violation of charge conjugation(C) and charge conjugation-parity(CP) symmetry

3. Departure from thermal equilibrium

The baryon number is defined as

$$B = \frac{1}{3}(n_q - n_{\bar{q}}) \tag{2.63}$$

with n_q and $n_{\bar{q}}$ being the number of quarks and anti-quarks, respectively. Processes that violate baryon number such as proton decay [48] and neutron-antineutron oscillation [49] are needed for an imbalance in quarks and anti-quarks. However, searches for such processes have been unsuccessful up until now. The second condition is needed to avoid baryon violating processes with their anti-matter partners to balance out the increase in matter. However, CP-violation has only been discovered in processes involving the weak interaction such as neutral kaon decay in 1964 [50] and B-meson decay in 2001 [51, 52]. Departures from thermal equilibrium is necessary to make sure that any process generating matter does not occur at the same rate in the reverse direction. Top quark pair data can be used to investigate the matter-antimatter asymmetry by measuring CP violation in top quark pair production and decay processes [53, 54].

2.5.4 Unifying Forces

In Sec. 2.3.3 we saw how two seemingly separate forces, the electromagnetic and weak, can be unified into one theory. The introduction of this new gauge theory reduced the unknown parameters in the Standard Model, introduced the W and Z gauge bosons and gave a picture more fitting with observations. Theories that unify all forces, Grand Unified Theories (GUTs) for all forces excluding gravity and Theories of Everything (ToEs) including gravity, could therefore give a more complete

and elegant picture of nature and pose solutions for the previous mentioned problems.

One of the simplest GUT theories is the Georgi-Glashow model [55] involving the gauge group $SU(5)$. However, limits set on proton life-times have ruled out the original proposed model [56]. Other GUT examples such as supersymmetric theories involving the $SU(5)$ group [57, 58, 59] or $E(6)$ [60] group have yet to be excluded. Many of these theories can be tested or constrained with top quark data [61, 62].
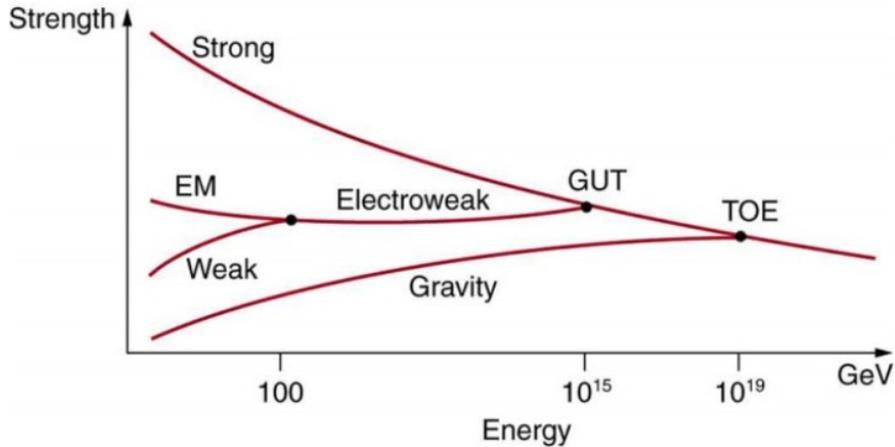
Figure 2.4: The relative strengths of the four fundamental forces vary with distance and, hence, energy is needed to probe small distances. From [63].

Gravity on larger scales is described by Einstein's theory of general relativity [64]. However, for physical scenarios such as black holes and the Big Bang a quantum mechanical description of gravity is needed [65] in which the Standard Model falls short. Theories such as loop quantum gravity [66, 67] or large extra dimensions [68] could be the needed extension of the Standard Model. Indirect searches of gravitons, i.e. gravity gauge bosons, [69, 70] and a recently proposed laser table-top experiment [71] continue the search for subtle signs of these theories. However, arguably the leading candidate for a theory of anything is M-theory [72, 73]. More generally also

known as String theory, this theory suggests that particles can be represented by tiny energy strings and gives a framework from which all the five forces naturally emerge. However, experimental setups have yet to find even subtle evidence [74, 75, 76] that supports this ToE.

2.6   Effective Field Theory

Many of the discussed BSM theories predict particles or manifest themselves as other measurable effects such as deviations of Standard Model parameters. The current available experimental evidence, be it from experiments of the LHC or other physics programmes, excluded many scenarios but does not point to one of these BSM theories as the obvious successor [77, 78, 79]. Effective field theory (EFT) supplies a model-independent framework that can be matched to many BSM theories. A single EFT analysis can therefore cover many BSM scenarios instead of systematically excluding phase-space one BSM theory at a time. Additionally, it can show effects of heavy undiscovered particles in the low-energy regime of our current experimental setups. In this section we will introduce the concept of an effective field theory and the Standard Model Effective Field Theory, the EFT used in this thesis.

2.6.1   Basics of EFT

Lets define some energy scale Λ and a Lagrangian $\mathcal{L}_{UV}$ describing physics at energies $E \gg \Lambda$. Recall from Sec. 2.2 that cross sections are calculated from n-point correlation functions which depend on functionals. Here we redefine Eq. (2.18) as [80, 81]

$$ Z_{UV}[J_\phi, J_H] = \int \mathcal{D}\phi \, \mathcal{D}X \, \exp\left[i/\hbar \int d^4x \left(\mathcal{L}_{UV}(\phi, X) + J_\phi \phi + J_X X \right) \right] \qquad (2.64) $$

where the distinction is made between *light fields* ϕ and *heavy fields* X. Now let us assume for now that X particles are not present in our physical process, e.g. because they are to heavy to create with our probing energies, and thus we can "integrate out" X. From this we get the effective functional

$$Z_{EFT}[J_\phi] = \int \mathcal{D}\phi \exp\left[i/\hbar \int d^4x \left(\mathcal{L}_{EFT}(\phi) + J_\phi\phi\right)\right]. \qquad (2.65)$$

The non-trivial task is to find $\mathcal{L}_{EFT}$ with the relevant degrees of freedom and field content. In general, an effective lagrangian is an expansion in operator dimension D

$$\mathcal{L}_{EFT} = \mathcal{L}^{D\leq 4} + \frac{1}{\Lambda}\mathcal{L}^{D=5} + \frac{1}{\Lambda^2}\mathcal{L}^{D=6} + \frac{1}{\Lambda^3}\mathcal{L}^{D=7} + \frac{1}{\Lambda^4}\mathcal{L}^{D=8} + ... \qquad (2.66)$$

with Λ being the cut-off scale and the Lagrangian terms $\mathcal{L}^D$ consisting of operators with dimension D constructed from the lighter fields ϕ. In Appendix A.6 it is shown with a simple toy scalar theory how an effective Lagrangian can be a good approximation for a full UV theory Lagrangian.

2.6.2 The Standard Model Effective Field Theory

The Standard Model Effective Field Theory (SMEFT) is a defined by an effective Lagrangian containing the Standard Model Lagrangian but expands into higher dimensional operators constructed from SM fields. Following Eq. (2.66) we can define the SMEFT Lagrangian as

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \frac{1}{\Lambda}\sum_i c_i \mathcal{O}_i^{D=5} + \frac{1}{\Lambda^2}\sum_j c_j \mathcal{O}_j^{D=6} + \frac{1}{\Lambda^3}\sum_k c_k \mathcal{O}_k^{D=7} + \frac{1}{\Lambda^4}\sum_l c_l \mathcal{O}_l^{D=8} + ...$$

$$(2.67)$$

The operators are constructed from SM fields and are invariant under the same $SU_C(3) \times SU_L(2) \times U_Y(1)$ gauge symmetry. Operators of dimension $D = 5$ vio-

late baryon and/or lepton number which does not align with current experimental evidence and are therefore omitted [82, 83]. The set of all operators of dimension $D = 6$ is very large but will contain many redundant operators. Operators are redundant when their contribution to the matrix element construction cancel each other out. Instead, the fields are redefined such that those terms already cancel out in the Lagrangian [84].

A minimal non-redundant set of operators is also known as a *basis*. The first non-redundant and most commonly used operator basis constructed for dimension $D = 6$ is the *Warsaw basis* [85] and amounts to 59 independent CP-even, B-conserving operators, excluding their Hermitian conjugates. There are other $D = 6$ operators sets which have been utilized such as the Strongly-Interacting Light Higgs (SILH) [86] or Hagiware-Ishihara-Szalapski-Zeppenfeld (HISZ) operator set. However, these sets are not bases as they are neither a complete nor a non-redundant set. The SMEFT model implementation [87] that is used in this thesis further simplifies the operator set by imposing a $U(2)_q \times U(2)_u \times U(3)_d$ flavour symmetry on the first two generations of left-handed quark doublets, up-type right-handed singlets and three generations of down-type right-handed singlets. This symmetry is suitable for this thesis because top quark observables at the LHC are mostly invariant under changes between light flavors with the same quantum numbers of the incoming quarks [88, 89, 90]. Chirality flipping or charged currents involving light quarks are therefore forbidden. Similarly, a $U(1)_l \times U(1)_e$ symmetry is enforced in the lepton sector. The CKM-matrix is taken to be a unit matrix, i.e. diagrams with charged currents coupling different generations of fermions do not contribute, and it forbids all fermion masses and Yukawa couplings except for the top quark.

2.6.3 SMEFT Top Quark Sector

The Standard Model Effective Field Theory can be used to interpret observations from many different physics processes. However, in view of its largest Yukawa coupling and mass it is believed that the top quark is closely connected to new physics [84, 91]. Studying the couplings of the top quark can therefore give unique insights on where to look beyond the Standard Model. In this section we will discuss the operators that affect the production of top quark pairs as produced at the LHC. Let us denote left-handed quark $SU(2)$ doublets with q and right-handed quark singlets of the up- and down-type with u and b. As introduced in Sec. 2.3.3 and 2.3.4 we denote the Pauli matrices with τ and $T^a = \frac{1}{2}\lambda^a$ with the Gell-mann matrices λ^a. Flavor indices are denoted with i, j, k and l and the Higgs doublet is denoted by ϕ. Some additional notations needed for the operator definitions are $\epsilon \equiv i\tau^2$, $\sigma_{\mu\nu} = \gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu$, $\tilde{\phi} = \epsilon\phi^*$, $\phi^\dagger i\bar{D}_\mu\phi \equiv \phi^\dagger(iD_\mu\phi) - (iD_\mu\phi^\dagger)\phi$ and $\phi^\dagger i\bar{D}_\mu^I\phi \equiv \phi^\dagger\tau(iD_\mu\phi) - (iD_\mu\phi^\dagger)\tau\phi$. The dimension six SMEFT operators relevant for top quark processes at the LHC can be divided into two groups; operators involving two quarks and bosons and operators involving four quarks. The bosonic operators are defined as

$$
\begin{aligned}
&{}^\ddagger\mathcal{O}_{u\phi}^{(ij)} = \bar{q}_i u_j \tilde{\phi}(\phi^\dagger\phi) && {}^\ddagger\mathcal{O}_{uW}^{(ij)} = (q_i\sigma^{\mu\nu}\tau u_j)\tilde{\phi}W_{\mu\nu} \\
&\mathcal{O}_{\phi q}^{1(ij)} = (\phi^\dagger i\bar{D}_\mu\phi)(\bar{q}_i\gamma^\mu q_j) && {}^\ddagger\mathcal{O}_{dW}^{(ij)} = (q_i\sigma^{\mu\nu}\tau u_j)\phi W_{\mu\nu} \\
&\mathcal{O}_{\phi q}^{3(ij)} = (\phi^\dagger i\bar{D}_\mu^I\phi)(\bar{q}_i\gamma^\mu\tau q_j) && {}^\ddagger\mathcal{O}_{uB}^{(ij)} = (q_i\sigma^{\mu\nu}u_j)\tilde{\phi}B_{\mu\nu} \\
&\mathcal{O}_{\phi u}^{(ij)} = (\phi^\dagger i\bar{D}_\mu\phi)(\bar{u}_i\gamma^\mu u_j) && {}^\ddagger\mathcal{O}_{uG}^{(ij)} = (q_i\sigma^{\mu\nu}T^a u_j)\tilde{\phi}G_{\mu\nu}^a \\
&{}^\ddagger\mathcal{O}_{\phi ud}^{(ij)} = (\tilde{\phi}^\dagger iD_\mu\phi)(\bar{u}_i\gamma^\mu d_j)
\end{aligned}
\tag{2.68}
$$

with $W_{\mu\nu}$, $B_{\mu\nu}$ and $G_{\mu\nu}^a$ being the electroweak and strong field-strength tensors introduced in Sec. 2.3.3 and 2.3.4. The operators involving four quarks are defined as

$$\mathcal{O}_{qq}^{1(ijkl)} = (\bar{q}_i\gamma^\mu q_j)(\bar{q}_k\gamma_\mu q_l)$$

$$\mathcal{O}_{uu}^{(ijkl)} = (\bar{u}_i\gamma^\mu u_j)(\bar{u}_k\gamma_\mu u_l)$$

$$\mathcal{O}_{qq}^{3(ijkl)} = (\bar{q}_i\gamma^\mu\tau q_j)(\bar{q}_k\gamma_\mu\tau q_l)$$

$$\mathcal{O}_{ud}^{1(ijkl)} = (\bar{u}_i\gamma^\mu u_j)(\bar{d}_k\gamma_\mu d_l)$$

$$\mathcal{O}_{qu}^{1(ijkl)} = (\bar{q}_i\gamma^\mu q_j)(\bar{u}_k\gamma_\mu u_l)$$

$$\mathcal{O}_{ud}^{8(ijkl)} = (\bar{u}_i\gamma^\mu T^a u_j)(\bar{d}_k\gamma_\mu T^a d_l) \qquad (2.69)$$

$$\mathcal{O}_{qu}^{8(ijkl)} = (\bar{q}_i\gamma^\mu T^a q_j)(\bar{u}_k\gamma_\mu T^a u_l)$$

$$^\ddagger\mathcal{O}_{quqd}^{1(ijkl)} = (\bar{q}_i u_j)\epsilon(\bar{q}_k d_l)$$

$$\mathcal{O}_{qd}^{1(ijkl)} = (\bar{q}_i\gamma^\mu q_j)(\bar{d}_k\gamma_\mu d_l)$$

$$^\ddagger\mathcal{O}_{quqd}^{8(ijkl)} = (\bar{q}_i T^a u_j)\epsilon(\bar{q}_k T^a d_l)$$

$$\mathcal{O}_{qd}^{8(ijkl)} = (\bar{q}_i\gamma^\mu T^a q_j)(\bar{d}_k\gamma_\mu T^a d_l).$$

Each operator introduces a degree of freedom with similar notation i.e. $\mathcal{O}_{qq}^{1(ijkl)}$ introduces the Wilson coefficient $C_{qq}^{1(ijkl)}$. However, linear combinations of these Wilson coefficients were introduced in [92] and since then adopted to reduce the number of relevant parameters and reduce cancelling effects on LHC observables. The four-quark Wilson coefficients can be redefined as

$$c_{QQ}^1 \equiv 2C_{qq}^{1(3333)} - \frac{2}{3}C_{qq}^{3(3333)}$$

$$c_{QtQb}^1 \equiv \mathrm{Re}\{C_{quqd}^{1(3333)}\} \qquad c_{td}^1 \equiv C_{ud}^{1(33ii)}$$

$$c_{QQ}^8 \equiv 8C_{qq}^{3(3333)}$$

$$c_{QtQb}^8 \equiv \mathrm{Re}\{C_{quqd}^{8(3333)}\} \qquad c_{td}^8 \equiv C_{ud}^{8(33ii)}$$

$$c_{Qt}^1 \equiv C_{qu}^{1(3333)}$$

$$c_{Qq}^{1,1} \equiv C_{qq}^{1(ii33)} + \frac{1}{6}C_{qq}^{1(i33i)} + \frac{1}{2}C_{qq}^{3(i33i)} \qquad c_{tq}^1 \equiv C_{qu}^{1(ii33)}$$

$$c_{Qt}^8 \equiv C_{qu}^{8(3333)}$$

$$c_{Qq}^{3,1} \equiv C_{qq}^{3(ii33)} + \frac{1}{6}(C_{qq}^{1(i33i)} - C_{qq}^{3(i33i)}) \qquad c_{Qu}^1 \equiv C_{qu}^{1(33ii)}$$

$$c_{Qb}^1 \equiv C_{qd}^{1(3333)}$$

$$c_{Qq}^{1,8} \equiv C_{qq}^{1(i33i)} + 3C_{qq}^{3(i33i)} \qquad c_{Qd}^1 \equiv C_{qd}^{1(33ii)}$$

$$c_{Qb}^8 \equiv C_{qd}^{8(3333)}$$

$$c_{Qq}^{3,8} \equiv C_{qq}^{1(i33i)} - C_{qq}^{3(i33i)} \qquad c_{tq}^8 \equiv C_{qu}^{8(ii33)}$$

$$c_{tt}^1 \equiv C_{uu}^{1(3333)}$$

$$c_{tu}^1 \equiv C_{uu}^{ii33} + \frac{1}{3}C_{uu}^{i33i} \qquad c_{Qu}^8 \equiv C_{qu}^{8(33ii)}$$

$$c_{tb}^1 \equiv C_{ud}^{1(3333)}$$

$$c_{tu}^8 \equiv 2C_{uu}^{i33i} \qquad c_{Qd}^8 \equiv C_{qd}^{8(33ii)}$$

$$c_{tb}^8 \equiv C_{ud}^{8(3333)}$$

$$(2.70)$$

with Q, t and b denoting coefficients involving the left-handed top-bottom doublet, right-handed top singlet and right-handed bottom singlet, respectively. Note that the imaginary parts of certain operators are not included as we will restrict ourselves to CP-even scenarios from now on. The i is an index corresponding to one of the two lighter quark generations. No distinction is made with the i index which is equivalent to imposing the flavor symmetry $U(2)_q \times U(2)_u \times U(3)_d$ introduced in the previous section. The Wilson coefficients involving the bosonic operators can also be redefined as

$$
\begin{aligned}
&c_{t\phi} \equiv \mathrm{Re}\{C_{u\phi}^{(33)}\} \\
&c_{\phi Q}^{-} \equiv C_{\phi q}^{1(33)} - C_{\phi q}^{3(33)} && c_{tW} \equiv \mathrm{Re}\{C_{uW}^{(33)}\} \\
&c_{\phi Q}^{+} \equiv C_{\phi q}^{1(33)} + C_{\phi q}^{3(33)} && c_{tZ} \equiv \mathrm{Re}\{-s_W C_{uB}^{(33)} + c_W C_{uW}^{(33)}\} \\
&c_{\phi Q}^{3} \equiv C_{\phi q}^{3(33)} && c_{tA} \equiv \mathrm{Re}\{c_W C_{uB}^{(33)} + s_W C_{uW}^{(33)}\} && (2.71) \\
&c_{\phi t} \equiv C_{\phi u}^{(33)} && c_{bW} \equiv \mathrm{Re}\{C_{dW}^{(33)}\} \\
& && c_{tG} \equiv \mathrm{Re}\{C_{uG}^{(33)}\} \\
&c_{\phi tb} \equiv \mathrm{Re}\{C_{\phi ud}^{(33)}\}
\end{aligned}
$$

with $s_W = \sin\theta_W$ and $c_W = \cos\theta_W$ and θ_W being the Weinberg angle. The redefined four-quark and bosonic Wilson coefficients give a reduced number of degrees of freedom of 36 of which 34 are independent. The Wilson coefficients of each operator will have an effect on $t\bar{t}$ cross-sections, total and differential. Figure 2.5 gives an example on how SMEFT Wilson coefficients can affect $t\bar{t}$ differential cross sections quite differently. The Wilson coefficients can have either a linear c_i/Λ^2 or a quadratic $c_i c_j/\Lambda^4$ contribution to cross sections. By comparing observed $t\bar{t}$ events with SMEFT Monte Carlo simulations we can put constraints on Wilson coefficients and further explore possible BSM theories. The decay channel and differential cross section observables will be further discussed in Ch. 6. The choice of the c_{tq}^{8} and $c_{Qq}^{8,1}$ Wilson coefficients

and its motivation will be further discussed in Ch. 7.
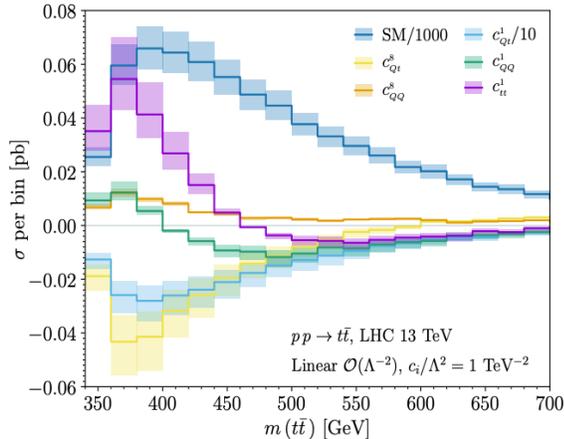
Figure 2.5: SMEFT predictions of $t\bar{t}$ invariant mass differential cross sections. The distributions show the linear $\mathcal{O}(\Lambda^{-2})$ SM-EFT interference effects of five four-quark Wilson coefficients. The error bands represent the renormalization/factorization scale uncertainty. From [87].

2.7   Top Quark Pair Production at the LHC

In this section we will discuss the theoretical framework used to make predictions of $pp \rightarrow t\bar{t}$ production cross sections, total and differential, as produced at the LHC. Additionally, we will include the software packages that implement parts of this framework and are used in this thesis to generate Monte Carlo simulated data. Figure 2.9 gives a schematic view of the Monte Carlo simulation chain.

2.7.1   Hard Scatter Processes

Interactions between composite particles, such as protons, can not be described from first principles. However, the factorization theorem [93, 94] states that we can factorize the total production cross section $\sigma_{pp \rightarrow t\bar{t}}$ according to

$$\sigma_{pp \to t\bar{t}} = \sum_{a,b} \int_0^1 dx_1 \, dx_2 \int f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) d\hat{\sigma}^{ab \to t\bar{t}}(\alpha_s(\mu_R), \mu_F)$$

$$= \sum_{a,b} \int_0^1 dx_1 \, dx_2 \int d\Phi \, f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \frac{1}{2x_1 x_2 \sqrt{s}} |\mathcal{M}_{ab \to t\bar{t}}|^2 (\Phi, \alpha_s(\mu_R), \mu_F)$$

$$(2.72)$$

where a is a parton of one of the protons, b is a parton of the other proton, x_1 and x_2 are the energy fractions of resp. parton a and b, μ_F is the factorisation scale and α_s is the coupling constant evaluated at the renormalisation scale μ_R and related to the coupling constant g_3 of the strong interaction via $\alpha_s = \frac{g_3^2}{4\pi}$. The functions $f_a(x_1, \mu_F^2)$ and $f_b(x_2, \mu_F^2)$ are the parton distribution functions (PDF) of a proton, $\hat{\sigma}^{ab \to t\bar{t}}$ is the partonic cross section and $\mathcal{M}_{ab \to t\bar{t}}$ is the matrix element for that partonic process evaluated at a particular point in final-state phase space Φ. One can either obtain the matrix element with a similar approach to Sec. 2.2.4 or use Feynman rules for a specific Lagrangian. The latter requires one to identify all the Feynman diagrams of the partonic process up to a fixed QCD order. At the LHC the partonic processes for $t\bar{t}$ production are gluon fusion $gg \to t\bar{t}$ and quark-antiquark annihilation $q\bar{q} \to t\bar{t}$. Figure 2.6 and 2.7 show all the LO and some NLO Feynman diagrams of $t\bar{t}$ partonic production processes, respectively. At the LHC the gluon fusion channel constitutes about 90% of the total $t\bar{t}$ production.
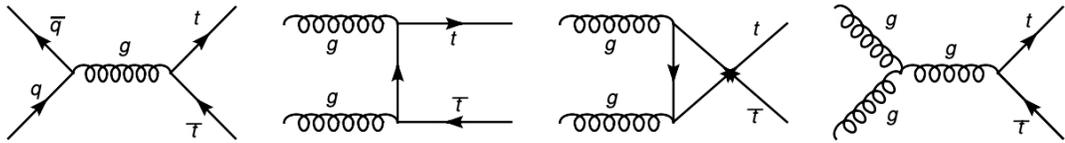
Figure 2.6: LO Feynman diagrams for the $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$ partonic processes. From [95].

The parton distribution functions give the probability to find a quark or gluon in a proton with a particular energy fraction of the proton. The functions are process

39
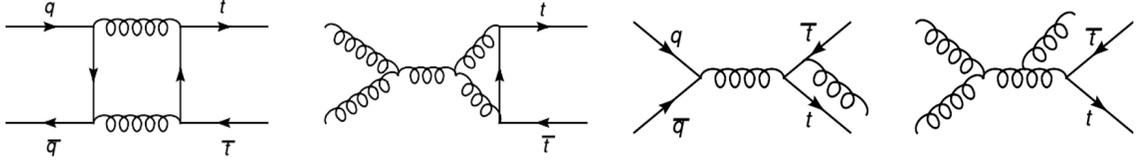
Figure 2.7: NLO Feynman diagrams with virtual one-loop corrections and real gluon/quark emissions for the $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$ partonic processes. From [95].
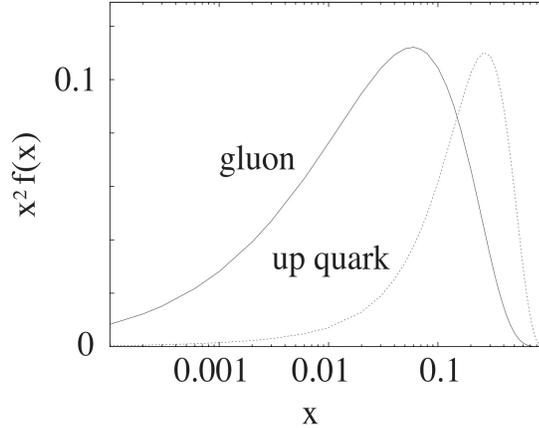
Figure 2.8: Gluon and up quark distributions in the proton according to the CTEQ3M [96] parton distribution set. From [97].

independent parameterizations fitted to experimental data of multiple scattering processes measured by various experiments. Most Monte Carlo software packages have several PDFs available from dedicated research projects like CT10 [98] and NNPDF [99] or have an interface to the LHAPDF PDF library [100]. Figure 2.8 shows two example PDFs published by the CTEQ collaboration [96].

Let us assume that we chose a PDF and calculated the matrix element. The last step to obtain the total production cross section $\sigma_{pp \to t\bar{t}}$ is to perform the multi-dimensional phase-space integral $\int_0^1 dx_1 dx_2 \int d\Phi$. Often, these are too complicated to approach analytically and are therefore calculated with Monte Carlo integration methods [101, 102]. The result is a fully differential cross section that also can be used to randomly generate events. Generator software packages used in this thesis to

calculate cross sections and generate events are:

- POWHEG-BOX [103, 104, 105, 104]

- MADGRAPH5_AMCNLO [106]

- SHERPA [107]

- COMIX [108]

- OPENLOOPS [109]

The input QFT models dictating the Feynman rules for the matrix element calcula-
tion are either available in libraries implemented in the generator software packages
directly or are linked to external libraries such as `FeynRules` [110]. Their QCD and
EW accuracy can vary alongside with the available scattering processes.

2.7.2 Underlying Event and Pile-up

In many collisions remaining partons of the incoming protons can result in secondary
hard or semi-hard scattering events. These are known as Underlying Events (UE)
and are characterized by both high (hard) and low (soft) p_T transfer between the
scattering particles. The latter is where perturbative QCD breaks down and other
phenomenological models need to be deployed such as the Multiple Interaction (MI)
model first implemented in PYTHIA 6 [111] and now also in PYTHIA 8 [112]. These or
other phenomenological models are often but not exclusively tuned to Minimum Bias
events, events selected with a less stringent trigger such that more soft interactions
pass. Additional collisions than the collision of interest are noted as *pile-up* events
[113]. Here we distinguish between *in-time* and *out-time* pileup where the secondary
collisions originate from the same and different bunch crossings, respectively. Both

types are commonly simulated with Pythia 8 [114] for an average number of interactions per bunch crossing $\langle \mu \rangle$ which depends on the luminosity and detector settings of the data taking campaign one wants to simulate.
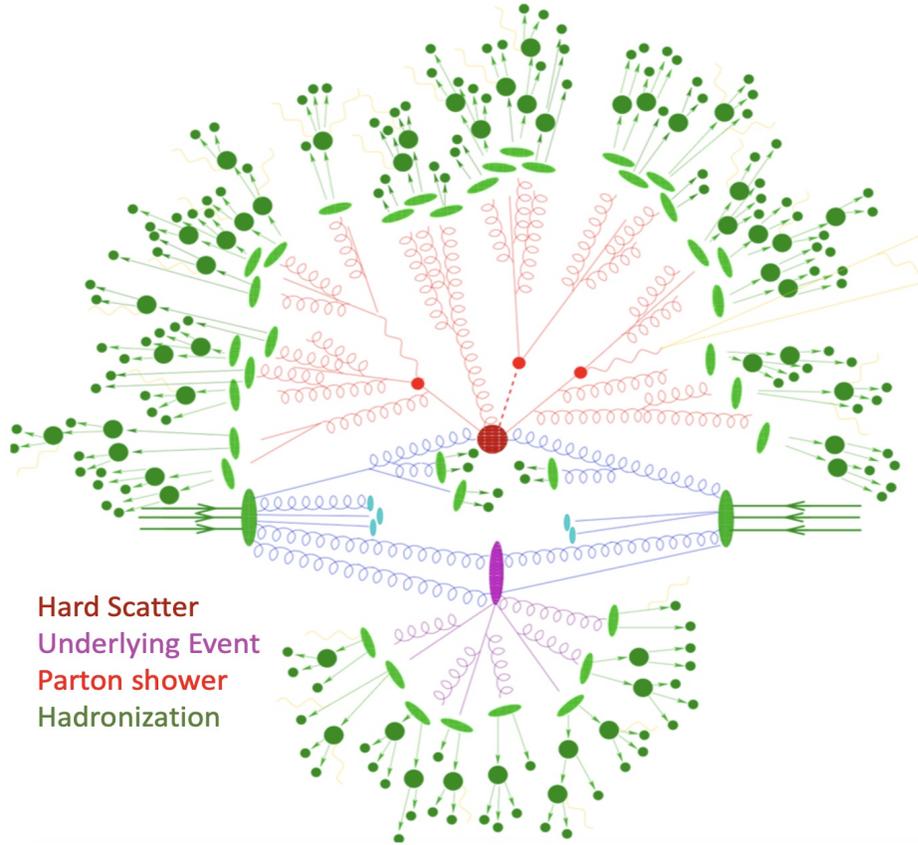
Figure 2.9: Schematic view of the physics implemented in standard multi-purpose event generators for proton-proton collisions. The colours indicate the different parts of the simulation chain. From [115].

2.7.3 Parton Shower and Hadronization

After events are generated on parton level the emerging constituents need to be connected to a hadronization model that works at energy scales where perturbative QCD no longer works. In this process high-energy partons can emit *QCD radiation* in the form of gluons. Here we make the distinction between radiation before and after the collision, also known as initial state (ISR) and final state radiation (FSR). Addition-

ally, gluons can split in pairs of quark and anti-quarks. The cascade of emerging quarks and gluons are known as *parton showers* and its evolution is dictated by the *DGLAP equations* [116]. In general terms, DGLAP equations describe how parton distribution functions and splitting kernel functions, i.e. functions that describe the parton splitting dynamics, evolve from one energy scale to another.

Once the energy scale of a parton shower is down to $1\,\mathrm{GeV}$ the partons are modelled to form hadrons, also known as hadronization. At these energies *colour confinement* dictates that colour charged partons cannot be isolated and observed but need to form colour neutral hadrons. Hadronization can be divided into cluster and string models. Cluster models split gluons non-perturbatively into quark anti-quark pairs and form singlet clusters which further decay into hadrons. String models assume partons from the parton shower to be connected into a string-like configuration assuming a potential linearly increasing with distance. The string eventually breaks up into quark anti-quark pairs that form hadrons. Softwares used in this thesis that model parton showers and their hadronization are:

- HERWIG 7 7 [117, 118]

- PYTHIA 8 [114]

- SHERPA [107, 119]

Chapter 3

Experimental Setup

This chapter discusses the experimental setup of the measurement carried out in this thesis which can be roughly divided into two parts. Section 3.1 and 3.2 describe the Large Hadron Collider and the ATLAS detector which are a collective of systems that acquire, filter and process pp collision data into a form that can be used for comparison with theory. Both are hosted by the European Organization for Nuclear Research (CERN), a research centre based in Geneva, Switzerland that houses 2660 staff members and 12,400 scientists from institutions in more than 70 countries [120]. Apart from the ATLAS experiment it has housed many other fruitful experiments and has constantly pushed the boundaries in engineering, physics and international collaboration since its founding in 1954.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular collider built in a 27 km long tunnel located 100 meters underground and has been operational since 2008 [121]. The accelerator is constructed of two near-circular rings that accelerate either heavy ions, such as lead (Pb) or xenon (Xe), or protons. The latter is the type of collisions used for this thesis. Both rings accelerate the particles in opposite directions and are brought

to intersect at several interaction points. Each of these points has one or multiple detectors in place to collect the collision data. The ATLAS detector is one of these experimental setups and the focus of this thesis.

The accelerated protons are created by stripping the electrons of hydrogen atoms with an electric field. The protons are first injected in the linear accelerator 2 (LINAC2) and then subsequently in the Booster, Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) to gradually ramp up to an energy of $450\,\text{GeV}$ [122]. Protons with these energies are injected into the LHC in bunches with a nominal number of protons of 1.15×10^{11} per bunch. A beam is created by injecting the bunches $25\,\text{ns}$ apart and can hold up to 2808 bunches per beam [123]. Once injected the beams are accelerated up to an energy of $7\,\text{TeV}$ per beam with the use of linear accelerators [124]. This analysis uses data generated from beams with $6.5\,\text{TeV}$ energy, i.e. $\sqrt{s} = 13\,\text{TeV}$ center-of-mass energy. To bend and focus the beam the LHC is equipped with superconducting niobium-titanium magnets which create magnetic fields of up to $8.33\,\text{T}$ when cooled down to a temperature of $1.9\,\text{K}$ with the use of liquid helium [125]. The beams collide at one of the four interaction points where the following experiments perform the data taking:

- **ATLAS** [126] is a general purpose detector that focuses on the discovery of new particles, interactions beyond the Standard Model and precision measurements in electroweak and QCD physics.

- **CMS** [127] is also a general purpose detector that focuses on scientific goals similar to the ones of ATLAS.

- **ALICE** [128] is a detector designed for heavy-ion collisions to study the physics of strongly interacting matter.

- **LHCb** [129] is a detector designed to investigate the difference between matter

and antimatter and study the bottom quark.

Each detector has a unique design and consists out of many subsystems fine-tuned to optimally perform for their specific physics program. In addition to these there is a whole collection of smaller experiments located in the interaction point caverns or on the above ground sites of CERN. After data taking the beam is safely dumped at one of the dumping caverns. The beam dump absorber consists of a 7 m long segmented carbon cylinder, water cooled and surrounded by tonnes of iron and concrete shielding to safely stop the beam [130]. An overview of the Large Hadron Collider and CERN complex is given in Fig. 3.1.
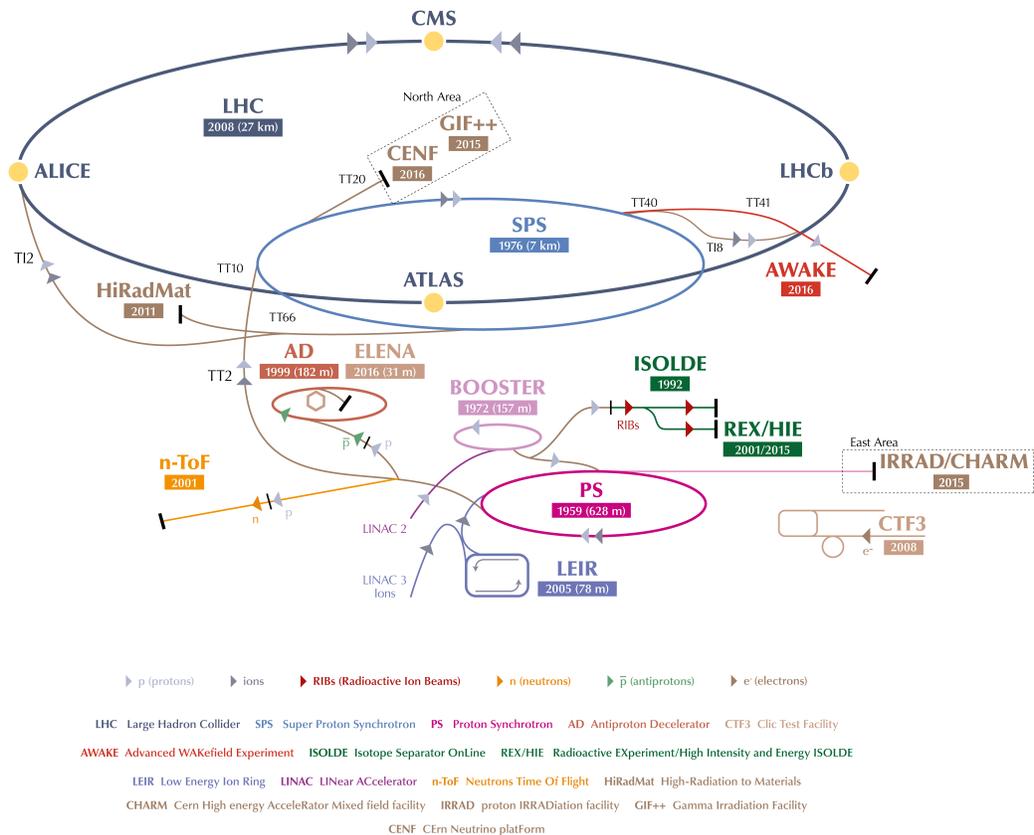
Figure 3.1: The CERN accelerator complex. From [131].
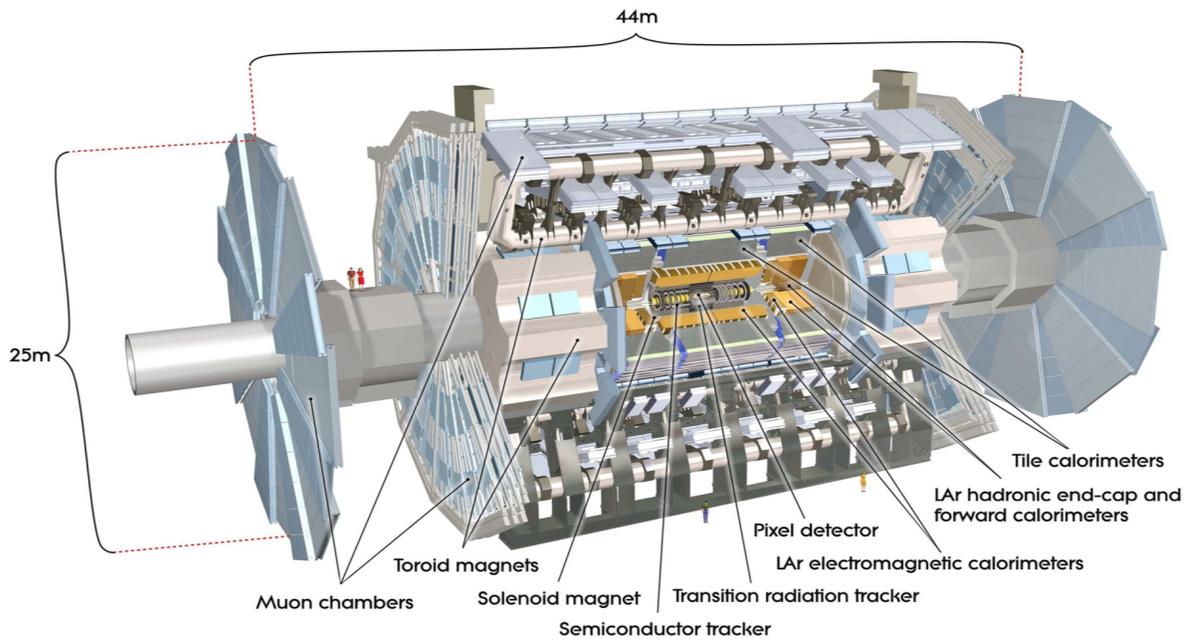
Figure 3.2: Schematic of the ATLAS detector. From [132].
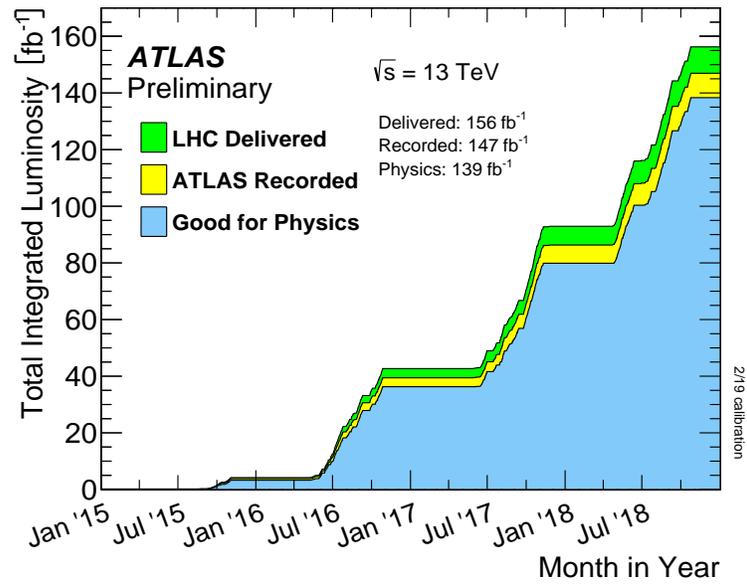
Figure 3.3: Total integrated luminosity of pp collisions at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ during the 2015 to 2018 data taking period. From [133].

Figure 3.3 shows that during the data taking period of 2015 to 2018 the Large Hadron Collider has delivered an integrated luminosity of $159\,\mathrm{fb}^{-1}$ in pp collision data at a center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$. The data recorded at the ATLAS detector, passing data quality criteria and used for this thesis amount to an integrated luminosity of $139\,\mathrm{fb}^{-1}$.

3.2   The ATLAS Detector

The ATLAS (A Toroidal LHC Apparatus) detector is a general purpose detector of $44\,\mathrm{m}$ long, $25\,\mathrm{m}$ in diameter and a weight of about $7.000.000\,\mathrm{kg}$ [134]. It is the largest detector at the LHC and is mainly focused on physics at the TeV scale. The detector has a cylindrical layer structure of subdetectors that each have their own specific task in particle detection and identification. A schematic representation of the detector is given in Fig. 3.2. The interaction point of the particle beams is the origin of both the Cartesian and the cylindrical coordinate systems. The Cartesian z-coordinate is aligned with the beam line with the positive direction pointing to Geneva, the x-coordinate directed from the interaction point to the centre of the LHC ring and the y-coordinate directed upwards. However, a cylindrical coordinate system is more practical in many cases and is defined as:

- Parameter z - The z-coordinate is defined along the beam axis.

- Parameter θ - The polar angle, $\theta \in [0, \pi]$, is the angle from the beam axis.

- Parameter ϕ - The azimuthal angle, $\phi \in [0, 2\pi]$, is the angle around the beam axis.

Other frequently occurring spacial metrics are the pseudorapidity $\eta = -\ln\left[\tan(\frac{\theta}{2})\right]$ instead of the polar angle θ and the distance coordinate $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. In the

upcoming sections each subdetector is described starting from the most inner parts to the outer side of the ATLAS detector.

3.2.1 Inner Detector

One of the most important characteristics of a collision event are the trajectories and interaction points of the different particles that are created at the collision. The purpose of the Inner Detector (ID) [135] is to reconstruct these trajectories and interaction points, also known as tracks and vertices. It is therefore located closest to the interaction point of all the ATLAS subdetectors. A strong magnetic field is created by the magnet system, further discussed in Sec. 3.2.4, such that the charged particles experience a Lorentz force that curves the particle tracks. Knowing the magnetic field and the curvature of a track one can calculate the tranverse momentum p_T of the traversing particle [136]. The Inner Detector has an outer radius of 115 cm, a total length of 7 m and is divided into four systems of which each uses a different but complementary technology. A schematic of the Inner Detector is given in Fig. 3.5.
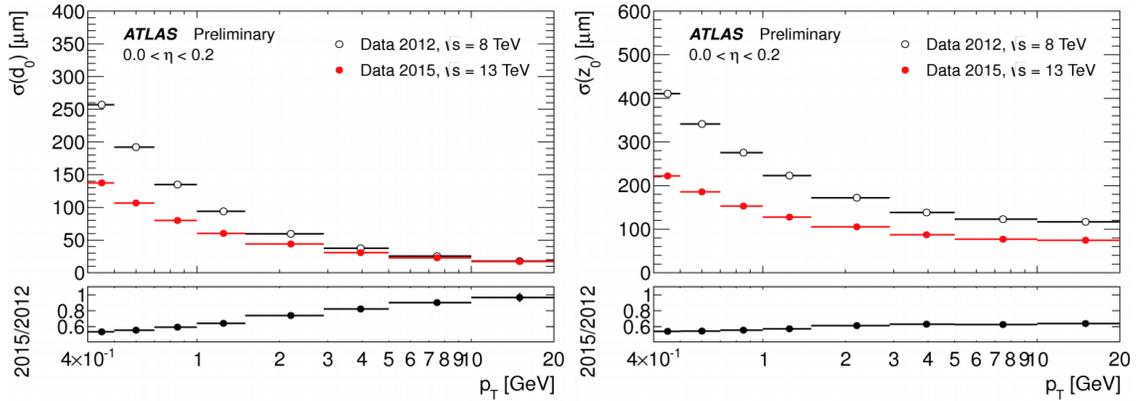
Figure 3.4: The resolutions of the track reconstruction parameters d_0 (left) and z_0 (right) as a function of the transverse momentum p_T with (Data 2015) and without the IBL (Data 2012). From [137].

Insertable B-layer

This Insertable B-layer (IBL) [138, 137] is the tracker layer that is the closest to the interaction point with a radial extension of $31.0 < r < 40.0\,\mathrm{mm}$ and a length of $332\,\mathrm{mm}$. The layer is made out of a grid of 12 million doped silicon cells, i.e. pixels, each with a size of $50 \times 250\,\mu\mathrm{m}^2$. When charged particles traverse the active silicon material, electron-hole pairs will be created and will create a current under an applied bias voltage. It was added to the Inner Detector during Long Shutdown 1 (2013-2015) to improve on radiation hardness and increased data rate foreseen with the increased instantaneous luminosity of the LHC Run-2 data taking. The IBL also improved the track reconstruction by reducing the resolution on important track reconstruction parameters. Figure 3.4 shows how the inclusion of the IBL reduced the resolution for the transverse impact parameter d_0 and the longitudinal impact parameter z_0. Section 4.1 will further discuss what these parameters actually mean and how they are used to reconstruct tracks and vertices.

Pixel Detector

The Pixel Detector (PD) [139] consists out of three concentric cylindrical layers and three end-cap disks at each side and has a pseudo rapidity coverage of $|\eta| < 2.5$. The layers are made of silicon cells that uses the same detection process as the IBL but have a bigger size of $50 \times 400\,\mu\mathrm{m} \times \mu\mathrm{m}$. Together with the IBL the Pixel Detector is able to distinguish vertices with a spatial resolution of $\delta r \delta \phi \approx 8\,\mu\mathrm{m}$ and $\delta z = 75\,\mu\mathrm{m}$ [140]. A schematic of the Pixel Detector is depicted in Fig. 3.6.

Semiconductor Tracker

The semiconductor tracker (SCT) [141] consists out of four concentric cylindrical layers and two end-caps which also have a pseudo rapidity coverage of $|\eta| < 2.5$. Instead of independent pixels the active silicon material has rectangular strips with
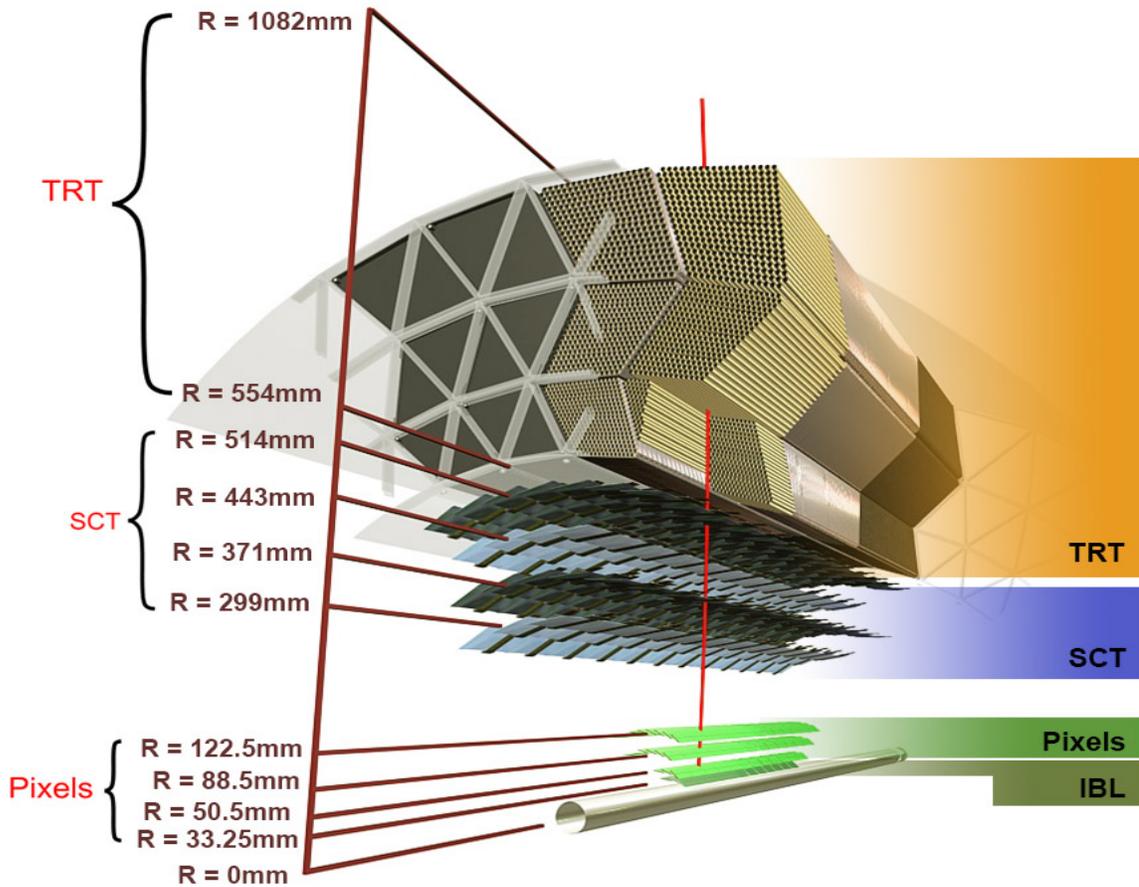
Figure 3.5: Schematic of the Inner Detector. From [132].

a $80\,\mu$m pitch mounted on them to read-out the charge signal.

Transition Radiation Tracker

The Transition Radiation Tracker (TRT) [142] consists of one barrel and one end-cap on each side covering a pseudo-rapidity range of $|\eta| < 2$. Instead of silicon cells, the barrel and end-caps consist of ~ 300000 thin-walled straws drift tubes that have a conductive outer coating, a golden plated tungsten sense wire in the centre and filled with a gas mixture of xenon, carbon dioxide and oxygen. When a charged particle passes they ionise the gas mixture which frees up electrons to move freely. The outer wall of the straws are put under a negative voltage which will accelerate the electrons and create a detectable charge current. Additionally, the TRT uses a
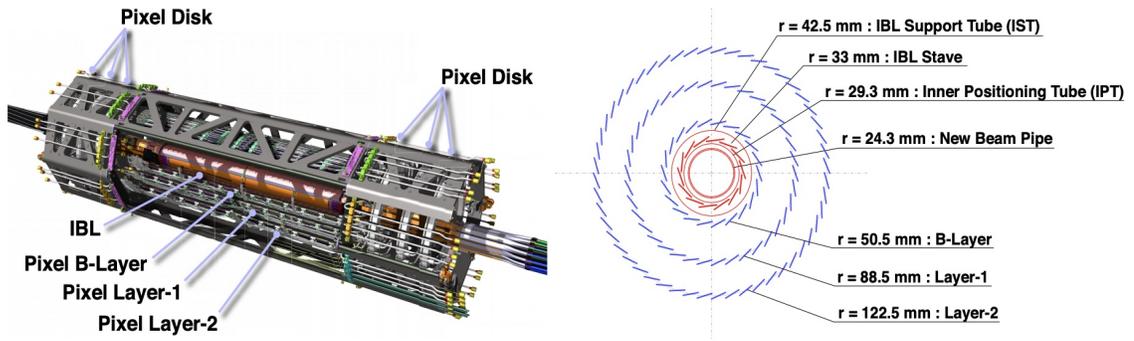
Figure 3.6: Schematic of the Pixel Detector including the IBL. From [140].

phenomenon known as transition radiation for particle identification. When a charged particle passes through the boundary between the outer straw and the gas mixture a photon will be emitted, i.e. transition radiation, of which its energy ionizes the gas in addition to the initial ionisation. Because different particles result in different transition radiation energies the signal strength in the wire can be used to distinguish between e.g. electrons and pions [143, 144].

3.2.2 Calorimeters

The goal of the calorimeter system of ATLAS is to measure the energy and direction of an incoming particle by absorbing particle showers. The calorimeter system consists out of electromagnetic and hadronic calorimeters, where the former focuses on electromagnetic showers, initiated by particles such as the electron or photon, and the latter on hadronic showers, initiated by particles such as charged pions or kaons. The calorimeter system of ATLAS can be divided into the Liquid Argon (LAr) calorimeter and the Tile calorimeter which use different triggers and measurement techniques. Both are sampling calorimeters i.e. calorimeters consisting out of alternating layers of a *passive* and *active* mediums. The passive medium is a high density material to trigger the particle showers and completely absorb the incoming particles while the active medium creates a detectable signal related to the energy lost by the particle.

A schematic of the calorimeter system is depicted in Fig. 3.7.

Liquid Argon Calorimeter

The Liquid Argon (LAr) calorimeter [145] is dedicated finding energy deposits of both electromagnetic and hadronic particles. The electromagnetic calorimeter (ECAL) consists of one barrel and an end-cap on either side covering a pseudo-rapidity range of $|\eta| < 3.2$. Both the barrel and the end-caps have an accordion-like structure with layers of electrodes to read-out the electric signal, layers of lead as passive material and with the gaps in between filled with liquid argon to act as active material. This specific geometry was chosen to avoid any cracks in the plates having any influence on the measurements and to decrease the signal rise time i.e. the time for a signal to rise past measurement threshold [145]. The ECAL is further segmented into cells with varying size $\Delta\eta \times \Delta\phi$ depending on its location in the calorimeter. See Fig. 3.9 for a sketch of the accordion structure.

The LAr calorimeter has also two hadronic end-caps (HEC) placed behind the ECAL end-caps that cover a pseudo-rapidity range of $1.5 < |\eta| < 3.2$. These end-caps are constructed from parallel layers of copper and liquid argon instead of the accordion structure of the previous calorimeters. Finally, each end is fitted with a forward calorimeter (FCAL) [146] that covers a pseudo-rapidity range of $3.1 < |\eta| < 4.9$. Each end consists of three layers; one electromagnetic calorimeter of copper and liquid argon and two hadronic calorimeters of tungsten and liquid argon. These have a special design of anode rods in cathode tubes with the space in between filled with liquid argon. These are placed in parallel to the beam line and held up by a matrix of the passive material i.e. copper or tungsten. Figure 3.11 shows a sketch of the cross-section of the FCAL.
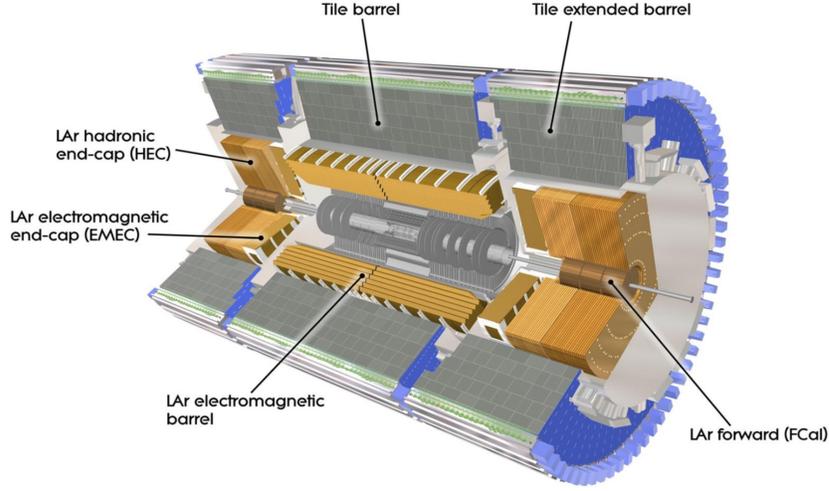
Figure 3.7: Schematic of the ATLAS calorimeter system. From [147].
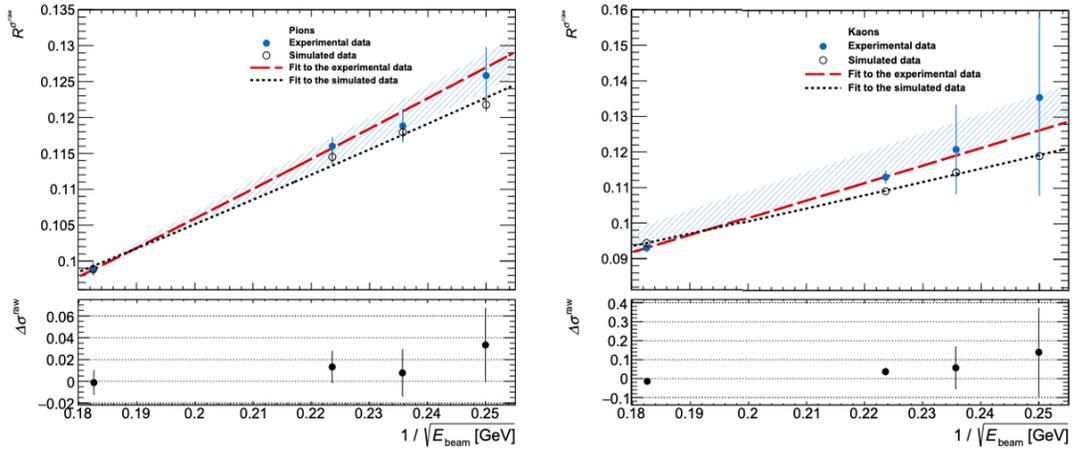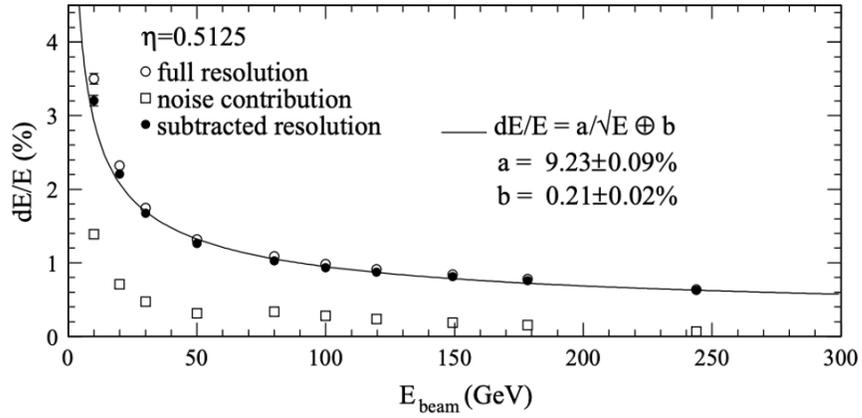
Figure 3.8: Energy resolutions relative to the beam energy E_beam for electrons in the LAr barrel ECAL (Top) and for pions (Bottom left) and kaons (Bottom right) in the Tile calorimeter. For the bottom plots the variables $R^{\sigma^\mathrm{raw}} = \frac{\sigma}{E_\mathrm{beam}}$ and $\Delta\sigma^\mathrm{raw} = \frac{\sigma}{\sigma_\mathrm{MC}} - 1$ are defined. From [148, 149].
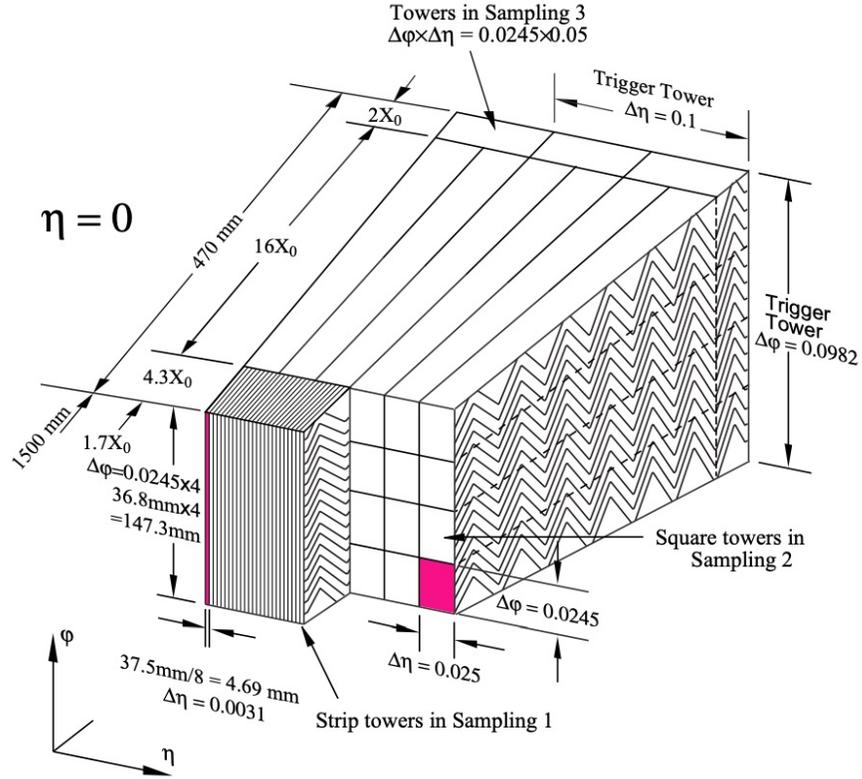
Figure 3.9: Sketch of the accordion structure of the EM calorimeter. The depths are expressed in both millimeters (mm) and material radiation length (X_0). From [145].

Tile Calorimeter

The Tile Calorimeter [150] is one barrel that focusses on measuring the energy deposits of hadronic particles. It can be subdivided into the central long barrel and the extended long barrel that together cover a pseudo-rapidity range of $|\eta| < 1.7$. The Tile Calorimeter is also a sampling calorimeter but uses scintillating plastic as active and low-carbon steel as passive material. The barrels are constructed from 64 wedges which each consist out of layers of plastic and steel tiles oriented along the $r\phi$-plane. Early Monte Carlo studies showed that this orientation, compared to the conventional $z\phi$-plane parallel to the beam line, does not affect the energy resolution [150]. This design was made with practical and cost-saving considerations for the construction and mounting of the Tile Calorimeter [151].
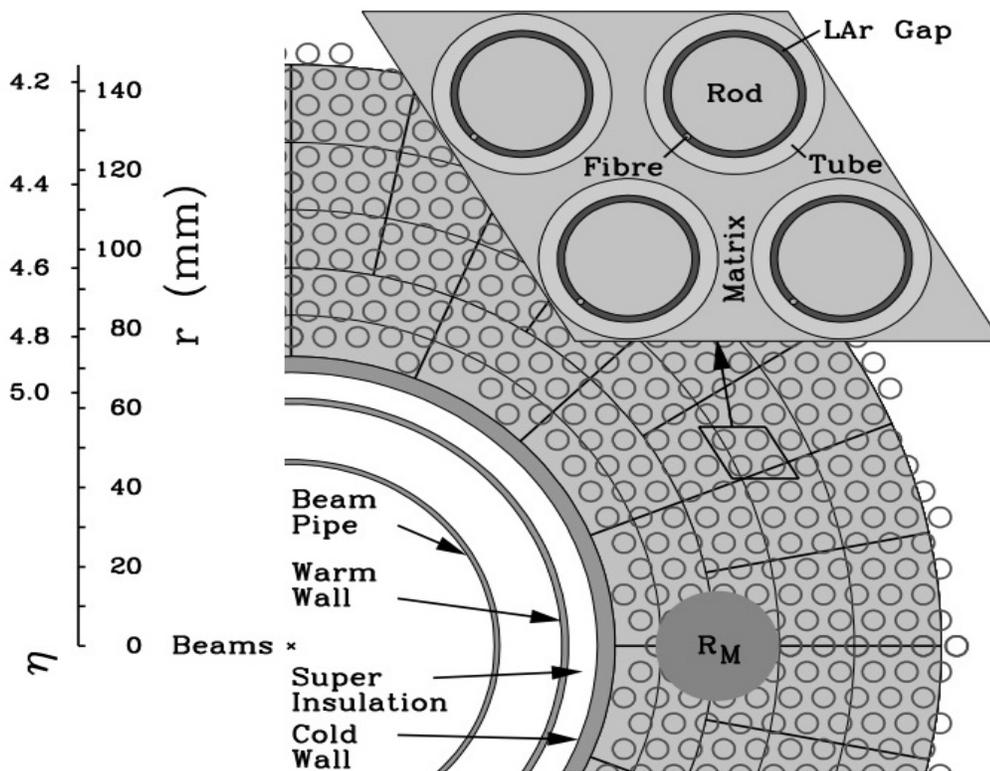
Figure 3.10: Sketch of the cross section of the FCAL. From [146].

3.2.3   Muon Spectrometer

The Muon Spectrometer (MS) [152] is responsible for the identification of muons and the measurement of their charge sign and momenta. In general, muons pass through the Inner Detector and calorimeters undetected. However, muons are a crucial part of many physics analyses which motivated the construction of the Muon Spectrometer. It is the outermost and largest part of the ATLAS detector and consists of three cylindrical layers, four vertical placed wheels on each side that function as end-cap and covers a pseudo-rapidity range of $|\eta| < 2.7$. A schematic of the Muon Spectrometer is given in Fig. 3.12. Figure 3.13 shows the relative momentum resolution and the muon reconstruction efficiency of the Muon Spectrometer. The Muon Spectrometer can be divided into four subsystems.
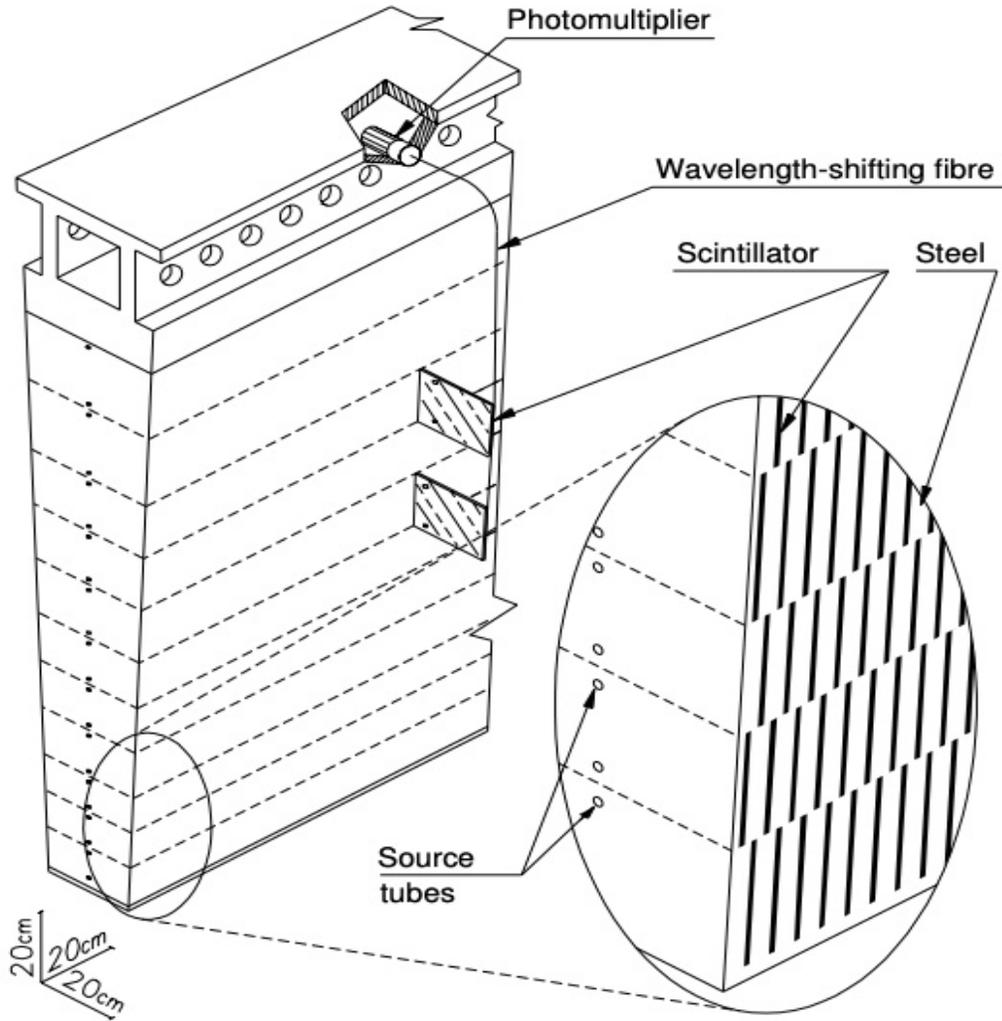
Figure 3.11: Sketch of a barrel wedge of the Tile Calorimeter. From [153].

Monitored Drift Tube chambers

The Monitered Drift Tube (MDT) chambers are used for the precision measurement of muon tracks in all the layers except for the most inner end-cap layer. The detection technology uses an aluminium tube with a tungsten-rhenium (W-Re) wire in the center and is filled with a argon-methane (Ar-CH$_4$) gas mixture. At the ends of the tubes there are oppositely charged electrodes that measure the drifting charges created by an ionizing muon. The spatial resolution of the tracking is $80\,\mu$m in the $r\phi$-plane and $35\,\mu$m in the z-direction.
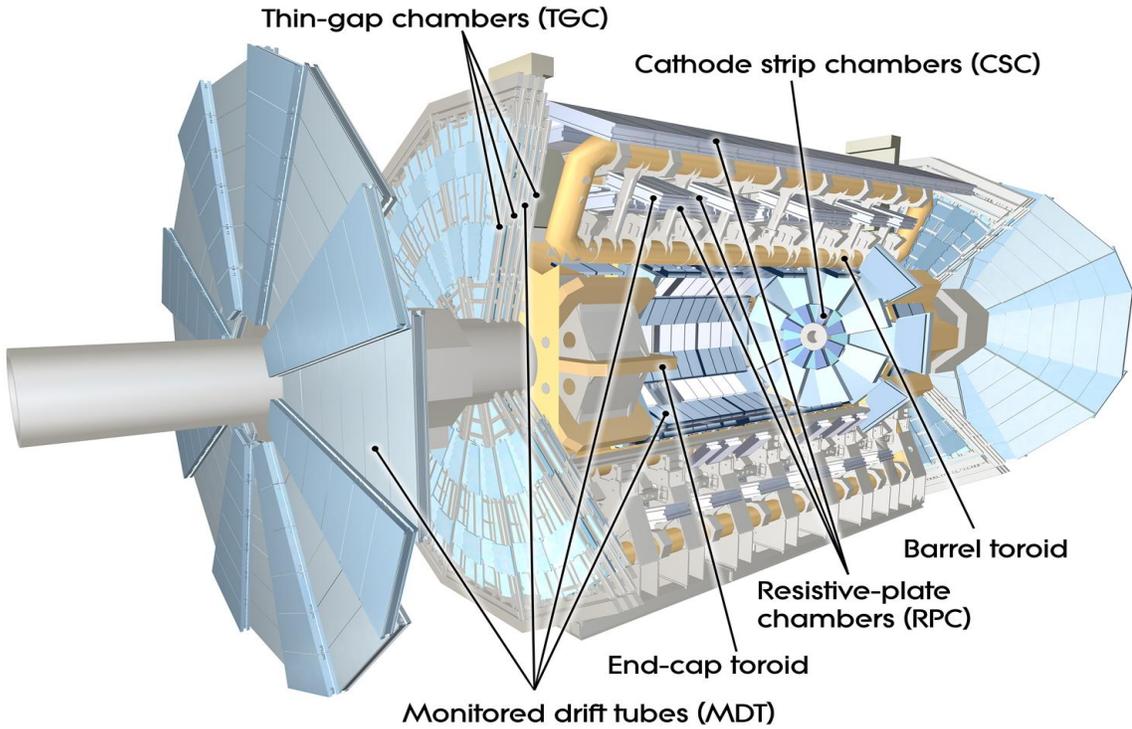
Figure 3.12: Schematic of the Muon spectrometer. From [132].
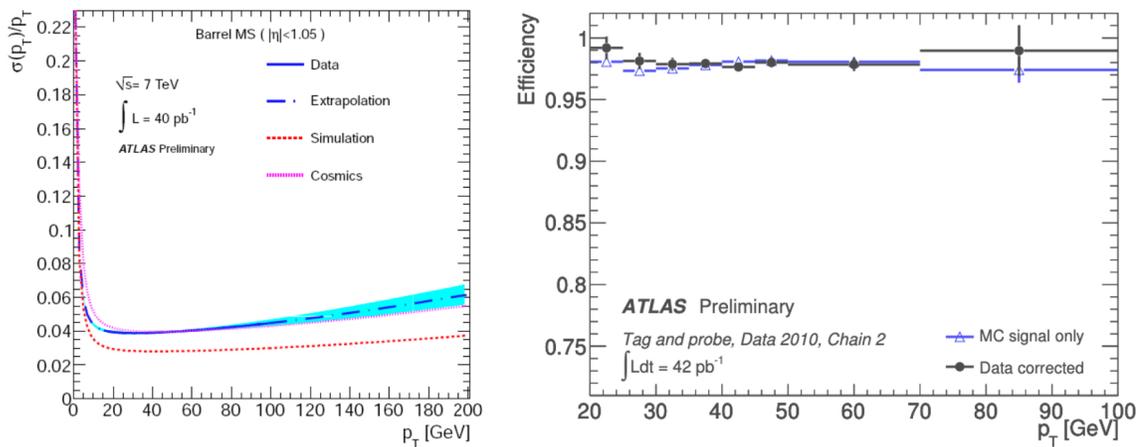
Figure 3.13: The relative muon momentum resolution in the MS barrel region (Left) and muon reconstruction efficiency (Right) both as a function of the muon momentum. From [154, 155].

Cathode Strip chambers

The Cathode Strip (CS) chambers are also used for muon tracking but are only present in the end-cap layer closest to interaction point for pseudo-rapidities of $2 < |\eta| < 2.7$. The CS chambers are multi-wire proportional chambers with a cathode strip and anode wire read-out of charges created by ionizing muons. The wire and gas material is the same as in the MDT chambers. The spatial resolution of all the CS chambers combined is $40\,\mu$m in η-direction and $4\,$mm in the ϕ-direction.

Resistive Plate chambers

The Resistive Plate (RP) chambers are wireless gas-chambers filled with a mixture of tetrafluoroethane ($C_2H_2F_4$) and isobutane (C_4H_{10}) with metal plates on either side for read-out. The RP chambers are only placed in the barrel with two on either side of the middle MDT chamber layer and one on the outside of the outer MDT chamber layer.

Thin Gap chambers

The Thin Gap (TG) chambers are only used in the end-cap and are all located near the first and second end-cap layer. The TG chambers have a similar structure as the Cathode Strip chambers but have a anode-wire pitch and anode-cathode distance that are not the same and use a gas mixture of CO_2 and n-pentane (n-C_5H_{12}).

3.2.4 Magnet System

A set of electromagnets are installed around the majority of the ATLAS detector to exert a Lorentz force on the traversing charged particles. A precise known magnetic field together with the curvature of the charged particle determined by the Inner Detector and Muon Spectrometer allows for a measurement of the particle's momentum

[136]. The magnet system consists out of the Barrel Toroid, Central Solenoid Magnet and the End-cap Toroids. A schematic of the magnet system is given in Fig. 3.14.
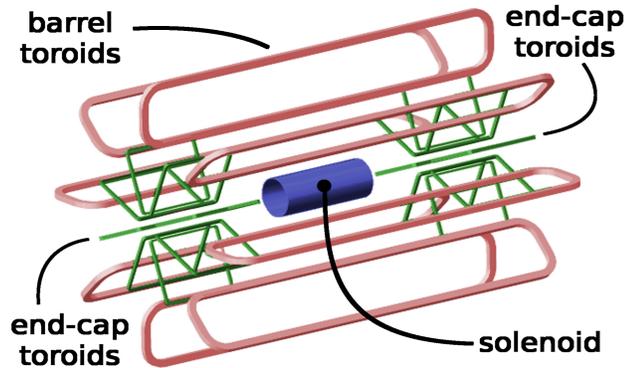
Figure 3.14: Schematic of the magnet system of the ATLAS detector. From [156].

Barrel and End-cap Toroids

The Barrel [157] and End-cap Toroids [158] consist each out of 8 superconducting coil magnets that provide a magnetic field for the Muon Spectrometer. The superconducting wires are made from a NbTi/Cu alloid that is cooled to a temperature of 4.5 K by a cryostat system with liquid helium. The magnetic system can deliver a magnetic field of up to 4 T.

Central Solenoid Magnet

The Central Solenoid Magnet [159] is a 2.4 m diameter coil that provides a magnetic field for the Inner Detector. The superconducting wires are made from NbTi/Cu with a aluminium stabilizer to minimize the thickness of the solenoid for the calorimeter system but keep electrical stability. Because of its vicinity to the ECAL it is integrated with the LAr cryostat system to keep the solenoid at a temperature of 4.5 K. The solenoid can deliver a magnetic field of up to 2 T.

3.2.5 Trigger and Data Acquisition System

The ATLAS detector has been designed to observe 1.7×10^9 *pp* collisions per second which corresponds to 60 TB of data per second [160]. This does not only result in unmanageable data flows but data on many events that will not be of interest for our physics programs. The goal of the trigger system is to select the events of interest, acquire the data of these events and discard the rest of the signals created in the sub detectors. The ATLAS detector has a trigger system that contains two levels.

Level 1 Trigger

The Level 1 triggers [161, 162, 163] are hardware-based triggers of custom made electronics which reduce the event rate to at least 100 kHz. Signals from Calorimeters and Muon Spectrometer are passed to front-end electronics (FEs) which digitise the signals. These signals are then passed with 1900 optical links to custom buffer hardware that apply selection filters. If an event passes the filters the digitized signals of the subdetectors, e.g. from the Pixel Detector, Calorimeters or Semiconductor Tracker, are passed from the FEs to the Read-Out Drivers (RODs). A subset of these events is passed to the HLT for further reduction. Additionally, the information from the custom buffer hardware is used to identify Regions of Interest (RoIs) in the Calorimeter and Muon Spectrometer which is also passed to the HLT. This level acts on an event in a decision time of $2.5\,\mu$s. A overview of the data flow, storage and processing nodes in the L1 and HL Trigger system is given in Fig. 3.15.

High-Level Trigger

The High-Level triggers [164] is a software-based trigger which takes the fragments of the digitized signals from all the subdetector systems and applies more sophisticated selection algorithms to reduce the event rate from 100 kHz to approximately 1 kHz. These algorithms are run on a processor farms and take at most 1 s per event.
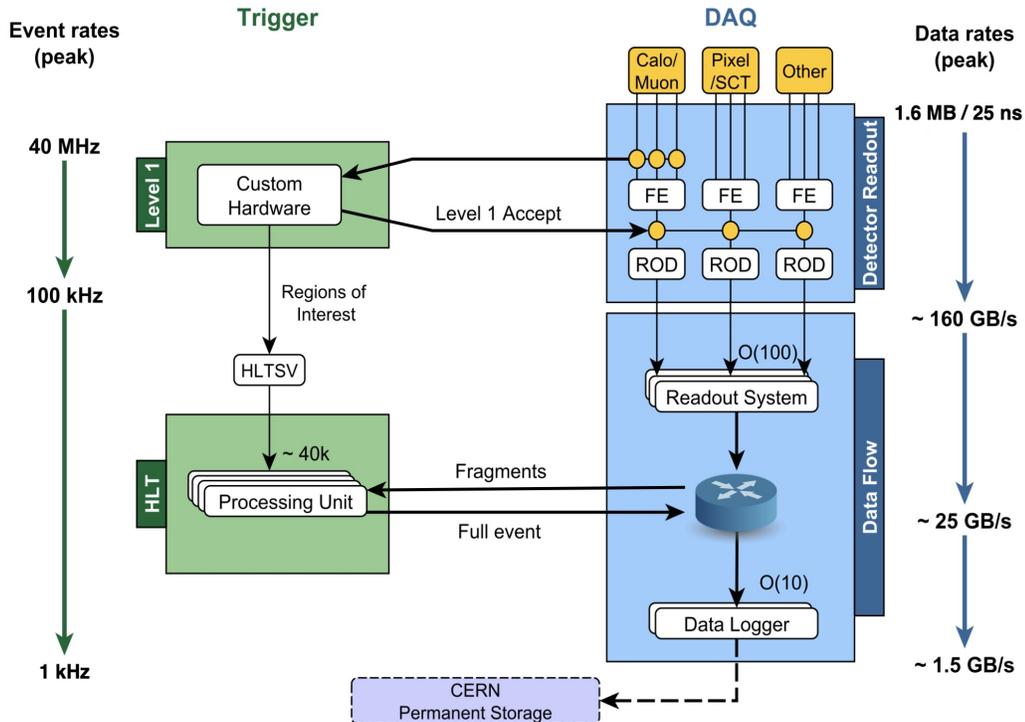
Figure 3.15: Schematic of the trigger and data acquisition system. From [163].

The criteria that are imposed on the events in these trigger levels can vary from kinematic thresholds to geometry or track selection. After the HLT the final selection of events is sent to permanent digital storage by the Data Logger. The data is then distributed across the Worldwide LHC Computing Grid (WLCG) [165], a global network of storage and processing nodes facilitated and used by all the research institutes and universities collaborating with CERN. The ATLAS data can then be pulled from this network for further reconstruction and analysis.

3.2.6 Detector Simulation

Chapter 2 describes the theoretical principles behind the Monte Carlo simulations of pp collisions up until hadronization where stable hadrons, leptons and photons remain. These proceed further away from the interaction point and interact with the

detector materials to induce measurement processes. The standard ATLAS detector simulation is build with the GEANT4 particle simulation toolkit [166] of which the output file format can be chosen to be identical to the raw data format of the ATLAS data acquisition system or in a different object-based format. The GEANT4 (G4) framework transports the generated particles through all the subdetector systems and models their interaction with the material including the read-out signals. This framework is the most complete and detailed model of the ATLAS detector and is therefore often denoted as *full simulation*. However, after identifying certain simulation processes that take the majority of computing time, e.g. 80% of the G4 simulation is spend on simulating particle and calorimeter interaction, efforts have been made to speed up certain parts. One of these efforts produced the ATLFAST-II [167], a simulation of the ATLAS detector that uses the full GEANT4 simulation for the inner detector and muon system but uses FastCaloSim [168] for the calorimeter systems. The latter is a model that uses geometrical simplifications and a parameterization of the particle energy response to speed up the calorimeter simulations. Table 3.1 shows the substantial improvement in computation times between the two. In this thesis ATLFAST-II was only used for preliminary studies. All Monte Carlo samples used in the final measurement were passed through the GEANT 4 simulation framework.

Sample	Geant4	AtlFast-II
Minimum Bias	551	31.2
$t\bar{t}$	1990	101
Jets	2640	93.6
Photon and jets	2850	71.4
$W^{\pm} \rightarrow e^{\pm}\nu_e$	1150	57
$W^{\pm} \rightarrow \mu^{\pm}\nu_{\mu}$	1030	55.1
Heavy ion	56,000	3050

Table 3.1: Simulation times per event for the GEANT 4 and the ATLFAST-II simulation of the ATLAS detector [169]. The times were normalized to kSI2K seconds i.e. the processing time the SpecInt2000 CPU benchmark system would need for the simulation [170].

Chapter 4

Object Reconstruction and Identification

In the previous chapter it is discussed how particles created during pp collisions create signals in the subdetectors which are then processed and saved on disks. However, this raw ATLAS data still needs to be processed and further compressed offline into useful information that can be used in physics analyses. The process of converting these raw electrical signals into physics objects such as electrons, jets and photons is called reconstruction and are saved in an Event Data Model (EDM) called an Analysis Object Data (xAOD) [171]. This is a data file format based on the ROOT library [172], a C++ open-source data analysis framework, and is used by physicists as input for their analyses. The ATLAS raw data can be used to reconstruct almost any particle. The physics objects of $t\bar{t}$ production in the l+jets decay channel that need to be reconstructed from detector signals are hadronic jets, electrons, muons and neutrinos. The top quark can also decay into τ-leptons but these are excluded from the analysis because they can decay hadronically, unlike electrons and muons. The reconstruction of photons and τ-leptons is not discussed in this chapter as they are not part of the l+jets decay channel.

4.1 Tracks and Vertices

One of the most important characteristics of a collision event are the tracks and interaction points, i.e. vertices, of the charged particles emerging from the *pp* collisions. These are crucial for the identification of the particles associated to the tracks and the interactions they undergo. The reconstruction of tracks is done by a sequence of algorithms [173, 174]. It starts of with an *inside-out* followed by an *outside-in* approach. The inside-out approach starts with constructing track seeds from the hits in the pixel detector and semiconductor tracker. A track seed is a set of points in separate layers which is set by default to three. Each track seed can be uniquely defined by three parameters if one assumes a perfect helical track and a constant magnetic field; the momentum p_T, the transverse impact parameter d_0 and the azimuthal angle ϕ_0.
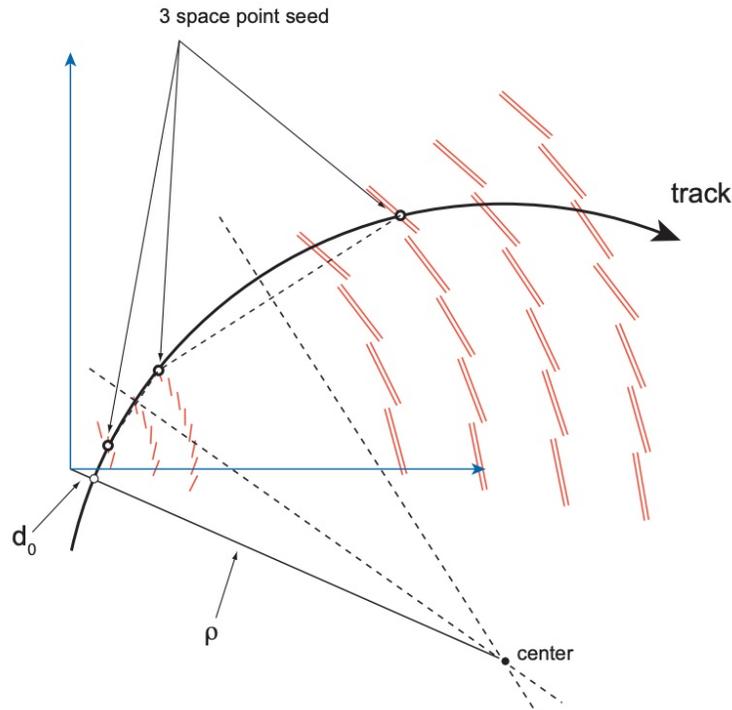
Figure 4.1: Sketch of a perfect helical track used to estimate the momentum p_T and transverse impact parameter d_0 of a track seed. From [173].

The transverse momentum is calculated using

$$p_T = \rho \times 3 \cdot 10^{-4} \times q \times B \tag{4.1}$$

with ρ being the radius of the track, q the charge of the particle and B the magnetic field strength along the z-axis. The impact parameter is determined by

$$d_0 = \sqrt{c_X^2 + c_Y^2} - \rho \tag{4.2}$$

with c_X and c_Y being the distances from a reference point along the x- and y-axis, respectively.

Track seeds passing minimum p_T and maximum d_0 criteria are passed on to a combinatorial Kalman Filter [175], an iterative procedure that forms track candidates by adding hits moving towards the transition radiation tracker. These tracks are subject to criteria such as each seed being part of only one track candidate or sufficient hits per track candidate. The track candidates are then passed to an ambiguity solver [173], an algorithm that firstly applies loose requirements on track candidate characteristics such as momentum or holes, missing hits in layers, and then applies a track scoring algorithm. A cut on the final score defines a set of tracks that is further extended with segments from the transition radiation tracker. The second outside-in approach starts off with seeds from energy deposit regions in the electromagnetic calorimeter. These seeds are extended with segments from the TRT and unused hits from the pixel detector and semiconductor tracker. Together with the inside-out approach these tracks result in the final set.

Using the tracks found in this procedure, *primary vertices* are reconstructed with an iterative χ^2-fit [173]. A primary vertices represent the interaction pionts (IP) of

pp collisions. The z-position of the reconstructed tracks on the beamline are used as seeds and the reconstructed vertices are required to be matched to at least two tracks. In case multiple vertices of an event are associated with a pp collision the vertex with the highest p_T is chosen as *the* primary vertex. All other vertices representing proton-proton interactions, either from the same or a different bunch of protons, are considered as *pile-up*. *Secondary vertices* represent decay points of particles originating from the primary vertex. The reconstruction of secondary vertices is of special importance for identifying b-quarks that come from a decaying top quark. More on this will be discussed in Sec. 4.4.

4.2   Electrons

An electron signature is characterized by clusters of energy deposits in the electromagnetic calorimeter and a curved particle track. Cluster candidates are formed from energy deposits by using a sliding-window algorithm [176]. As mentioned in Ch. 3 the electromagnetic calorimeter consists out of cells. A frame of 3×5 cells in the $\eta \times \phi$ plane scans both the η and ϕ direction for energy deposits exceeding $2.5\,\mathrm{GeV}$. In case of cluster candidates overlapping, i.e. if the frames are within a distance of $\Delta\eta \times \Delta\phi = 5 \times 9$ cells, the cluster with the lowest transverse energy is discarded if the energy difference is bigger than 10%. If the difference is less than 10% then the cluster with the lowest transverse energy in the central tower is discarded. The remaining seed clusters are then further reconstructed by extending the window to $\eta \times \phi = 3 \times 7$ in the barrel region and $\eta \times \phi = 5 \times 5$ in the end caps. The extended window clusters are then matched to a reconstructed track which is refit to take brehmsstrahlung of the electron into account. Identification algorithms are then applied to the reconstructed electrons to distinguish between electrons originating from the signal physics process, or *prompt electrons*, or from backgrounds such as photon

conversion or hadron misidentification. A likelihood is constructed from probability distribution functions of various variables from tracking and calorimeter information [177]. This likelihood is fitted to data and Monte Carlo samples and used to construct a discriminant optimal for identification of prompt electrons. Discriminant cut-offs can be chosen based on their signal to background ratio. The working points *Tight*, *Medium* and *Loose* correspond to a ratio of 80%, 88% and 93%, respectively.
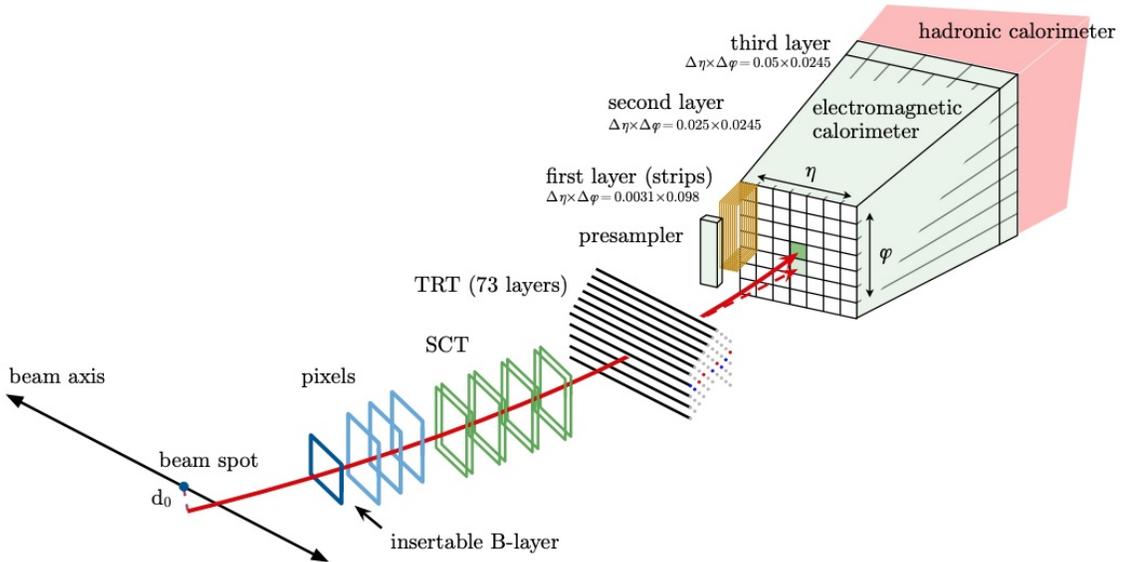
Figure 4.2: A schematic illustration of the path of an electron through the detector. From [177].

4.3 Muons

As mentioned in Ch. 3, muons will leave curved tracks in the inner detector but are very unlikely to interact with the electromagnetic and hadronic calorimeter other than creating ionization. Muons are therefore reconstructed and identified with additional information from the muon spectrometer. An additional track reconstruction in the muon spectrometer starts by applying a Hough transform [178] on hits in the MS. A Hough transform takes points in Cartesian space and transforms them into a

line in line parameter space. Features of these lines are used to reconstruct lines from the points in Cartesian space [179]. Track candidates are formed from the segments by fitting a parabolic trajectory originating from the interaction point of that event. A final χ^2 fit finds a track that takes misalignments in the muon spectrometer and interactions with the detector material into account. After removing holes and ambiguities, tracks from the inner detector and muon spectrometer are matched and used in a combined track fit. If no matches can be found, the muon spectrometer segments are extrapolated to the beamline. Another χ^2 fit is perfomed using information of the reconstructed muons to identify prompt muons from muons originating from pion or kaon decays. Variables used in the fit are the χ^2 value of the track fit, the p_T difference as measured in the inner detector and muon spectrometer and charge over momentum ratio q/p [180, 181]. *Loose*, *Medium* and *Tight* working points are available to give analysis teams various prompt muon reconstruction efficiencies to choose from.

4.4   Jets

The majority of particles produced at pp collisions at the LHC will be gluons and quarks. When emerging these particles will hadronize because of colour confinement and produce a spray of particles known as a jet. Topologically clustered calorimeter cells, or topo-clusters [182], are used as inputs for jet reconstruction. Topo-clusters are formed with the use of a growing-volume algorithm [182] that starts with collecting seed cells that have an energy deposit 4 times bigger than the average cell noise. Adjacent cells or other topo-clusters are merged if surpassing an energy 2 times bigger than the average cell noise. Finally, all cells with an energy deposit bigger than zero and on the perimeter of the topo-clusters are added as well. The topo-clusters are used in the anti-k_t algorithm [183] with distance parameter $R = 0.4$ to reconstruct

small R jets and $R = 1.0$ for *large R jets*. Additionally, the Cambridge-Aachen (C/A) algorithm [184] is used to reconstruct jets with $R = 1.2$. Figure 4.3 compares the results of both algorithms for a Monte Carlo event. The final reconstructed jets are then matched to tracks from the inner detector and muon spectrometer.
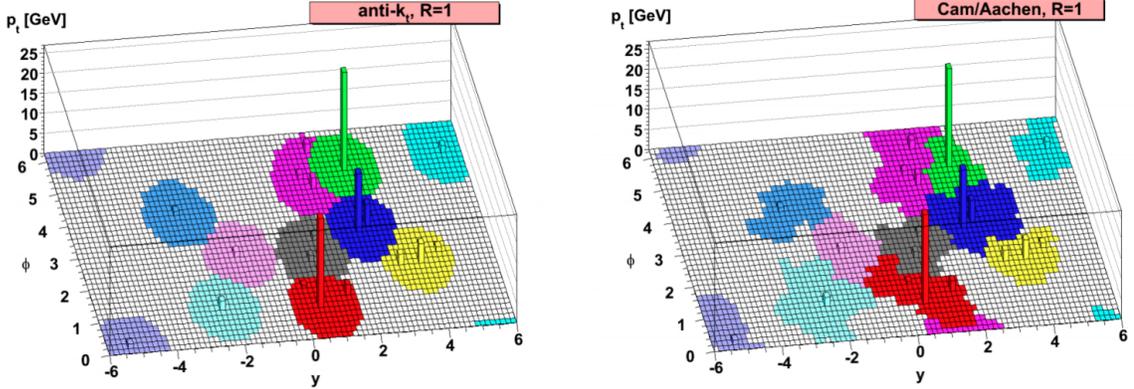
Figure 4.3: An illustration of calorimeter cells clustered with the anti-k_t (Left) and Cambridge-Aachen (Right) algorithm for $R = 1.0$. The inputs come from a Monte Carlo event generated with Herwig. From [183].

To reduce the effects of pile-up on the jet reconstruction the jet-area [183] and jet-vertex tagging [185] methods are used. The jet-area method draws cylinder shapes around the particles that make up a jet and uses the bottom-view area to identify jets originating from the interaction of interest or from a pile-up or Underlying Event. The jet-vertex tagging (JVT) method uses track information from the tracks of the particles used to construct the jet to distinguish between jets from the primary vertex and pile-up vertices.

Identifying what type of particle initiated the jet is known as tagging. The top quark decays before hadronization and almost always into a W boson and a b quark. Identifying jets initiated by a b-quark is known as *b-tagging* and is vital for identifying the $t\bar{t}$ signal events. Jets initiated by u, d or s-quarks are labelled as *light jets*. The DL1 algorithm [186] uses a deep neural network (DNN) to label a jet as coming from

either a *b*, *c* or light-flavor quark. The DNN is trained with jets from simulated $t\bar{t}$ and
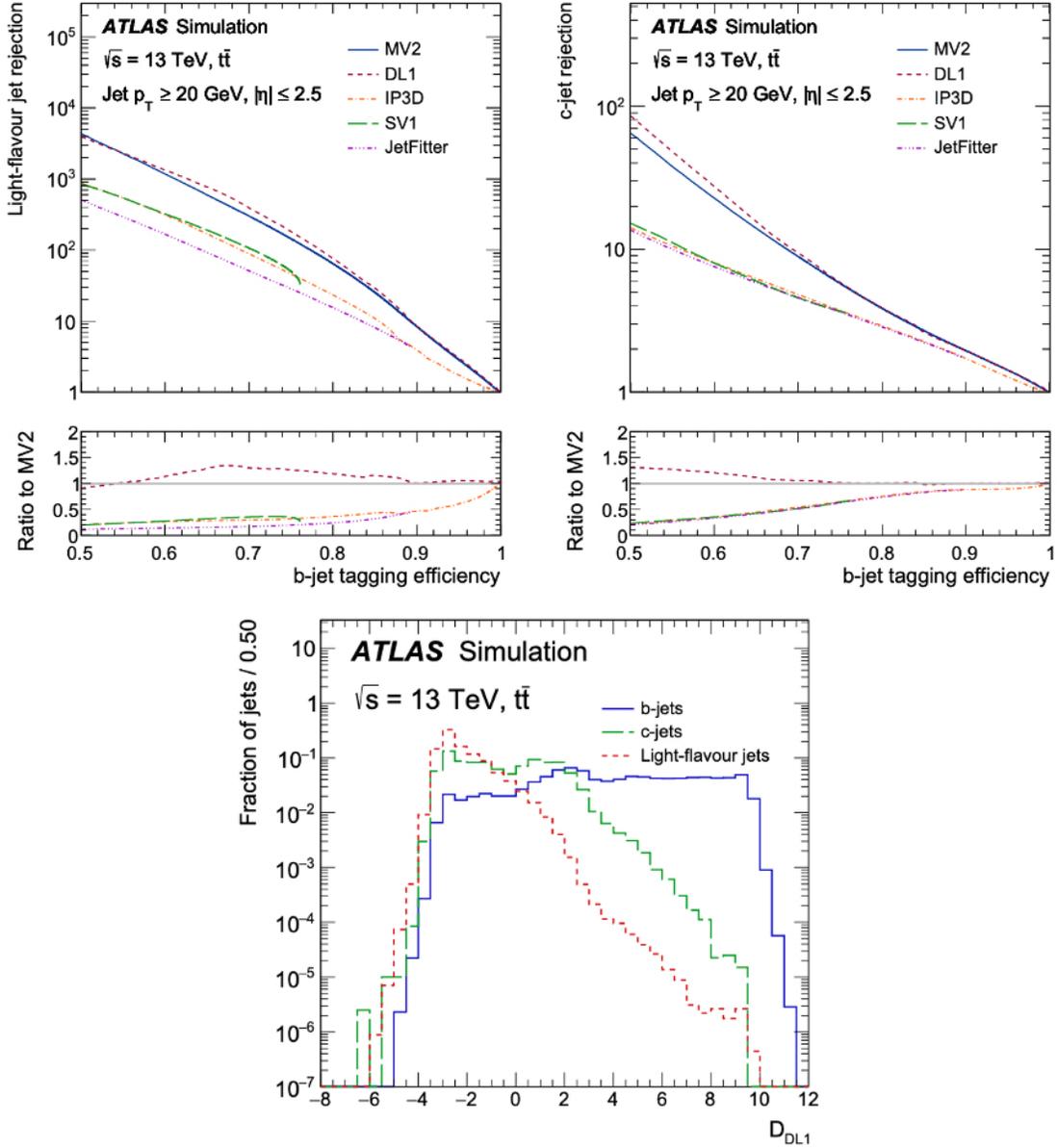
Figure 4.4: Light-flavour and *c*-jets rejection as a function of *b*-tagging efficiency for the DL1 algorithm compared with previous *b*-tagging algorithms (Top). Distribution of the output discriminant of the DL1 b-tagging algorithms for *b*-jets, *c*-jets and light-flavour jets on $t\bar{t}$ simulated events (Bottom). From [187].

Z' events to ensure good identification of jets over a wide range of jet p_T. The input variables are kinematic properties of the jets such as p_T, η, vertex displacements or invariant masses and gives as output three probabilities for a b, c and light-flavor jet. The probabilities are used to define the single discriminant

$$D_{DL1} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\text{light}}}\right) \tag{4.3}$$

with p_b, p_c and p_{light} the output probabilities for the jet being b, c or light flavored and f_c the c-jet fraction in the training sample. Figure 4.4 shows rejection rates of light flavoured and c-jets and the DL1 discriminant outputs of the trained DNN for a $t\bar{t}$ sample. This discriminant is then used to define working points on either a 60%, 70%, 77% or 85% b-tagging efficiency.

4.5 Missing Transverse Energy

The missing transverse energy (MET) is an event quantity based on momentum conservation and is used to identify particles that escape the detector without any interaction in any of the subdetectors. In this thesis it is used to reconstruct neutrinos (see Sec. 4.6). The MET is calculated from hard objects, such as reconstructed jets or muons, and soft signals consisting of charged particle tracks not associated to any reconstructed objects. It is defined as a 2D vector in the xy-plane, i.e.

$$E_{x(y)}^{\text{miss}} = -\sum_{i \in \text{hard objects}} p_{x(y),i} - \sum_{j \in \text{soft signals}} p_{x(y),j}. \tag{4.4}$$

If all particles are accounted for in the both the hard objects and soft signals the sum should amount to zero. Reconstructing MET is very challenging because of the dependence on the performance of all the other measuring and reconstruction procedures. The MET is therefore reconstructed with reconstructed objects subject

to specific quality criteria and kinematic selections [42]. Figure 4.5 shows E_x^{miss} and E_y^{miss} distributions for data and Monte Carlo simulated $Z \to \mu\mu$ events.
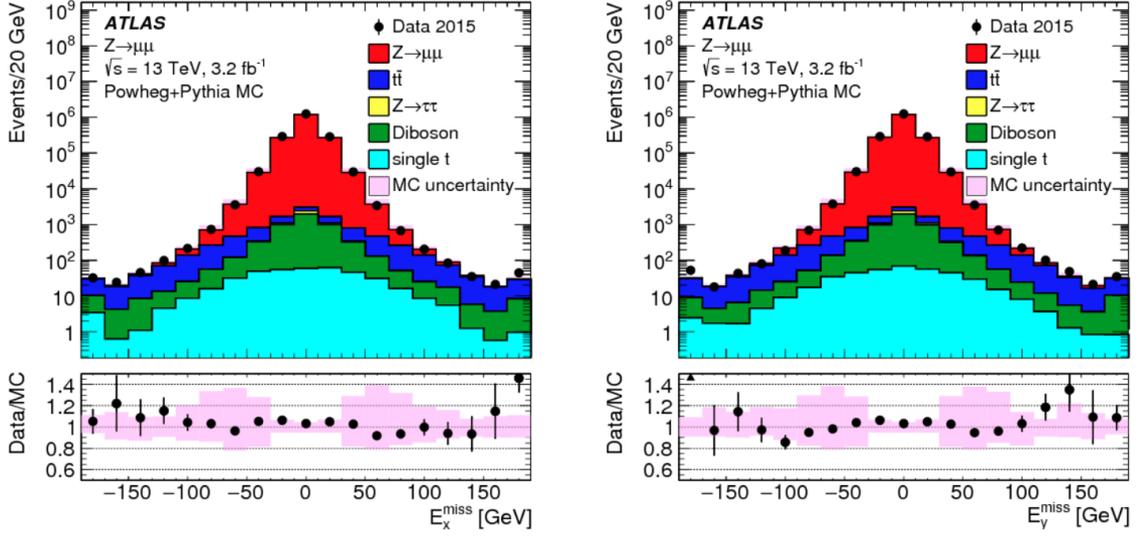
Figure 4.5: Comparison of E_x^{miss} (Left) and E_y^{miss} (Right) distributions filled with both data and Monte Carlo $Z \to \mu\mu$ events. From [42].

4.6 Top Quark

The top quark has a very short lifetime and decays before hadronization occurs. This means that the top quark decays before it reaches the detector and therefore can not be directly be measured. Instead, the top quark is reconstructed from its decay products using the pseudo-top algorithm [188]. A top quark decays almost exclusively into a W boson and a b quark. The W boson can either decay into a charged lepton and neutrino or two light flavor quarks (see Fig. 4.6). The pseudo-top algorithm therefore defines a leptonic decaying and a hadronic decaying top. The four-momentum of the leptonic top is reconstructed from the four-momenta vectors of a b-tagged jet, a muon or electron and reconstructed neutrino. Unlike the muon and electron can the τ lepton also decay hadronically. Because of the difficulty to disentangle these hadrons from the ones of the W boson or other QCD backgrounds

this decay mode is not included in the algorithm.
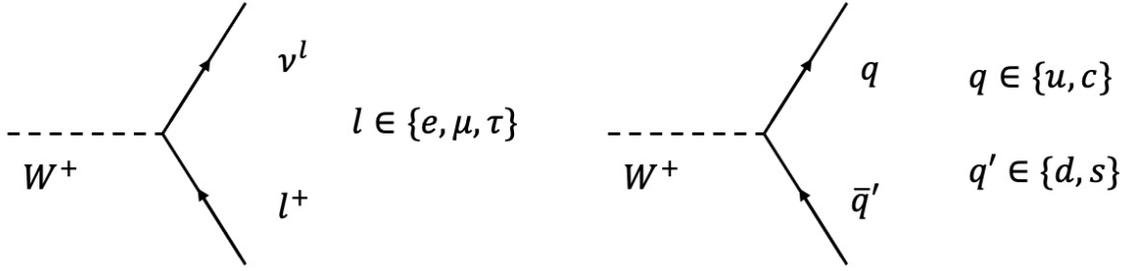
Figure 4.6: The leptonic and hadronic decay channels of a W boson.

The leptonic decaying top quark is reconstructed from a b-jet and W boson four-momentum. The two b-jets with the highest p_T are chosen as decay products of the two top quarks. The one with the smallest $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ w.r.t. the lepton is assigned to the leptonic top quark. The W boson is reconstructed from the lepton and neutrino vector. However, the neutrino is not directly measured so needs to be reconstructed from the MET and the invariant mass of the W boson. The transverse momentum components of the neutrino $p_{x,\nu}$ and $p_{y,\nu}$ are taken from the reconstructed MET. The z-component of neutrino uses the invariant mass definition

$$m_W^2 = (E_l + E_\nu)^2 - (p_{x,\nu} + p_{x,l})^2 - (p_{y,\nu} + p_{y,l})^2 - (p_{z,\nu} + p_{z,l})^2. \qquad (4.5)$$

This equation can be solved for $p_{z,\nu}$

$$p_{z,\nu} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad (4.6)$$

with

$$a = E_l^2 - p_{z,l}^2, \qquad b = -2kp_{z,l}, \qquad c = E_l^2 p_{T,\nu}^2 - k^2 \qquad (4.7)$$

and

Decay mode	Branching Ratio (Γ_i/Γ) [%]
$e\nu_e$	10.75 ± 0.13
$\mu\nu_\mu$	10.57 ± 0.15
$\tau\nu_\tau$	11.25 ± 0.20
Hadronic	67.60 ± 0.27

Table 4.1: Branching ratios for W boson decay modes [189]. The remaining decay modes with branching ratios $< 1\%$ are omitted.

$$k = \frac{m_W^2 - m_l^2}{2} + p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu}. \tag{4.8}$$

If both solutions to the quadratic equations are real the smallest p_z^ν solution is used. In case $\sqrt{b^2 - 4ac} < 0$ then the z-component of the neutrino's momentum is taken to be $p_z^\nu = -b/2a$. The other W boson is reconstructed from the two remaining jets with the highest p_T and which should not be b-tagged. Together with the remaining b-jet the four-momenta are used to reconstruct the hadronic top quark. Figure 4.7 shows a comparison between top quarks reconstructed with the pseudo-top algorithm and parton level top quarks. The big absolute difference is caused by the jet and single lepton selection. The ratio plot of the normalized distributions shows some disagreements. However, the kinematic properties of the pseudo-top quark still carry a lot of information on the top quark.
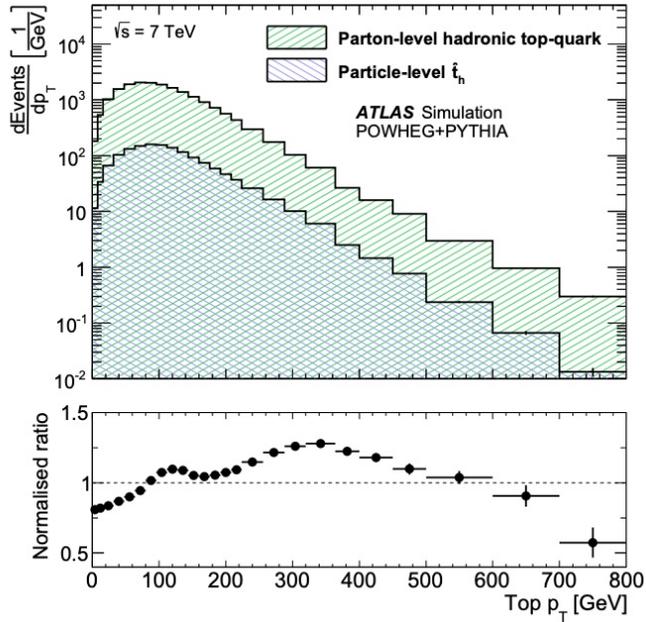
Figure 4.7: Transverse momentum distributions of simulated top quarks before parton shower, i.e. on parton level (green), and top quarks reconstructed with the pseudo-top algorithm, i.e. on particle level (blue). In both cases, the hadronically decaying top-quark is chosen. The distributions are evaluated for the same event sample based on POWHEG+PYTHIA at $\sqrt{s} = 7$ TeV. The ratio plots is of the normalized distributions. From [188].

Chapter 5

Unfolding Methods

In High Energy Physics (HEP) and in many other fields one often measures distributions of quantities such as particle energies or other characteristics of observed events. Because the experimental apparatus (the "detector") inevitably has a limited resolution, the measured (or "reconstructed") value of the quantity in question will differ in general from its true value. This results in a distortion or smearing of the measured distribution relative to what would be obtained if the detector had perfect resolution. The statistical procedure of estimating the true distribution from the directly measured one is usually called unfolding in HEP, or deconvolution in many other fields. In case of this thesis, unfolding distributions of top quark observables is a crucial step. Section 5.1 discusses the mathematics behind unfolding. Section 5.2 discusses the different unfolding algorithms used in this thesis as implemented in the software package `RooUnfold` [190]. Section 5.3 discusses figures of merit that can be used to optimize the regularization parameter. Section 5.4 presents a study comparing different unfolding algorithms using some of these figures of merit. Section 5.5 and 5.6 present novel techniques to estimate the response matrix and the covariance matrix, respectively. Section 5.7 finishes with some conclusions and discussion points.

5.1 Mathematical Formulation

Lets assume we have a variable of interest x following a true distribution $h(x)$. In practice we do not have direct access to the true distribution but only its measured value y following a different distribution $g(y)$. These two distributions are related through the Fredholm integral equation of the first kind [191],

$$g(y) = \int K(y, x) h(x) \, dx, \tag{5.1}$$

with $K(y, x)$ being the kernel function encoding the measurement process. In practice, this kernel function is a very complicated function and is not known a priori. However, we will see later on that it can be approximated with Monte Carlo simulations under the assumption of some model $h(x)$.

For experiments that count the number of observed events of interest, e.g. $pp \to t\bar{t}$ events, it makes more sense and to discuss the discretized case of histograms instead of continuous distributions. In that case Eq. (5.1) changes into

$$\nu_i = \sum_{j=1}^{M} R_{ij} \mu_j, \tag{5.2}$$

where $\boldsymbol{\nu} = (\nu_1, ..., \nu_N)$ is known as the reconstructed histogram and $\boldsymbol{\mu} = (\mu_1, ..., \mu_M)$ as the true histogram that we would like to estimate. They are related to Eq. (5.1) via

$$\nu_i = \nu_{\text{tot}} \int_{\text{bin i}} g(y) \, dy, \qquad \mu_j = \mu_{\text{tot}} \int_{\text{bin j}} h(x) \, dx \tag{5.3}$$

with ν_{tot} and μ_{tot} being the total number of true and reconstructed events, respectively.

The kernel function changes into the response matrix R_{ij} which is defined as [192]

$$R_{ij} = \frac{\int_{\text{bin j}} dx \int_{\text{bin i}} dy \, K(y,x)h(x)}{\int_{\text{bin j}} dx \, h(x)} \tag{5.4}$$

The response matrix can be interpreted as the probability an event will be found with measured value y in bin i given that the true value x was in bin j,

$$R_{ij} = P(\text{measured in bin } i | \text{true in bin } j). \tag{5.5}$$

Note that the response matrix depends on the true distribution $h(x)$. However, if the bins of $\boldsymbol{\mu}$ are small enough such that the kernel function $K(y,x)$ is constant over each bin in x the dependency will cancel out. One should also take into account that in case of counting experiments the events of the *signal process* are often accompanied by events of one or more irreducible *background processes*. The expected bin values for the total reconstructed background are denoted by $\boldsymbol{\beta} = (\beta_1, ..., \beta_N)$ which changes equation (5.2) into

$$\nu_i = \sum_{j=1}^{M} R_{ij}\mu_j + \beta_i. \tag{5.6}$$

Now let us assume we have some set of measured values binned in a histogram $\boldsymbol{n} = (n_1, ..., n_N)$ with the expected values $E[\boldsymbol{n}] = \boldsymbol{\nu}$. The goal of unfolding is to estimate the parameters $\boldsymbol{\mu}$ from the measured histogram $\boldsymbol{n} = (n_1, ..., n_N)$. For now, let us also assume we have determined the response matrix R_{ij} with infinite precision. We will discuss in upcoming sections how the response matrix can be estimated and how to include uncertainties of the response matrix. The most obvious approach would be invert the response matrix and multiply it with the background subtracted data:

$$\hat{\boldsymbol{\mu}} = R^{-1}(\boldsymbol{n} - \boldsymbol{\beta}). \tag{5.7}$$

However, this will often give unsatisfactory results. The sample of measured events used to fill the histogram $\boldsymbol{n}$ has a finite size and are subject to statistical fluctuations. The response matrix R is often ill-conditioned, i.e. will have a large condition number $||R||\,||R^{-1}||$, which means that even small statistical fluctuations in $\boldsymbol{n}$ can result in large fluctuations in the estimator $\hat{\boldsymbol{\mu}}$ [193]. These fluctuations and variances are more apparent with response matrices with large off-diagonal elements or with measured histograms $\boldsymbol{n}$ filled with a small number of statistics. A toy example of an ill-conditioned response matrix and a matrix inversion solution using this response matrix that shows large estimator fluctuations and variances is given in Fig. 5.2 and Fig. 5.1, respectively.
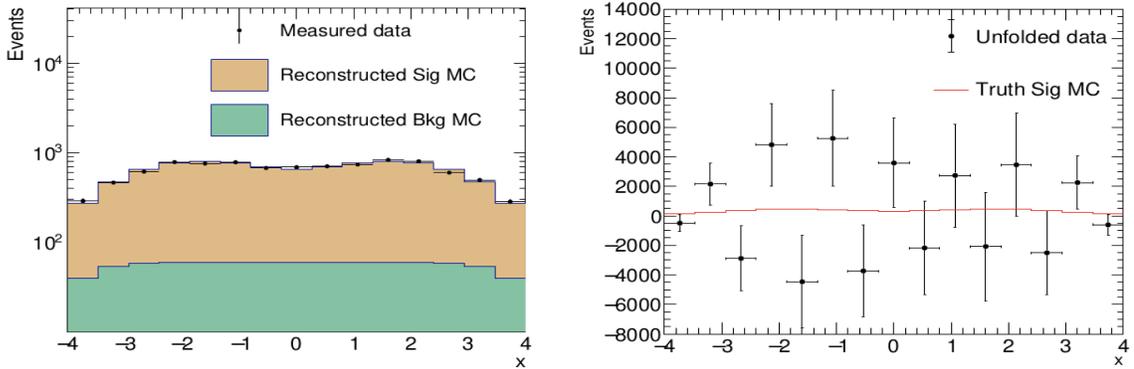
Figure 5.1: An example of how statistical fluctuations in the $\boldsymbol{n}$, an ill-conditioned response matrix R and an estimator $\hat{\boldsymbol{\mu}}$ constructed with matrix inversion can result in large fluctuations in the estimator. The figure shows the measured data compared with the simulated Monte Carlo distributions on reconstructed level (Left) and the unfolded data and truth distributions (Right).

Often, each measured bin n_i can be regarded as an independent Poisson variable i.e. distributed according to a Poisson probability distribution function (p.d.f.) $f(n_i|\nu_i)$ with mean ν_i. However, we know that $\boldsymbol{\nu}$ depends on $\boldsymbol{\mu}$ according to Eq. (5.6) so we can write $f(n_i|\boldsymbol{\mu})$. By taking the product of the p.d.f.s $f(n_i|\boldsymbol{\mu})$ we get the likelihood function
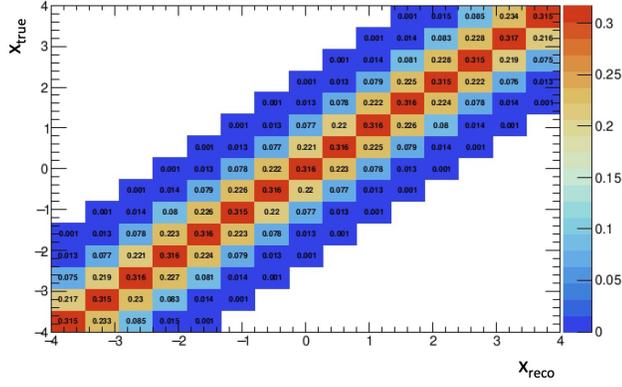
Figure 5.2: An example of an ill-conditioned response matrix as a result of detector smearing. This response matrix is estimated using simulated Monte Carlo events.

$$L(\boldsymbol{\mu}) = \prod_{i=1}^{N} f(n_i|\boldsymbol{\mu}) = \prod_{i=1}^{N} \frac{\nu_i^{n_i}}{n_i!} e^{-\nu_i}. \tag{5.8}$$

Alternatively, one can treat n_i as a continuous variable following a Gaussian p.d.f. with mean ν_i and standard deviation σ_{n_i}. In this case the likelihood would be

$$L(\boldsymbol{\mu}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_{n_i}} \exp\left[\frac{(n_i - \nu_i)^2}{\sigma_{n_i}^2}\right]. \tag{5.9}$$

Maximizing Eq. (5.8) or (5.9) w.r.t. $\boldsymbol{\mu}$ will give the Maximum-Likelihood Estimators (MLE) of the true histogram. If one sets the derivative of either likelihoods to zero and solves for $\boldsymbol{\mu}$ one gets the same solution as in Eq. (5.7). The matrix inversion is equivalent to the ML estimators and are both known to be unbiased estimators with minimum variance [192]. One should note that there are some situations where the matrix inversion solution is still suitable, e.g. if the unfolded histogram and covariance are used in a least-squares fit or with very little detector smearing. However, in general one looks for estimators that reduce the (co-)variances by constraining the possible solution space of $\hat{\boldsymbol{\mu}}$ with some prior knowledge on what the estimators should "look like". Constraining the solution space $\hat{\boldsymbol{\mu}}$ with prior knowledge is known as regularization and inherently introduces bias. The subtle game of unfolding is finding an ideal balance in this *bias-variance trade-off*. For purposes of the upcoming section

81

we define the bias of an estimator $\hat{\mu}_i$ as

$$b[\hat{\mu}_i] = E[\hat{\mu}_i] - \mu_i \qquad (5.10)$$

and the covariance between two estimators $\hat{\mu}_i$ and $\hat{\mu}_j$ as

$$U[\hat{\mu}_i, \hat{\mu}_j] = E[(\hat{\mu}_i - E[\hat{\mu}_i])((\hat{\mu}_j - E[\hat{\mu}_j])]. \qquad (5.11)$$

One way of explaining the large (anti-)correlations is by estimating the covariance matrix of the ML estimators. Because the ML estimators are unbiased estimators one can use the unbiased Cramer-Rao Bound (RCB) [194, 195], also known as the Minimum Variance Bound (MVB), to calculate the covariance matrix. This bound states that that the covariance between two estimators $\hat{\lambda}_i$ and $\hat{\lambda}_j$ has a lower bound set by the following inequality

$$U[\hat{\mu}_i, \hat{\mu}_j] \geq \left((\mathcal{I} - B) I^{-1} (\mathcal{I} - B)^T \right)_{ij} \qquad (5.12)$$

with $\mathcal{I}$ being the identity matrix and B being the bias gradient matrix with $B_{ij} = \frac{\partial b_{\hat{\mu}_i}}{\partial \mu_j}$. The Fisher information matrix I is defined by

$$I_{ij} = E\left[\frac{\partial^2 \log L}{\partial \mu_i \partial \mu_j} \right] \qquad (5.13)$$

with L being the likelihood function. In the case of negligible bias we see that Eq. (5.12) reduces to the unbiased RCB,

$$U[\hat{\mu}_i, \hat{\mu}_j] \geq \left(I^{-1} \right)_{ij}. \qquad (5.14)$$

Under the large sample approximation the covariance is assumed to equal the zero-bias RCB. If one can estimate the matrix of second derivatives of the log-likelihood, also known as the *Hessian matrix*, one can take the inverse of this matrix as an

estimate for the covariance. Under the large sample approximation one can also assume the log-likelihood is parabolic shaped around its maximum. In this case one can numerically approximate the second derivatives, e.g. with finite differences,

$$U[\hat{\lambda}_i, \hat{\lambda}_j] = \left(\frac{\partial^2 \log L}{\partial \lambda_i \partial \lambda_j} \bigg|_{\hat{\boldsymbol{\lambda}}} \right)^{-1}. \tag{5.15}$$

By taking the second derivative of Eq. (5.8) w.r.t. $\boldsymbol{\mu}$ and solving for zero one can get an expression for the covariance matrix [192],

$$U[\hat{\mu}_i, \hat{\mu}_j] = (H^{-1})_{ij} = \sum_{k=1}^{N} (R^{-1})_{ik} (R^{-1})_{jk} \nu_k. \tag{5.16}$$

We can see that the off-diagonal elements of the covariance matrix depend largely on the off-diagonal elements of the response matrix R.

5.2   RooUnfold Unfolding Methods

This section gives an overview of the unfolding methods, i.e. how they construct their estimators $\boldsymbol{\mu}$, as available in the software package `RooUnfold` [190] and how they impose regularization on the estimators. The matrix inversion and maximum likelihood (ML) estimators have been excluded as they are discussed in the previous section.

5.2.1   Bin-by-bin

The bin-by-bin unfolding method applies correction factors and ignores the off-diagonal elements of the response matrix. The estimator is found to be as

$$\hat{\mu}_i = C_i(n_i - \beta_i) \tag{5.17}$$

with the correction factor estimated with Monte Carlo simulated events,

$$C_i = \frac{\mu_i^{\mathrm{MC}}}{\nu_i^{\mathrm{MC}}}. \tag{5.18}$$

The covariance is defined as

$$U[\hat{\mu}_i, \hat{\mu}_j] = C_i^2 \nu_i \delta_{ij}. \tag{5.19}$$

5.2.2 Tikhonov

Instead of maximizing the likelihood $L(\boldsymbol{\mu})$ one can maximize a linear combination of the likelihood and some regularization function $S(\boldsymbol{\mu})$. This regularization function constrains $\boldsymbol{\mu}$-space according to some prior knowledge encoded in the function $S(\boldsymbol{\mu})$. One can find the estimators $\hat{\boldsymbol{\mu}}$ by maximizing the expression

$$\Phi(\boldsymbol{\mu}) = L(\boldsymbol{\mu}) + \tau S(\boldsymbol{\mu}) \tag{5.20}$$

with regularization parameter τ determining how strongly the analyst wants to constrain the solutions. For $\tau = 0$ we retrieve the ML estimators with minimum bias and maximum variance. For $\tau > 0$ the estimators will have a larger bias but smaller variance w.r.t. the ML estimators. Many choices for $S(\boldsymbol{\mu})$ are possible such as L_1 (Lasso) [196, 197] and L_2 (Ridge) [198, 199] regularization that constrain on absolute sizes of $\boldsymbol{\mu}$. It is possible to adjust these regularization functions by taking the differences between $\boldsymbol{\mu}$ and some reference histogram $\boldsymbol{\mu}_0$ instead of the absolute sizes of $\boldsymbol{\mu}$. However, a very common choice in HEP is Tikhonov regularization [200] which constraints the solution on its smoothness. The discretized Tikhonov regularization function is

$$S(\boldsymbol{\mu}) = -\sum_{i=1}^{M-2} (-\mu_i + 2\mu_{i+1} - \mu_{i+2})^2. \tag{5.21}$$

Note that this is for equal bin widths. An expression that works for both constant

and variable bin widths can be written down as

$$S(\boldsymbol{\mu}) = -\sum_{i=1}^{M-2}(-\mu_i^{\text{dens}} + 2\mu_{i+1}^{\text{dens}} - \mu_{i+2}^{\text{dens}})^2, \qquad \mu_i^{\text{dens}} = \frac{\mu_i}{\delta\mu_i} \qquad (5.22)$$

with $\delta\mu_i$ being the bin width of bin i. The TUnfold [201] and TSVD [202] unfolding algorithms available in `RooUnfold` both use Tikhonov regularization with Eq. (5.9) as likelihood. They construct closed form solutions for $\boldsymbol{\mu}$ by setting the derivative of the likelihood including the regularization term to zero. However, both have their own additional features to choose the regularization parameter τ. We will return to this in the upcoming section.

5.2.3 Richardson-Lucy (Iterative Bayes)

The Richardson-Lucy algorithm [203, 204], or iterative bayes unfolding as described by D'Agostini in [205], is an unfolding procedure that updates the estimator $\hat{\boldsymbol{\mu}}$ with each iteration. The estimator is defined as

$$\hat{\mu}_j^{r+1} = \hat{\mu}_j^r \sum_{i=1}^{N} \frac{R_{ij}n_i}{\sum_{k=1}^{M} R_{ik}\hat{\mu}_k^r} \qquad (5.23)$$

for the j-th bin and $r+1$ iterations. The regularization is introduced by setting an initial truth distribution $\hat{\boldsymbol{\mu}}^0$. For a low number of iterations r the estimator $\hat{\boldsymbol{\mu}}^r$ will be very biased towards the initial truth distribution $\hat{\boldsymbol{\mu}}^0$ but have relatively low covariances. In the limit of many iterations it is emperically shown that the estimator will converge to the ML estimator with zero bias but large covariances.

5.2.4 Iterative Dynamically Stabilized

Iterative Dynamically Stabilized (IDS) unfolding is another iterative procedure as proposed in [206]. Just like the Richardson-Lucy algorithm the IDS method starts

from some initial truth distribution $\boldsymbol{\mu}^0$ to construct an estimator which then is improved with each iteration. The estimator is defined as

$$\hat{\mu}_i = \mu_i^0 C + \Delta\beta_i^{\text{true}} + \sum_{j=1}^{N} \left(f(\Delta n_i, \sigma_{\Delta n_i}, \lambda) R_{ji} \Delta n_i - (1 - f(\Delta n_i, \sigma_{\Delta n_i}, \lambda)) \Delta n_i \delta_{ij} \right) \quad (5.24)$$

with a MC normalization coefficient C that corrects for differences in MC and data, $\Delta\beta_i^{\text{true}}$ the expected number of truth events associated with fluctuations in the background subtracted from the data, $\Delta n_i = n_i - \Delta\beta_i^{\text{reco}} - C\nu_i$, σ_{n_i} being the uncertainty on the measured data n_i, R_{ji} the response matrix and $f(\Delta n_i, \sigma_{\Delta n_i}, \lambda)$ a function with a value between 0 and 1. With each iteration C, R_{ji} and $f(\Delta n_i, \sigma_{\Delta n_i}, \lambda)$ are adjusted. Again, similar to the Richardson-Lucy method, low iterations is equivalent to high regularization and high iterations to low regularization.

5.2.5 Gaussian Process Unfolding

Gaussian Process (GP) Unfolding as proposed in [207] is a form of Bayesian regression that defines the mode of a posterior distribution as an estimator. The posterior distribution is constructed with Bayes' theorem from a Gaussian likelihood and a Gaussian Process prior. By maximizing the posterior a closed form solution can be found,

$$\hat{\boldsymbol{\mu}} = K \left[K + R^{-1} V (R^{-1})^T \right] \left(R^{-1} \boldsymbol{n} - \boldsymbol{m} \right) + \boldsymbol{m} \quad (5.25)$$

with response matrix R, kernel matrix K defined by the kernel function and some reference histogram $\boldsymbol{m}$. In [207] it is shown that the choice of $\boldsymbol{m}$ does not substantially influence the mode of the posterior and can therefore be set to zero. Regularisation is introduced via the kernel function of the GP defining K. Different choices of kernel can be made such as the radial basis function (RBF) kernel,

$$k(\boldsymbol{x}, \boldsymbol{x}') = A \exp\left(\frac{-||\boldsymbol{x} - \boldsymbol{x}'||^2}{2l^2}\right) \tag{5.26}$$

with hyperparameters A and l, or the Gibbs function kernel,

$$k(\boldsymbol{x}, \boldsymbol{x}') = \prod_{i=1}^{M}\left(\frac{2l_i(\boldsymbol{x})l_i(\boldsymbol{x}')}{l_i(\boldsymbol{x})^2 + l_i^2(\boldsymbol{x}')^2}\right)^{\frac{1}{2}} \exp\left(-\sum_{i=1}^{M}\frac{(x_i - x_i')^2}{l_i(\boldsymbol{x})^2 + l_i^2(\boldsymbol{x}')^2}\right). \tag{5.27}$$

with its hyperparameters defined in the distance functions $l_i(\boldsymbol{x})$. The kernel function and hence its hyperparameters completely determine the amount of regularization. There are various methods of choosing these hyperparameters of which some are proposed in [207].

5.3 Determining the regularisation parameter

In the previous section we discussed various unfolding methods that construct an estimator $\hat{\boldsymbol{\mu}}$ in different ways. One thing they all have in common is that they have one or more parameters that define the amount of regularization that defines the bias-variance trade-off. A common misconception is that there is one value for each regularization parameter that is the best in any unfolding scenario. However, in reality does each new setting require a different value. This section defines several figures of merit that can be used to not only to compare different unfolding methods but also help determining the regularization parameters. In terms of statistical learning this is known as *hyperparameter optimization*.

5.3.1 Mean Squared Error

Both variance and bias of can be seen as a form of error on an estimator. A well known figure of merit that takes both into account is the mean squared error defined

as

$$\text{MSE} = \frac{1}{M} \sum_{i=1}^{M} (U(\hat{\mu}_i, \hat{\mu}_i) + b[\hat{\mu}_i]^2) \tag{5.28}$$

with variance $U(\hat{\mu}_i, \hat{\mu}_i)$ and bias $b[\hat{\mu}_i]$. An approach could be to choose a regularization parameter that minimizes this quantity. Histograms were the bin entries differ by a couple of scales, e.g. for an exponential distribution, will have the bins with the highest entries have the largest weight in defining the regularization parameter. In these cases one could redefine Eq. (5.28) to

$$\begin{aligned}
\text{MSE} &= \frac{1}{M} \sum_{i=1}^{M} (U_{\text{rel}}(\hat{\mu}_i, \hat{\mu}_i) + b_{\text{rel}}[\hat{\mu}_i]^2) \\
U_{\text{rel}}(\hat{\mu}_i, \hat{\mu}_i) &= \frac{U(\hat{\mu}_i, \hat{\mu}_i)}{\hat{\mu}_i} \\
b_{\text{rel}}[\hat{\mu}_i] &= \frac{b[\hat{\mu}_i]}{\hat{\mu}_i}
\end{aligned} \tag{5.29}$$

5.3.2 Coverage Probability

Another criterion one could base their choice of regularization parameter on is the coverage probability of the confidence interval constructed for each bin, as proposed in [208]. A 68% confidence interval can be constructed as $[\hat{\mu}_i - \sigma_{\hat{\mu}_i}, \hat{\mu}_i + \sigma_{\hat{\mu}_i}]$ with standard deviation $\sigma_{\hat{\mu}_i} = \sqrt{U(\hat{\mu}_i, \hat{\mu}_i)}$ and estimator $\hat{\mu}_i$ for bin i. If one would repeat many measurements and construct a confidence interval for each measurement $\boldsymbol{n}$ then 68.3% of these intervals will cover the true value μ_i. It is possible to estimate the probability that the confidence interval will have 68% coverage of some true value μ_i.

$$P_{\text{cov}} = \frac{1}{M} \sum_{i=1}^{M} P(\hat{\mu}_i - \sigma_{\hat{\mu}_i} < \mu_i < \hat{\mu}_i + \sigma_{\hat{\mu}_i} | \mu_i) \tag{5.30}$$

In the limit of no regularization the coverage probability should converge to its nominal value $P_{\text{nom}} = 0.683$. However, when bias is increased and variance reduced with

increasing regularization the coverage probability will inevitably go down. One approach is to choose a regularization parameter that gives a minimum $P_{\text{cov}} = P_{\text{nom}} - \epsilon$ defined by a tolerance parameter ϵ. One trades the ambiguous regularization parameter τ with an interpretable parameter ϵ that controls for a statistical property, i.e. coverage, of the constructed estimator.

5.3.3   L-curve

The *L-curve criterion*, as proposed in [209], selects the regularization parameter τ based on the curvature of the L-curve. The L-curve is a plot of the log-likelihood $\log L(\boldsymbol{\mu})$ and the regularization constraint term $S(\boldsymbol{\mu})$ for various regularization strengths(See Fig. 5.3) after maximization. In case of $\tau \to 0$ the values of $\log L(\boldsymbol{\mu})$ will be very small, $S(\boldsymbol{\mu})$ very big and the variance will be large. However, for large τ it will be the other way around and the bias will be very large. In previous studies[210, 211] it is shown that the L-curve criterion will give a good bias-variance trade-off and would result in a robust choice for τ. The TUnfold algorithm has a build-in feature that scans the L-curve and returns the τ corresponding to the largest curvature.

5.3.4   Global Correlation Coefficient

Global correlation coefficients [201] estimate the correlation between an estimator $\hat{\mu}_i$ and a linear combination of estimators of the remaining truth bins. In case of large regularization the bins will be positively correlated because of the constraint on the shape of the unfolded distribution. However, in the limit of $\tau \to 0$ the correlations will become negative as is known of the ML estimator. A minimum average correlation coefficient will be located in between these two scenarios and can be taken as a quantity to choose an optimal regularization parameter τ.
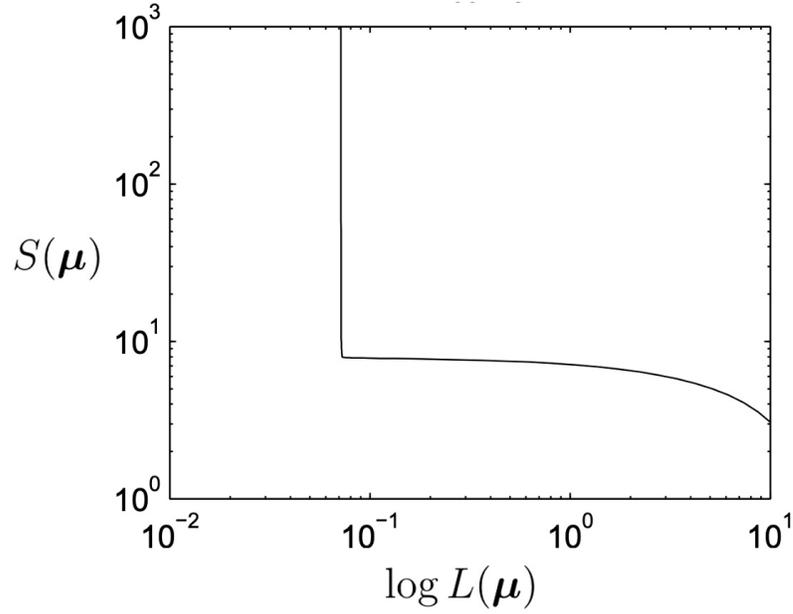
Figure 5.3: An example of an L-curve. From [209].

$$\text{Average } \rho = \frac{1}{N} \sum_{i=1}^{N} \sqrt{1 - ((U)_{ii}(U)_{ii}^{-1})^{-1}} \qquad (5.31)$$

5.3.5 χ^2

The χ^2 is used in many cases as a goodness-of-fit measure. The same can be done here by defining

$$\chi^2/\text{n.d.f.} = \frac{1}{N}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)U^{-1}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^T \qquad (5.32)$$

with U being the covariance matrix of the estimators $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{\mu}_0$ some reference truth distribution. One can choose τ that gives a $\chi^2/n.d.f.$ that is closest to 1 that signifies a good agreement w.r.t. the reference truth distribution.

All of the proposed figures of merit can be suitable for choosing a regularization parameter. Additionally, one can optimize the regularization parameter w.r.t. your chosen figure of merit with cross validation methods [212, 213, 214]. Cross validation

methods can reduce the dependency of the regularization parameter on the training data and give a more realistic validation of the unfolding estimator it constructs.

5.4   Comparison Study

In this section we present a comparison study [215] of the different `RooUnfold` unfolding algorithms as described in Sec. 5.2. In this setup we use analytical expressions for the physics process and detector response model. This gives a benchmark unfolding problem that can be easily controlled but still maintains a realistic level of complexity. A *bimodal* and *exponential* model is used which both have a different underlying truth distribution and Gaussian smearing function.

5.4.1   Exponential Example

We define the model exponential model with an exponential decay distribution as underlying truth distribution $f_{\text{truth}}(x|\zeta)$, smeared with a resolution function that is loosely inspired on a calorimeter response:

$$
\begin{aligned}
f(x|\zeta) &= f_{\text{truth}}(x_{\text{true}}|\zeta) * f_{\text{detector}}(x_{\text{true}}, x) \\
&= (\zeta \cdot \exp(-\zeta \cdot x_{\text{true}})) * \text{Gauss}(x - x_{\text{true}}, 7.5, 0.5 \cdot \sqrt{x_{\text{true}}} + 2.5)
\end{aligned}
\tag{5.33}
$$

where $*$ denotes the convolution operator. The event generation is simulated by sampling values for x_{true} from the exponential distribution with parameter ζ. Here we denote the 'SM' and 'BSM' distribution sampled from an exponential with $\zeta = 0.035$ and $\zeta = 0.05$, respectively. The detector response is subsequently simulated by smearing the sampled values with an additive noise value sampled from a Gauss function with mean 7.5 and a standard deviation of $0.5 \cdot \sqrt{x_{\text{true}}} + 2.5$.

The response matrix is estimated under the assumption of the 'SM' scenario, as is
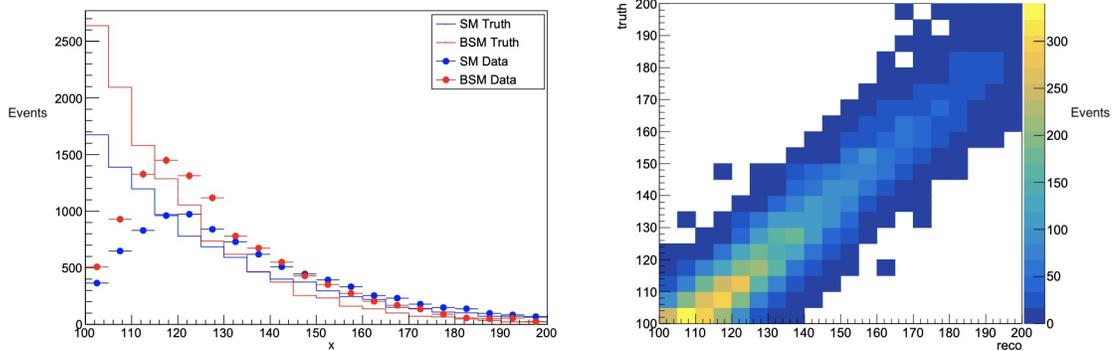
Figure 5.4: Left: The true distributions for the SM and BSM models and the corresponding smeared datasets. Right: the transfer matrix for the SM model, which is populated by the same events as the SM true distribution shown. From [215].

normally done for unfolding frameworks in HEP. The response matrix is used to unfold data generated with both the 'SM' and 'BSM' model. The 'BSM' model data is to check that truth models deviating from the 'SM' case can still be retrieved with 'SM' assumptions in the unfolding framework. The following unfolding algorithms from `RooUnfold` are compared:

- Richardson-Lucy (Iterative Bayes)

- Iterative Dynamically Stabilized (IDS)

- Singular Value Decomposition (TSVD)

- TUnfold

- Gaussian Processes (GP)

- Matrix Inversion

- Bin-by-bin

For each unfolding algorithm the bias, variance and coverage probability are estimated for the constructed truth estimator $\hat{\boldsymbol{\mu}}$. For both the 'SM' and the 'BSM' model we have some truth histogram $\boldsymbol{\mu}$ filled with the values generated with the before mentioned procedure. With this we can construct detector-level histogram by applying

the 'SM' response matrix, i.e. $\boldsymbol{\nu} = R\boldsymbol{\mu}$. Under the assumption that each observed bin count n_i is a random variable following a Poisson distribution with as mean ν_i one can sample new values $\boldsymbol{n}_k$ from the Poisson distributions. One sampled histogram $\boldsymbol{n}_k$ is known in HEP as a "toy MC" dataset. Here we use $K = 1000$ toy MCs which each is unfolded with the chosen method to construct an estimator $\hat{\boldsymbol{\mu}}_k$. The set of estimators is used to estimate the variance of the original estimator $\hat{\boldsymbol{\mu}}$ with the sample variance,

$$U[\hat{\mu}_i, \hat{\mu}_i] = \frac{1}{K-1} \sum_{k=1}^{K} (\hat{\mu}_{i,k} - \bar{\hat{\mu}}_i)^2, \qquad \bar{\hat{\mu}}_i = \frac{1}{K} \sum_{k=1}^{K} \hat{\mu}_{i,k}. \tag{5.34}$$

The bias can be estimated with

$$b[\hat{\mu}_i] = \frac{1}{K} \sum_{k=1}^{K} (\hat{\mu}_{i,k}) - \mu_i. \tag{5.35}$$

To estimate the coverage probability one could take one of the following two approaches. The first would be to construct many confidence intervals and take the fraction of confidence intervals that contain the truth value μ_i as the coverage probability. The confidence intervals can be constructed for additional sampled $\boldsymbol{n}$ either from the exponential distribution or via the toy MC approach. However, one can define a closed form estimate for the coverage probability for bin i based on bias and variance estimates,

$$P_{\text{cov},i} = \Phi\left(\frac{b[\hat{\mu}_i]}{\sqrt{U(\hat{\mu}_i, \hat{\mu}_i)}} + 1 \right) - \Phi\left(\frac{b[\hat{\mu}_i]}{\sqrt{U(\hat{\mu}_i, \hat{\mu}_i)}} - 1 \right) \tag{5.36}$$

with Φ is the Standard Gaussian cumulative distribution function. This definition holds under the assumptions that the estimators $\hat{\boldsymbol{\mu}}$ are linear dependent on the data $\boldsymbol{n}$ and that the data $\boldsymbol{n}$ follow a Gaussian distribution. The derivation and further details can be found in [208].
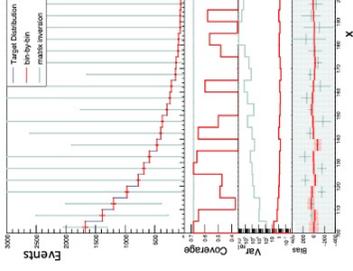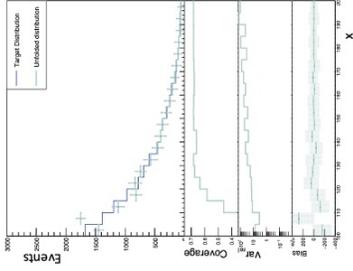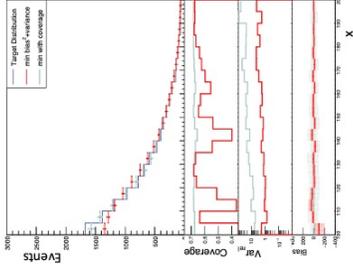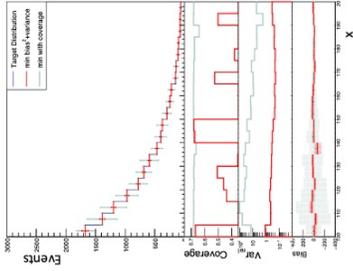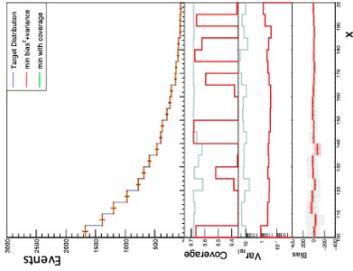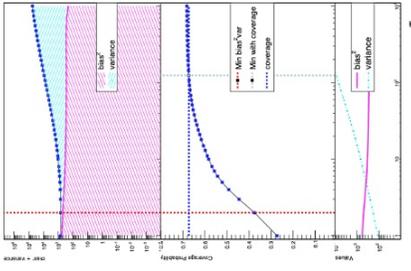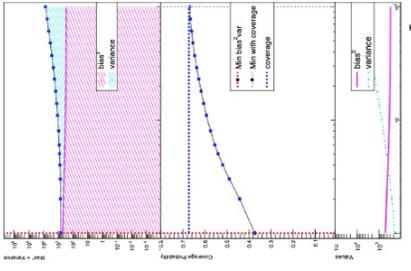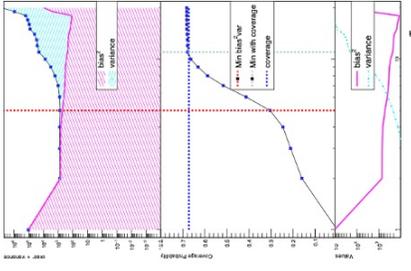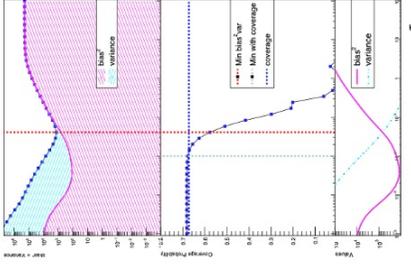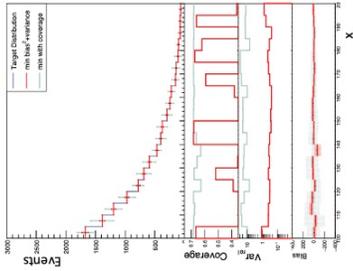
Figure 5.5: Comparison of the unfolding performance for different algorithms. For each method, the top plot shows the unfolded SM data compared to the SM true distribution, where SM response matrix was used as input to the unfolding method. For the methods with a tuneable regularisation strength two unfolding solutions are shown: (1) with a regularisation strength corresponding to the unconditionally minimised bin-averaged MSE (red), and (2) with a conditionally minimised bin-averaged MSE with the requirement that the bin-averaged coverage reaches the target coverage of 67.3%(grey). The bottom panels of the top row plots display the coverage probability, variance and bias estimates per bin. For the bias estimate, the statistical error of the unfolded toy MC used to estimate the bias(i.e the RMS of the distribution of $\hat{\mu}$) is included. For the tuneable methods, the bottom row shows the bin-averaged bias, variance and coverage probabilities of estimators $\hat{\mu}$ constructed for varying regularization parameters τ. The dotted vertical lines represent the unconditional bin-averaged MSE minimum(red) and the conditional bin-averaged MSE minimum with the target bin-averaged coverage of 67.3%(grey). From [215].

94

Figure 5.6: Comparison of the unfolding performance for different algorithms. For each method, the top plot shows the unfolded BSM data compared to the BSM true distribution, where SM response matrix was used as input to the unfolding method. For the methods with a tuneable regularisation strength two unfolding solutions are shown: (1) with a regularisation strength corresponding to the unconditionally minimised bin-averaged MSE (red), and (2) with a conditionally minimised bin-averaged MSE with the requirement that the bin-averaged coverage reaches the target coverage of 67.3%(grey). The bottom panels of the top row plots display the coverage probability, variance and bias estimates per bin. For the bias estimate, the statistical error of the unfolded toy MC used to estimate the bias(i.e the RMS of the distribution of $\hat{\mu}$) is included. For the tuneable methods, the bottom row shows the bin-averaged bias, variance and coverage probabilities of estimators $\hat{\mu}$ constructed for varying regularization parameters τ. The dotted vertical lines represent the unconditional bin-averaged MSE minimum(red) and the conditional bin-averaged MSE minimum with the target bin-averaged coverage of 67.3%(grey). From [215].

95

With the use of the bias, variance and coverage probability estimation methods one can compare the performance of the unfolding methods with a 'SM' response matrix on the 'SM' (See Fig. 5.5) and 'BSM' (See Fig. 5.6) data for varying regularization parameters. The unfolded distributions are shown for a regularization parameter that:

1. unconditionally minimizes the bin-averaged MSE (See Eq. (5.28))

2. conditionally minimizes the bin-averaged MSE with the requirement that the bin-averaged coverage (See Eq. (5.30)) reaches a minimum of 0.683-ϵ with $\epsilon = 0.01$.

Additionally, the bin-averaged bias, variance and coverage probability is shown for a large range of regularization parameters. The matrix inversion and bin-by-bin methods are unregularized and therefore only give one solution. The Gaussian Processes method was used with a RBF kernel which has two tuneable regularization parameters. Instead of scanning the 2D parameter space, a different data-driven method was used to choose their optimal values known as marginalized likelihood maximization. This method is implemented in `RooUnfold` as described in [207] and is used to find one optimal set of regularization parameters. A more detailed and general description of the method can be found in [216].

The plots in Fig. 5.5 show that for the SM case all the unfolding algorithms are able to reasonably retrieve the truth distribution. However, when using the unconditional minimized MSE unfolding solution many bins show a substantial undercoverage. Especially for the iterative methods (Bayes and IDS) it should be noted that the number of iterations needed to get a coverage probability close to nominal coverage of 68.3% is of the scale 10^2. This is substantially higher than some of the default values adopted by certain parts of the HEP community. The Gaussian Processes method shows bad

coverage and high bias in the low x-region which is caused by boundary effects. The matrix inversion method shows perfect coverage as expected whilst the bin-by-bin method has low bias but also low coverage. The coverage panels for the Bayes, IDS, TSVD and TUnfold shows that the coverage can peak in certain regions. This is because the bias to variance ratio is higher in that region relative to the other regions (See Eq. (5.36)). Possible explanation for these peaks could be a slight bias in the truth distribution used to calculate the bias relative to (See (5.10)). However, this increase in the bias-variance ratio can also just be the behaviour of the constructed estimator.

The plots in Fig. 5.6 show the unfolded results of observed data from a BSM model but with a response matrix estimated with a SM model. As mentioned in the beginning of this chapter the choice of the underlying model should not have an effect on the response matrix if the binning is chosen carefully. However, when regularization with a SM assumption is introduced the unfolded distribution can be heavily biased towards the SM truth. We see that all algorithms reasonably retrieve the shape of the BMS truth distribution. We see that some bins have a slight increase in bias and again undercoverage for the unconditional minimized MSE unfolding distributions. For the conditional minimized MSE we see a large increase in variance for many bins in the IDS and Gaussian Processes unfolded distributions.

5.4.2 Bimodal Example

The bimodal example is used to investigate the response matrix model dependence even further. This is done by unfolding data from models increasingly different from the model used to estimate the response matrix. We define the bimodal model as

$$
\begin{aligned}
f(x|\zeta) &= f_{\text{truth}}(x_{\text{true}}|\zeta) * f_{\text{detector}}(x_{\text{true}}, x) \\
&= ((0.5 f_{\text{CB}}(x_{\text{true}}|\mu = 2.4, \sigma = 0.48, \zeta, n = 1) + \\
&\quad (0.5 f_{\text{CB}}(x_{\text{true}}|\mu = 5.6, \sigma = 0.48, \zeta, n = 1)) \\
&\quad * \text{Gauss}(x - x_{\text{true}}|0, 0.4)
\end{aligned}
\tag{5.37}
$$

where f_{CB} is the Crystal Ball probability density function which has a Gaussian core determined by the parameters μ and σ and an additional power law tail controlled by the parameters n and ζ. The data is sampled from a bimodal model with $n = 1$, $\zeta = [0.5, 1, 1.5, 2, 2.5, 3]$ and the response matrix is generated with $n = 1$ and $\zeta = \infty$ which reduces the Crystal Ball function to a Gaussian. Figure 5.7 shows the truth distributions of the bimodal model for different ζ. Figure 5.8 shows the comparison of bin-averaged bias for unfolded distributions of the Bayes, IDS, SVD and TUnfold algorithms. The regularization parameters are chosen again such that they unconditionially minimize the MSE and once for regularization parameters that conditionally minimize the MSE with a minimal coverage of $0.683 - \epsilon$ with $\epsilon = 0.01$. Figure 5.9 shows the bin-averaged bias for Gaussian Processes, matrix inversion and bin-by-bin unfolding
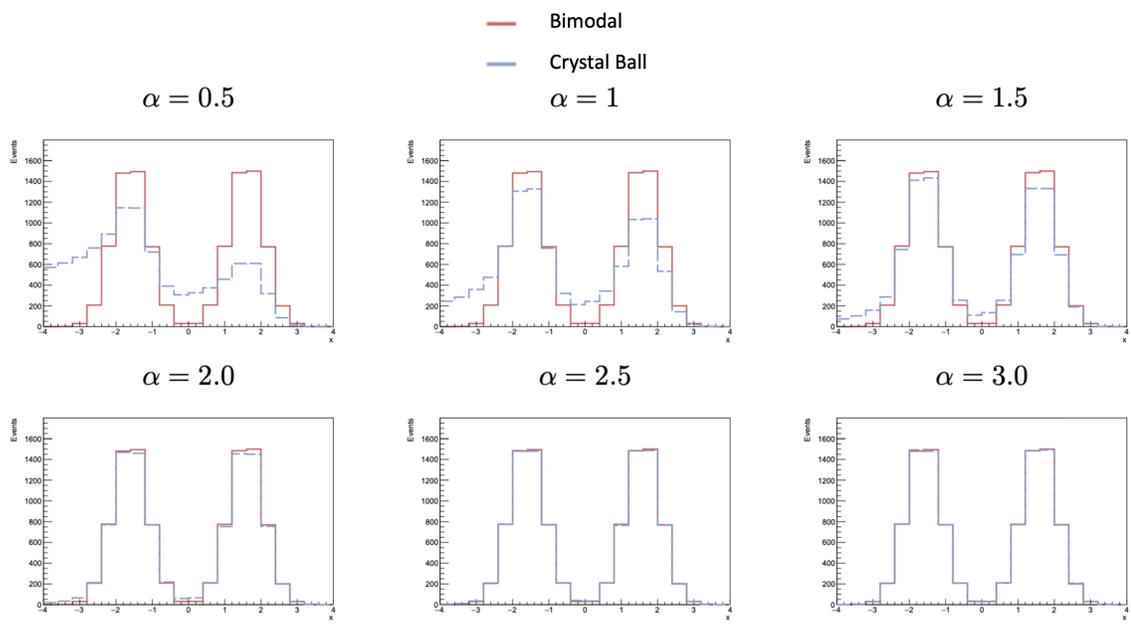
Figure 5.7: The truth distributions of the bimodal model compared with the Crystal Ball models for $\zeta = [0.5, 1, 1.5, 2, 2.5, 3]$. From [215].

(a)   (b)
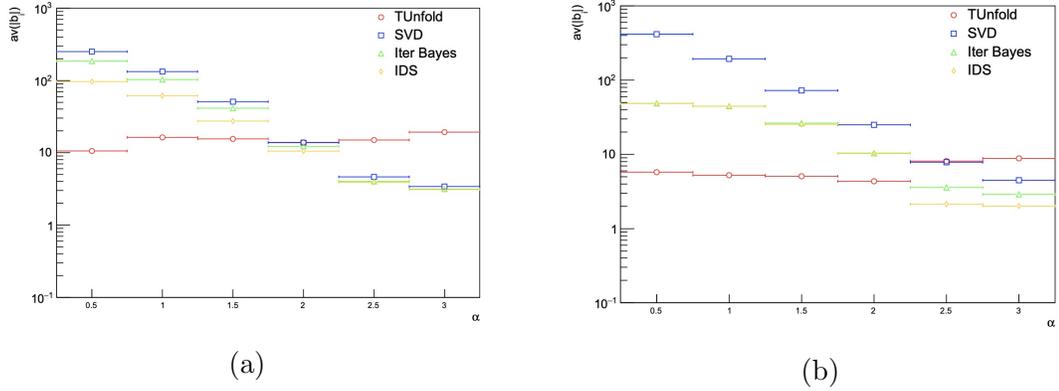
Figure 5.8: The bin-averaged bias for the unfolded distributions for varying α for the TUnfold, SVD, Bayes and IDS method. The regularization parameters are chosen such that they a) unconditionally minimize the MSE and b) conditionally minimize the MSE with a minimal coverage of $0.683 - \epsilon$ with $\epsilon = 0.01$. From [215].
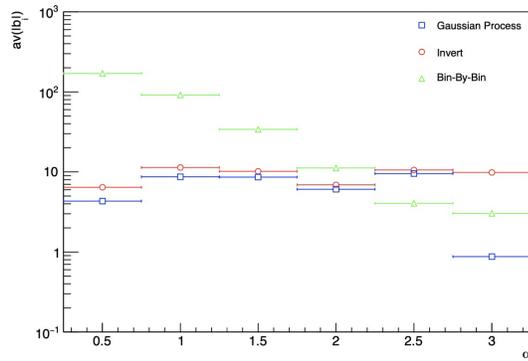
Figure 5.9: The bin-averaged bias for the unfolded distributions for varying α for the Gaussian Processes, matrix inversion and bin-by-bin method. From [215].

methods. Quite notable is the bin-averaged bias of the TUnfold, Gaussian Processes and matrix inversion methods which indicate very little dependence on the α parameter. This indicates that they are slightly more robust w.r.t. datasets originating from models deviating from the model used to estimate response matrix. The other methods clearly show an increase in bias for lower α i.e. for bigger data model deviations.

100

5.5 Response Matrix Estimation

Up until now we assumed that the response matrix R has been known. However, in general this is not the case and we need to use Monte Carlo simulations with model assumptions to estimate it. In this section we will firstly introduce the conventional way of estimating the response matrix. After that we will introduce a novel method to estimate the response matrix and compare it with the first. As example unfolding setup we define the exponential model,

$$f(x|\alpha) = f_{\text{truth}}(x_{\text{true}}|\alpha) * f_{\text{detector}}(x_{\text{true}}, x)$$

$$= (0.2 \cdot \exp(-0.2 \cdot x_{\text{true}})) * \text{Gauss}(x - x_{\text{true}}, \mu = 0, \sigma = 2.5 + x_{\text{true}}/25).$$

$$(5.38)$$
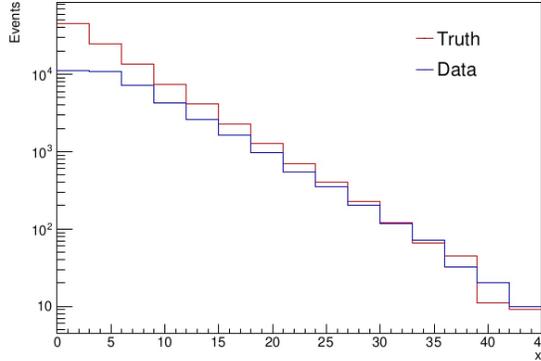
Figure 5.10: Truth and reconstructed histogram of the exponential model.

Note the big difference in number of events between the first and the last bins. These differences are very common in HEP, i.e. like in the p_T distributions of the top quark.

5.5.1 Normalizing the transfer matrix

Here we discuss the conventional approach to estimating a response matrix. We start off by defining the binning of the truth ($\boldsymbol{\mu}$) and reconstructed ($\boldsymbol{\nu}$) histograms for

some observable of interest. One can use these binnings to construct a 2D histogram N commonly known as the transfer matrix. If one has a Monte Carlo simulated data sample with the truth and reconstructed events one can fill the transfer matrix with the truth and reconstructed value of the observable of interest on the x- and y-axis, respectively. One can then estimate the response matrix by normalizing the filled transfer matrix N.

$$R_{ij} = P(\text{measured in bin i|true in bin j}) = \frac{N_{ij}}{\sum_{k=1}^{M} N_{ik}} \tag{5.39}$$

It is also possible that certain truth events are not reconstructed because of reconstruction efficiencies or kinematic selections. For such events one can use the underflow bins i.e. for a truth event in bin i one can add an entry in the underflow bin N_{i0} and extend the sum in the denominator such that the underflow bins are included. One should note that a sufficiently large sample is needed such that the statistical uncertainties on each response matrix element estimate is negligible. Especially in distributions such as exponentially decaying distributions one can have trouble populating certain bins e.g. at the end of the tail. Alternatively one could increase certain bin sizes to increase the event population in the transfer matrix. Otherwise one would have to propagate the uncertainties onto the unfolded distribution which will be further discussed in Sec. 5.6. Figure 5.11 shows estimates of the response matrix by normalizing the transfer matrix but with different numbers of generated events. We can clearly see the deterioration of the estimates for lower numbers. Especially in the high-x regions. Also, one should note that multidimensional distributions will need even higher numbers to populate all the bins to get a reasonable response matrix estimate.

Figure 5.11: Response matrices estimated with the conventional transfer matrix method with a) $N = 10^6$, b) $N = 10^5$ and c) $N = 10^4$ simulated events.
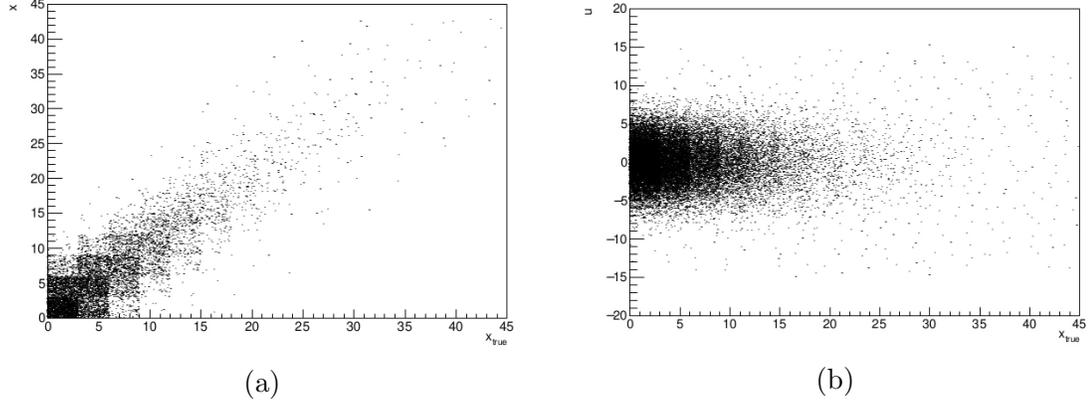
Figure 5.12: Scatter plots of a) the truth and reconstructed values (x_{true}, x) and b) the truth and deviation between the truth and reconstructed value (x_{true}, u).

5.5.2 Fitting a likelihood

Preferably one would like to have a "smooth" description of the response matrix that does not depend so heavily on the number of generated events in you MC sample. The following approach has been largely based on [217]. One could start by looking at the deviation between the reconstructed and the true value of the observable of interest,

$$u = x - x_{\text{true}}. \tag{5.40}$$

A $x_{\text{true}}u$-scatter plot is shown in Fig. 5.12. From this one could reasonably assume that u follows a Gaussian distribution with its mean centered around zero and standard deviation only slowly varying with x_{true}, e.g.

$$g(u|x_{\text{true}}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[\frac{(u-\mu)^2}{2\sigma^2}\right] \tag{5.41}$$

with

$$\mu = \theta_1 + \theta_2 x_{\text{true}} \tag{5.42}$$

$$\sigma = \theta_3 + \theta_4 x_{\text{true}} + \theta_5 x_{\text{true}}^2. \tag{5.43}$$

104

Here the parameters $\boldsymbol{\theta}$ completely parameterize the detector response. According to Bayes' theorem one can define the joint probability density function of u and x_{true} as

$$g(u, x_{\text{true}}) = g(u|x_{\text{true}})g(x_{\text{true}}) \tag{5.44}$$

where $g(x_{\text{true}})$ is the marginal probability distribution function of x_{true}. We can use this to construct a log-likelihood,

$$\log L(\boldsymbol{\theta}) = \sum_i \log g(u_i, x_{\text{true},i}|\boldsymbol{\theta}) = \sum_i \log g(u_i|x_{\text{true},i}, \boldsymbol{\theta}) + C \tag{5.45}$$

where $(u_i, x_{\text{true},i})$ are the deviation and the true value of the observable of interest of the i-th generated MC event. Note the last C term which depends on x_{true} but is constant in $\boldsymbol{\theta}$. Maximizing Eq. (5.45) w.r.t. $\boldsymbol{\theta}$ will therefore not be affected by this term. Let's assume that we found the maximum-likelihood estimators $\boldsymbol{\theta}_{\text{ML}}$. We can transform the probability distribution function back from u to x,

$$f(x|x_{\text{true}}, \boldsymbol{\theta}_{\text{ML}}) = g(u(x)|x_{\text{true}}, \boldsymbol{\theta}_{\text{ML}}) \left| \frac{du}{dx} \right| = g(x - x_{\text{true}}|x_{\text{true}}, \boldsymbol{\theta}_{\text{ML}}). \tag{5.46}$$

Now let us recall the definition of the response matrix,

$$
\begin{aligned}
R_{ij} &= P(\text{measured in bin i}|\text{true in bin j}) = \frac{P(\text{measured in bin i} \wedge \text{true in bin j})}{P(\text{true in bin j})} \\
&= \frac{\int_{\text{bin i}} dx \int_{\text{bin j}} dx_{\text{true}} f(x, x_{\text{true}}|\boldsymbol{\theta}_{\text{ML}})}{\int dx \int_{\text{bin j}} dx_{\text{true}} f(x, x_{\text{true}}|\boldsymbol{\theta}_{\text{ML}})} \\
&= \frac{\int_{\text{bin j}} dx_{\text{true}} f(x_{\text{true}}|\boldsymbol{\theta}_{\text{ML}}) \int_{\text{bin i}} dx f(x|x_{\text{true}}, \boldsymbol{\theta}_{\text{ML}})}{\int_{\text{bin j}} dx_{\text{true}} f(x_{\text{true}}|\boldsymbol{\theta}_{\text{ML}})} \\
&\approx \int_{\text{bin i}} dx f(x|\langle x_{\text{true}}\rangle_j, \boldsymbol{\theta}_{\text{ML}}) = \int_{\text{bin i}} dx g(x - \langle x_{\text{true}}\rangle_j|\langle x_{\text{true}}\rangle_j, \boldsymbol{\theta}_{\text{ML}})
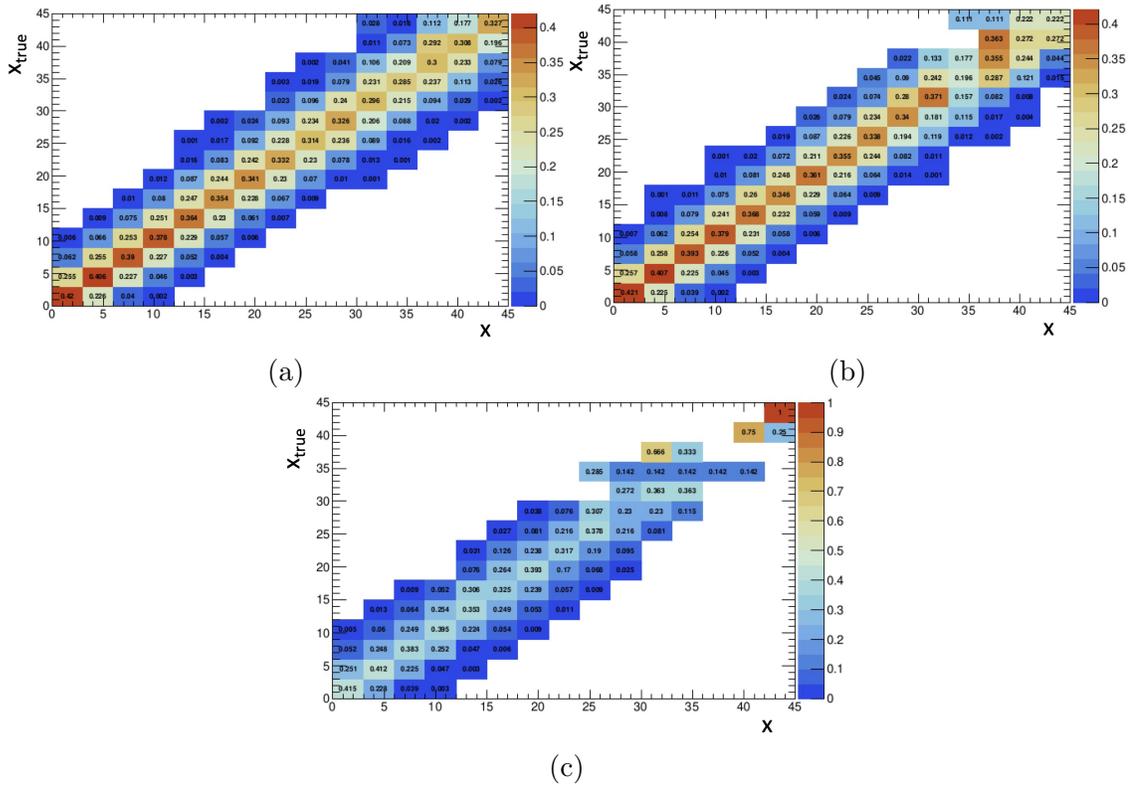\end{aligned}
$$

$$\tag{5.47}$$
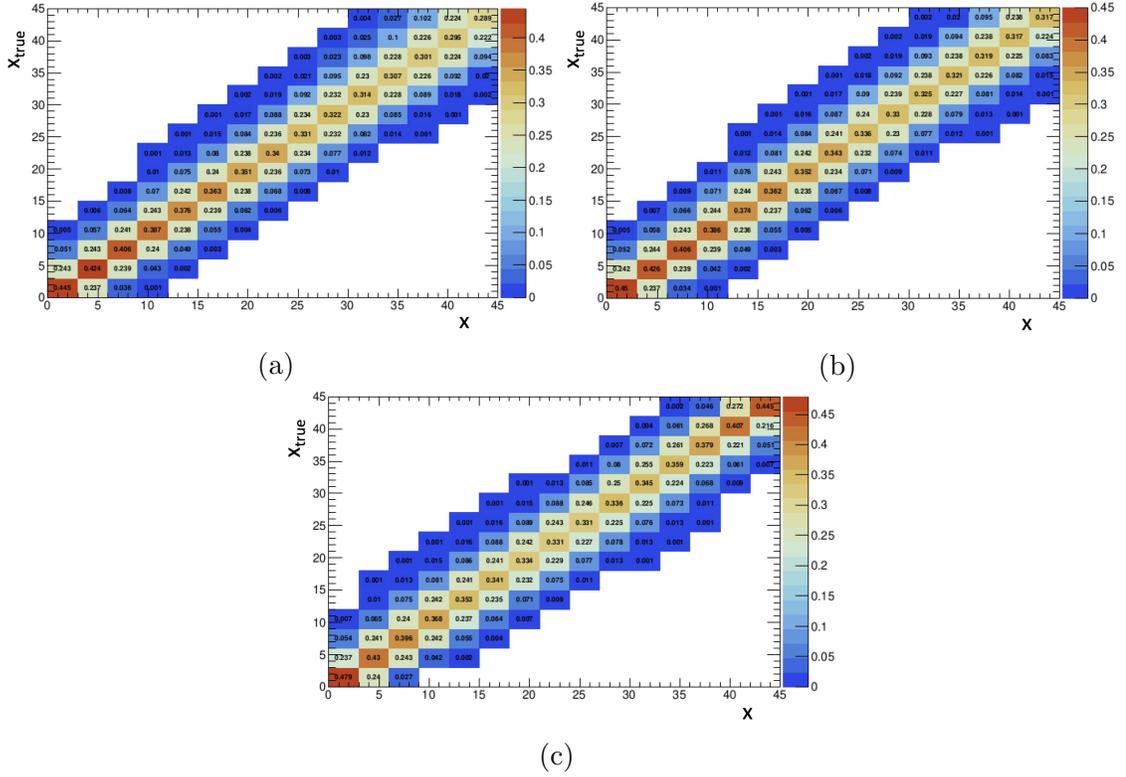
Figure 5.13: Response matrices estimated with the conventional transfer matrix method with a) $N = 10^6$, b) $N = 10^5$ and c) $N = 10^4$ simulated events.

with $\langle x_{\text{true}} \rangle_j$ being the mean value of x_{true} in bin j. Here we assume that $f(x | x_{\text{true}}, \boldsymbol{\theta}_{\text{ML}})$ is constant in x_{true} over bin j. So for sufficiently small bins we can evaluate $g(x - x_{\text{true}} | x_{\text{true}}, \boldsymbol{\theta}_{\text{ML}})$ for x_{true} at the center of bin j and integrate over bin i to get an estimate for R_{ij}.

Figure 5.13 shows that for less MC events the response matrix estimate changes. However, there are no missing elements or big jumps in value between neighbouring matrix elements i.e. like with the conventional transfer matrix method. Lets apply the response matrix to quantify their quality even further. Recall the relation $\boldsymbol{\nu} = R\boldsymbol{\mu}$. Here we generate a test sample of 10^6 events to fill both a truth $\boldsymbol{\mu}$ and reconstructed histogram $\boldsymbol{\nu}$. If the response matrix is well estimated then it should retrieve the reconstructed histogram when multiplied with the truth histogram. Lets

note a response matrix estimated using the conventional transfer matrix procedure of the previous section with $\hat{R}_{\mathrm{conv.}}$ and a response matrix using the likelihood fit with $\hat{R}_{\mathrm{lik.}}$. The reconstructed histograms, $\hat{\boldsymbol{\nu}}_{\mathrm{conv.}} = \hat{R}_{\mathrm{conv.}}\mu$ and $\hat{\boldsymbol{\nu}}_{\mathrm{lik.}} = \hat{R}_{\mathrm{lik.}}\mu$, are then both compared to the histogram $\boldsymbol{\nu}$ filled with the reconstructed events of the test sample. One can find all three histograms in Fig. 5.14 and the corresponding ratio plots $\hat{\boldsymbol{\nu}}/\boldsymbol{\nu}$ in Fig. 5.15. We can see in the ratio plots that for a relatively high number of estimation events $N = 10^6$ the conventional method performs slightly better. We also see that $\hat{R}_{\mathrm{lik.}}$ has a positive bias which means that the fitted probability distribution functions $g(x - \langle x_{\mathrm{true}} \rangle_j | \langle x_{\mathrm{true}} \rangle_j, \boldsymbol{\theta}_{\mathrm{ML}})$ does not model the data perfectly. One way to improve this could be to change the probability distribution function in Eq. (5.41) or the parameterization in Eq. (5.42) and Eq. (5.43). However, when going one or two scales down in events shows a deterioration in the high x bins, i.e. the bins with low counts, for the conventional method. Especially here the response matrices estimated with the likelihood fit perform better. The low x bins, i.e. the high statistics bins, both methods perform similar.

(a)   (b)   (c)
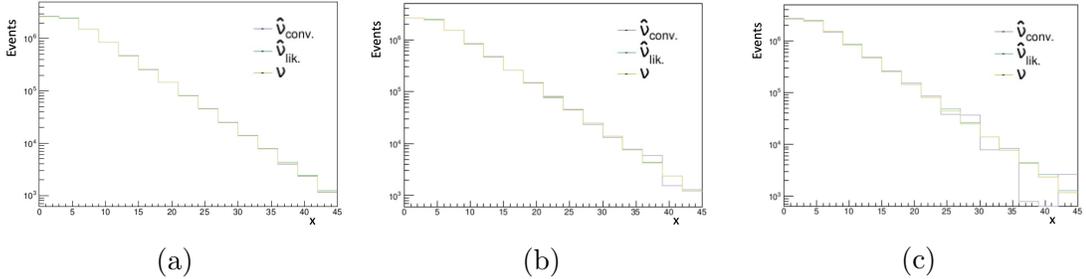
Figure 5.14: The reconstructed histograms $\hat{\nu}_{\mathrm{conv.}} = \hat{R}_{\mathrm{conv.}}\mu$, $\hat{\nu}_{\mathrm{lik.}} = \hat{R}_{\mathrm{lik.}}\mu$ and $\boldsymbol{\nu}$. The response matrices of both estimation methods are estimated with a) $N = 10^6$, b) $N = 10^5$ and c) $N = 10^4$ simulated MC events.
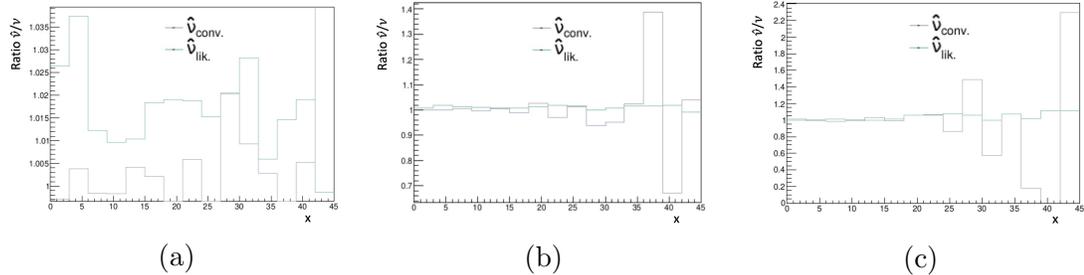
Figure 5.15: The reconstructed histogram ratio plots $\hat{\nu}_{\text{conv.}}/\nu$ and $\hat{\nu}_{\text{lik.}}/\nu$. The response matrices of both estimation methods are estimated with a) $N = 10^6$, b) $N = 10^5$ and c) $N = 10^4$ simulated MC events.

5.6   Covariance Estimation

Estimating a true distribution will result in one or more estimators which have an uncertainty and non-trivial correlations between them. Estimating these uncertainties and correlations is an important part of any unfolding framework. Many unfolding algorithms supply an estimator for statistical covariances (See Sec. 5.2) i.e. covariances as a result of the stochastic nature of the data and the bin-to-bin migrations of events. However, in HEP it is very common to also have systematic sources of error which induce additional uncertainties and correlations. These systematic sources of error originate from estimates or assumptions that enter the unfolding framework, eg. energy resolution or reconstruction efficiencies, but have limited accuracy. The goal is to include all statistical and systematic sources of error in unfolding covariance estimation. A common approach within HEP is to model the sources of systematic uncertainty with *nuisance parameters*.

5.6.1 Nuisance parameters

In general, the response matrix R and backgrounds $\boldsymbol{\beta}$ depend on additional parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_K)$ introduced by properties of the detector response also known as *nuisance parameters*. One influence this will have is that the increased likelihood model flexibility will reduce the bias but increase the variance. One should therefore take care when removing, also known as pruning, or introducing nuisance parameters into the model.

Up until now, we assumed that these nuisance parameters, and therefore the response matrix and background distribution, were known with negligible uncertainty. However, in practice this is not valid and one needs to propagate these into the uncertainty on the unfolded distribution. By expressing the response matrix and background distribution as a function of $\boldsymbol{\theta}$ one includes the nuisance parameters in the likelihood function and thus incorporating systematic uncertainties in the model. Their correlations with the *parameters of interests* $\boldsymbol{\mu}$ will inflate the variance of the estimators $\hat{\boldsymbol{\mu}}$. Additionally, the best estimates $\tilde{\theta}_k$ of θ_k are treated as an *auxiliary measurement* which follow some probability distribution function $g(\tilde{\theta}_k|\theta_k)$ that assumes some value θ_k. A common choice for these p.d.f.s is a Gaussian which will result in the extended likelihood function

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}) = \prod_{i=1}^{N} f(n_i|\boldsymbol{\mu}, \boldsymbol{\theta}) \prod_{k=1}^{K} g(\tilde{\theta}_k|\theta_k) = \prod_{i=1}^{N} \frac{\nu_i^{n_i}}{n_i!} e^{-\nu_i} \prod_{k=1}^{K} \frac{1}{\sigma_{\tilde{\theta}_k}\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(\theta_k - \tilde{\theta}_k)^2}{\sigma_{\tilde{\theta}_k}^2}}. \quad (5.48)$$

with ν_i now depending on both $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$. However, other choices such as a log-normal or Student's t distribution are also possible as p.d.f. for the auxiliary measurements. One can construct a new linear combination of Eq. (5.48) and some regularization function $S(\boldsymbol{\mu})$,

$$\Phi(\boldsymbol{\mu}, \boldsymbol{\theta}) = L(\boldsymbol{\mu}, \boldsymbol{\theta}) + \tau S(\boldsymbol{\mu}). \tag{5.49}$$

In the rest of this section we will assume that we have constructed estimators $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\theta}}$ by maximizing Eq. (5.49) with regularization function (5.21) w.r.t. $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$.

5.6.2 Inverse Hessian

In Sec. 5.2 we discussed the Cramer-Rao Bound (RCB), also known as the Minimum Variance Bound (MVB). We can extend Eq. (5.12) to include estimators for both the parameters of interest $\boldsymbol{\mu}$ and the nuisance parameters $\boldsymbol{\theta}$,

$$U[\hat{\lambda}_i, \hat{\lambda}_j] \geq \left((\mathcal{I} - B) I^{-1} (\mathcal{I} - B)^T \right)_{ij} \tag{5.50}$$

where $\hat{\boldsymbol{\lambda}}$ denotes the estimators of all parameters, i.e. $\boldsymbol{\lambda} = (\boldsymbol{\mu}, \boldsymbol{\theta})$. Under the large sample approximation and for negligble bias the covariance is assumed to equal the unbiased RCB. If one can estimate the matrix of second order derivatives of the log-likelihood, also known as the *Hessian matrix*, one can take the inverse of this matrix as an estimate for the covariance. Under the large sample approximation one can also assume the log-likelihood is parabolic shaped around its maximum. In this case one can numerically approximate the second derivatives with finite differences.

$$U[\hat{\lambda}_i, \hat{\lambda}_j] = \left(\frac{\partial^2 \log L}{\partial \lambda_i \partial \lambda_j} \bigg|_{\hat{\boldsymbol{\lambda}}} \right)^{-1} \tag{5.51}$$

However, this approach has two important caveats. Firstly, this method is only suitable for the special case of no regularization i.e. $\tau = 0$. In general, regularization is needed which introduces non-zero bias and thus calls for the non-trivial task of estimating the bias gradient matrix B.

Secondly, a common misconception is assuming Eq. (5.49) can be treated as a likeli-

hood and its Hessian matrix can be used as a covariance matrix estimate, i.e.,

$$U[\hat{\lambda}_i, \hat{\lambda}_j] \neq \left(\frac{\partial^2 \Phi}{\partial \lambda_i \partial \lambda_j} \bigg|_{\hat{\boldsymbol{\lambda}}} \right)^{-1}. \tag{5.52}$$

Again, only in the special case of no regularization can Eq. (5.52) be used as a covariance matrix estimate as it will reduce back to Eq. (5.51). In the upcoming sections we will denote Eq. (5.52) as the *inverse Hessian* method and show when this method can hold or break.

5.6.3 Frequentist Pseudo-Experiments

An alternative approach would be to use pseudo experiments, or *toy MC*, to estimate the covariance. A similar method was described in Sec. 5.4. We assumed that the measured bin values follow a Poisson distribution $n_i \sim f(n_i|\boldsymbol{\mu}, \boldsymbol{\theta})$ and the auxiliary measurements a Gaussian distribution $\tilde{\theta}_k \sim g(\tilde{\theta}_k|\theta_k)$. One can set the p.d.f. parameters to the estimates of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\theta}}$ constructed with the observed data $\boldsymbol{n}$ i.e. $f(n_i|\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\theta}})$ and $g(\tilde{\theta}_k|\hat{\theta}_k)$. From these it is possible to sample new values $\boldsymbol{n}^t$ and $\tilde{\boldsymbol{\theta}}^t$ for the data and auxiliary measurements [218, 219, 220]. Each new sample is what is known as a *pseudo-experiment* or *toy MC*. For each t-th pseudo-experiment one can then evaluate Eq. (5.48) with the sampled values and construct new estimators $\hat{\boldsymbol{\lambda}}^t$ by maximizing w.r.t $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$. For T pseudo-experiments one can use the set of many estimators to estimate the covariance matrix

$$U[\hat{\lambda}_i, \hat{\lambda}_j] = \frac{1}{T-1} \sum_{t=1}^{T} (\hat{\lambda}_i^t - \bar{\hat{\lambda}}_i)(\hat{\lambda}_j^t - \bar{\hat{\lambda}}_j) \tag{5.53}$$

with

$$\bar{\hat{\lambda}}_i = \frac{1}{T} \sum_{t=1}^{T} (\hat{\lambda}_i^t) \tag{5.54}$$

This covariance estimation approach will work for both biased and unbiased estimators even if the log-likelihood function is not parabolic.

5.6.4 Frequentist-Bayes Hybrid Pseudo-Experiments

For some unfolding algorithms an explicit definition of a likelihood is not obvious, e.g. like for some iterative unfolding algorithms described in Sec. 5.2. This makes the definition of nuisance parameters and thus the previous two covariance estimation methods nonviable. In this section a covariance estimation method is presented that includes both statistical and systematic effects but does not need an explicit likelihood definition or an alteration of the chosen unfolding algorithm to include nuisance parameters. The algorithm consists out of the following steps:

1. Sample new nuisance parameter values from a prior e.g. $\theta_k \sim \text{Gaus}(\tilde{\theta}_k, \sigma_{\tilde{\theta}_k})$.

2. Compute a new response matrix $R(\boldsymbol{\theta})$ and consequently new expected values $\boldsymbol{\nu}(\boldsymbol{\theta})$.

3. Sample new data $n_i \sim \text{Pois}(\nu_i)$ with the newly calculated means $\boldsymbol{\nu}(\boldsymbol{\theta})$.

4. Repeat many times and use the set of evaluated estimators to calculate the sample covariance.

Note that instead of sampling auxiliary measurements $\tilde{\boldsymbol{\theta}}$ one samples parameter values $\boldsymbol{\theta}$ from a prior which introduces the Bayesian aspect of this treatment. However, we expect the methods proposed in Sec. 5.6.3 and 5.6.4 to be equivalent in certain scenarios. Frequentist-Bayesian hybrid methods have been used before for upper limit setting [221] but never within unfolding frameworks.

5.6.5 Bimodal Example

We divide a bimodal physics model into a two parts. The signal model is defined as

$$
\begin{aligned}
f^{\text{sig}}(x|\theta_1,\theta_2,\theta_3) &= f^{\text{sig}}_{\text{truth}}(x_{\text{true}}) * f_{\text{detector}}(x_{\text{true}}, x|\theta_1,\theta_2,\theta_3) \\
&= ((0.5 \cdot \text{Gauss}(x_{\text{true}}|\mu = 1.5, \sigma = 0.12) \\
&\quad + (0.5 \cdot \text{Gauss}(x_{\text{true}}|\mu = -1.5, \sigma = 0.12)) \\
&\quad * f_{\text{detector}}(x_{\text{true}}, x|\theta_1,\theta_2,\theta_3).
\end{aligned}
\tag{5.55}
$$

The detector response function $f_{\text{detector}}(x_{\text{true}}, x|\theta_1,\theta_2,\theta_3)$ is Gaussian smearing function but has an additional dependency on the nuisance parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$.

Figure 5.16: A plot of a) the input distributions on reconstructed level and b) corresponding response matrix of the bimodal model

These parameters are a scale factor applied to the reconstructed value x, a parameter determining the smearing and a parameter determining the efficiency, respectively. A background process is included modelled by uniform probability distribution function,

$$
\begin{aligned}
f^{\text{bkg}}(x|\theta_1,\theta_2,\theta_3) &= f^{\text{bkg}}_{\text{truth}}(x_{\text{true}}) * f_{\text{detector}}(x_{\text{true}}, x|\theta_1,\theta_2,\theta_3) \\
&= \text{Uniform}(x_{\text{true}}) * f_{\text{detector}}(x_{\text{true}}, x|\theta_1,\theta_2,\theta_3)
\end{aligned}
\tag{5.56}
$$

The filled truth and reconstructed histograms range between $[-4, 4]$ and both have constant bin size $\Delta x = 1.6$. The sampled toy data that is to be unfolded contains

113

$N_{sig} = 5000$ signal events and $N_{bkg} = 500$ background events. The sampled toy data for all the nuisance parameter variations used to estimated the response matrices each contain $N_{sig} = 10^8$.

5.6.6 Exponential Example

We divide also the exponential physics model into a two parts. The signal model is defined as

$$f^{\text{sig}}(x|\alpha, \theta_1, \theta_2, \theta_3) = f^{\text{sig}}_{\text{truth}}(x_{\text{true}}|\alpha) * f_{\text{detector}}(x_{\text{true}}, x|\theta_1, \theta_2, \theta_3)$$
$$= \alpha \exp[-\alpha x_{\text{true}}] * f_{\text{detector}}(x_{\text{true}}, x|\theta_1, \theta_2, \theta_3). \tag{5.57}$$

with $\alpha = 0.14$.
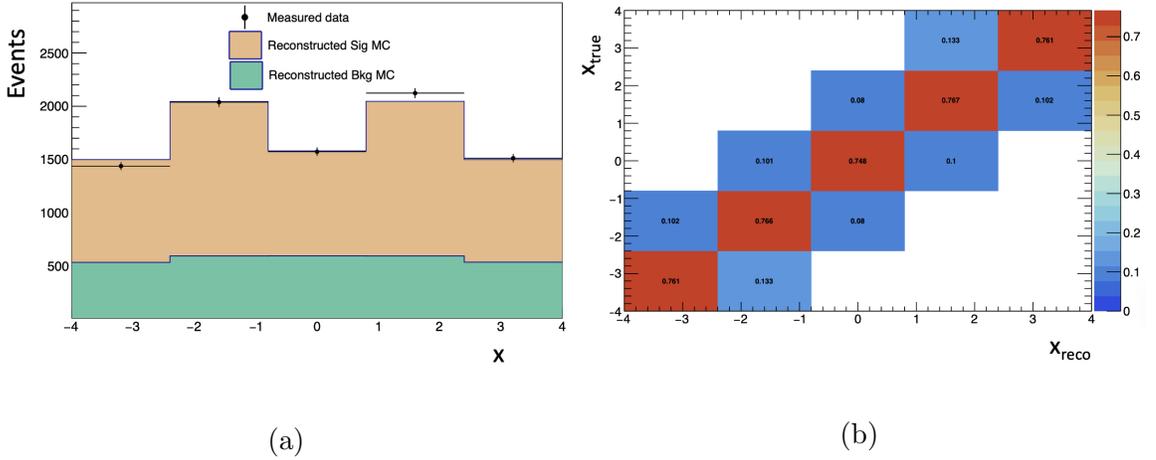

(a)   (b)

Figure 5.17: A plot of a) the input distributions on reconstructed level and b) corresponding response matrix of the exponential model

The background distribution is defined by another exponential model defined as

$$f^{\text{bkg}}(x|\alpha, \theta_1, \theta_2, \theta_3) = f^{\text{bkg}}_{\text{truth}}(x_{\text{true}}|\alpha) * f_{\text{detector}}(x_{\text{true}}, x|\theta_1, \theta_2, \theta_3)$$
$$= \alpha \exp[-\alpha x_{\text{true}}] * f_{\text{detector}}(x_{\text{true}}, x|\theta_1, \theta_2, \theta_3). \tag{5.58}$$

114

with $\alpha = 0.15$. The filled truth and reconstructed histograms range between $[0, 60]$ and variable bin widths with bin edges $x_{\text{true}} = \{0, 2, 4, 6, 8, 10, 12, 14, 18, 25, 35, 60\}$ and $x = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 25, 30, 35, 45, 60\}$. The sampled toy data that is to be unfolded contains $N_{sig} = 10000$ signal events and $N_{bkg} = 40000$ background events. The sampled toy data for all the nuisance parameter variations used to estimated the response matrices each contain $N_{sig} = 10^8$.

5.6.7 Covariance Estimates

The toy data is unfolded by maximizing the regularized full likelihood of Eq. (5.49) w.r.t. $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$ for different amounts of regularization $\tau \in \{0, 10^{-6}, 10^{-5}, 5 \times 10^{-5}\}$. For each unfolded distribution the covariance matrix was estimated once with each of the three methods with 10000 pseudo-experiments for the frequentist and frequentist-Bayes method. We would like to stress here that the inverse Hessian method is only valid to use for $\tau = 0$. The covariance matrix estimates for the bimodal and exponential model are plotted on Fig. 5.18 and 5.19, respectively. Additionally, the relative differences between the inverse hessian and frequentist toys and between the pseudo-hybrid and frequentist toys methods are shown in Fig. 5.20 and 5.21, respectively. Covariance matrix estimates can be further compared with summary statistics that are common in HEP such as the global correlation coefficient, the relative errors on the unfolded distribution and χ^2 as presented in Sec. 5.3. The data in each of these summary statistics is the unfolded toy data. Their comparison at the different regularisation strengths τ in Fig. 5.22, 5.23 and 5.24 make the (dis)agreements more obvious. We observe that for the bimodal distribution with no regularization all of the methods seem to agree reasonably. However, as expected, the inverse Hessian method diverges when regularization is introduced. The covariance estimates of the exponential distribution show a complete disagreement between the inverse hessian method and the other two methods. The inverse Hessian method also shows

a drop in covariance estimates for $\hat{\mu}_3$ in the exponential example. This is because of a low dependency of the likelihood on the parameter μ_3. The frequentist-Bayes hybrid pseudo-experiments method agrees well on all quantities with the frequentist pseudo-experiments method.

(a) Frequentist $\tau = 0$.   (b) Hybrid $\tau = 0$.   (c) Hessian $\tau = 0$.

(d) Frequentist $\tau = 10^{-6}$ (e) Hybrid $\tau = 10^{-6}$ (f) Hessian $\tau = 10^{-6}$

(g) Frequentist $\tau = 10^{-5}$ (h) Hybrid $\tau = 10^{-5}$ (i) Hessian $\tau = 10^{-5}$

(j) Frequentist $\tau = 5 \times 10^{-5}$ (k) Hybrid $\tau = 5 \times 10^{-5}$ (l) Hessian $\tau = 5 \times 10^{-5}$
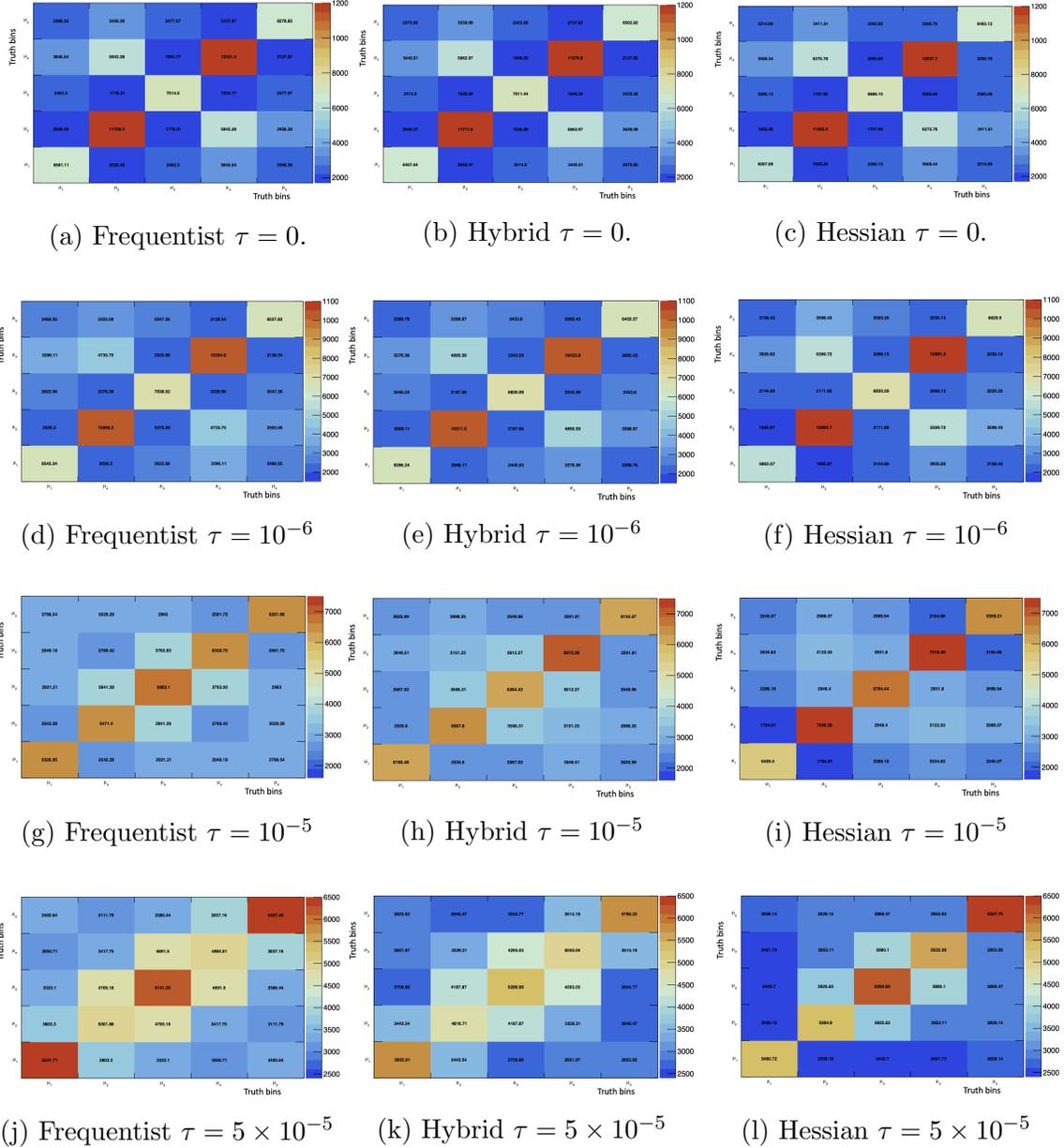
Figure 5.18: Covariance matrix for the bimodal distribution estimated with the frequentist pseudo-experiments, inverse hessian and frequentist-Bayes hybrid pseudo-experiments method for various regularization strengths τ

(a) Frequentist $\tau = 0$.   (b) Hybrid $\tau = 0$.   (c) Hessian $\tau = 0$.

(d) Frequentist $\tau = 10^{-6}$ (e) Hybrid $\tau = 10^{-6}$ (f) Hessian $\tau = 10^{-6}$

(g) Frequentist $\tau = 10^{-5}$ (h) Hybrid $\tau = 10^{-5}$ (i) Hessian $\tau = 10^{-5}$

(j) Frequentist $\tau = 5 \times 10^{-5}$ (k) Hybrid $\tau = 5 \times 10^{-5}$ (l) Hessian $\tau = 5 \times 10^{-5}$
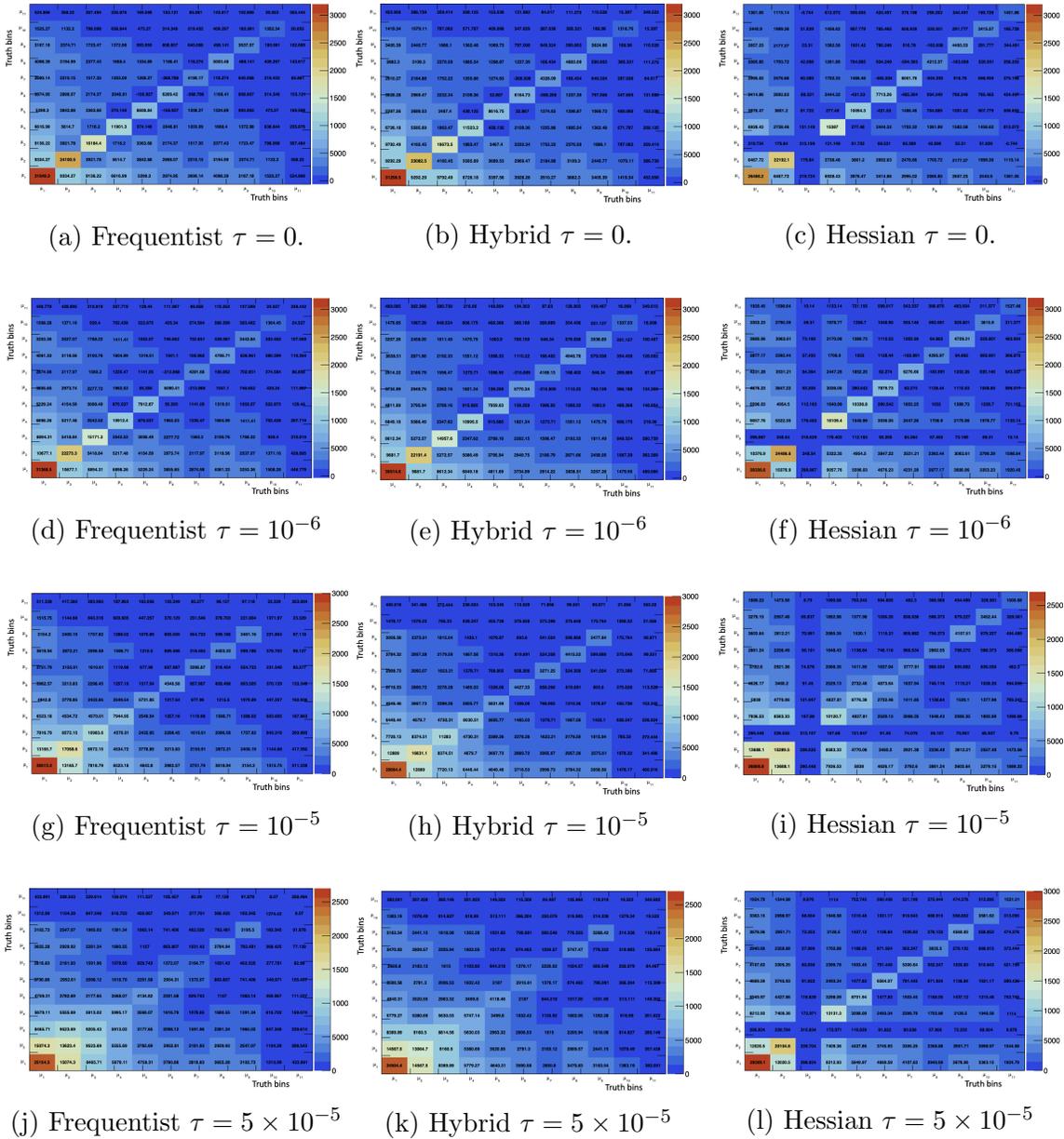
Figure 5.19: Covariance matrix for the exponential distribution estimated with the frequentist pseudo-experiments, inverse hessian and frequentist-Bayes hybrid pseudo-experiments method for various regularization strengths τ

117
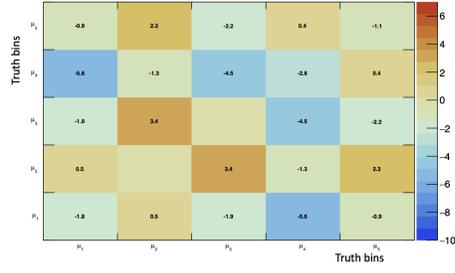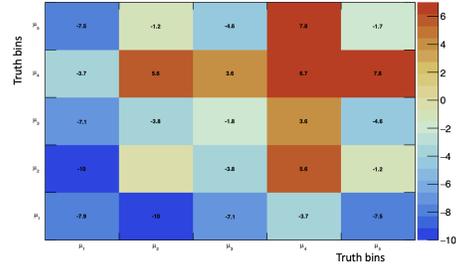
(a) Hybrid $\tau = 0$.

(b) Hessian $\tau = 0$.
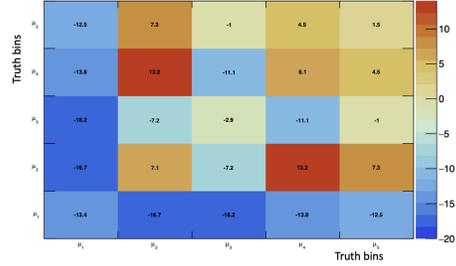
(c) Hybrid $\tau = 10^{-6}$
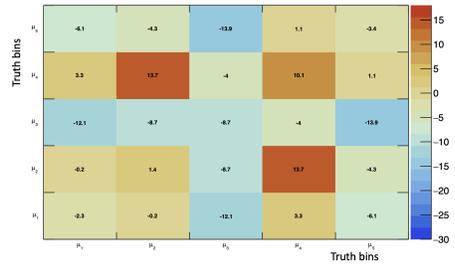
(d) Hessian $\tau = 10^{-6}$

(e) Hybrid $\tau = 10^{-5}$
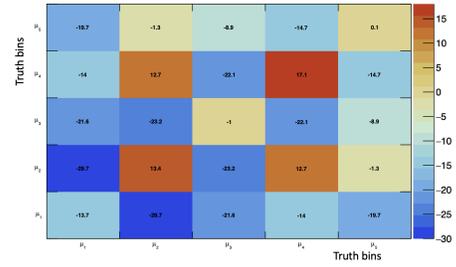
(f) Hessian $\tau = 10^{-5}$
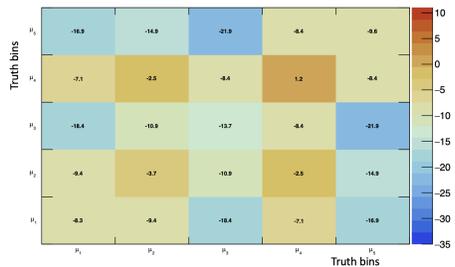
(g) Hybrid $\tau = 5 \times 10^{-5}$
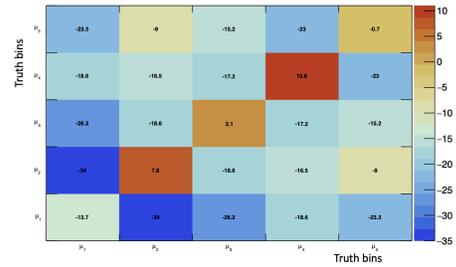
(h) Hessian $\tau = 5 \times 10^{-5}$

Figure 5.20: Covariance matrix differences in percentage for the bimodal distribution between the fully frequentist toy method and the inverse hessian method(1st column) and hybrid toy method(2nd column) for various regularization strengths τ
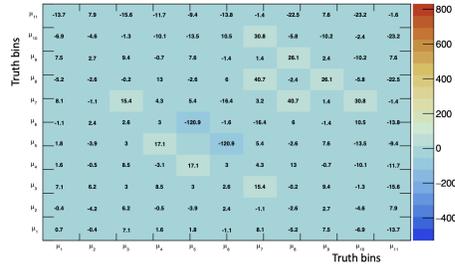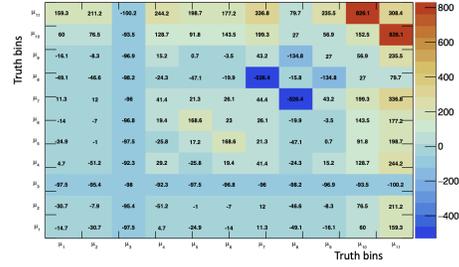
(a) Hybrid $\tau = 0$.
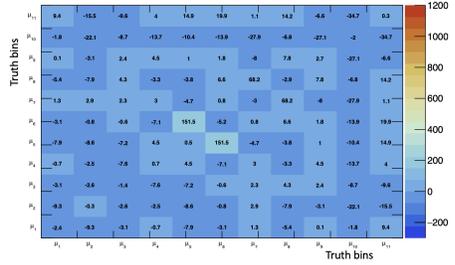
(b) Hessian $\tau = 0$.
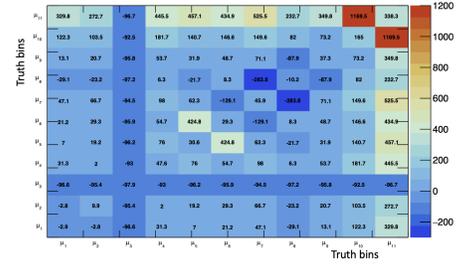
(c) Hybrid $\tau = 10^{-6}$
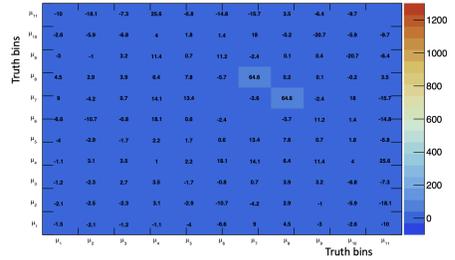
(d) Hessian $\tau = 10^{-6}$

(e) Hybrid $\tau = 10^{-5}$
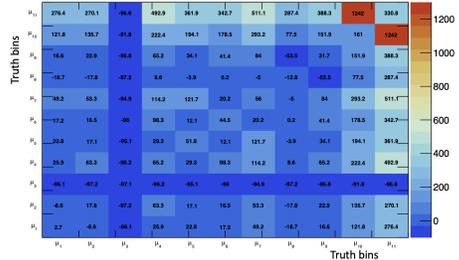
(f) Hessian $\tau = 10^{-5}$
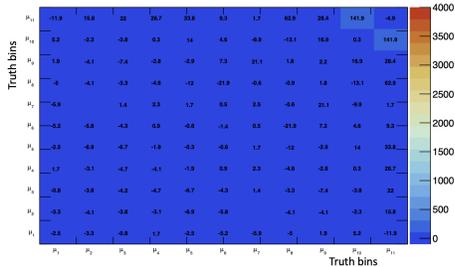
(g) Hybrid $\tau = 5 \times 10^{-5}$
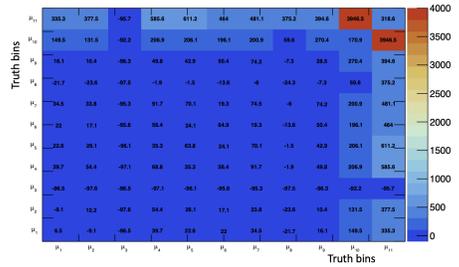
(h) Hessian $\tau = 5 \times 10^{-5}$

Figure 5.21: Covariance matrix differences in percentage for the exponential distribution between the fully frequentist toy method and the inverse hessian method(1st column) and hybrid toy method(2nd column) for various regularization strengths τ

(a) Double Gaussian distribution

(b) Exponential distribution
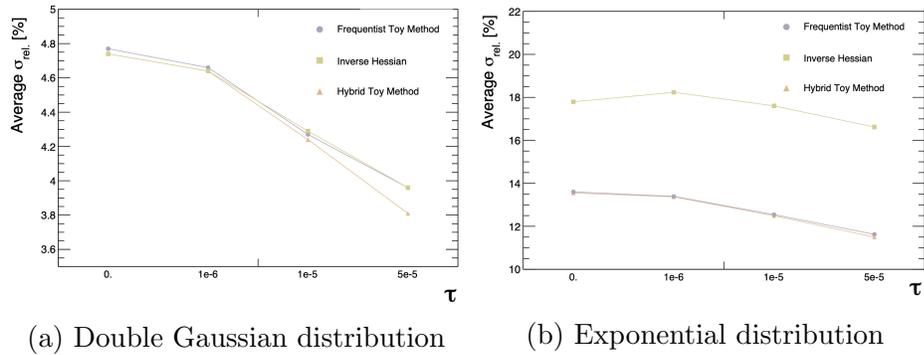
Figure 5.22: Average relative unfolding errors for a) the double Gaussian and b) the exponential distribution for all covariance estimation methods and various regularization strengths τ

(a) Bimodal distribution

(b) Exponential distribution
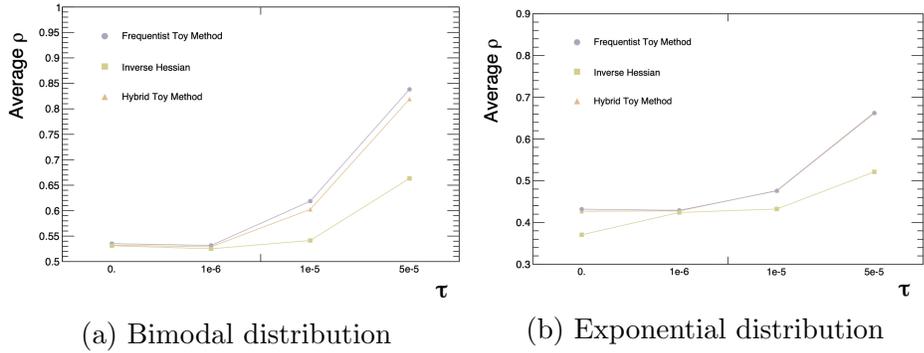
Figure 5.23: Average global correlation coefficient for a) the bimodal and b) the exponential distribution for all covariance estimation methods and various regularization strengths τ

(a) Bimodal distribution

(b) Exponential distribution
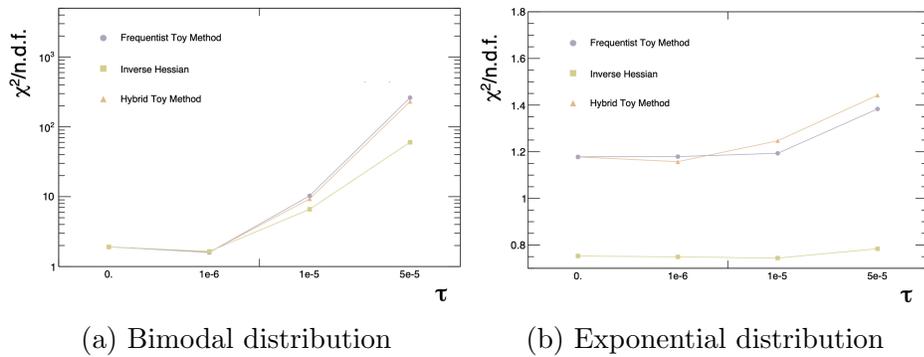
Figure 5.24: χ^2/n.d.f. for a) the bimodal and b) the exponential distribution for all covariance estimation methods and various regularization strengths τ

5.7 Discussion

There are many different unfolding algorithms available that provide a solution for the discretized inverse problem of which various are implemented in the `RooUnfold` software package. Many of them impose some form of regularization of which the regularization parameters determine the bias-variance trade-off of the unfolded distribution. Section 5.1 presents several figures of merit that can help the analyst to choose a regularization parameter that corresponds to an optimal bias-variance trade-off. Section 5.4 shows that a regularization parameter that minimizes the mean squared error can result in low coverage probabilities. One can increase the coverage closer to nominal coverage by reducing the amount of regularization w.r.t. the amount that minimizes the mean squared error. One should use these metrics not only to compare regularization parameters but also different unfolding algorithms. Different unfolding algorithms will introduce different forms of regularization and will therefore result in estimators with different biases, variances and coverage probabilities. There is not one solution that fits all unfolding scenarios which means that each analyst should carefully compare all the different options available.

Section 5.5 shows that low Monte Carlo statistics can have a detrimental effect on conventional response matrix estimation. An alternative likelihood-based estimation can give a smoother description of the response matrix that is less reliant on the Monte Carlo sample size. This can be very valuable for unfolding scenarios were Monte Carlo simulation is computationally very expensive. Possible future studies could involve exploring different distributions to construct the likelihood such as the Poisson or Student-t distribution.

Section 5.6 presents a novel technique for estimating the covariance matrix of the unfolded distribution including all statistical and systematic sources of error. The

frequentist-Bayes hybrid toy method gives similar results to the frequentist toy method for various amounts of regularization. The inverse hessian method deviates from both methods quite substantially, especially in the high regularization regime. The frequentist-Bayes hybrid toy method can be used with any of the `RooUnfold` unfolding algorithms, i.e. not just the ones with an explicit likelihood definition. Possible future studies could involve methods that estimate the covariances between separately unfolded distributions. A "bootstrap method" has been applied in previous ATLAS analyses [222, 223] in case one has access to each event of the measured dataset. However, in many physics analyses the published results only constitute the unfolded distribution and its covariance matrix. A first step in combining covariances would be to report the covariance contribution for each nuisance parameter. One could then estimate the covariances between the two unfolded distribution for each nuisance parameter to get a total covariance and sum each of these total covariances. However, the mathematical justification is yet to be developed. Also, it would need to be compared with covariances estimated with an unfolding framework that unfolds the two distributions as one, i.e. with the distributions laid end to end, to test the accuracy of the method.

Chapter 6

Measurement of $t\bar{t}$ Differential Cross Sections in the l+jets Decay Channel

In this chapter, the measurement of $t\bar{t}$ production cross sections as a function of kinematic variables such as $m^{t\bar{t}}$, $H_T^{t\bar{t}}$, $p_T^{t\bar{t}}$ and $|y^{t\bar{t}}|$, i.e. differential cross sections, is presented. This analysis focuses on the lepton+jets decay channel and presents the final results as differential cross sections in a fiducial phase space, i.e. fiducial differential cross sections, unfolded to particle level. Section 6.1 and 6.2 summarize the used data and Monte Carlo samples. Section 6.3 discusses the data-driven estimation of fake lepton background process. Section 6.4 discusses the object and event selection applied on the data and Monte Carlo samples. Section 6.5 summarize the chosen observables and Sec. 6.6 the agreement between data and Monte Carlo samples. Section 6.7 presents the unfolding setup including the fiducial differential cross sections unfolded to particle level. Section 6.9 closes with conclusions and some discussion points.

6.1 Data Samples

The data samples used in this measurement were collected by the ATLAS detector at the LHC from pp collisions during the full Run-2 running period from 2015 to 2018. The LHC operated at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ and the data amounts to a total integrated luminosity of $139 \pm 2.4\,\mathrm{fb}^{-1}$ [224]. All data samples were acquired as described in Ch. 3 and reconstructed as described in Ch. 4.

6.2 Monte Carlo Samples

Monte Carlo samples are needed to simulate both the $t\bar{t}$ signal process as well as all the background processes. The QFT and phenomenological principles used to define input models for the Monte Carlo simulations are presented in Ch. 2. The detector and detector simulation is described in Ch. 3. Here we simply list the samples, the software used to produce them and important input parameters of the used models. A summary of all samples is given in Table 6.1.

Pile-up events from the same and neighbouring bunch crossings were modelled with PYTHIA 8.186 with its parameters set to the values of the ATLAS A3 [225] tune set and the `NNPDF2.3 LO` PDF set [226].

6.2.1 $t\bar{t}$ Signal Process

The signal sample comprises $t\bar{t}$ events that where generated using matrix element calculation at NLO QCD as implemented in POWHEG-BOX v2 with the `NNPDF3.0NLO` PDF [227]. The parton shower, hadronization and underlying events were simulated with PYTHIA 8.230 [114, 112] with the `NNPDF23LO` PDF [226] with parameters set to the ATLAS A14 tune set [228]. The top quark mass m_t and width were set to $172.5\,\mathrm{GeV}$ and $1.32\,\mathrm{GeV}$ [229], respectively. The sample was normalized with the

cross section $\sigma_{t\bar{t}} = 832^{20}_{-30} \pm 35(\text{PDF}, \alpha_s) \pm 23(m_t)$ calculated with the TOP++2.0 program [230] at NNLO QCD [229], including soft-gluon resummation to next-to-next-to leading-log order (NNLL) [231]. This MC sample will be denoted from here on as 'Pwg+Py8'.

6.2.2 Background Processes

There are several processes that can produce the same final state as the $t\bar{t}$ in the l+jets decay channel. The background contributions need to be estimated either with Monte Carlo simulation or other data-driven methods. The background contributions can then be subtracted from the observed data such that the signal process remains. Here we will list the background processes estimated with Monte Carlo simulations.

Single Top

Single top events, as illustrated by Table 6.4, is the biggest background contribution relative to $t\bar{t}$ events and consists of three background processes: t-channel, s-channel and W-associated production. Figure 6.1 shows leading order Feynman diagrams of the single top processes.
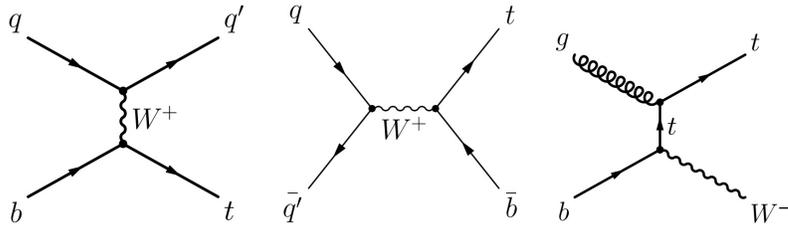
Figure 6.1: Tree level diagrams of the electroweak single top quark production in the t-channel(left), s-channel(centre) and the W-associated production(right), commonly denoted as tW. From [232].

The matrix elements of the three background processes were calculated at NLO QCD using POWHEG-BOX v2. The parton shower and hadronization were simulated with PYTHIA 8.230 with parameters set to the A14 tune set with the NNPDF2.3 LO PDF

set. The Feynman diagrams of tW at NLO QCD interfere with $t\bar{t}$ diagrams i.e. accurate simulation for both would include this contribution to the matrix element calculations as well. Instead, because the two processes are simulated separately, other modelling techniques need to be deployed. For these samples the Diagram Removal approach is used to correct for this [233, 234]. A more detailed description of the problem and other modelling approaches can be found in [235].

W+jets and Z+jets

The matrix element of a single $W^{\pm}$ or Z^0 boson production in association with jets was calculated at NLO QCD with Sherpa 2.2.1 for up to 2 jets and at LO QCD with Comix and OpenLoops for up to 4 jets. The parton shower and hadronization was modelled with Sherpa and the NNPDF3.0NNLO PDF set. The W+jets and Z+jets samples were normalized with cross sections calculated at NNLO QCD [236, 237].

Diboson

The diboson background denotes the production of two $W^{\pm}$ or two Z^0 bosons where one of the bosons decays leptonically and the other hadronically. Their matrix elements were calculated at NLO QCD with Sherpa 2.2.2 with the NNPDF3.0NNLO PDF set. The parton shower and hadronization was modelled by Sherpa.

$t\bar{t}V$

The $t\bar{t}V$ background comprises the production of $t\bar{t}$ in association with a $W^{\pm}$ or Z^0 boson. Figure 6.2 shows corresponding leading order diagrams. The matrix element calculation was performed at NLO QCD with MadGraph5_aMCNLO 2.3.3 and the NNPDF3.0NLO PDF set. The parton shower and hadronization is modelled with Pythia 8.230 with parameters set to the A14 tune set and using the NNPDF2.3LO PDF set.
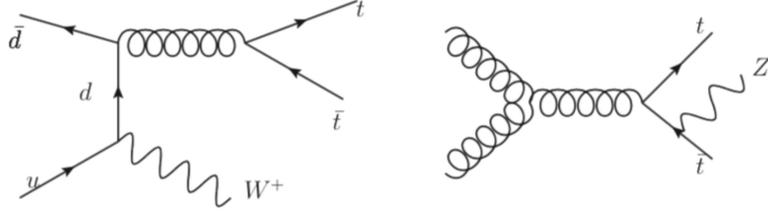
Figure 6.2: Feynman diagrams for the $t\bar{t}W$(left) and $t\bar{t}Z$(right) production processes. From [92]

Physics Process	ME + Event Generation	ME PDF set	PS + Hadr.
$t\bar{t}$	Powheg-Box v2	NNPDF3.0NLO	Pythia 8.230
Single Top	Powheg-Box v2	NNPDF3.0NLO	Pythia 8.230
W+jets and Z+jets	Sherpa 2.2.1 + Comix + OpenLoops	NNPDF3.0NNLO	Sherpa
Diboson	Sherpa 2.2.2	NNPDF3.0NNLO	Sherpa
$t\bar{t}V$	MadGraph5_aMCNLO 2.3.3	NNPDF2.3LO	Pythia 8.230

Table 6.1: Summary of the simulated Monte Carlo samples for all signal and background processes.

Other background processes such as $t\bar{t}t\bar{t}$, tZ, $t\bar{t}WW$, $t\bar{t}WZ$ and $t\bar{t}H$ were considered but their background to signal ratio were deemed negligible.

6.3 Fake Lepton Estimation

The fake lepton background, denoted as simply *fakes*, comprises events with non-prompt leptons and jets or photons that "fake" a lepton. These various different origins are hard to estimate with Monte Carlo simulations and are therefore instead estimated with a data-driven method known as the matrix element method [238, 239]. The method starts of by taking a dataset where each event has exactly one lepton. Events where the lepton pass the *Tight* likelihood-selection criterion, as introduced in Ch. 4, are denoted with N^T and events passing the *Loose* criterion with N^L. With these we can define the linear combinations

$$N^L = N_r^L + N_f^L \tag{6.1}$$

127

$$N^T = \epsilon_r N_r^L + \epsilon_f N_f^L \tag{6.2}$$

where N_r^L and N_f^L are the leptons that pass the loose criterion and are in reality a resp. real lepton or a non-prompt/fake lepton. The coefficients ϵ_r and ϵ_f denote the fraction of events with leptons that pass the loose but also the tight criterion. The quantity we want to estimate are the number of events that pass the tight criterion but are fakes $N_f^T = \epsilon_f N_f^L$. One can solve the set of linear equations 6.1 and 6.2 to get the estimator

$$N_f^T = \frac{\epsilon_f}{\epsilon_r - \epsilon_f}(\epsilon_r N^L - N^T). \tag{6.3}$$

From this an event weight is computed for each data event passing the tight selection to correct for the number of fakes in the tight event selection,

$$w_i = \frac{\epsilon_f}{\epsilon_r - \epsilon_f}(\epsilon_r - \delta_{i \in N^T}). \tag{6.4}$$

with $\delta_{i \in N^T}$ being 1 if the event passes the tight criterion and otherwise 0. The lepton efficiencies ϵ_r and ϵ_f can be estimated with data from control regions, i.e. regions outside the signal region defined by the object and event selection. These regions have either a large contribution of real leptons or non-prompt/fake leptons. The real lepton efficiency ϵ_r is estimated with events containing $Z \to ll$ using a tag and probe method. If one of the leptons passes the loose requirement (tag) then the other lepton is subjected to the tight requirement (probe). The ratio of probes passing the tight requirement and tags passing loose gives an estimate for ϵ_r. The fake efficiency ϵ_f is estimated with events from a control region with many non-prompt/fake leptons. The real leptons in this region are estimated with Monte Carlo simulation and subtracted. The ratio between remaining events with a lepton passing the tight criterion and remaining events with a lepton passing the loose criterion is used as an estimate for ϵ_f. The final estimation of the non-prompt/fake lepton background in bin j is the

sum of weights defined in Eq. 6.4 for the data events in bin j, i.e.

$$\beta_{\text{fake},j} = \sum_{i}^{\text{events in bin } j} w_i \tag{6.5}$$

6.4 Object and Event Selection

This section describes the object and event selection criteria constructed to increase the yield of $t\bar{t}$ to e/μ+jets events relative to the background processes in the measured data. Here we define *detector level* and *particle level* object and event selection criteria were the detector level criteria are imposed on the detector level Monte Carlo and data samples. The trigger requirements of the samples are summarized in Table 6.2. The particle level selection is imposed on the Monte Carlo samples that are produced up until detector simulation. The selection on MC samples are applied such that the data, both on detector level and after unfolding to particle level, can be compared properly with theory.

6.4.1 Detector level

Here we summarize the object and event selection criteria applied to detector level Monte Carlo and data samples.

- The vertex with at least 2 reconstructed tracks and the highest sum of $\sum p_T$ over those tracks with $\sum p_T > 0.4$ is chosen to be the primary vertex.

- The electrons must pass the *Tight* likelihood-criterion, as introduced in Ch. 4. This ensures a 80% *prompt electron* identification efficiency.

- The electron must have a transverse momentum of $p_T > 27\,\text{GeV}$.

- The track of the electron must have a pseudo rapidity of $|\eta| < 2.47$ and $1.37 < |\eta| < 1.52$ were the latter is because of the gap between the barrel and end-cap

calorimeters.

- The longitudinal impact parameter z_0, the distance between the lepton track and the beam spot in the longitudinal plane, has to satisfy $|\Delta z_0 \sin \theta| < 0.5 \, \mathrm{mm}$, where θ is the polar angle of the track.

- The transverse impact parameter d_0, as introduced in Ch. 4, of the lepton must satisfy $|d_0/\sigma(d_0)| < 5$, with σ_{d_0} being the uncertainty on d_0.

- The track of the muon must have a pseudo rapidity of $|\eta| < 2.5$ and correspond to a transverse momentum of $p_T > 27 \, \mathrm{GeV}$.

- The *FCTight* isolation working points were used for both the muon and the electron.

- The electron isolation requirements are $E^T_{topo}(R = 0.2)/p_T < 0.06$ and $p_T^{TightTTVA}(R = 0.2)/p_T < 0.06$, where E^T_{topo} is the energy of the calorimeter topo-cluster and $p_T^{TightTTVA}$ the pile-up robust track isolation variable.

- The muon isolation requirements are $E^T_{topo}(R = 0.2)/p_T < 0.15$ and $p_T^{TightTTVA}(R = 0.2)/p_T < 0.04$.

- If a muon shares a track with an electron the electron is not selected.

- Jets are corrected for pile-up effects with the jet-area [183] method.

- The jet-vertex tagging [185] method, as described in Ch. 4, is applied to jets satisfying $p_T < 60 \, \mathrm{GeV}$ and $|y| < 2.4$ to reduce the number of jets originating from pile-up.

- All jets must satisfy $p_T > 25 \, \mathrm{GeV}$ and $|y| < 2.5$. The p_T cut supresses the non-$t\bar{t}$ background and reduces the uncertainty of the JES.

- For the b-tagging algorithm DL1r the working point of 70% was chosen, as introduced in Ch. 4.

- Jets need to have sufficient distance from electrons i.e. $\Delta R(\text{jet}, e) \geq 0.2$ to reduce the number of jets faking electrons.

- Jets need to have sufficient distance from muons, i.e. $\Delta R(\text{jet}, \mu) \geq 0.4$. If not, the jet is removed if it has less then three tracks and otherwise the muon is removed. This is to reduce the number of jets faking a muon.

- Exactly one electron or muon passed the above requirements is required. This is to exclude $t\bar{t}$ production in the dilepton decay channel.

- At least 4 jets passing the above requirements is required, i.e. 2 jets from the hadronically decaying top quark and 2 for each b quark from the decayed top quarks.

- At least 2 of those jets need to be b-tagged.

6.4.2 Particle level

Here we summarize the object and event selection criteria applied to the particle level Monte Carlo samples.

- Leptons (e or μ) are required not to originate from a hadron.

- The four-momenta of the *bare* lepton, i.e. after final state radiation, and all photons within a cone of $\Delta R = 0.1$ are summed to correct for *bremsstrahlung*. The resulting four-momentum represents what is known as a *dressed* lepton.

- The dressed lepton must satisfy $p_T > 27\,\text{GeV}$ and $|\eta| < 2.5$.

- Particle level jets are also reconstructed with the anti-k_t algorithm with $R = 0.4$ using particle level particles except for leptons and neutrinos not originating from a hadron.

- A particle level jet is b-tagged if a hadron with $p_T > 5\,\text{GeV}$ is matched to the jet with the ghost matching technique [183].

- Electrons and muons within a distance of $\Delta R < 0.4$ of jets are removed.

- Exactly one electron or muon passed the above requirements is required.

- At least 4 jets passing the above requirements is required.

- At least 2 of those jets need to be b-tagged.

Running Period	2015	2016 - 2018
Electrons	HLT_e24_lhmedium_L1EM20VH .OR. HLT_e60_lhmedium .OR. HLT_e120_lhloose	HLT_e26_lhtight_nod0_ivarloose .OR. HLT_e60_lhmedium_nod0 .OR. HLT_e140_lhloose_nod0
Muons	_HLT_mu20_iloose_L1MU15 .OR. HLT_mu50	HLT_mu26_ivarmedium .OR. HLT_mu50

Table 6.2: Trigger configuration for the different running periods and type of lepton. A logical OR of the configurations listed was used.

6.5 Observables

This section summarizes the chosen observables of which the differential cross sections are measured and unfolded to particle level. The observables were chosen based on past top quark pair differential cross section measurements [240] and their sensitivity to SMEFT Wilson coefficients. The four momentum of the top quark pair is reconstructed by summing the four momenta of the hadronic and leptonic decaying top quarks reconstructed with the pseudo-top quark algorithm. The invariant mass is defined as

$$m = \sqrt{E^2 - ||p||^2} \qquad (6.6)$$

Observable	Description
p_T^{tt}	Transverse momentum of the $t\bar{t}$ system
$m^{t\bar{t}}$	Invariant mass of the $t\bar{t}$ system
$|y^{t\bar{t}}|$	Absolute rapidity of the $t\bar{t}$ system
$H_T^{t\bar{t}}$	The scalar sum of the transverse momenta of the two top quarks

Table 6.3: Kinematic observables for which differential cross sections are measured.

(a) $p_T^{t\bar{t}}$

(b) $m^{t\bar{t}}$

(c) $H_T^{t\bar{t}}$
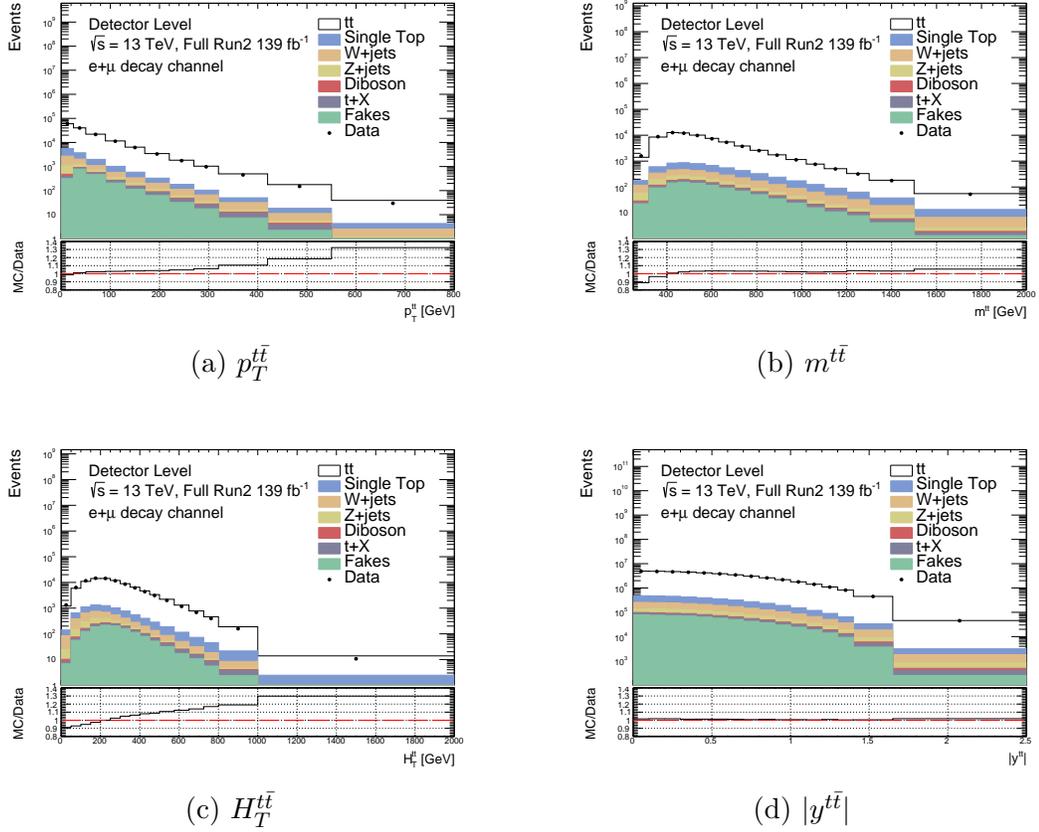
(d) $|y^{t\bar{t}}|$

Figure 6.3: The data and Monte Carlo agreement for the total rate, $p_T^{t\bar{t}}$, $H_T^{t\bar{t}}$, $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$.

6.6 Data-MC agreement

This section summarizes the agreement between data and Monte Carlo simulation on detector level in the fiducial phase space defined by the kinematic cuts introduced in Sec. 6.4. Figure 6.3 shows the events as a function of the kinematic observables introduced in Sec. 6.5. Table 6.4 shows the total event yield of the data and Monte Carlo for each physics process. We see a slight overestimation of the total rate of 1% and the absolute rapidity $|y^{t\bar{t}}|$. The transverse momenta show a slight underestimation in the low p_T-region and an overestimation in the high p_T-region.

Process	Event Yield	MC Ratio [%]
$t\bar{t}$	4095230	90.789
Single Top	193658	4.260
$W+$ jets	102069	2.262
Fakes	63908	1.455
$Z+$ jets	36338	0.812
$t + X$	14658	0.289
Diboson	5925	0.131
Total MC	4510710	100
Data	4451760	98.693

Table 6.4: The total event yield of the data and Monte Carlo on detector level in fiducial phase space for each physics process together with their ratio w.r.t. the total MC event yield.

6.7 Unfolded Results

In this section we summarize the unfolding framework setup used to unfold the data distributions to particle level. This includes the unfolding algorithms used to unfold, training and testing split in the Monte Carlo simulated $t\bar{t}$ samples, response matrix estimation, fine-tuning of regularization parameters based on MSE and coverage, closure and stress test on Monte Carlo data and finally the unfolded data distributions and covariances.

6.7.1 Unfolding Setup

In this analysis four unfolding algorithms were used as implemented in `RooUnfold` and discussed in Ch. 5. Here we introduce their short-hand notation used throughout the rest of the section and some additional details:

- **Invert** - The unregularized matrix inversion method as introduced in Sec. 5.1.

- **GP** - The Gaussian Processes unfolding method with a radial kernel function as introduced in Sec. 5.2.5. The kernel regularization parameters are optimized by maximizing the marginal likelihood as proposed in [207].

- **Poisson** - Maximizing the linear combination of a likelihood and a Tikhonov regularization function of Eq. 5.49 as introduced in Sec. 5.2.2. The likelihood is constructed from Poisson probability distribution functions for each measured bin and uses the discretized Tikhonov function of Eq. 5.22.

- **Bayes** - The Richardson-Lucy method also known as Iterative Bayes as introduced in Sec. 5.2.3.

Only the latter two unfolding algorithms have tunable regularization parameters as the first is unregularized and the second has an automated data-driven method that sets its regularization parameters. The binning of detector-level and particle level distributions were taken the same as in [240]. The efficiencies were included in the response matrix estimation.

The signal sample of $t\bar{t}$ events was split in two even-sized subsamples which, after applying the object and event selection criterions introduced in the previous section, each contain approximately 15×10^6 simulated events. One half, the training sample, is used to estimate the response matrix. The other half, the testing sample, is used to create a pseudo dataset to test the unfolding framework with. A dataset with values set to their expactation values and with negligble statistical fluctuations is also known
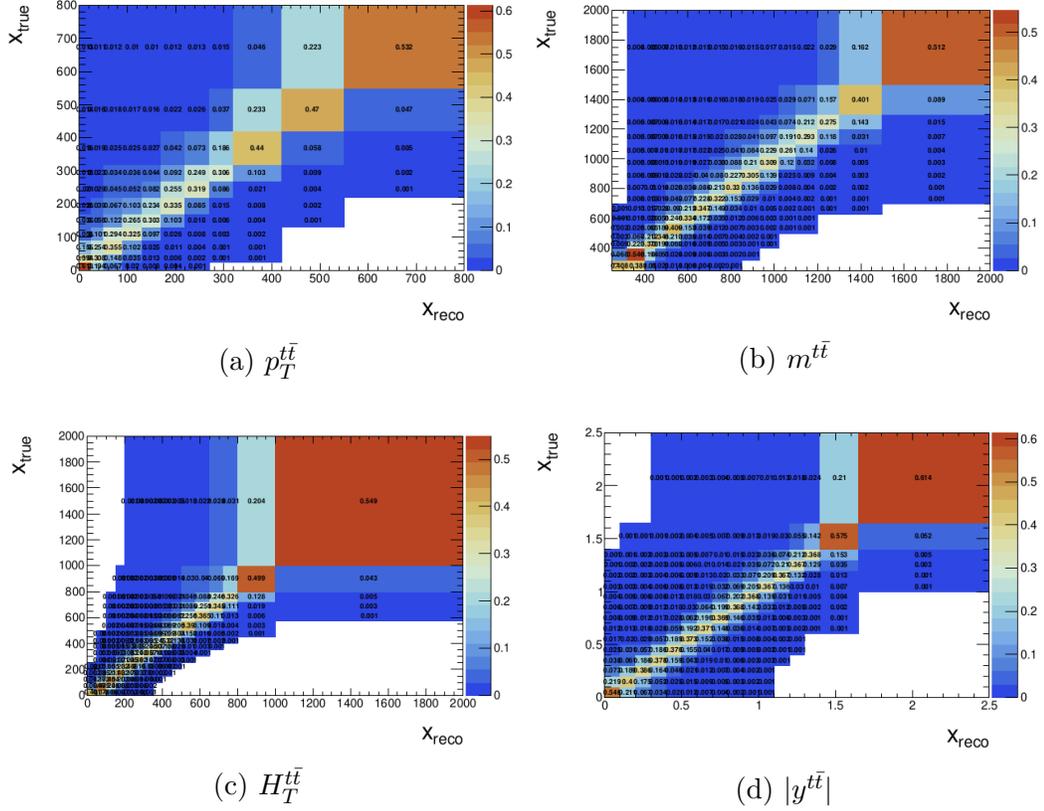
Figure 6.4: Estimated response matrices for $p_T^{t\bar{t}}$, $H_T^{t\bar{t}}$, $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$.

as *Asimov dataset*. One could label the pseudo dataset as such as the largest relative statistical error of all the observables is still below 1%. The truth events of the testing subsample are also used to compare the unfolded pseudo data with.

The response matrices of all the observables were estimated using the conventional method as discussed in Sec. 5.5.1. The reason for choosing this method over the novel method is that the Monte Carlo simulated $t\bar{t}$ sample used to estimate the response matrix, i.e. the training sample, contains more than 15×10^6 simulated events. However, any uncertainties introduced by the limited size of these MC samples are accounted for as will be discussed in the upcoming sections. Figure 6.4 shows the response matrix estimates for all the kinematic observables. Figure 6.5 shows the response matrix elements for certain truth bin slices. Many of the slices have a Gaussian-like shape as expected. However, the $p_T^{t\bar{t}}$ and $m^{t\bar{t}}$ show large non-Gaussian
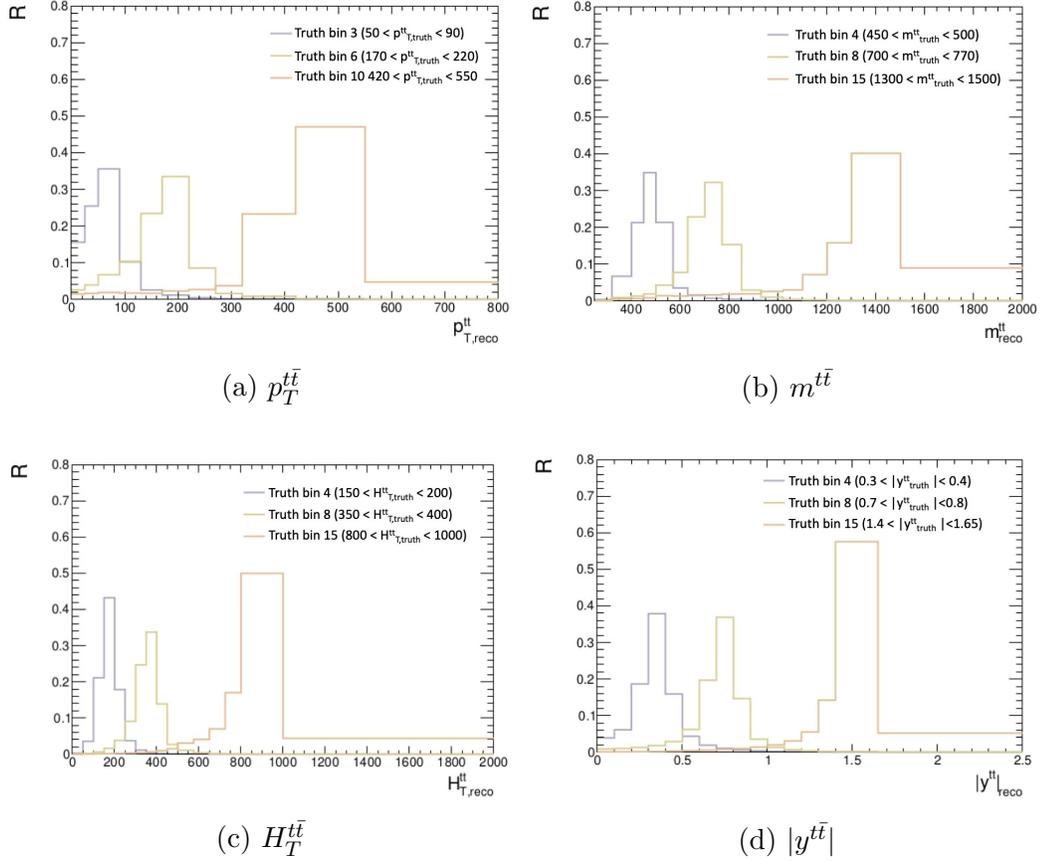
Figure 6.5: Estimated response matrix elements for truth bin slices for $p_T^{t\bar{t}}$, $H_T^{t\bar{t}}$, $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$.

tails on the left side of the peaks for the slices of truth bins 10 and 15, respectively. This should be taken into account if one would parameterize the response matrix with probability distributions as proposed in Sec. 5.5.

6.7.2 Systematic Uncertainties

This section summarizes the sources of systematic uncertainty included in this thesis and how they are propagated onto the unfolded distribution. The covariances, and therefore also the total errors on the unfolded distribution, are estimated with the frequentist-Bayes hybrid pseudo-experiment method as introduced in Sec. 5.6.4. Systematic uncertainties involving both signal and background processes were considered. Here we make the distinction between three types of nuisance parameters

that are used in modelling of the systematic uncertainties.

Detector

Detector nuisance parameters $\boldsymbol{\theta}$ are defined as all parameters that would affect events on detector level but not particle level. The nuisance parameters $\boldsymbol{\theta}$ follow a Gaussian probability distribution function $\theta_i \sim \mathrm{Gaus}(\tilde{\theta}_i, \sigma_{\tilde{\theta}_i})$ with its mean $\tilde{\theta}_i$ and standard deviation $\sigma_{\tilde{\theta}_i})$ set to their best known values, e.g. supplied by experts or calibration studies. Note that we are taking a Bayesian approach here instead of a more common frequentist approach, i.e. we are sampling nuisance parameters $\boldsymbol{\theta}$ instead of auxiliary measurements $\tilde{\boldsymbol{\theta}}$, just as described in Sec. 5.6.4. These nuisance parameters also model the systematic uncertainties from the background contributions. A summary of all detector related nuisance parameters included in this thesis are given in Table 6.5.

Response Matrix MC Stat.

Each transfer matrix bin has a nuisance parameter γ_i assigned that follows a Poisson probability distribution function $\gamma_i \sim \mathrm{Pois}(\tilde{\gamma}_i)$. The mean of the Poisson p.d.f. $\tilde{\gamma}_i$ is set to Monte Carlo event count in that bin. One can sample new transfer and thus response matrices from these p.d.f.s.

Theory

Theory nuisance parameters are defined as all parameters that would affect events on particle level and therefore subsequently on detector level. No theory nuisance parameters were included in this study. However, it is possible to model them in a similar way as the detector nuisance parameters.

Parameterizing the Response Matrix and Background

The procedure described in 5.6.4 depends on the parameterization of the response matrix R and the background $\boldsymbol{\beta}$ in terms of $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$. Let us define the nuisance parameters of all categories with the vector $\boldsymbol{\alpha} = (\boldsymbol{\theta}, \boldsymbol{\gamma})$. Each bin is parameterized by a piecewise linear interpolation function fitted to values estimated with Monte Carlo simulation samples generated at varying points in nuisance parameter space. A common approach is to generate one MC sample with all the nuisance parameters set to their *nominal* value $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$ including two MC samples per nuisance parameter with their parameter set to their *up* $\alpha_{i,\text{up}} = \tilde{\alpha}_i + \sigma_{\tilde{\alpha}_i}$ and *down* value $\alpha_{i,\text{down}} = \tilde{\alpha}_i - \sigma_{\tilde{\alpha}_i}$. For K nuisance parameters in $\boldsymbol{\theta}$ and L nuisance parameters in $\boldsymbol{\gamma}$ the response matrix and background piece-wise linear interpolation functions are defined as,

$$R_{ij}(\boldsymbol{\alpha}) = R_{ij}(\boldsymbol{\alpha}_{\text{nom}}) + \sum_{l=1}^{K+L} R_{ijl}(\boldsymbol{\alpha}) \tag{6.7}$$

$$\beta_i(\boldsymbol{\theta}) = \beta_i(\boldsymbol{\theta}_{nom}) + \sum_{k=1}^{K} \beta_{i,k}(\boldsymbol{\theta}) \tag{6.8}$$

with

$$R_{ijl}(\boldsymbol{\alpha}) = \begin{cases} \alpha(R_{ij}(\alpha_{1,\text{nom}}, ..., \alpha_{l,\text{up}}, ..., \alpha_{K+L,\text{nom}}) - R_{ij}(\boldsymbol{\alpha}_{\text{nom}})), & \text{for } \alpha_l \geq \alpha_{l,\text{nom}} \\ \alpha(R_{ij}(\boldsymbol{\alpha}_{\text{nom}}) - R_{ij}(\alpha_{1,\text{nom}}, ..., \alpha_{l,\text{down}}, ..., \alpha_{K+L,\text{nom}})), & \text{for } \alpha_l < \alpha_{l,\text{nom}} \end{cases} \tag{6.9}$$

$$\beta_{ik}(\boldsymbol{\theta}) = \begin{cases} \theta(\beta_i(\theta_{1,\text{nom}}, ..., \theta_{k,\text{up}}, ..., \theta_{K,nom}) - \beta_i(\boldsymbol{\theta}_{\text{nom}})), & \text{for } \theta_k \geq \theta_{k,\text{nom}} \\ \theta(\beta_i(\boldsymbol{\theta}_{\text{nom}}) - \beta_i(\theta_{1,\text{nom}}, ..., \theta_{k,\text{down}}, ..., \theta_{K,\text{nom}})), & \text{for } \theta_k < \theta_{k,\text{nom}} \end{cases} \tag{6.10}$$

Note that the statistical uncertainties of the background contributions are not included. Lets assume we filled bin j of some background histogram with events with each a weight w_i. One can estimate the statistical uncertainty with the sum of the

squares of the weights,

$$\sigma_{\text{bin } j}^2 = \sum_i^{\text{events in bin } j} w_i^2.$$ (6.11)

Even for the single top background, the largest background relative to the $t\bar{t}$ signal, the largest $\sigma_{\text{bin } j}$ relative to the signal $t\bar{t}$ contribution is below 1%. They are therefore assumed negligible and not modelled with nuisance parameters. The parameterizations of the response matrix and backgrounds as described above are implemented `RooUnfold` together with the frequentist-Bayes hybrid pseudo experiment method described in Sec. 5.6.4. One only needs to supply the transfer matrices and the background MC estimates for the different points in nuisance parameter space. The MC samples to fill the transfer matrices and backgrounds are generated with the same Monte Carlo simulation setup as described in Sec. 6.2 with the exception that the detector simulation has been performed with ATLFAST-II.

| Category	#NPs	Description
B-tagging	19	Parameters from the DL1r b-tagging algorithm as introduced in Sec. 4.4. Their nominal values and standard deviations account for the disagreement between simulation and data for b-, c- and light-flavour jet tagging.
Jet Energy Scale & Resolution	38	Parameters from the jet energy scale and resolution calibration performed on test beam data and simulations. This includes correcting for effects such as pile-up and jet-flavour composition.
Missing Transverse Energy	3	Correction for miscalibration of soft-track components of the missing transverse energy.
Electron Energy Scale & Resolution	3	Parameters from the electron energy scale and resolution calibration to correct for disagreement between simulation and data.
Muon Energy Scale & Resolution	5	Parameters from the muon energy scale and resolution calibration to correct for disagreement between simulation and data.
Electron Reconstruction & Identification	4	Parameters correcting for the disagreement between simulation and data for electron reconstruction, identification, isolation and trigger performance.
Muon Reconstruction & Identification	10	Parameters correcting for the disagreement between simulation and data for muon reconstruction, identification, isolation and trigger performance.

Table 6.5: Summary of detector related $\boldsymbol{\theta}$ nuisance parameters divided into categories [241]. A full list of included nuisance parameters can be found in App. B.1.

6.7.3   MSE & Coverage Scan

To choose the regularization parameter for the regularized unfolding methods, i.e. Poisson and Bayes, the same approach is used as presented in Sec. 5.4. The MSE is minimized but under the condition of sufficient coverage. The regularization parameters are the τ parameter in Eq. 5.49 for Poisson unfolding and the number of iterations r in Eq. 5.23 for Bayes unfolding. The minimum coverage is set to $P_{\mathrm{cov}} = P_{\mathrm{nom}} - \epsilon$ with P_{nom} being the nominal coverage probability of 0.683 and the tolerance ϵ set to 0.002. The regularization parameter that satisfies this condition and also minimizes the MSE is chosen as optimum [208, 215]. The estimated variance used to calculate the MSE uses the frequentist-Bayes hybrid pseudo experiment method and includes all the above mentioned systematic uncertainties. Figure 6.6 and 6.7 shows the MSE including its decomposition in bias and variance and the coverage probability for varying regularization parameters for all the observables. Notice that high values for r result in no regularization whilst high values for τ result in high regularized solutions. Table 6.6 gives the optimized regularization parameters for both unfolding algorithms and all the observables.

Observable	$r_{\mathrm{opt.}}$	$\tau_{\mathrm{opt.}}$
$p_T^{t\bar{t}}$	1	0.01
$m^{t\bar{t}}$	11	5
$H_T^{t\bar{t}}$	100	10
$|y^{t\bar{t}}|$	5000	6×10^{-13}

Table 6.6: Regularization parameters optimized w.r.t. coverage and MSE for both the Bayes($r_{\mathrm{opt.}}$) and Poisson($\tau_{\mathrm{opt.}}$) unfolding methods for all observables.
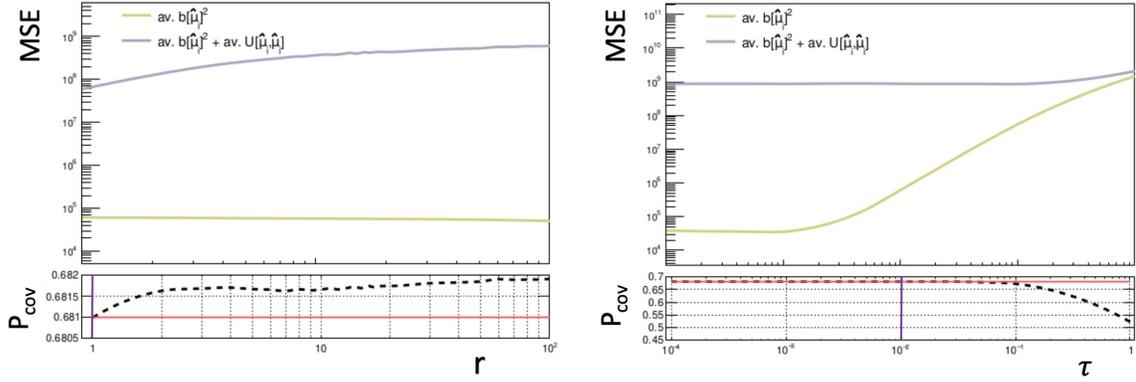
(a) $p_T^{t\bar{t}}$ unfolded with Bayes (Left) and Poisson (Right).
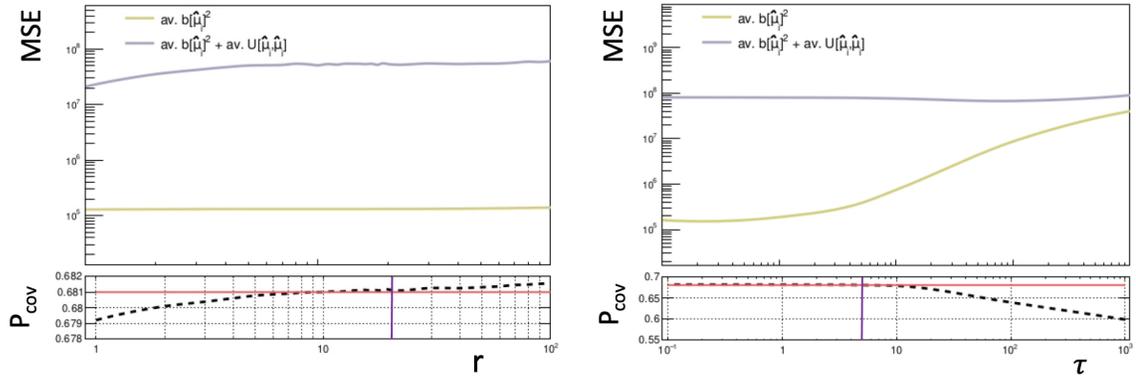
(b) $m^{t\bar{t}}$ unfolded with Bayes (Left) and Poisson (Right).

Figure 6.6: The MSE and coverage of the unfolded $p_T^{t\bar{t}}$ and $m^{t\bar{t}}$ distributions for varying regularization parameters. The vertical purple line in the bottom panel indicates the optimized value of the regularization parameters. The top panels shows the $b[\hat{\mu}_i]^2$ and the MSE $= b[\hat{\mu}_i]^2 + U[\hat{\mu}_i, \hat{\mu}_i]$ averaged over all bins. The bottom panels show the coverage averaged over all the bins with the minimal coverage indicated with a horizontal red line.
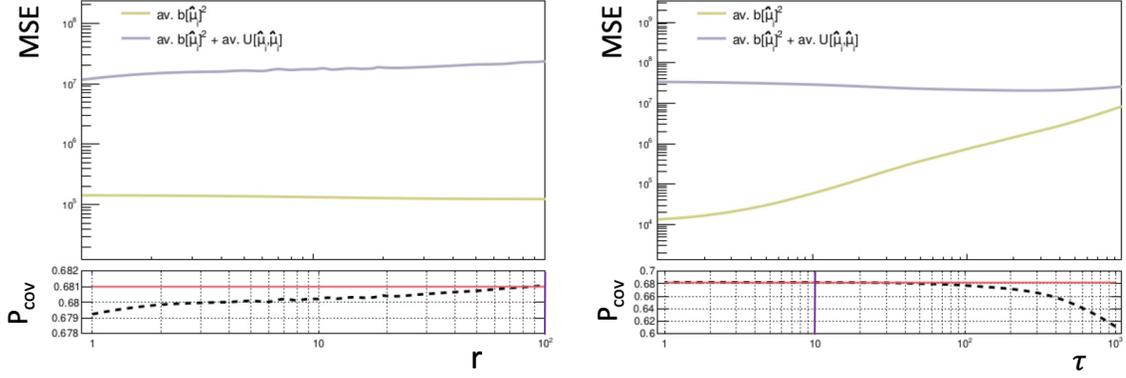
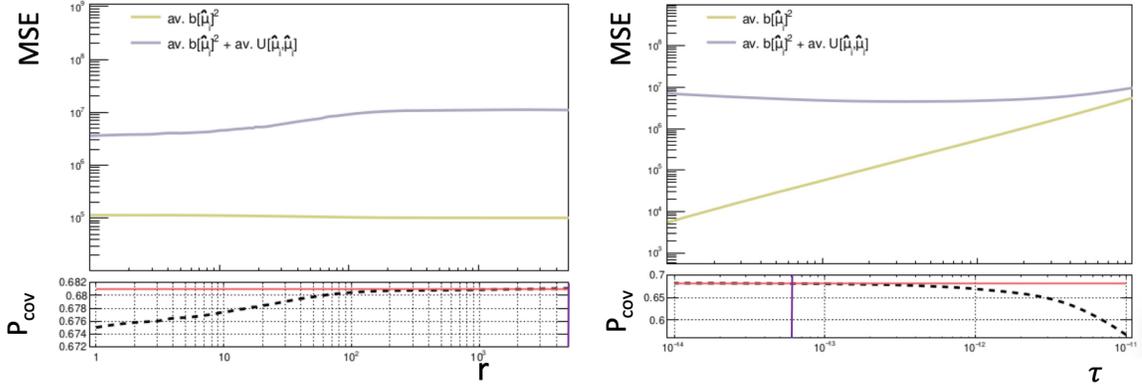(a) $H_T^{t\bar{t}}$ unfolded with Bayes (Left) and Poisson (Right).

(b) $|y^{t\bar{t}}|$ unfolded with Bayes (Left) and Poisson (Right).

Figure 6.7: The MSE and coverage of the unfolded $H_T^{t\bar{t}}$ and $|y^{t\bar{t}}|$ distributions for varying regularization parameters. The vertical purple line in the bottom panel indicates the optimized value of the regularization parameters. The top panels shows the $b[\hat{\mu}_i]^2$ and the MSE $= b[\hat{\mu}_i]^2 + U[\hat{\mu}_i, \hat{\mu}_i]$ averaged over all bins. The bottom panels show the coverage averaged over all the bins with the minimal coverage indicated with a horizontal red line.

We see that for both unfolding methods the coverage probability converges to its nominal value of 0.683 in the limit of no regularization and that it diminishes for increased regularization. One can also see that the average coverage probability is relatively high, i.e. compared to the same plots in Fig. 5.5 and 5.6. This can be explained by the inclusion of systematic uncertainties which inflates the variance and thus the coverage probability of the confidence interval set by that variance. Note that for the same MSE-coverage requirement different observables can require a very different regularization parameter. For example, the optimized number of iterations for the Bayes unfolding method applied on the $p_T^{t\bar{t}}$ distribution is $r_{\text{opt.}} = 1$ whilst the $|y^{t\bar{t}}|$ distribution needed $r_{\text{opt.}} = 5000$ for the same MSE-coverage requirement. This should fortify the argument against a single regularization parameter fits all approach as mentioned earlier in Ch. 5. The differences in values for the $\tau_{\text{opt.}}$-parameter can be explained by the different shapes of the unfolded observables. For example, $|y^{t\bar{t}}|$ distribution is naturally smoother than the $p_T^{t\bar{t}}$ distribution and therefore will need less regularization i.e. will have a lower regularization parameter $\tau_{\text{opt.}}$. Figure 6.5 also shows that the $p_T^{t\bar{t}}$ and $m^{t\bar{t}}$ distributions, despite their wider binning w.r.t. the $|y^{t\bar{t}}|$ and $H_T^{t\bar{t}}$ distributions, has wider and longer tails which indicate more smearing i.e. they will require more regularization. We see that the unfolded distributions for all the unfolding methods with all observables close nicely. We do see that the Bayes method on $m_{t\bar{t}}$ and $p_T^{t\bar{t}}$ have a better agreement with the truth with a smaller total uncertainty. However, the former is because of the relatively low number of iterations ($r_{\text{opt.}} = 1$ and $r_{\text{opt.}} = 11$) which strong regularizes towards the truth distribution in the Bayes algorithm that is statistically independent from the truth used for comparison but is generated with the same underlying model. The latter is because the relatively strong regularization reduces the variance in all bins.

6.7.4 Closure Test

A closure test is unfolding pseudo data, often generated with Monte Carlo simulation, and compare it with their corresponding truth. The unfolded pseudo data and their total uncertainties should overlap the truth. The unfolded pseudo data is the same testing or *Asimov* dataset used in the previous section to optimize the regularization parameters.

(a) $p_T^{t\bar{t}}$ unfolded with Matrix Inversion

(b) $p_T^{t\bar{t}}$ unfolded with Gaussian Processes

(c) $p_T^{t\bar{t}}$ unfolded with Bayes

(d) $p_T^{t\bar{t}}$ unfolded with Poisson

Figure 6.8: Unfolded Asimov data for $p_T^{t\bar{t}}$ including all statistical and systematic uncertainties compared with their corresponding truth distribution.

(a) $H_T^{t\bar{t}}$ unfolded with Matrix Inversion

(b) $H_T^{t\bar{t}}$ unfolded with Gaussian Processes

(c) $H_T^{t\bar{t}}$ unfolded with Bayes
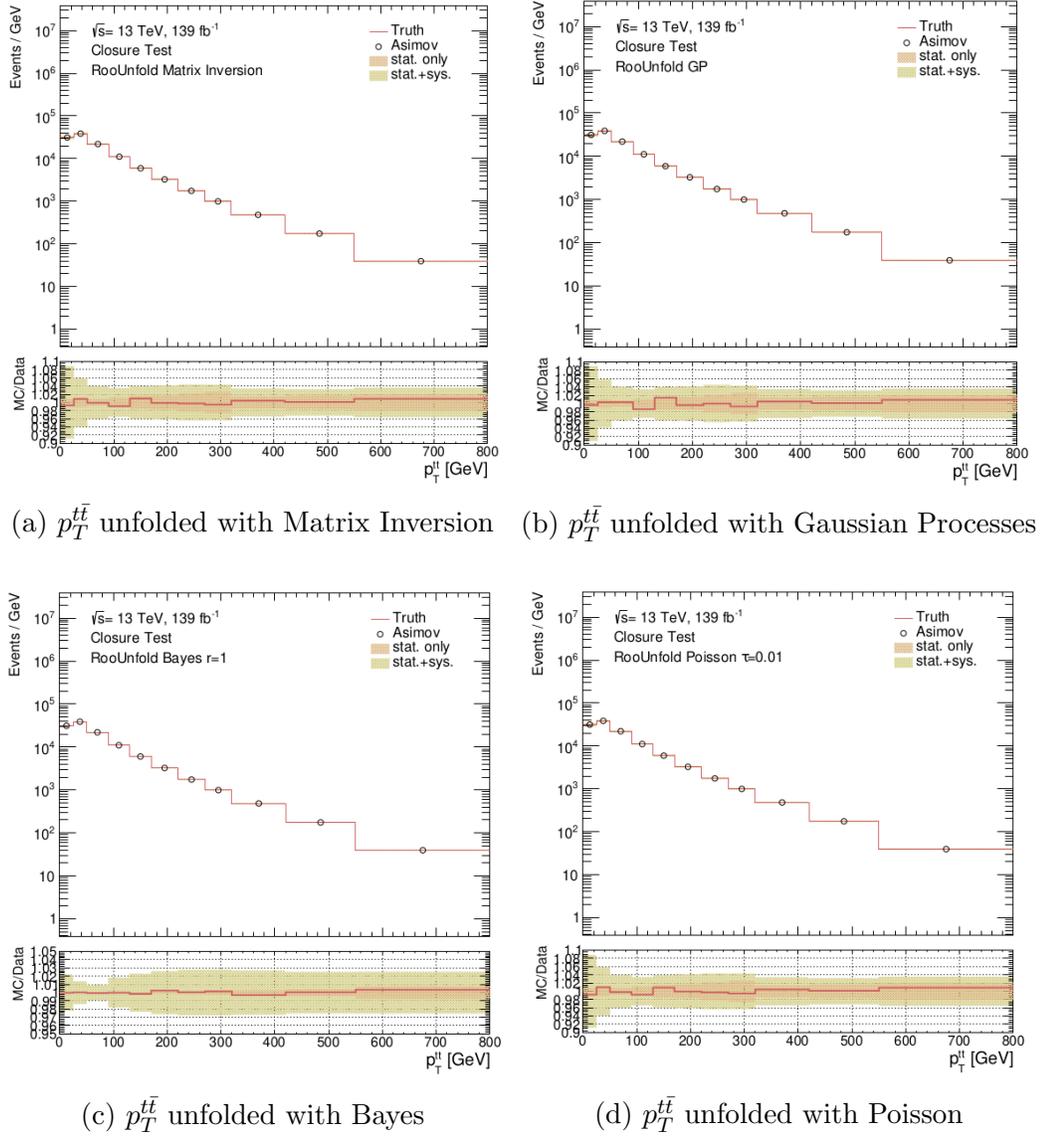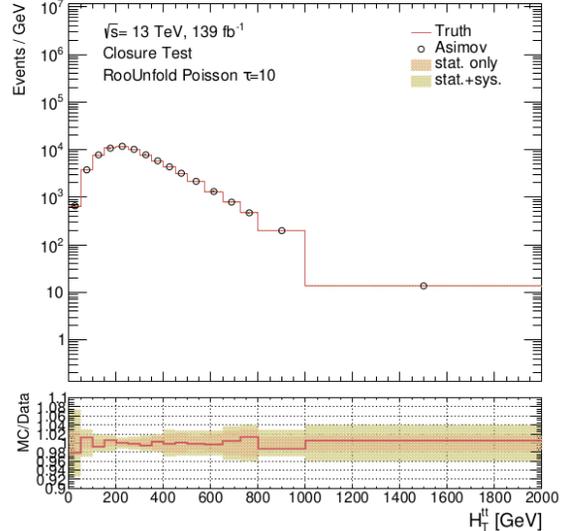
(d) $H_T^{t\bar{t}}$ unfolded with Poisson

Figure 6.9: Unfolded Asimov data for $H_T^{t\bar{t}}$ including all statistical and systematic uncertainties compared with their corresponding truth distribution.
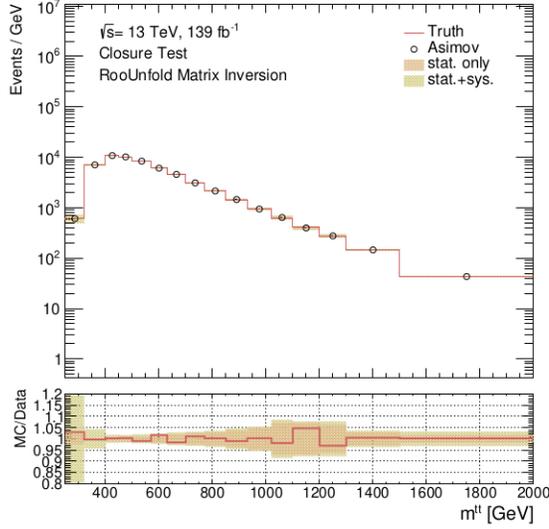
(a) $m^{t\bar{t}}$ unfolded with Matrix Inversion
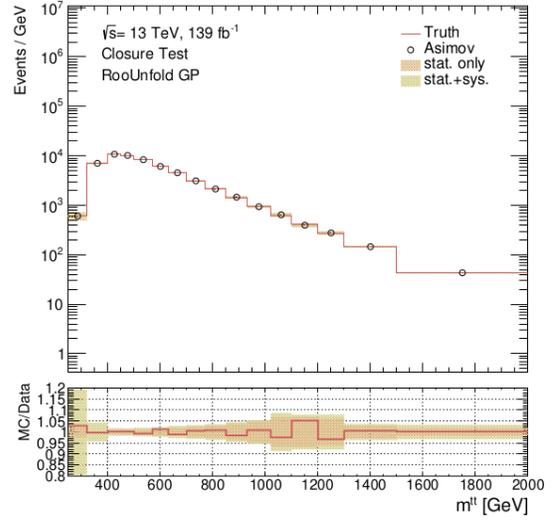
(b) $m^{t\bar{t}}$ unfolded with Gaussian Processes
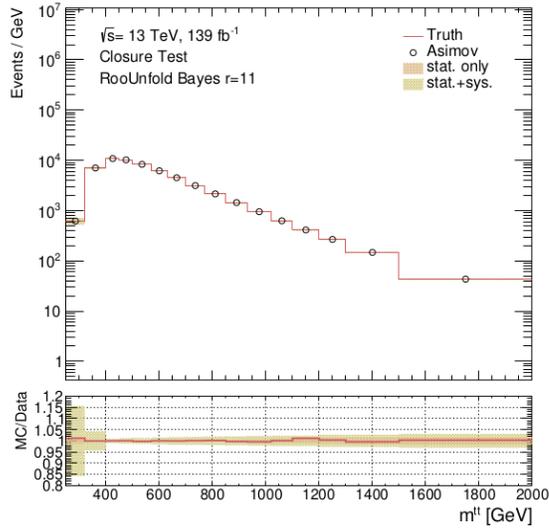
(c) $m^{t\bar{t}}$ unfolded with Bayes
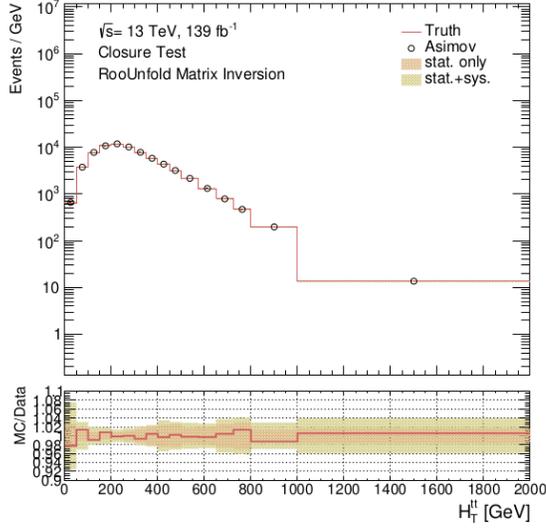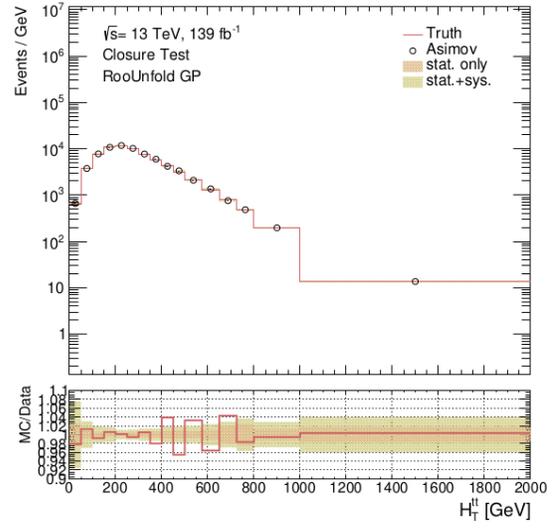

(d) $m^{t\bar{t}}$ unfolded with Poisson

Figure 6.10: Unfolded Asimov data for $m^{t\bar{t}}$ including all statistical and systematic uncertainties compared with their corresponding truth distribution.

(a) $|y^{t\bar{t}}|$ unfolded with Matrix Inversion

(b) $|y^{t\bar{t}}|$ unfolded with Gaussian Processes

(c) $|y^{t\bar{t}}|$ unfolded with Bayes
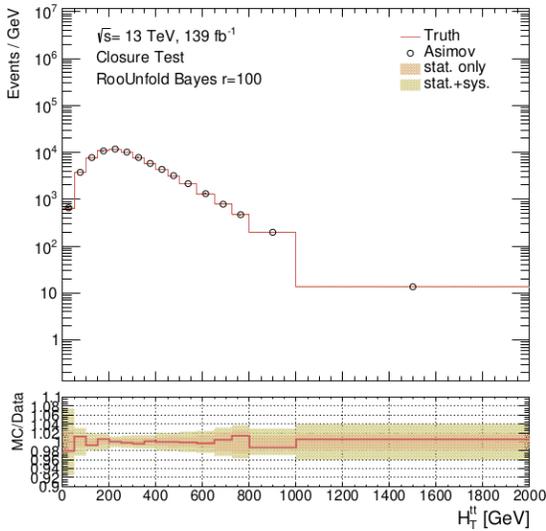
(d) $|y^{t\bar{t}}|$ unfolded with Poisson

Figure 6.11: Unfolded Asimov data for $|y^{t\bar{t}}|$ including all statistical and systematic uncertainties compared with their corresponding truth distribution.
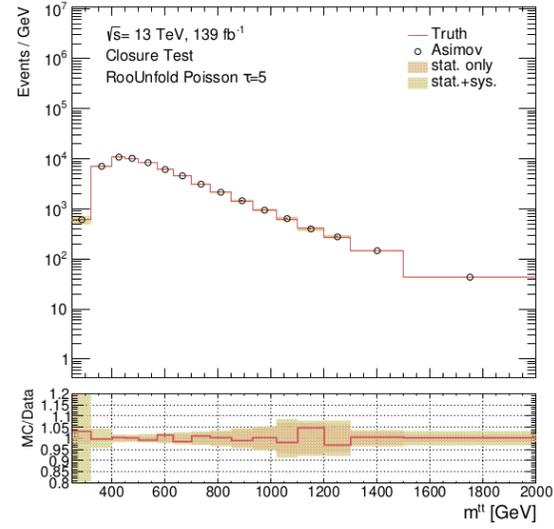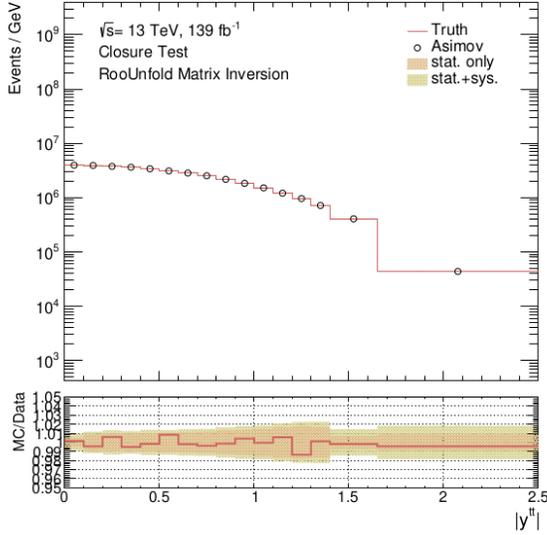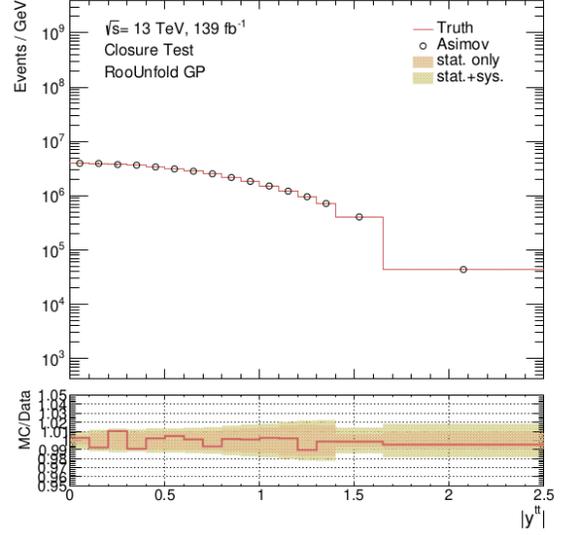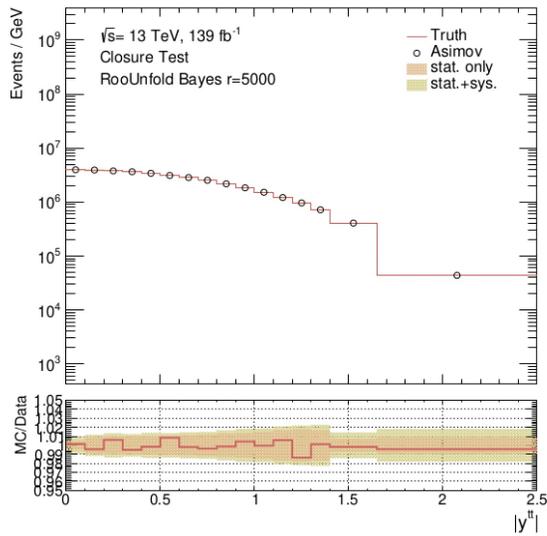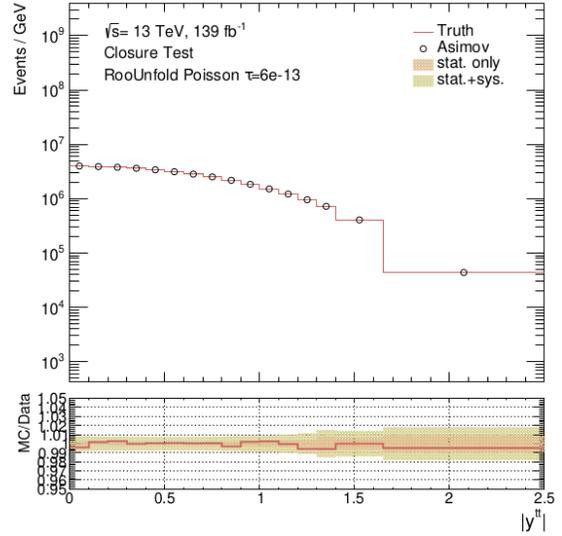
6.7.5 Fiducial Differential Cross Sections

This section presents the final measurement of the fiducial differential cross sections unfolded to particle level for the Full Run 2 dataset. The fiducial differential cross section is calculated using the equation,

$$\frac{d\sigma_{\text{fid.}}}{dX_i} = \frac{1}{\mathcal{L}_{\text{int.}} \delta X_i} \hat{\mu}_i \tag{6.12}$$

with $\mathcal{L}_{\text{int.}}$ being the total integrated luminosity of the Full Run 2 data set, δX_i the bin width and $\hat{\mu}_i$ the estimated truth bin i.e. number of unfolded events in bin i. Additionally, we include a $\chi^2/\text{n.d.f.}$ from Eq. 5.32 in the plot to quantify the agreement between data and prediction. The χ^2 is a statistic that follows the χ^2 probability density function with the number degrees of freedom as only parameter. This probability density function is defined as

$$f(x|k) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}, \qquad \Gamma(k/2) = \int_0^\infty t^{k/2-1} e^{-t} dt \tag{6.13}$$

with the number degrees of freedom here denoted with k and $\Gamma(k/2)$ known as the gamma function. For a measured χ^2 and $k = \text{n.d.f.}$ one can define the p-value as,

$$p = \int_{\chi^2}^\infty f(x|k) dx. \tag{6.14}$$

The p-value therefore gives the probability to observe the measured χ^2 or one more extreme, i.e. larger in value, under the assumed prediction $\boldsymbol{\mu}$ used to calculate the χ^2. Figure 6.12 shows that the fiducial differential cross sections and Figure 6.13 and 6.14 show the corresponding covariance matrix estimates used for the χ^2. The fiducial differential cross sections and covariance matrix estimates unfolded with the GP, Bayes and Poisson methods can be found in Appendix B. The high-p_T regions of the $p_T^{t\bar{t}}$ and $H_T^{t\bar{t}}$ observables show a substantial disagreement between the POWHEG-

Box V2+Pythia8 $t\bar{t}$ signal MC sample and the data. The p-values of the $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$ observables are still above the 5σ significance boundary of 3×10^{-7}. We see the same MC/data ratio shapes as in the detector level data and Monte Carlo comparison of Sec. 6.6. Either the signal $t\bar{t}$ or one of the background Monte Carlo samples is overestimating. One should also note that no theoretical uncertainties were taken into account here. These will definitely decrease the χ^2 and thus increase the low p-values. However, significant disagreement can still remain after that. The unfolded results using the matrix inversion are used as the nominal results in the rest of this thesis. This is to avoid any dependence on regularization parameters.
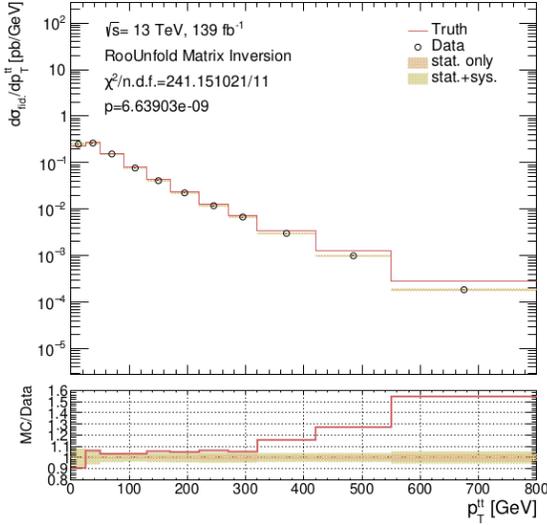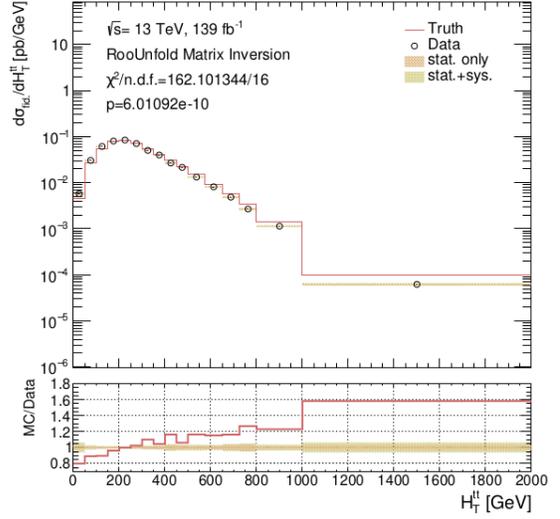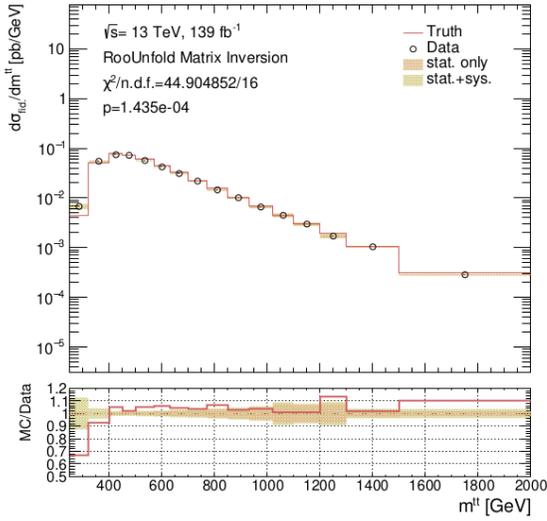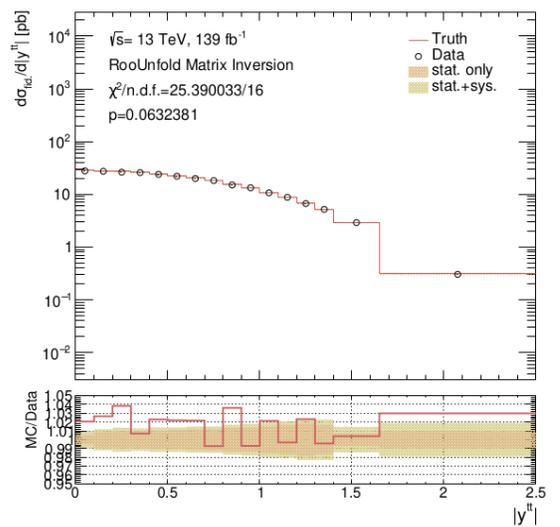
Figure 6.12: Measured fiducial differential cross sections with the Full Run 2 dataset of $139\,\mathrm{fb}^{-1}$ unfolded to particle level with the matrix inversion method including all statistical and systematic uncertainties. The unfolded data is compared with the testing subsample of $t\bar{t}$ events from the Powheg-Box v2+Pythia8 signal sample described in Sec. 6.2 and includes a χ^2 with corresponding p-value.

(a)
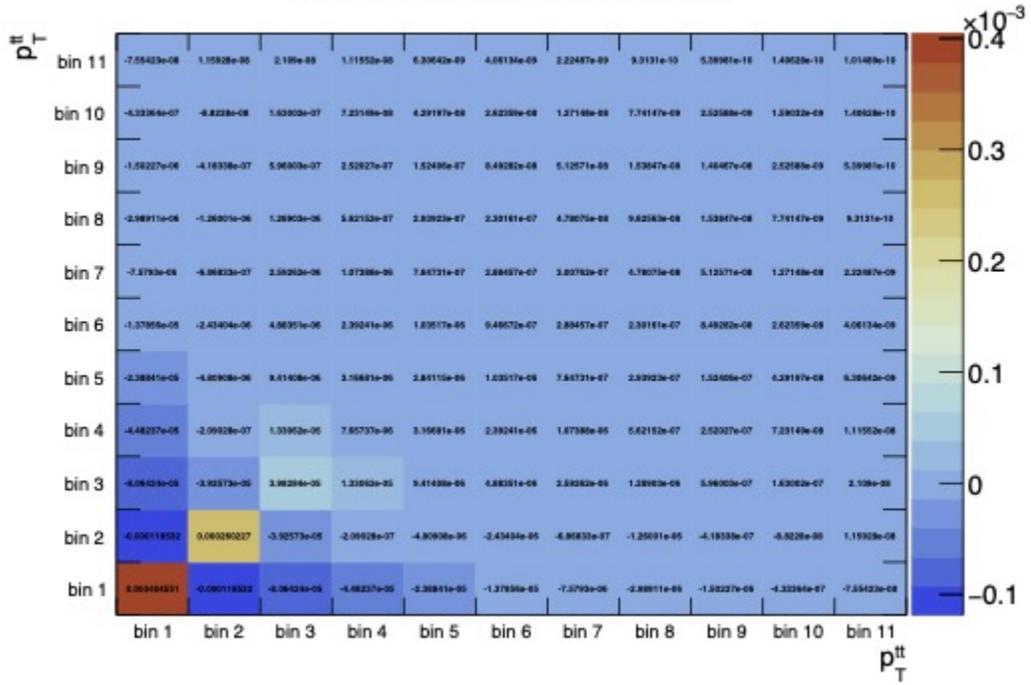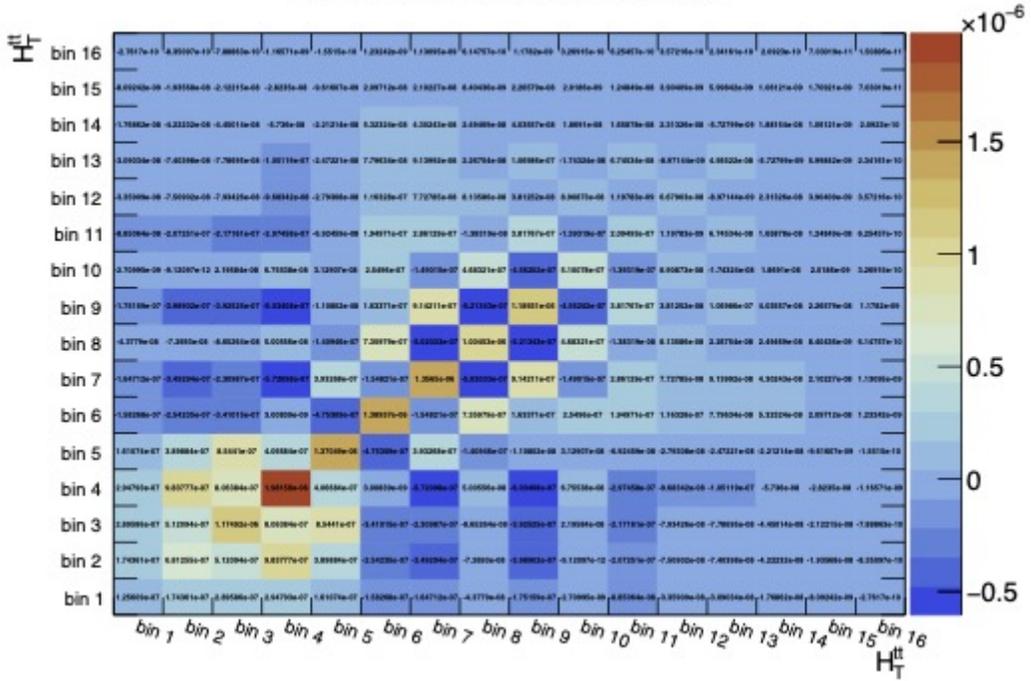
(b)

Figure 6.13: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $p_T^{t\bar{t}}$ and $H_T^{t\bar{t}}$ unfolded with the matrix inversion method.

(a)
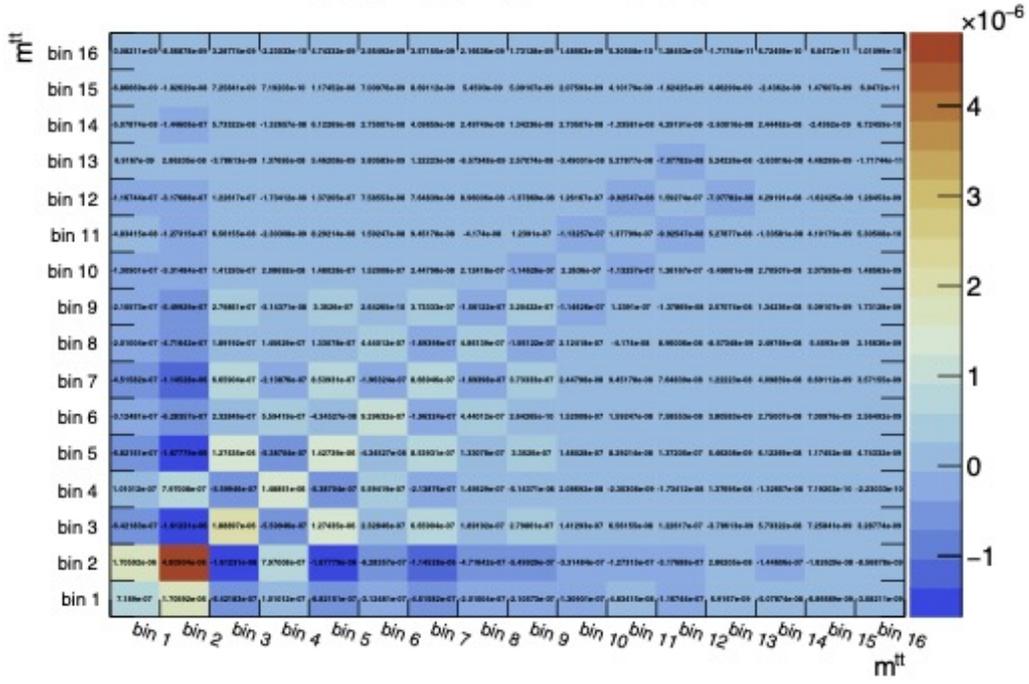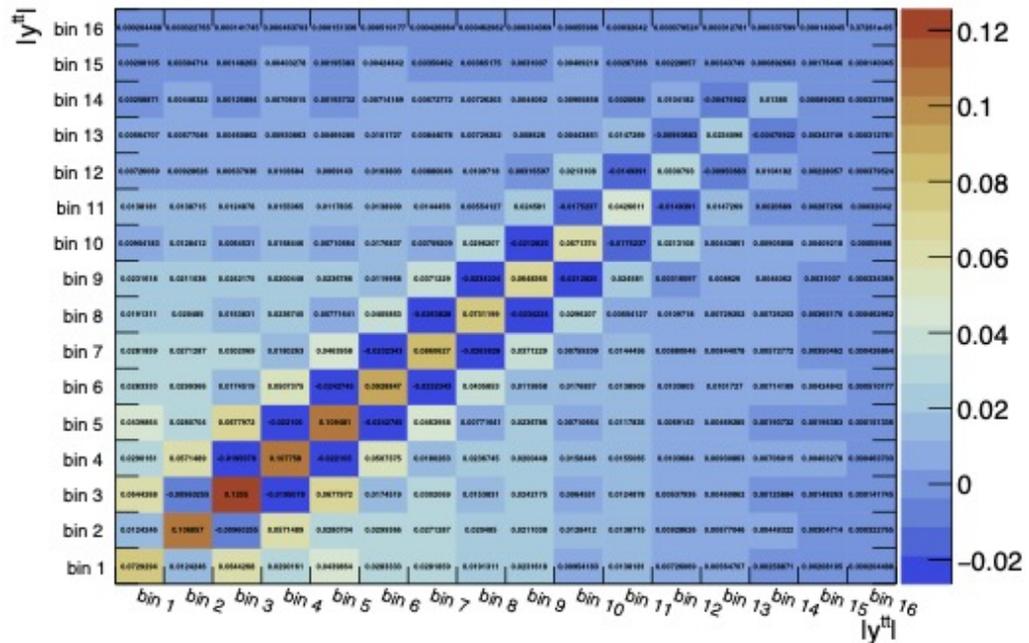
(b)

Figure 6.14: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$ unfolded with the matrix inversion method.

154

6.8 Comparisons with Previous Measurements

In 2019 a similar analysis was published presenting as well top-quark pair differential cross-sections in the l+jets decay channel using pp collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ using the ATLAS detector [240]. Differential cross sections of the $p_T^{t\bar{t}}$, $m^{t\bar{t}}$ and $H_T^{t\bar{t}}$ kinematic observables were presented here as well. However, this previous analysis contains some differences with respect to the analysis presented in this thesis.

The measured data of this previous analysis was collected during the 2015 and 2016 LHC running periods and amounts to an integrated luminosity of $36\,\text{fb}^{-1}$. This is less than the $139\,\text{fb}^{-1}$ of the 2015 to 2018 running period used in this analysis. The Monte Carlo samples used in this thesis have also been produced with different or upgraded versions of event generators, hadronization and parton shower software packages. For example, this thesis used PYTHIA 8.230 and modelled the single top background with POWHEG-BOX v2. The previous analysis used PYTHIA 6.428 or PYTHIA 8.186 and modelled single top with POWHEG-BOX v1. Also, software packages for the detector simulation, i.e. GEANT 4 and ATLFAST-II, and for the object reconstruction and identification, i.e. ATHENA, have improved since this previous analysis.

The event and object selection has remained largely the same with the most notable difference being the change from the MV2c10 to the DL1r b-tagging method. Additionally, the previous analysis used the Iterative Bayes unfolding algorithm with $r = 4$ iterations for all differential cross sections in contrast to the several unfolding algorithms with optimized regularization parameters used in this thesis. Lastly, less sources of statistical and systematic sources of error were included and propagated with a different statistical method in this thesis. Figure 6.15 and 6.16 show the data-MC comparison and the unfolded fiducial differential cross sections as presented in [240]. We can see that the ratio plots in Fig. 6.16 show the same disagreement be-

tween data and the Powheg-Box v2+Pythia8 predictions for all variables.

When evaluating a difference between data and MC one could also argue that the data taking at the ATLAS experiment introduces a bias in the $t\bar{t}$ differential cross section measurements. However, similar differences between data and MC have also been observed at the CMS experiment. Figure 6.17 shows $t\bar{t}$ differential cross sections in the l+jets decay channel with pp collisions recorded at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ at the CMS experiment. Several p_T variables are shown of which Powheg-Box v2+Pythia8 predictions shows an overestimation w.r.t. CMS data unfolded to particle level. However, one should take into account their detection methods, data reconstruction and selection algorithms are different from the ones presented in this thesis.
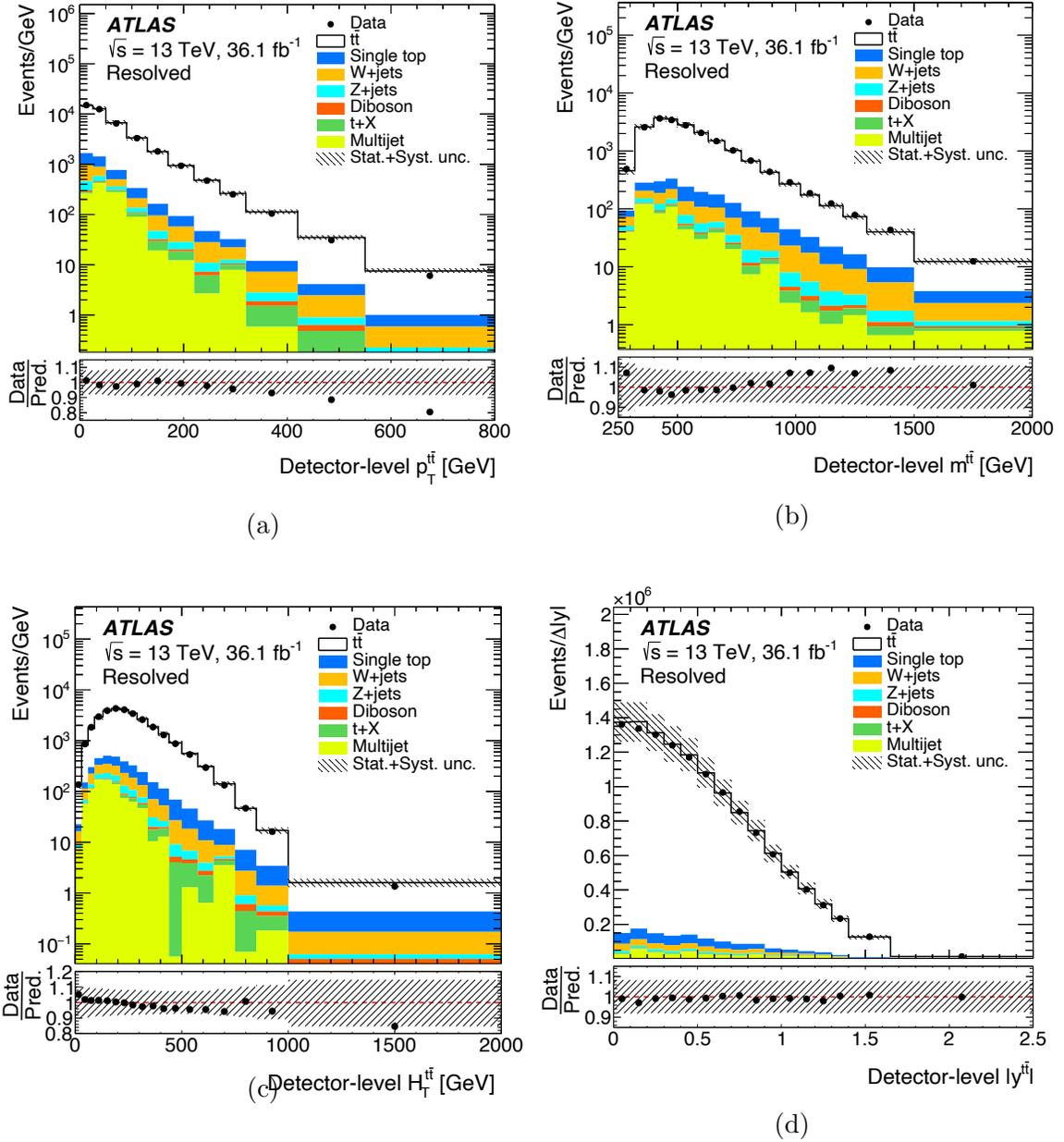
Figure 6.15: Data-MC comparison from a previous $t\bar{t}$ differential cross section measurement in the l+jets decay channel measured at the ATLAS detector. From [240].
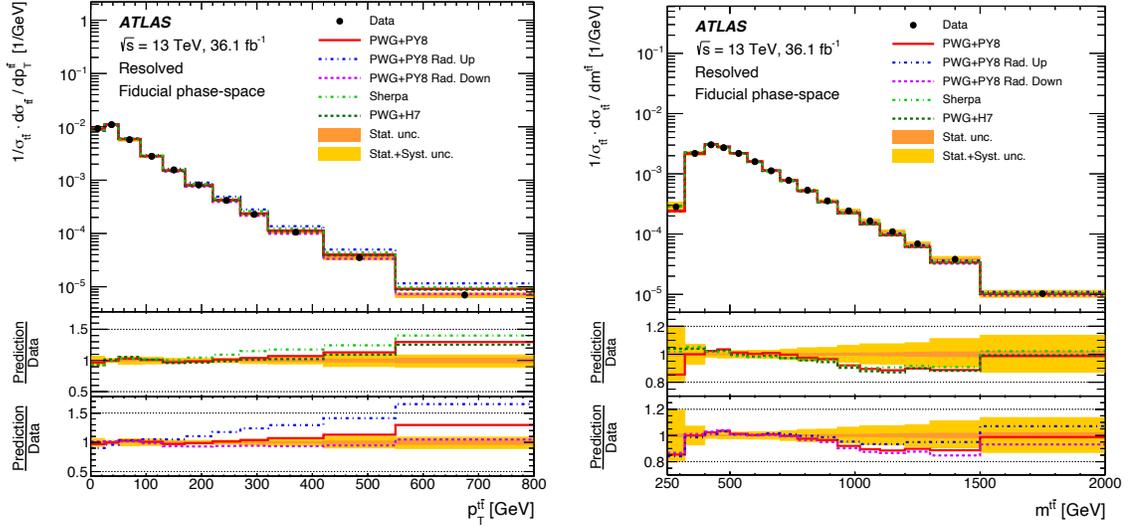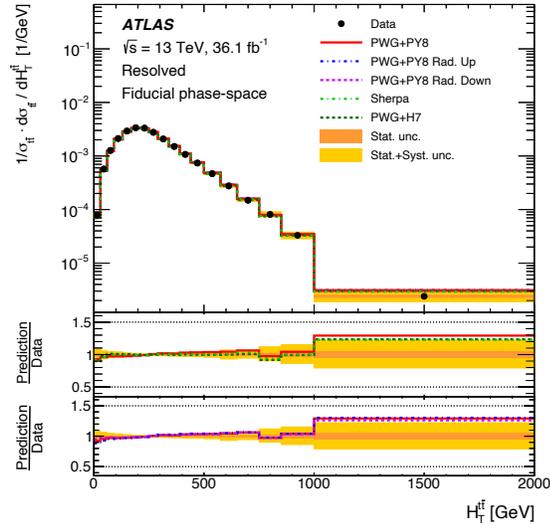
Figure 6.16: Unfolded distributions from a previous $t\bar{t}$ differential cross section measurement in the l+jets decay channel measured at the ATLAS detector. From [240].
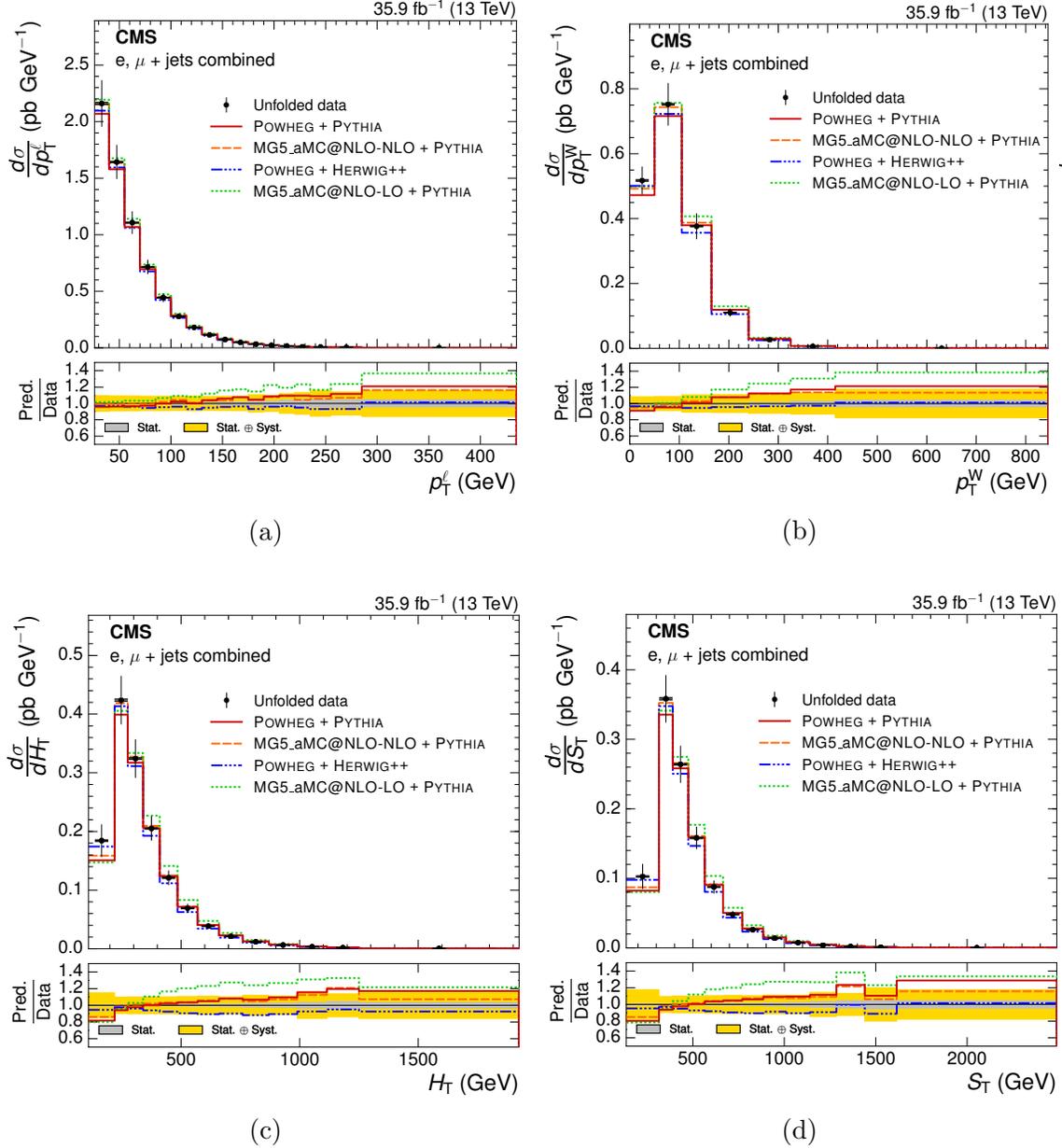
Figure 6.17: Unfolded distributions from a $t\bar{t}$ differential cross section measurement in the l+jets decay channel measured at the CMS detector. Differential cross sections are shown as a function of the p_T of the lepton coming from the top quark (p_T^l), the p_T of the leptonically decaying W boson coming from the top quark (p_T^W), the scalar sum of the p_T of all jets (H_T) and the scalar sum of the p_T of all particles (S_T). From [242].

6.9   Discussion

The fiducial differential cross-sections for $t\bar{t}$ production in the l+jets decay channel are presented as a function of the $p_T^{t\bar{t}}$, $m^{t\bar{t}}$, $H_T^{t\bar{t}}$ and $|y^{t\bar{t}}|$ kinematic variables. The differential cross sections were unfolded to particle level including various sources of statistical and systematic uncertainty and compared with a previous measurement of $t\bar{t}$ differential cross sections at the ATLAS experiment. The $p_T^{t\bar{t}}$ and $H_T^{t\bar{t}}$ show a disagreement with the POWHEG-BOX V2+PYTHIA8 prediction in the high p_T region. This is a known disagreement and is also seen in other analyses with both the ATLAS [240] and the CMS experiment [242]. Future studies could entail investigating any of the aspects of the Monte Carlo simulation. Also, there are many detector and theory systematic uncertainties that could be added to the analysis that might cover the discrepancies.

The measurement presented in [240] compares $t\bar{t}$ predictions from multiple event generators, hadronization and parton shower implementations. One of the possible improvements to this analysis would be to compare the unfolded data to $t\bar{t}$ predictions from other event generators such as MADGRAPH5_AMCNLO or SHERPA in combination with different hadronization and parton showering implementations such as PYTHIA or HERWIG. Additionally, one could fine tune the parameters of these models to investigate their effects on $t\bar{t}$ differential cross section predictions.

The highest QCD order at which events were generated for any of the MC samples was at NLO with possibly a NNLO QCD scaling factor. However, an event generator with parton showering at NNLO QCD could give a substantial improvement to this analysis. Developments in this area are being made [243, 244, 245] and could improve both the $t\bar{t}$ particle level predictions and reduce theoretical uncertainties. Future studies could explore these NNLO QCD implementations. Additionally, higher EW

160

corrections could be included but one would expect these to have a lower impact on a QCD dominated process such as $t\bar{t}$ production [246].

Lastly, the inclusion of theoretical sources of uncertainty would also be an improvement for the analysis presented in this thesis. This could include sources involving parton distribution functions, QCD radiation effects, renormalization and factorization scales. Some of these nuisance parameters are listed at the end of App. B.1. It has been shown that each of these sources could contribute up to a few percent of relative uncertainty in the final unfolded distribution [241, 240]. For example, final state radiation effects can contribute a systematic uncertainty of up to 8% [240]. In addition to these sources, one could also investigate new sources of uncertainty.

Chapter 7

Constraining Effective Field Theories

In this chapter, the measurements of the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients of the Standard Model Effective Field Theory, as introduced in Sec. 2.6.2, are presented. Frequentist confidence bounds are set by comparing SMEFT predictions at particle level with the unfolded fiducial differential cross sections presented in Ch. 6. The comparison of measured data with EFT theories is also known as EFT interpretation and is considered by more and more analyses within ATLAS as a valuable addition to the measurement of cross sections. Some ATLAS analyses that included this extra perspective can be found in Higgs [247, 248, 249, 250, 251], electroweak [252, 253, 254, 255] and top physics [256, 257, 258, 259, 260]. The results presented in this chapter contribute to this new paradigm in both methods and results. Some points on EFT interpretation strategy are included that were developed in collaboration with the ATLAS Top community to guide other analyses in their EFT interpretation and prepare for any possible future efforts that will simultaneously fit more Wilson coefficients to datasets from multiple analyses, also known as a *global EFT approach*.

7.1 Strategy

The Standard Model Effective Field Theory is a complex model of which the Wilson coefficients can have a myriad of effects on physics processes and their observables. From this it seems evident that one needs datasets collected by many experiments looking at multiple physics processes and observables to truly know how SMEFT manifests itself in nature. Additionally, estimates of Wilson coefficients will have non-negligble correlations. A simultaneous fit involving all correlated Wilson coefficients would make the number of parameters of interest quickly quite large. A generally adopted approach within the ATLAS collaboration has therefore become to focus each analysis on a small number of observables and a subset of Wilson coefficients and leave the results in such a way that they can be combined at a later moment in a *global fit*. The goal of this measurement is therefore to use the available phase-space spanned by the $t\bar{t}$ l+jets fiducial differential cross sections presented in the previous chapter, measure a subset of Wilson coefficients and prepare the results for later combination.

As mentioned before, the comparison of SMEFT predictions with data can be done either on detector, particle or parton level. Performing the EFT interpretation on detector-level has advantages such as straightforward handling of systematic uncertainties, inclusion of background dependencies or likelihood function and nuisance parameter combination for a global EFT fit approach. However, a detector level approach does introduce a heavy dependence on the collaboration software from detector simulation and reconstruction that is not available for physicists outside the collaboration. On top of that, these software packages require large computing resources for an already large set of Monte Carlo samples needed for an interpretation. It is for these reasons that the gold standard within the ATLAS Top community has become to compare SMEFT predictions with data unfolded to particle level. This level has a

reduced detector dependency but does apply corrections that include model depen-
dent assumptions on hadronization and parton showering such is with unfolding to
parton level. The strategy for an EFT interpretation on particle level can be defined
in the following steps:

1. Unfold data to particle level, either with or without regularization, including
   covariance estimation

2. Define a likelihood for the unfolded data, e.g. $\hat{\boldsymbol{\mu}} \sim \text{Gaus}(\boldsymbol{\mu}, U)$ with mean $\boldsymbol{\mu}$
   and covariance U

3. Choose relevant operators

4. Parameterize the likelihood in terms of the Wilson coefficients, i.e. $\boldsymbol{\mu} \to \boldsymbol{\mu}(\boldsymbol{c})$

5. Set confidence bounds by maximizing and varying the likelihood w.r.t. the
   Wilson coefficients

The first step has already been presented in Ch. 6. We will therefore dedicate this
chapter to the steps 2 to 5.

7.2   Statistical Framework

7.2.1   Likelihood & Confidence Intervals

The final results of an EFT interpretation are confidence bounds set by maximizing
and varying a likelihood function w.r.t. the Wilson coefficients. If we assume that
the unfolded data follow a multivariate Gaussian probability distribution then we can
define the log-likelihood as

$$\log L(\boldsymbol{c}) = -\frac{1}{2}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}(\boldsymbol{c}))^T U^{-1}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}(\boldsymbol{c})), \tag{7.1}$$

with the unfolded distribution $\hat{\boldsymbol{\mu}}$, the corresponding covariance $U_{ij} = \text{cov}[\hat{\mu}_i, \hat{\mu}_j]$ including all the statistical and systematic sources of error, as introduced in Ch. 5, and the particle level model prediction $\boldsymbol{\mu}(\boldsymbol{c})$ as a function of the Wilson coefficients $\boldsymbol{c}$. Note that the multivariate Gaussian probability distribution function is chosen to ensure that all the correlations between the bins introduced in the unfolding are taken into account. The large number of events in the fiducial phase space make it reasonable to assume that the large sample limit holds and that the unfolded data follows a Gaussian instead of a Poisson distribution. Even for low-event bins in the high-p_T region enough events are measured for this assumption to hold. Maximizing the log-likelihood w.r.t. the Wilson coefficients will give their respective maximum-likelihood estimators:

$$\hat{\boldsymbol{c}}_{\text{MLE}} = \arg\max_{\boldsymbol{c} \in \mathcal{C}} \log L(\boldsymbol{c}). \tag{7.2}$$

In the large sample limit and in case of one or uncorrelated Wilson coefficients, one can set the confidence interval $[\hat{c}_{i,\text{MLE}} - N\sigma_{\hat{c}_i}, \hat{c}_{i,\text{MLE}} + N\sigma_{\hat{c}_i}]$ for MLE $\hat{c}_{i,\text{MLE}}$ by varying the parameters $\boldsymbol{c}$ until the log-likelihood decreases by $N^2/2$ in value from its maximized value [192], i.e.

$$\log L(c_i + N\sigma_{c_i}) = \log L_{\text{max}} - \frac{N^2}{2}, \tag{7.3}$$

with Wilson coefficient c_i with $N \in \mathcal{N}$ being a positive integer defining the confidence level and error on the MLE $\sigma_{\hat{c}_i}$. For example, $N = 1$ gives a 1σ confidence interval at a confidence level of 68.3%. However, in cases of multiple correlated Wilson coefficients, Eq. (7.3) changes into

$$\log L(\boldsymbol{c}) = \log L_{\text{max}} - \frac{Q_{\alpha,n}}{2} \tag{7.4}$$

with

$Q_{\alpha,n}$			
$1-\alpha$	$n=1$	$n=2$	$n=3$
0.683	1	2.3	3.53
0.90	2.71	4.61	6.25
0.95	3.84	5.99	7.82
0.99	6.63	9.21	11.3

Table 7.1: Values for χ^2 quantiles for various confidence levels $1-\alpha$ and number of parameters (Wilson coefficients) n. [192]

$$Q_{\alpha,n} = F_{\chi^2}^{-1}(1-\alpha, n) \tag{7.5}$$

where $F_{\chi^2}^{-1}(1-\alpha, n)$ are the quantiles of a χ^2 distribution at confidence level $1-\alpha$ for n degrees of freedom corresponding to Wilson coefficients [192]. Values of these quantiles for various confidence levels and number of Wilson coefficients are given in Table 7.1. The only thing needed for setting confidence intervals is the parameterization of the truth distribution $\boldsymbol{\mu}(\boldsymbol{c})$ in terms of Wilson coefficients.

7.2.2 Quadratic Regression

Recall from Eq. (2.67) that the Lagrangian of the Standard Model Effective Field Theory follows a power series in $D \geq 4$ operators. This allows for cross sections, differential and total, to be written in quadratic form in terms of Wilson coefficients of dimension 6 operators,

$$\sigma_{\text{SMEFT}} = \sigma_{\text{SM}} + \sum_i^{N_{\text{dim6}}} \sigma_i \frac{c_i}{\Lambda^2} + \sum_i^{N_{\text{dim6}}} \sum_j^{N_{\text{dim6}}} \sigma_{ij} \frac{c_i c_j}{\Lambda^4}, \tag{7.6}$$

where σ_{SM} is the Standard Model prediction, c_i the Wilson coeffient of the i-th dimension 6 operator, σ_i the linear contribution arising from EFT operators interfering with the SM amplitude and σ_{ij} is the quadratic contribution originating from the squared amplitudes of the EFT operators. We can therefore define the multidimensional quadratic function

$$\mu_k(\boldsymbol{c}) = \mu_{\mathrm{SM},k} + \sum_i^{N_{\mathrm{dim6}}} a_{ik} c_i + \sum_i^{N_{\mathrm{dim6}}} \sum_j^{N_{\mathrm{dim6}}} b_{ijk} c_i c_j, \tag{7.7}$$

where $\mu_k(\boldsymbol{c})$ is the SMEFT prediction for bin k of some observable at particle level and $\mu_{\mathrm{SM},k}$ is the SM prediction for bin k of the same observable at particle level. Both can be estimated with Monte Carlo simulation. By estimating $\mu_k(\boldsymbol{c})$ at various points in $\boldsymbol{c}$-space one can find values for the coefficients a_{ik} and b_{ijk} by minimizing residuals of the function $\mu_k(\boldsymbol{c})$ and Monte Carlo estimates. Additionally, the Jacobians of the residuals are estimated with the use of finite differences to estimate the covariances and thus errors $\sigma_{a_{ik}}$ and $\sigma_{b_{ijk}}$ on the estimated parameters a_{ik} and b_{ijk}.

7.3 Monte Carlo simulation

Monte Carlo simulations were used to determine predictions of SMEFT for both the $t\bar{t}$ signal process as well as all the background processes. Almost all Monte Carlo generators have some EFT implementations available but they are often only available for a limited set of physics processes. More general SMEFT model implementations with a large set of physics processes are available in UFO format [261] and can be used in combination with MADGRAPH5_AMC@NLO. The UFO models available that implement matrix calculation and event generation with SMEFT operators in the Warsaw basis are:

- SMEFTsim3.0 [262, 263] - Possible choices of fermion flavor symmetries and electroweak parameter value sets are available. Both CP violating and conserving operators are implemented. Leading order QCD accuracy only. Depending on the chosen flavor symmetry, this implementation has between 81 and 129 free parameters.

- SMEFTatNLO [87] - The model implements a $U(2)_q \times U(3)_d \times U(2)_u$ symmetry in

the quark sector and a $U(1)_l \times U(1)_e$ in the lepton sector. Only CP conserving operators available. Wilson coefficients involving a b-quark are set to their lighter generation down-type equivalent e.g. $c_{Qb}^8 \to c_{Qd}^8$ or to zero if not present. Accuracy up to Next-to-leading order QCD available. This implementation contains 73 free parameters.

- dim6top [88] - The model implements a $U(2)_q \times U(2)_d \times U(2)_u$ quark flavor symmetry but also includes flavor-changing neutral current(FCNC) interactions. Leading order QCD accuracy only. This implementation contains 122 free parameters.

The Monte Carlo SMEFT simulations of this thesis use the SMEFTatNLO UFO model with MADGRAPH5_AMC@NLO v3.1 because of its available NLO QCD accuracy. The parton shower and hadronization were simulated with PYTHIA 8.230 [114, 112] with the NNPDF23LO PDF [226] and top quark mass $m_t = 172.5\,\text{GeV}$.

7.4 Choice of operators

As mentioned before, a set of operators need to be chosen to focus this measurement on. The choice is based on an operator sensitivity check, a brief review of recent global fits and differential distribution sensitivity check.

7.4.1 Operator contribution

It is possible to deduce per operator what type of physics processes it can affect based on their field contents and imposed flavor symmetries. However, it is easier to look at each physics process of the analysis and investigate to which operators they are sensitive. It is possible to set all the Wilson coefficients to a non-zero value, calculate the matrix element with MADGRAPH5_AMC@NLO and produce a list of

all the operators included in that matrix calculation. In the physics process output directory of the MadGraph5_aMC@NLO one can find the list in the file:

```
Source/MODEL/coupl.inc
```

The couplings are encoded but can be translated to their conventional notation by comparing the codes to a model file in the Madgraph source directory:

```
models/SMEFTatNLO/couplings.py
models/SMEFTatNLO/CT_couplings.py
```

By generating the process on both Leading Order and Next-to-Leading Order QCD one can see for a specific process at what accuracy which operators contribute. Figure 7.1 shows which bosonic, two-fermion and four-quark operators contribute to which physics processes involved in this thesis. The used operator and Wilson coefficient notation is defined for `SMEFTatNLO` in [264]. Notice that many operators only contribute from NLO QCD. This should be a strong argument for the choice of the `SMEFTatNLO` implementation as `SMEFTsim3.0` and `dim6top` would miss these contributions at LO QCD. A subset of operators can be defined from Fig. 7.1 that show sensitivity to the $t\bar{t}$ signal process:

$$\{c_{Qt}^8, c_{tt}^1, c_{Qt}^1, c_{QQ}^1, c_{QQ}^8, c_{td}^1, c_{tu}^1, c_{Qd}^1, c_{tq}^1, c_{Qu}^1, c_{Qq}^{1,1}, c_{Qq}^{3,1}, c_{td}^8, c_{tu}^8, c_{Qd}^8, c_{tq}^8, c_{Qu}^8, c_{Qq}^{1,8}, c_{Qq}^{3,8}, c_{tG}\}$$
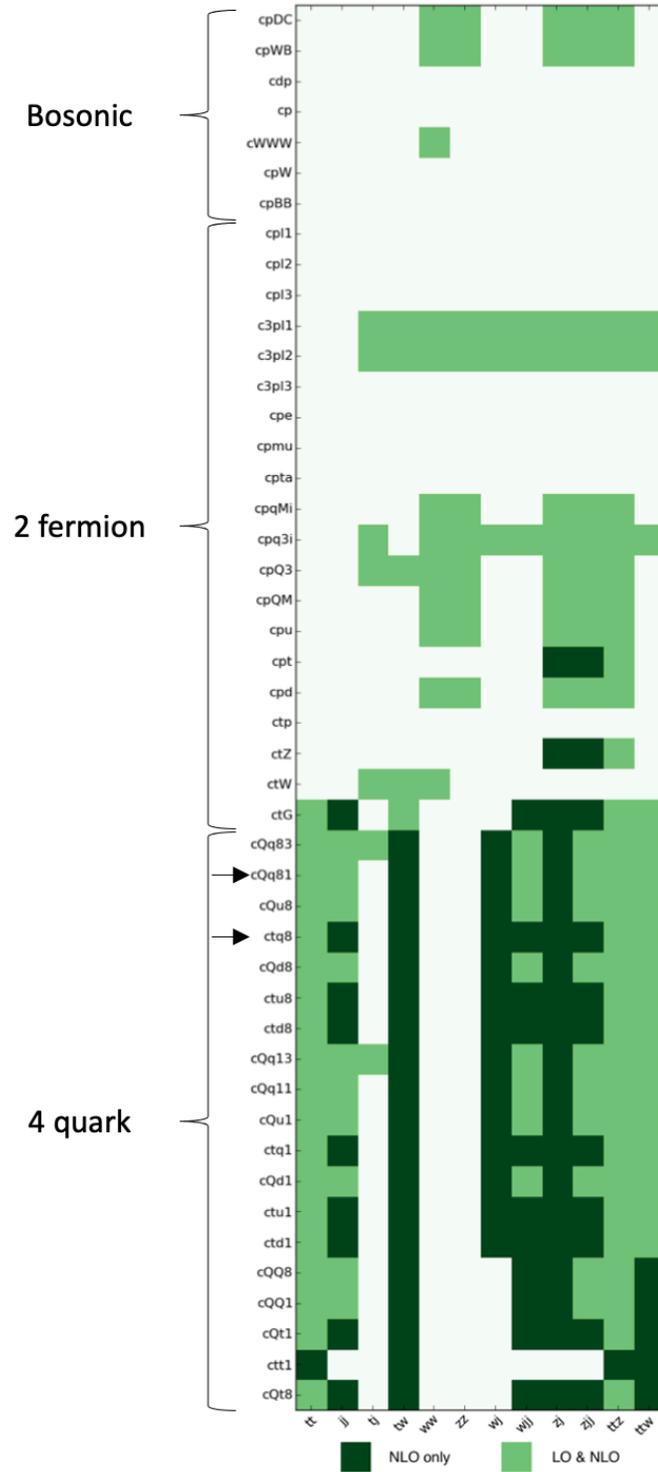
$$(7.8)$$

Figure 7.1: Wilson coefficients of `SMEFTatNLO` bosonic, two fermion and four-quark dimension-6 operators and their contribution to several physics processes. The color coding distinguishes between operators that contribute at LO and NLO QCD or NLO-only. See [264] for all the operator and Wilson coefficient definitions.

We see that most of the bosonic and two-fermion operators are not of interest for this thesis. One should note that many of these operators also contribute to background processes which could introduce a degeneracy in EFT effects. Luckily, figure 6.3 shows that the background contributions in this analysis are small relative to the signal process which makes it reasonable to assume that the contributions of these operators to the backgrounds will be negligible. Figure 7.1 does not include decay processes which would introduce sensitivity to more operators such as the $\mathcal{O}_{tW}$ operator.

7.4.2 Correlations, Flat Directions and Confidence Bounds

In this section we will briefly discuss correlations between Wilson coefficients, flat fit directions and set confidence intervals to motivate a choice in operators even further. Recent results on these subjects can be found in the global fits presented in [265, 92, 266].

Correlations are crucial to take into account when choosing a set of Wilson coefficients to measure. If the correlation between two parameters is small to zero than a simultaneous fit will not present any additional information w.r.t. an individual fit i.e. the limits will be the same. An estimate of Pearson correlations between 34 Wilson coefficients is shown in Fig. 7.2. The correlation coefficient estimate uses pseudo data and is defined as

$$\rho(c_i, c_j) = \frac{\frac{1}{N_{pseudo}} \sum_{k=1}^{N_{pseudo}} \hat{c}_i^{(k)} \hat{c}_j^{(k)} - \langle c_i \rangle \langle c_j \rangle}{\delta c_i \delta c_j} \tag{7.9}$$

with N_{pseudo} being the number of pseudo data samples, $\hat{c}_i^{(k)}$ the Wilson coefficient estimate for the k-th pseudo data sample, $\langle c_i \rangle$ the average of all the estimates and δc_i the estimated uncertainty. Note that most of the significant correlations are between the four-quark Wilson coefficients. This makes four-quark operators good candidates

for this thesis.

We know that certain directions in Wilson coefficient parameter space can result in a net zero change in the observable $\boldsymbol{\mu}_k$. This will leave the likelihood in Eq. (7.1) constant and consequently result in flat fit directions which leave the parameter space unconstrained. One would like to choose operators with expected flat directions different from those of other analyses. This would benefit all analyses in a prospective combined EFT effort. Flat fit directions are difficult to know a priori. However, the relative contribution of four-quark operators to $t\bar{t}$ production cross sections at leading order QCD has been identified for charge symmetric and asymmetric observables [266]. Figure 7.3 shows how the combination of a these observable types for a particular subset of Wilson coefficients can be complementary. Another categorization of observables can be made based on the physics processes they originate from. Figure 7.4 shows how $t\bar{t}$ observables can be complementary to observables from boosted $t\bar{t}$, $t\bar{t}Z$ or Higgs production for certain Wilson coefficients.
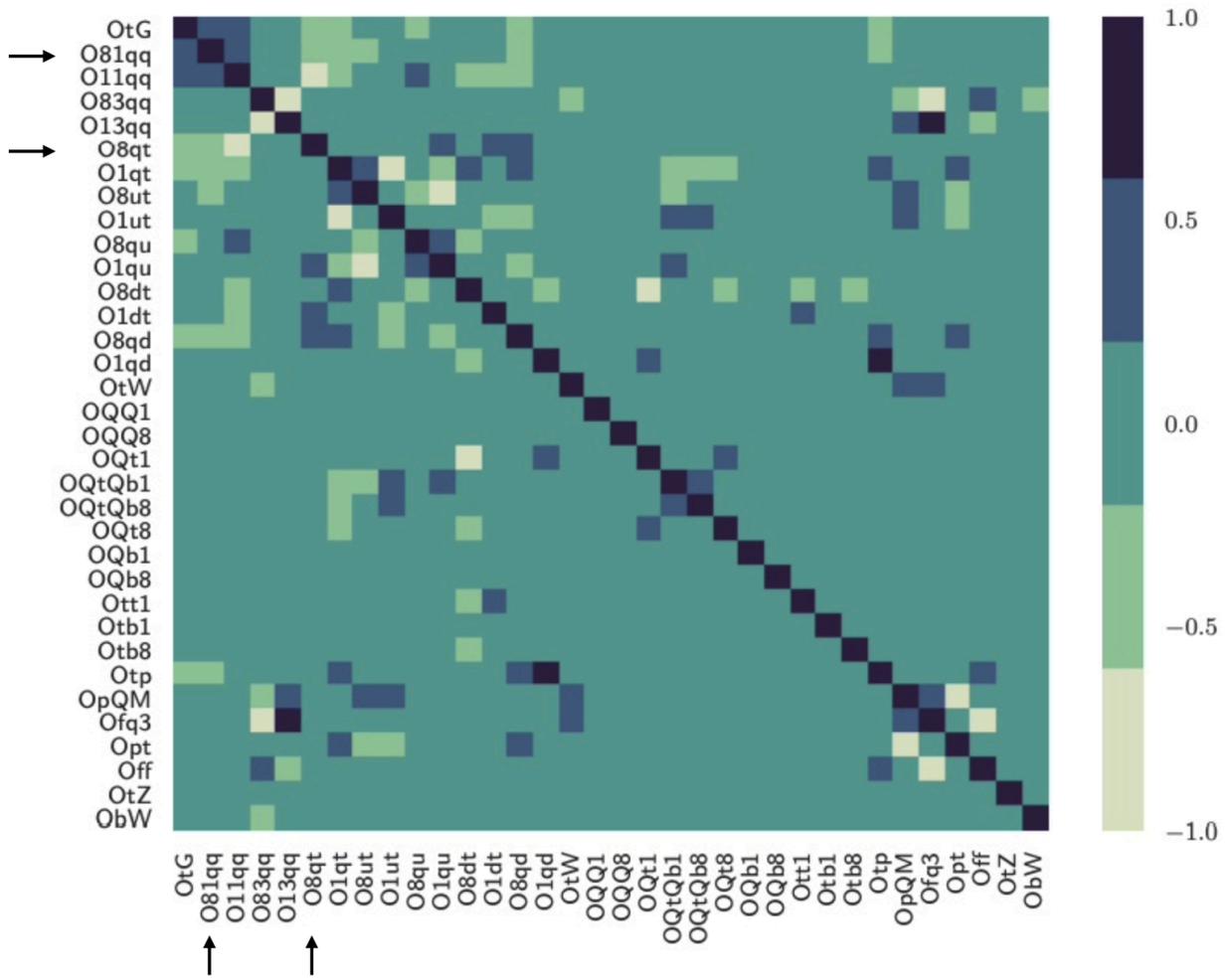
Figure 7.2: Estimates of the correlation coefficients between the 34 SMEFT Wilson coefficients. From [92].
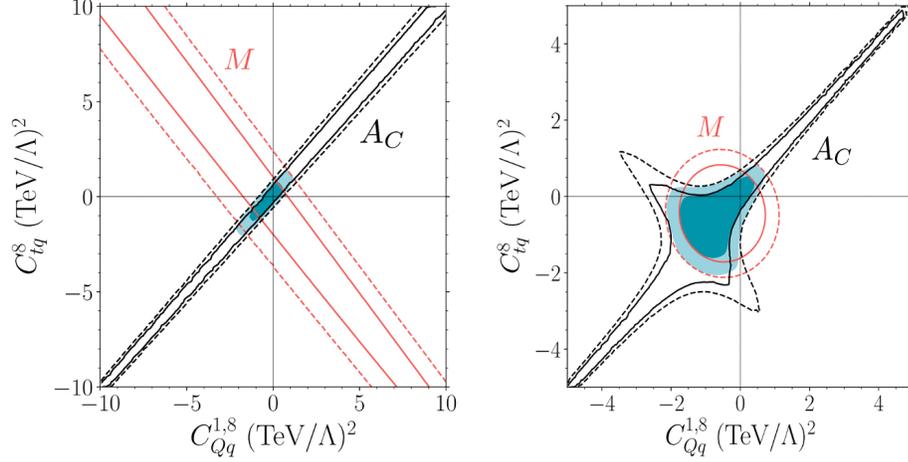
Figure 7.3: 2D simultaneous fit of the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients to an $t\bar{t}$ invariant mass differential cross section (charge symmetric (M)) and charge asymmetry measurement (A_C). The solid and dashed line represent the 68% and 95% confidence bounds, respectively. SMEFT predictions were estimated at LO QCD up to linear Λ^{-2}(left) and quadratic Λ^{-4}(right) order. From [266].
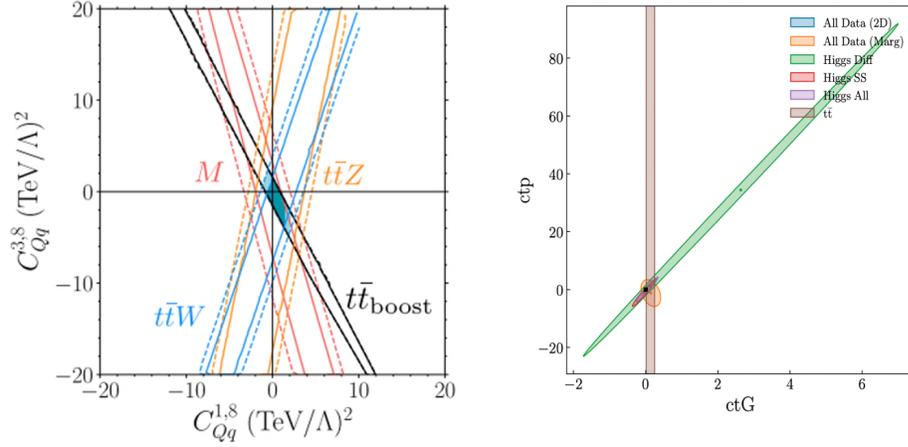
Figure 7.4: 2D simultaneous fits of the $c_{Qq}^{8,3}$ and $c_{Qq}^{8,1}$ and the $c_{t\phi}$ and c_{tG} Wilson coefficients. The $c_{Qq}^{8,3}$-$c_{Qq}^{8,1}$ fit is done with a $t\bar{t}$ invariant mass differential cross section and total rates of $t\bar{t}Z$ and $t\bar{t}W$. The $c_{t\phi}$-c_{tG} fit is done to many variables such as signal strengths(SS), differential cross sections and total rates from Higgs, top and diboson physics processes. The solid and dashed line in the left plot represent the resp. 68% and 95% confidence bounds and the lines in the right plot represent the 95% confidence bounds. SMEFT predictions were estimated up to linear Λ^{-2} order at LO(left) and NLO(right) QCD. From [266, 265].
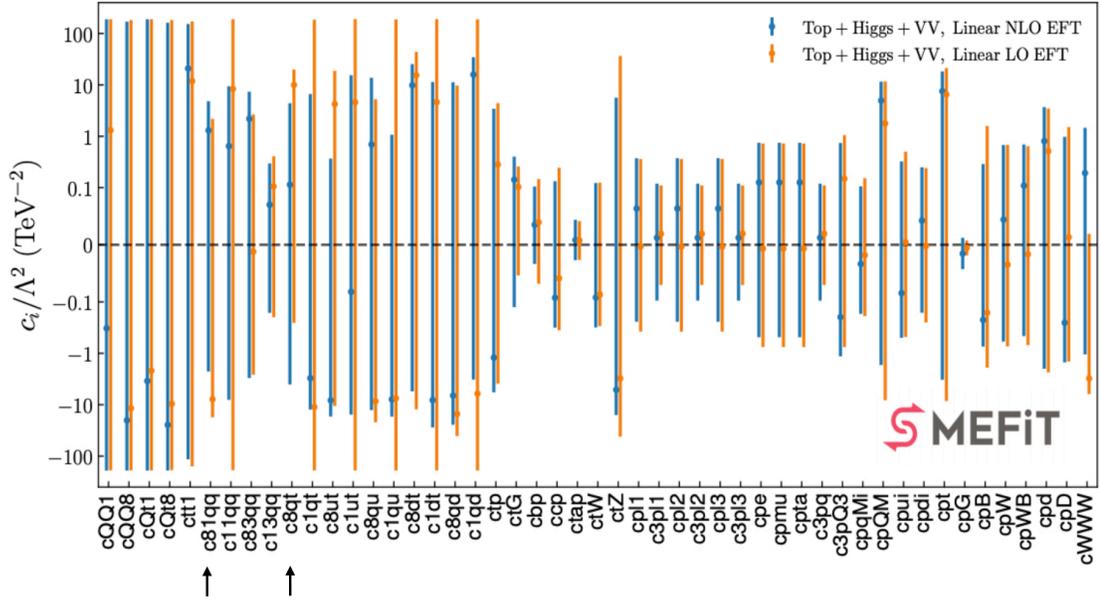
Figure 7.5: 95 % confidence intervals from a global fit on top, Higgs and diboson data taken at the LHC using SMEFT predictions up to the linear Λ^{-2} order for LO and NLO QCD. From [265].
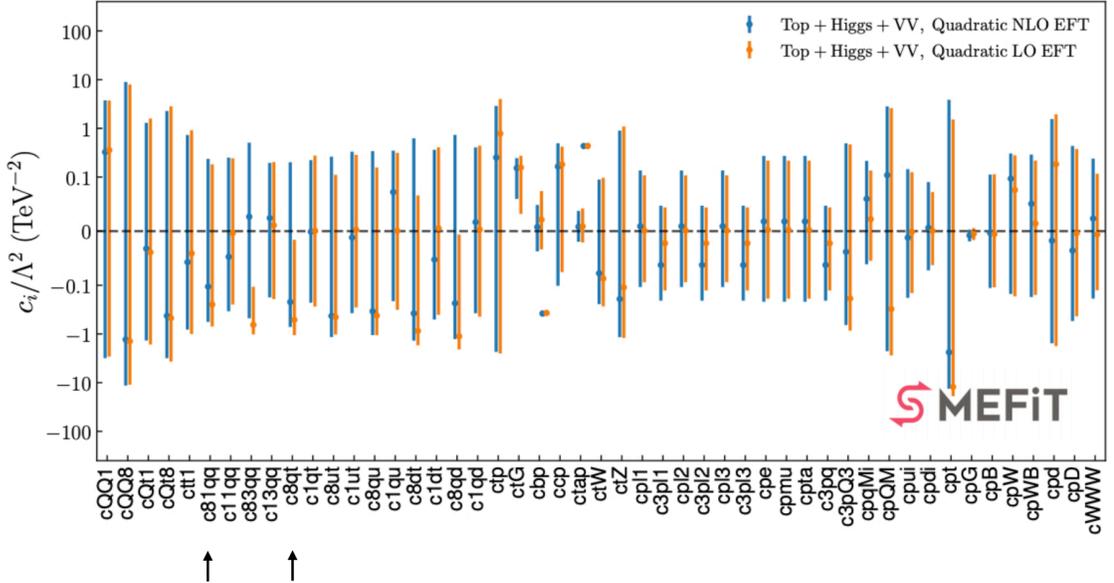
Figure 7.6: 95 % confidence intervals from a global fit on top, Higgs and diboson data taken at the LHC using SMEFT predictions up to the quadratic Λ^{-4} order for LO and NLO QCD. From [265].

A last way to motivate the choice of SMEFT operators is by reviewing recent published confidence bounds set on SMEFT Wilson coefficients in global fits. Preferably one would choose a set of Wilson coefficients that are not already heavily constrained and to which the available observables can improve on. Two notable publications can be found in [92, 266] which present global fits with top quark datasets taken at the LHC. However, the global fit with the largest set of Wilson coefficients that also includes LHC Higgs and diboson datasets is presented in [265]. The confidence bounds are shown in Fig. 7.5 and 7.6 and were set using different SMEFT predictions i.e. EFT orders and QCD accuracy. One can see from the figures that the inclusion of the quadratic $\Lambda^{-4}$ EFT order substantially reduces the size of the confidence bounds for most Wilson coefficients. Also notice that the inclusion of NLO QCD SMEFT predictions can both increase and decrease the confidence bounds substantially in some cases.

Based on the above considerations the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients were chosen. Figure 7.1 and 7.2 showed that they both contribute to $t\bar{t}$ production in a correlated way. Additionally, EFT interpretation combination plans are being made within the ATLAS Top community to improve constraints and explore methods needed for combinations. The c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients could provide a good candidate combination of $t\bar{t}$ differential cross sections and charge asymmetries as is shown in 7.3. This would be more feasible than a combination with other LHC physics groups (e.g. Higgs or diboson) or experiments (e.g. CMS) as the ATLAS Top community charge asymmetry and differential cross section analysts already work closely together. Also, $t\bar{t}$ differential cross sections and charge asymmetries use $t\bar{t}$ events recorded with the ATLAS detector which will result in many shared systematic uncertainty definitions. This would make the proper handling of systematic uncertainties and correlations, which is a non-trivial task, easier. Lastly, the global fit presented in Fig. 7.5 and 7.6

show that these two Wilson coefficients are not already heavily constrained by other datasets.

7.4.3 Parton level distributions

This section gives an indication of the sensitivity of $t\bar{t}$ differential cross sections to changes in the c_{tq}^8 and $c_{Qq}^{8,1}$ four-quark Wilson coefficients. One can generate events with MADGRAPH5_AMC@NLO and analyze them with MADANALYSIS5 [267, 268, 269, 270, 271, 272]. Not including parton shower and hadronization will provide only a proxy for the EFT effects on particle level observables but for significant reduced computing time. In combination with event reweighting [273] one could easily explore the effects of many other Wilson coefficients on total cross sections and the kinematics of generated events. Figures 7.7, 7.8 and 7.9 show comparisons of histograms of the top quark transverse momentum filled with events generated with SMEFT Monte Carlo setup presented in Sec. 7.3. The histograms denoted as "SM" are SMEFT predictions with all Wilson coefficients set to zero. One notable feature in all SMEFT predicitons is the increase of sensitivity for higher energies indicated by the large EFT-SM ratios of the high-p_T bins.

7.5 SMEFT Monte Carlo Samples

This section summarizes the particle level $t\bar{t}$ Monte Carlo samples that were created with the setup presented in Sec. 7.3. Events were generated, parton showered and hadronized with $\Lambda = 1\,\mathrm{TeV}^{-2}$ and different values for the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients (See Table 7.2). All SMEFT Monte Carlo samples were subsequently subjected to the particle level object and event selection criteria presented in Sec. 6.4 using RIVET 3.1 [274] and the validated Rivet analysis ATLAS_2019_I1750330 [275] of the 2019 $t\bar{t}$ differential cross section measurements in the l+jets decay channel with an

integrated luminosity of $L_{\text{int.}} = 36\,\text{fb}^{-1}$ at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ [276]. The analyzed events are used to fill histograms of the observables presented in Ch. 6. These histograms are then used in the quadratic regression method of Sec. 7.2.2 to parameterize the particle level observables in terms of the Wilson coefficients. Figure 7.10, 7.11 and 7.12 show some examples of the fitted functions and the SMEFT predictions used in the fit.

c_{tq}^8	$c_{Qq}^{8,1}$
1.0	0.0
0.0	1.0
-1.0	0.0
0.0	-1.0
0.5	0.5
-0.5	0.5
0.5	-0.5
-0.5	-0.5

Table 7.2: Points in Wilson coefficient parameter space at which SMEFT particle level samples were created. All samples were created at both LO and NLO QCD accuracy including both the linear Λ^{-2} and quadratic Λ^{-4} EFT order. Only the first 4 rows were also generated including the linear Λ^{-2} EFT order only.
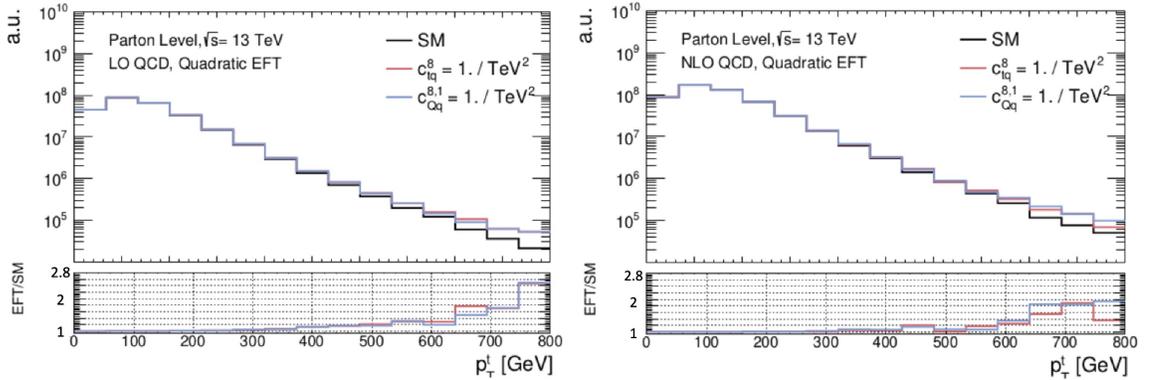
Figure 7.7: Parton level distributions of SMEFT predictions at LO (left) and NLO (right) QCD including both the linear Λ^{-2} and quadratic Λ^{-4} EFT order with the c_{tq}^8 and $c_{Qq}^{8,1}$ four-quark Wilson coefficients set to 1.

Figure 7.8: Parton level distributions of SMEFT predictions at LO (left) and NLO (right) QCD comparing the inclusion of the linear Λ^{-2} EFT order only(Lin. EFT) and with the inclusion of both the linear Λ^{-2} and quadratic Λ^{-4} EFT order(Quad. EFT) with the $c_{Qq}^{8,1}$ four-quark Wilson coefficient set to 1.
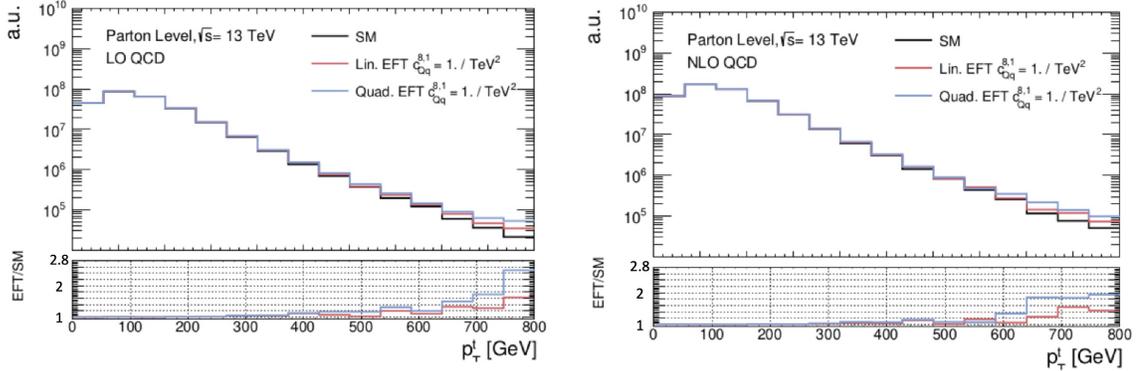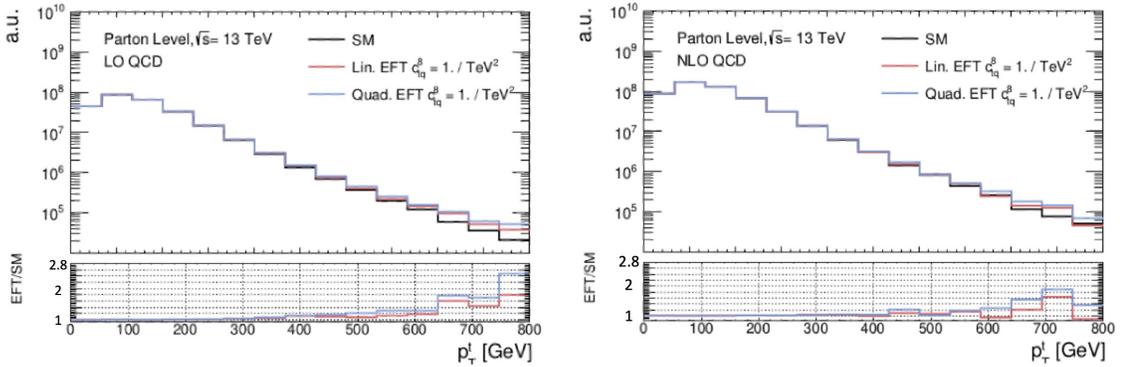

Figure 7.9: Parton level distributions of SMEFT predictions at LO (left) and NLO (right) QCD comparing the inclusion of the linear Λ^{-2} EFT order only(Lin. EFT) and with the inclusion of both the linear Λ^{2} and quadratic Λ^{-4} EFT order(Quad. EFT) with the c_{tq}^{8} four-quark Wilson coefficient set to 1.
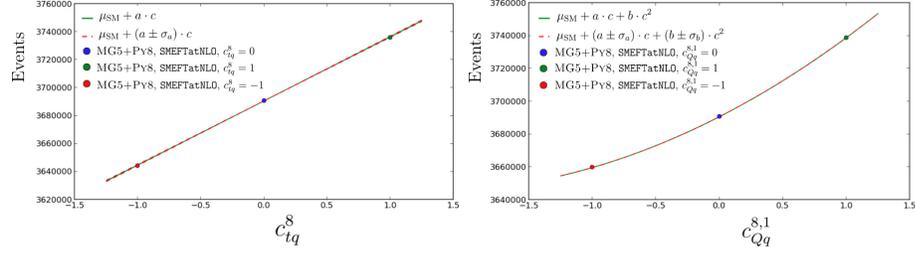
Figure 7.10: The particle level total rate predictions for various points in Wilson coefficient parameter space and the fitted function with the parameters evaluated at their best-fit and $\pm\sigma$ value. The predictions are generated at LO QCD with the linear Λ^{-2} EFT order only (left) and NLO QCD with both the linear Λ^{-2} and quadratic Λ^{-4} EFT order included (right).
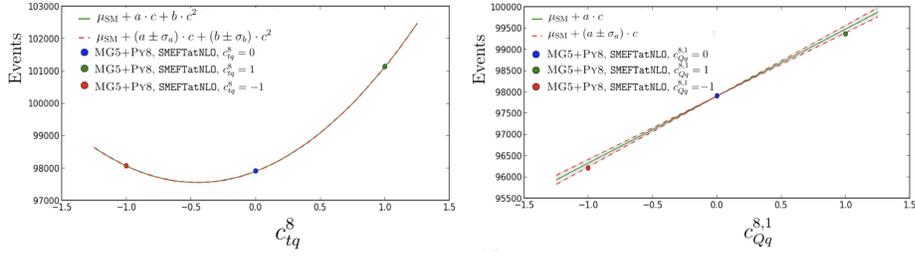
Figure 7.11: The particle level $H_T^{t\bar{t}}$ predictions for various points in Wilson coefficient parameter space and the fitted function with the parameters evaluated at their best-fit and $\pm\sigma$ value. The predictions are for the 12th bin ($575 < H_T^{t\bar{t}} < 650$ GeV) generated at NLO QCD with both the linear Λ^{-2} and quadratic Λ^{-4} EFT order included (left) and 3th bin ($100 < H_T^{t\bar{t}} < 150$ GeV) generated at LO QCD with the linear Λ^{-2} EFT order only included (right).
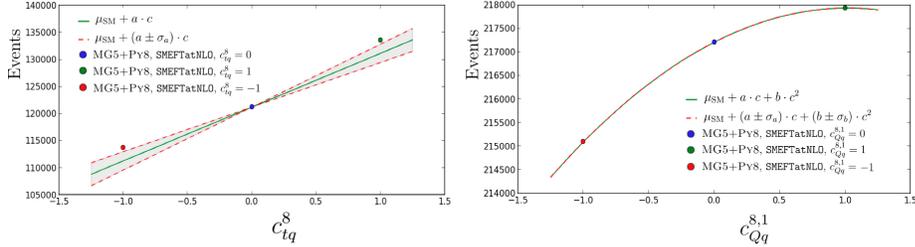
Figure 7.12: The particle level $|y^{t\bar{t}}|$ predictions for various points in Wilson coefficient parameter space and the fitted function with the parameters evaluated at their best-fit and $\pm\sigma$ value. The predictions are for the 9th bin ($0.8 < H_T^{t\bar{t}} < 0.9$) generated at NLO QCD with the linear Λ^{-2} EFT order only included (left) and 15th bin ($1.4 < H_T^{t\bar{t}} < 1.65$) generated at LO QCD with both the linear Λ^{-2} and quadratic Λ^{-4} EFT order included (right).

7.6 Regularization Effects

This section presents how the regularization of an unfolding framework can have effect on the Wilson coefficient confidence bounds. We use the SMEFT "SM" particle level prediction of the $p_T^{t\bar{t}}$, $m^{t\bar{t}}$, $H_T^{t\bar{t}}$ and $|y^{t\bar{t}}|$ distributions instead of unfolded data as input for the likelihood of Eq. (7.1) ($\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}_{\mathrm{SM}}$). The SMEFT "SM" particle level prediction is SMEFT at NLO QCD including all EFT orders and all Wilson coefficients set to 1×10^{-5}, as advised by the authors of the `SMEFTatNLO` model [277]. The covariance matrices are estimated with the same set of statistical and systematic sources of uncertainty and the frequentist-bayes hybrid pseudo-experiment method as introduced in Ch. 6. The unfolding for the covariance estimation is done with the Richardson-Lucy (Bayes) and the Poisson-likelihood with Tikhonov regularization (Poisson) algorithms for different values for the regularization parameters.

The covariance matrices and "SM" particle level prediction are used as input for the likelihood of Eq. (7.1). This is to isolate the effects of the covariance matrix and simulate an ideal unfolding scenario. The quadratic regression was done with NLO QCD SMEFT predictions including all EFT orders. The likelihood is then maximized w.r.t. the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients. One can use the maximum likelihood and Eq. (7.1) to define $\Delta \log L = 2(\log L_{\mathrm{max}} - \log L)$. By varying the values of the Wilson coefficients one can find the values for $\Delta \log L$ that equate to confidence intervals for the Wilson coefficients. Following Table 7.1 for $n = 2$ we can see that confidence intervals at a confidence level of 95% can be found at values $\Delta \log L = 5.99$.

Figure 7.13: The log-likelihood difference $\Delta \log L$ with $L_{\max}$ maximized w.r.t. the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients simultaneously. The $\Delta \log L$ is scanned for values of c_{tq}^8 with $c_{Qq}^{8,1}$ kept at its maximum-likelihood estimate. The likelihoods were evaluated with Asimov $H_T^{t\bar{t}}$ unfolded distributions and covariance estimates for various amounts of regularization.

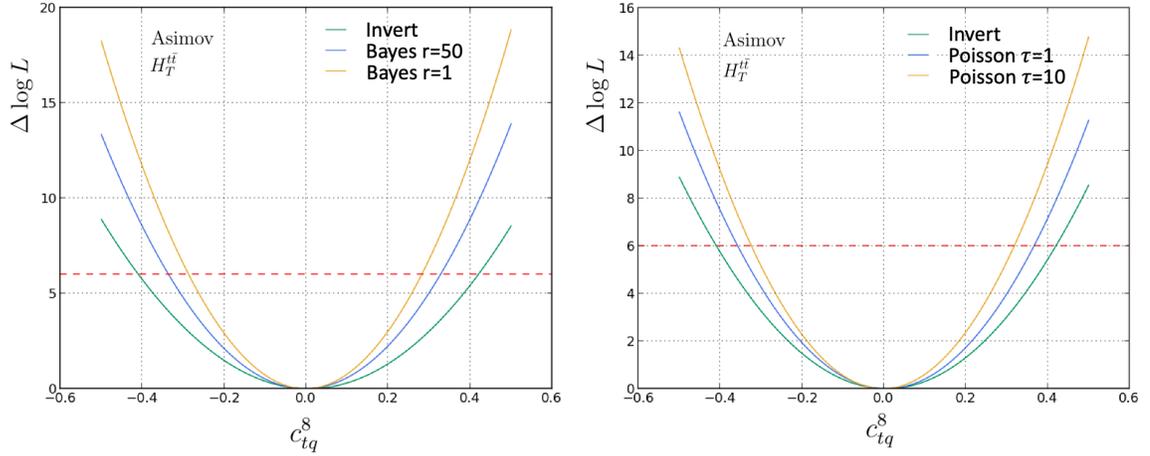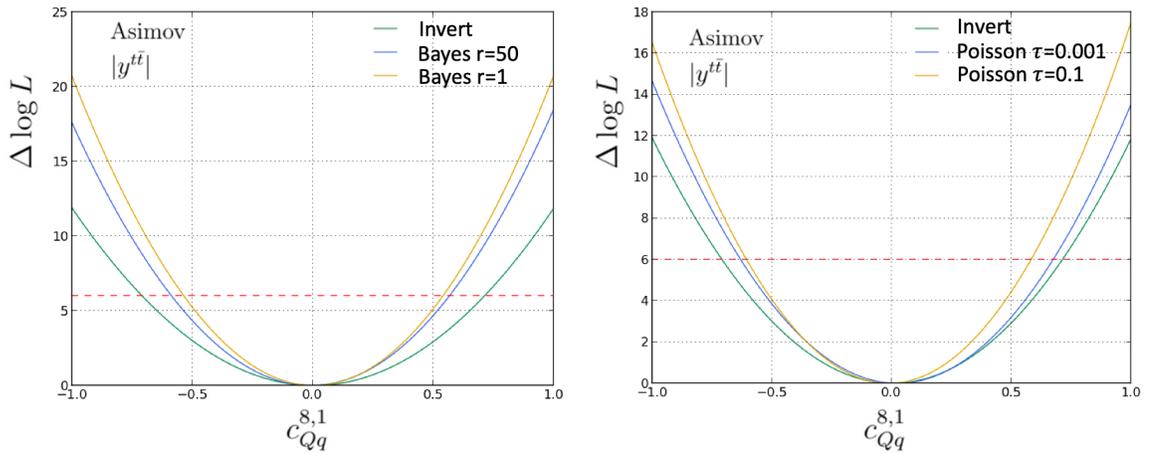Figure 7.14: The log-likelihood difference $\Delta \log L$ with $L_{\max}$ maximized w.r.t. the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients simultaneously. The $\Delta \log L$ is scanned for values of $c_{Qq}^{8,1}$ with c_{tq}^8 kept at its maximum-likelihood estimate. The likelihoods were evaluated with Asimov $|y^{t\bar{t}}|$ unfolded distributions and covariance estimates for various amounts of regularization.

Figure 7.13, 7.14 and 7.15 show clearly that with increasing regularization, i.e. $r \to 1$ for Richardson-Lucy and $\tau \to \infty$ for Poisson-likelihood with Tikhonov regularization, the confidence intervals for the same confidence level become smaller. This can be explained by the reduction of the variances that is characteristic for increased regularization. The smaller variances will results in bigger changes in the $\log L$ for the same change in the Wilson coefficients. This shows that one should be aware that the size of confidence intervals set with unfolded data depends on the amount of regularization introduced in the unfolding framework.
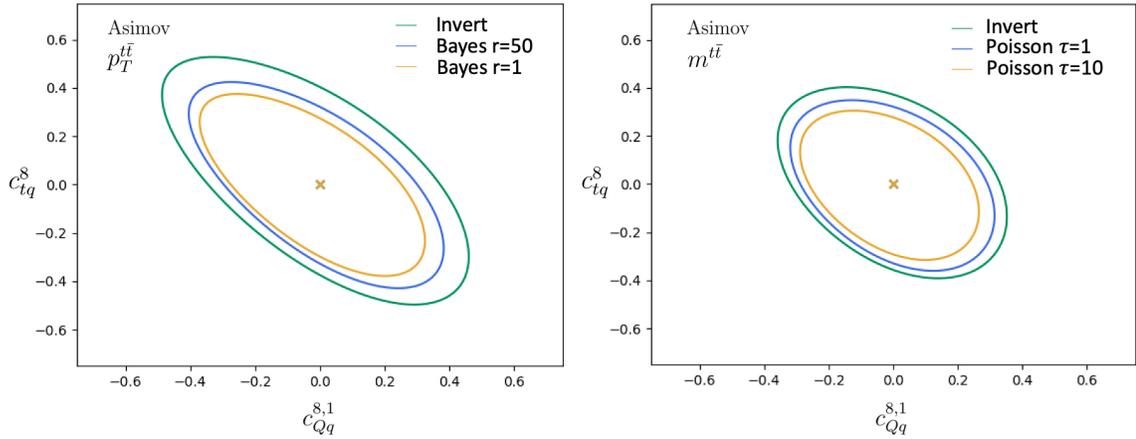
Figure 7.15: The confidence regions at a confidence level of 95% for the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients. The likelihoods were evaluated with Asimov $p_T^{t\bar{t}}$(Left) and $m^{t\bar{t}}$(Right) unfolded distributions and covariance estimates for various amounts of regularization.

7.7 Results

This section presents the estimates and confidence bounds set on the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients using the Full Run-2 dataset corresponding to an integrated luminosity of $L = 139\,\mathrm{fb}^{-1}$. The likelihoods were evaluated with data unfolded with the matrix inversion method to avoid the confidence bounds having any regularization

dependencies. The covariance matrices were estimated with the frequentist-bayes hybrid pseudo-experiment method including all statistical and systematic sources of error presented in Ch. 6. No theoretical uncertainties were included. Individual fits vary only one Wilson coefficient whilst keeping the other at its SM value 0. Two-parameter fits vary both Wilson coefficients simultaneously.

Figure 7.16 and 7.17 compares the unfolded data with several NLO QCD particle level predictions including pre-fit ($\boldsymbol{\mu}_{\mathrm{SM}}$) and post-fit ($\boldsymbol{\mu}(\hat{c}_{tq}^8, \hat{c}_{Qq}^{8,1})$) SMEFT predictions. Figure 7.18 and 7.19 give confidence regions of the two-parameter fits. Table 7.4 and Fig. 7.20 give the confidence bounds from the individual and two-parameter fits using both LO and NLO QCD and linear and quadratic EFT order SMEFT predictions. A clear correlation can be seen in all the two-parameter confidence regions. This also is shown in Table 7.4 with confidence bounds of the two-parameter fit being larger than the ones set by the individual fits. The $H_T^{t\bar{t}}$ observable sets the most stringent constraints whilst the $|y^{t\bar{t}}|$ observable leaves the Wilson coefficients the least constrained. Lastly, one can see that inclusion of NLO QCD and quadratic Λ^{-4} EFT orders often deflates the confidence bounds.

QCD	Observables			
	$p_T^{t\bar{t}}$	$H_T^{t\bar{t}}$	$m^{t\bar{t}}$	$|y^{t\bar{t}}|$
LO	-0.342	-0.425	-0.562	-0.103
NLO	-0.482	-0.614	-0.223	-0.098

Table 7.3: Correlation coefficient estimates $\rho\left[\hat{c}_{tq}^8, \hat{c}_{Qq}^{8,1}\right]$ from the two-parameter fits using both LO and NLO QCD SMEFT predictions including both linear Λ^{-2} and quadratic Λ^{-4} EFT orders.
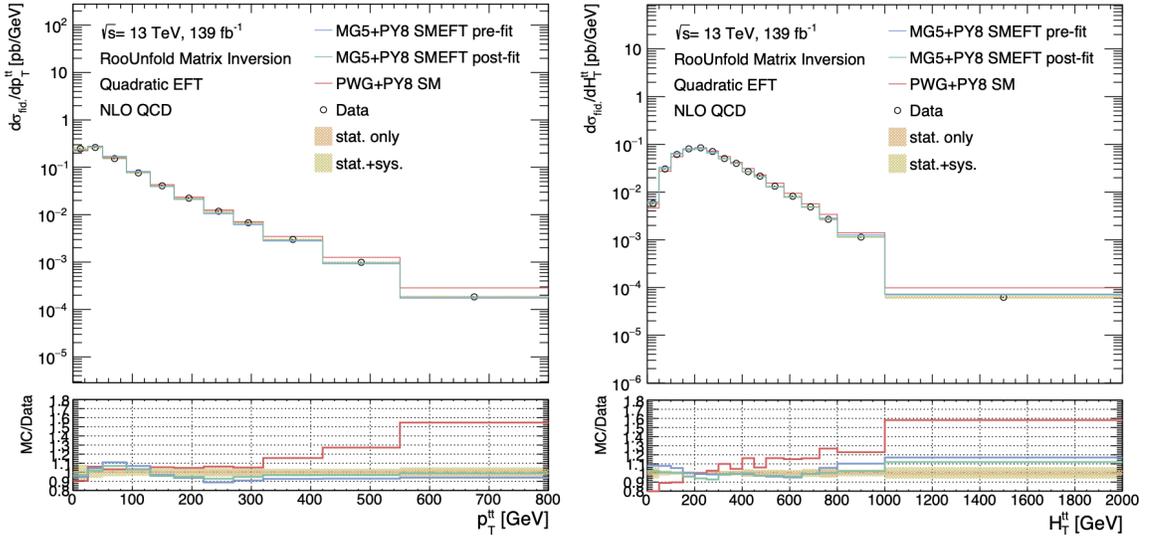
Figure 7.16: Comparison of the unfolded data including statistical and systematic uncertainties, Powheg-Box v2 SM, MadGraph5_aMC@NLO SMEFT pre-fit and MadGraph5_aMC@NLO SMEFT post-fit predictions for the $p_T^{t\bar{t}}$ and $H_T^{t\bar{t}}$ observable. The fit uses NLO QCD SMEFT predictions including both the linear and quadratic EFT order.
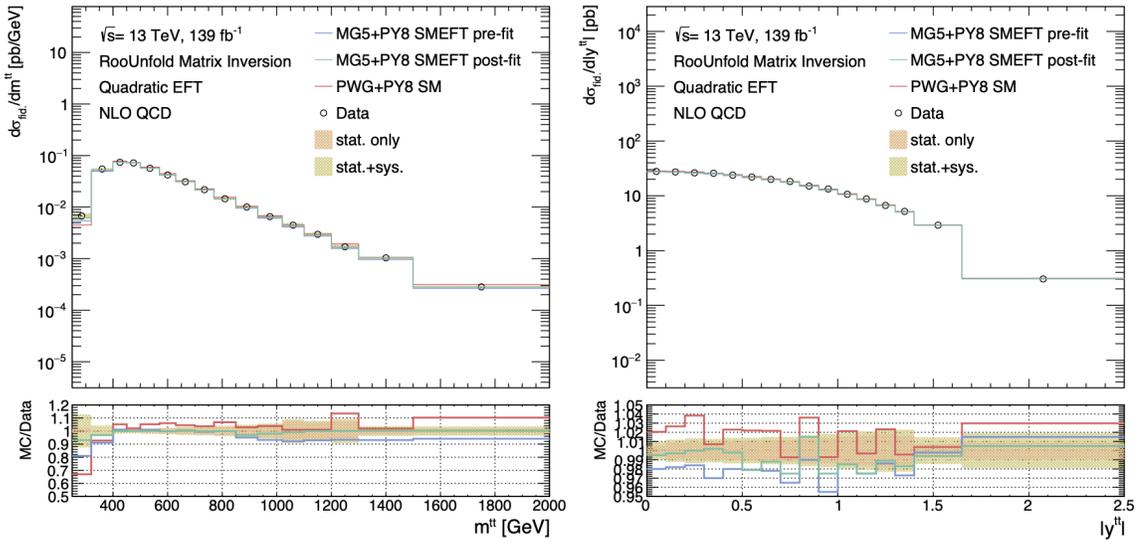
Figure 7.17: Comparison of the unfolded data including statistical and systematic uncertainties, Powheg-Box v2 SM, MadGraph5_aMC@NLO SMEFT pre-fit and MadGraph5_aMC@NLO SMEFT post-fit predictions for the $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$ observable. The fit uses NLO QCD SMEFT predictions including both the linear and quadratic EFT order.
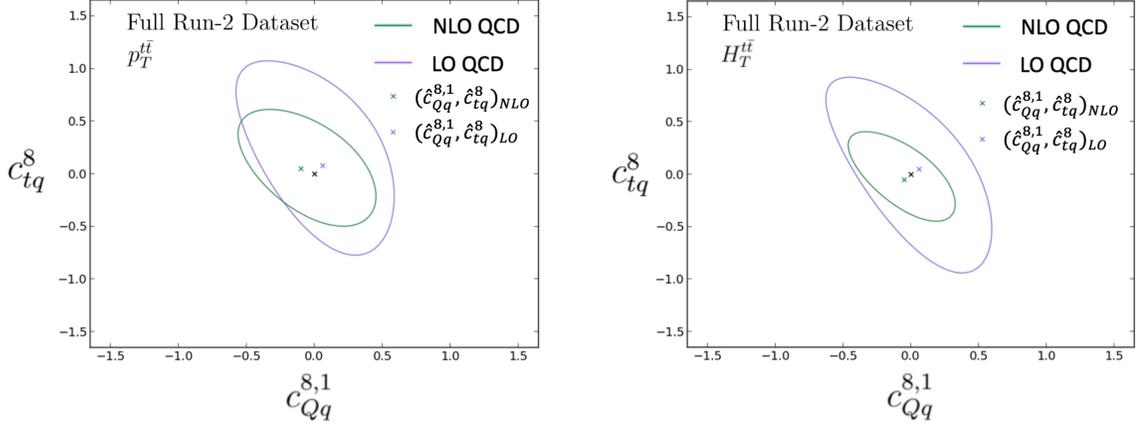
Figure 7.18: The confidence regions at a confidence level of 95% for the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients. The likelihoods were evaluated with the $p_T^{t\bar{t}}$(Left) and $H_T^{t\bar{t}}$(Right) unfolded distributions and covariance estimates using the matrix inversion method and the Full Run-2 dataset. The parameterization of the likelihood is constructed with the quadratic regression to either LO or NLO QCD SMEFT predictions.
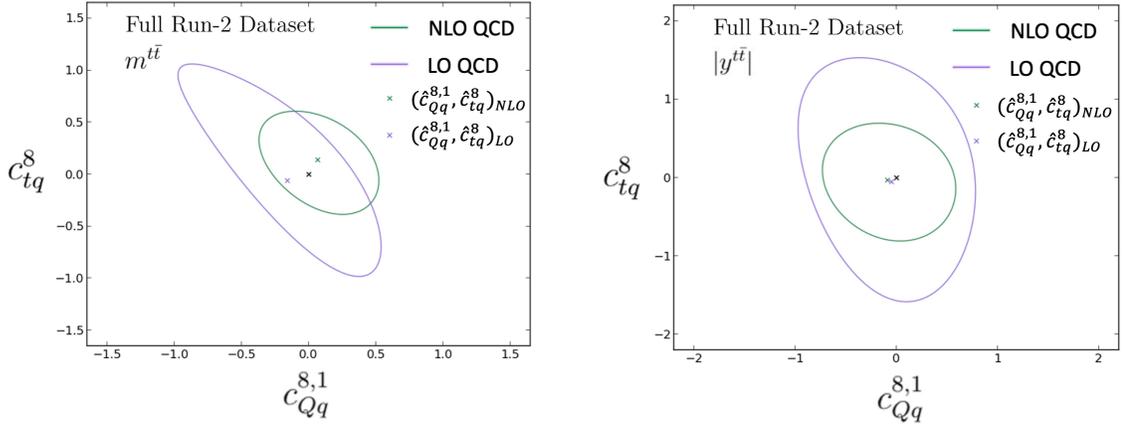
Figure 7.19: The confidence regions at a confidence level of 95% for the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients. The likelihoods were evaluated with the $m^{t\bar{t}}$(Left) and $|y^{t\bar{t}}|$(Right) unfolded distributions and covariance estimates using the matrix inversion method and the Full Run-2 dataset. The parameterization of the likelihood is constructed with the quadratic regression to either LO or NLO QCD SMEFT predictions.

Parameter	QCD	Observable	Individual $\mathcal{O}(\Lambda^{-2})$ EFT MLE	95% CL Range	$\mathcal{O}(\Lambda^{-4})$ EFT MLE	95% CL Range	Two-parameter $\mathcal{O}(\Lambda^{-4})$ EFT MLE	95% CL Range
c_{tq}^8	LO	$p_T^{t\bar{t}}$	1.454	[-2.273,3.441]	0.012	[-0.687,0.919]	0.066	[-0.773,1.07]
		$H_T^{t\bar{t}}$	-0.848	[-1.377,0.799]	-0.0312	[-0.61,0.473]	0.0493	[-0.747,0.73]
		$m^{t\bar{t}}$	1.981	[-3.112,5.44]	0.0109	[-0.584,0.679]	-0.0702	[-0.988,1.057]
		$|y^{t\bar{t}}|$	-2.333	[-10.637,8.201]	0.021	[-1.558,1.455]	-0.099	[-1.604,1.535]
	NLO	$p_T^{t\bar{t}}$	0.203	[-0.715,1.031]	0.0391	[-0.37,0.533]	0.0543	[-0.516,0.602]
		$H_T^{t\bar{t}}$	-0.936	[-1.436,0.801]	0.016	[-0.352,0.275]	-0.0518	[-0.464,0.404]
		$m^{t\bar{t}}$	0.173	[-0.89,1.09]	0.0789	[-0.344,0.567]	0.146	[-0.395,0.602]
		$|y^{t\bar{t}}|$	-3.414	[-8.724,6.045]	-0.0577	[-0.802,0.676]	-0.0612	[-0.825,0.687]
$c_{Qq}^{8,1}$	LO	$p_T^{t\bar{t}}$	2.777	[-3.03,5.786]	0.022	[-0.399,0.552]	0.061	[-0.578,0.585]
		$H_T^{t\bar{t}}$	0.359	[-0.311,1.621]	-0.0155	[-0.286,0.319]	0.0628	[-0.425,0.413]
		$m^{t\bar{t}}$	-0.514	[-1.51,0.907]	0.0107	[-0.519,0.366]	-0.1781	[-0.978,0.546]
		$|y^{t\bar{t}}|$	0.523	[-7.031,7.555]	-0.0205	[-0.905,0.772]	-0.0877	[-0.976,0.781]
	NLO	$p_T^{t\bar{t}}$	0.286	[-0.711,1.244]	0.002	[-0.486,0.393]	-0.0898	[-0.566,0.459]
		$H_T^{t\bar{t}}$	-0.191	[-1.31,1.513]	0.0296	[-0.353,0.259]	-0.048	[-0.466,0.326]
		$m^{t\bar{t}}$	0.267	[-0.947,0.609]	0.083	[-0.339,0.486]	0.0607	[-0.373,0.527]
		$|y^{t\bar{t}}|$	-2.654	[-5.099,3.193]	-0.0513	[-0.79,0.686]	-0.103	[-0.802,0.736]

Table 7.4: The maximum likelihood estimators (MLE) and confidence intervals at a confidence level (CL) of 95% from individual and two-parameter fits using LO and NLO QCD SMEFT predictions with up to linear Λ^{-2} or quadratic Λ^{-4} EFT order.

Wilson Coefficient	Global Fit Results Hartland et al.(2019) [92]	Brivio et al.(2019) [278]	Ethier et al.(2021) [265]
c_{tq}^8	$[-3.7, 4.1]$	$[-1.32, 0.44]$	$[-0.687, 0.186]$
$c_{Qq}^{8,1}$	$[-4.7, 7.8]$	$[-1.00, 0.52]$	$[-0.555, 0.236]$

Table 7.5: Recently published confidence intervals at a confidence level of 95% from global fits to top, Higgs and diboson datasets collected at the LHC and LEP. All use NLO QCD SMEFT predictions including both the linear $\mathcal{O}(\Lambda^{-2})$ and quadratic $\mathcal{O}(\Lambda^{-4})$ EFT order.
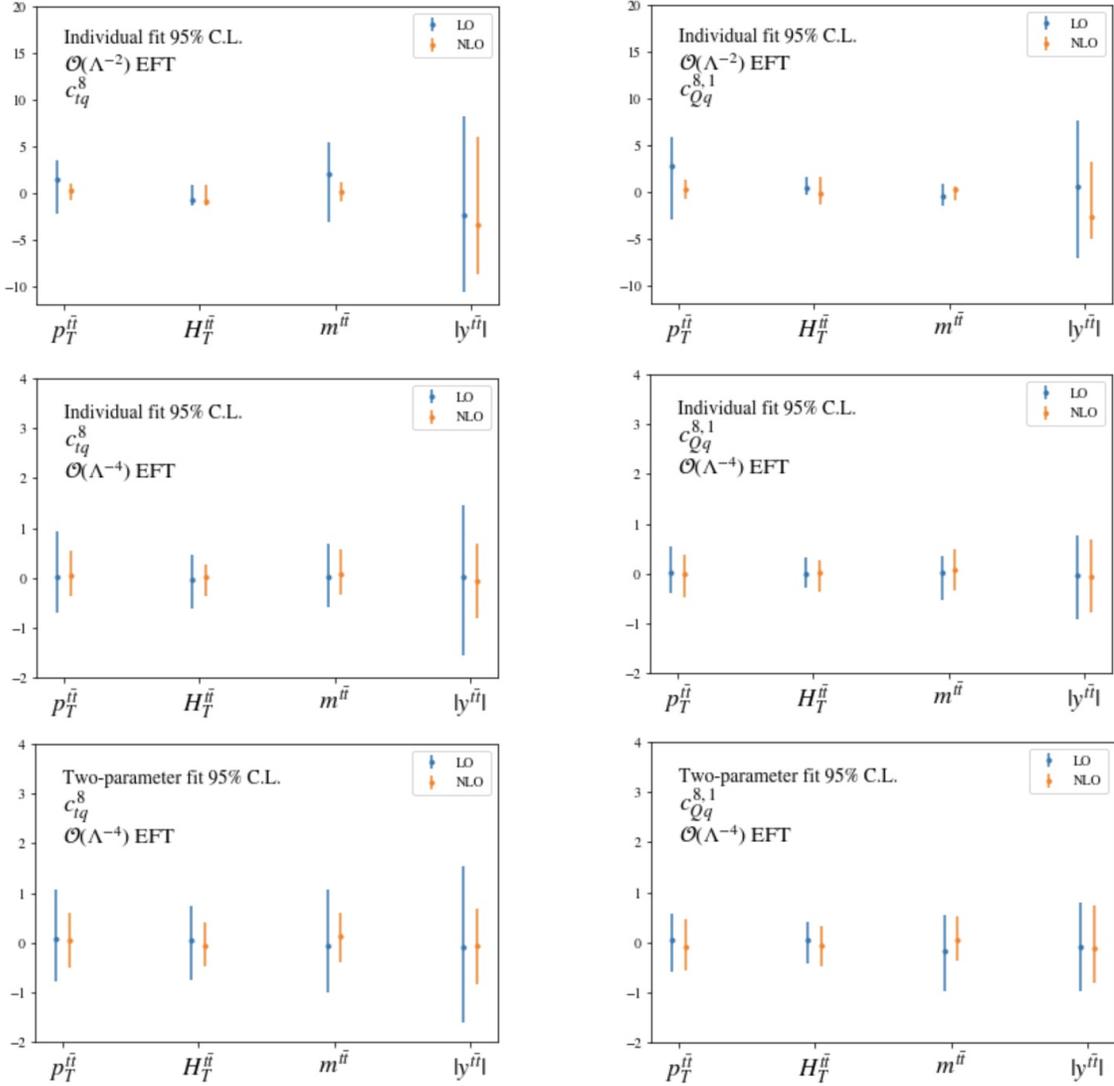
Figure 7.20: The maximum likelihood estimators (MLE) and confidence intervals at a confidence level (CL) of 95% from individual and two-parameter fits using LO and NLO QCD SMEFT predictions with up to linear Λ^{-2} or quadratic Λ^{-4} EFT order.

7.8   Discussion

The confidence bounds set on the c_{tq}^8 and $c_{Qq}^{8,1}$ Wilson coefficients of the Standard Model Effective Field Theory show no significant disagreement with the Standard Model. When confidence bounds in Table 7.4 are compared with the global fit results of Table 7.5 we can see that the sizes of the confidence bounds are comparable. However, one should note that all three global fits used considerably more observables,

included more parameters in the fit and included theoretical systematic uncertainties. The results also show that the inclusion of NLO QCD and quadratic Λ^{-4} EFT orders or the amount of regularization can have a significant impact on the confidence bounds. However, there are still many prospects for future studies. Future efforts should therefore try to include NLO QCD and quadratic Λ^{-4} EFT orders and avoid regularization whenever they can.

EFT effects in background processes have been neglected which can be justified with high signal-to-background ratios. However, many analyses or observables have much more sizable background contributions for which this assumption does not hold. Future studies should therefore explore the possibility of not subtracted the background from data before unfolding and performing a combined process EFT interpretation e.g. using $t\bar{t} + tW$ differential cross section measurements to constrain EFT parameter space. Another analysis adjustment for future EFT efforts should be closure tests with detector simulated EFT samples with non-SM Wilson coefficient values. This is to ensure that possible EFT effects in the data are not lost when applying object and event selection or the unfolding corrections.

Another notable improvement for future efforts would be to explore particle level MC sample reweighting techniques. Currently, MADGRAPH5_AMC@NLO supports reweighting techniques [273] for parton level MC samples. However, parton showering and hadronization need at least if not more computation time than matrix element calculation and event generation. The latest version of the particle level MC sample file format `HepMC3` [279] contains event weights which should make it conceptually possible. This would make the expansion of EFT efforts to higher-dimensional fits significantly easier. Also notable should be the use of Rivet routines which make it easy to reproduce particle level fiducial phase spaces for new particle level MC sam-

ples possibly needed in future EFT efforts.

One of the important hurdles in future EFT combinations is the proper handling of correlations between unfolded observables of different analyses. Recall that different analyses will have shared sources of error, e.g. ATLAS analyses will share detector systematics, which means that unfolded bins of two different analyses will be correlated. Combining the covariances of different analyses into one total covariance for a combined fit that includes all these correlations is a non-trivial task. One approach for this would be to estimate the covariance matrix contribution for each source of error for each analysis, estimate the covariances between bins of different analyses for each source of error and sum all the total covariances of each source of error into one final total covariance for the combined fit. However, very little work has been done that supports such an approach. Especially estimating covariances between different analyses for even a single source of error without the full datasets and ATLAS simulation machinery is not obvious and needs to be explored more.

Lastly, the estimates and confidence intervals of the c_{tq}^8 and $c_{Qq}^{8,1}$ SMEFT Wilson coefficients are used to quantify deviations from the SM in a model-independent way. However, one can use the constraints on a Wilson coefficient to constrain parameters of specific BSM theories such as SUSY or 2HDM. The procedure of converting Wilson coefficient constraints into constraints on UV theory parameters is also known as *matching*. It entails obtaining an effective theory of a specific UV theory by integrating out heavy fields, expressing the generated UV operators into the Warsaw basis and matching the UV parameters to the SMEFT Wilson coefficients. A simple toy scalar example of this process is given in A.6. However, in general, matching to specific UV theories and excluding BSM scenarios is a non-trivial task. The experimental HEP community has mostly been focusing on publishing estimates and

constraints on Wilson coefficients but should also eventually be able to include this final interpretation step. Especially in the case of significant deviations from the SM. Analyses are emerging that include or even solely dedicate to the matching of SMEFT to specific UV theories [280, 281] and even tools are being developed that facilitate this process [282, 283]. However, much more work on this final step of interpretation is needed and is therefore left for future studies.

Chapter 8

Conclusions

This thesis presented a study on novel unfolding algorithms, a top quark pair differential cross section measurement and a SMEFT interpretation including two Wilson coefficients. Most unfolding algorithms impose some form of regularization that reduces the variance but increases the bias of the constructed estimators. Optimizing the regularization parameters w.r.t. estimator characteristics such as the mean squared error or coverage probability can help to find the optimal bias-variance trade-off and compare different unfolding algorithms. This could be instructive for many analysis teams and make the choice of the regularization parameters less arbitrary. Additionally, novel statistical methods for estimating the response matrix and estimating the covariances of the unfolded distribution including statistical and systematic sources of error were presented. The former supplies a valid response matrix estimation method for analyses for which Monte Carlo simulation is computationally expensive. The frequentist-bayes hybrid pseudo experiment covariance estimation method shows good agreement with the more conventional frequentist pseudo experiment estimation method. However, the hybrid method can also be used by unfolding algorithms lacking a clear likelihood definition such as the iterative bayes unfolding algorithms. Some possible future studies could entail trying different likelihoods for the response

matrix estimation or investigating methods to estimate covariances between unfolded distributions after they have been unfolded separately.

Secondly, the measurement of $t\bar{t}$ differential cross section in the l+jets decay channel was presented. The differential cross sections were measured as a function of $p_T^{t\bar{t}}$, $m^{t\bar{t}}$, $H_T^{t\bar{t}}$ and $|y^{t\bar{t}}|$ and were unfolded to particle level. The covariances of the unfolded distributions were estimated with the frequentist-bayes hybrid pseudo-experiment method including various statistical and systematic sources of uncertainty. The SM predictions of POWHEG-BOX V2+PYTHIA8 overestimated many of the high-p_T differential cross section bins w.r.t. the unfolded data. Other top quark pair differential cross section measurements have seen this disagreement as well in roughly the same shape and magnitude. This fortifies the motivation to revise our models, detectors, reconstruction or analysis methods.

Finally, estimates and confidence intervals of the c_{tq}^8 and $c_{Qq}^{8,1}$ SMEFT Wilson coefficients were presented. The unfolded $t\bar{t}$ differential cross sections and covariances were used in a χ^2-fit to obtain maximum likelihood estimators and set 95% confidence intervals. The fit used both LO and NLO QCD SMEFT predictions including either up to linear $\mathcal{O}(\Lambda^{-2})$ or up to quadratic $\mathcal{O}(\Lambda^{-4})$ EFT order. The confidence intervals set on both were comparable in size with confidence intervals set in recently published global fits and showed no significant deviation from the Standard Model. Future studies could involve combining several top quark measurements, e.g. this one and a charge asymmetry measurement, to perform a combined fit. Investigation in estimating covariances between these separately unfolded distributions is crucial. Also, studies into reweighting techniques for Monte Carlo samples including hadronization and parton showering would be very valuable for future combined fits.

Appendix A

Theory Appendix

This appendix contains additional details and mathematics in support of Ch. 2.

A.1 Canonical Quantization

Quantization is a procedure of changing a classical theory to a quantum mechanical one. The *first quantization* is the change of Newtonian classical mechanics to non-relativistic quantum mechanics where a system is described by a quantum wave function $|\Psi(t)\rangle$ and their evolution is defined by time dependent Schrödingers equation[11].

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle \tag{A.1}$$

Additionally, one introduces quantum mechanical operators that correspond to a measurable quantity of the system. In Eq. (A.1) $\hat{H}$ denotes the Hamiltonian operator which generates the time evolution but also corresponds to the total energy of the system. Similarly there are operators for the position $\hat{x}$, momentum $\hat{p}$ or spin $\hat{S}_z$ that, when applied to the wave function $|\Psi(t)\rangle$, give all the possible values for the corresponding observable. A measurement will result in the observable taking on one of these possible outcomes[284, 285]. The *second quantization* is the procedure that

expands the above to many-particle systems described by *Fock states* and quantum operator valued fields.[286, 287]

The quantization procedures give a nice step-by-step transition from Newtonian mechanics to quantum field theory and is definitely crucial for a more general description of quantum field theory. However, the path integral narrative fits better in the context of this thesis as it couples easier to our observables such as the cross section.

A.2 n-point time-ordered correlation functions

With Eq. (2.10), Eq. (2.11) and Eq. (2.12) we can define a 2-point correlation function as

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)|0\rangle = \frac{1}{i}\Delta(x_2 - x_1') \tag{A.2}$$

If we want to generalize to $n > 2$ then we need to Taylor expand Eq. (2.10)[288] to get

$$Z_0[J] = 1 + \sum_{n=1}^{\infty} \frac{1}{n!}\left[\frac{-i}{2\hbar}\int d^4x d^4x' J(x)\Delta(x - x')J(x')\right]^n \tag{A.3}$$

For $n = 3$ we get

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)|0\rangle = 0 \tag{A.4}$$

and for $n = 4$

$$\begin{aligned}\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)|0\rangle = \frac{1}{i^2}\Big[&\Delta(x_1 - x_2)\Delta(x_3 - x_4)\\ &+\Delta(x_1 - x_3)\Delta(x_2 - x_4)\\ &+\Delta(x_1 - x_4)\Delta(x_2 - x_3)\Big]\end{aligned} \tag{A.5}$$

For the correlation functions of Eq. (A.2) and A.5 we get resp. Fig. A.1 and A.2.

Figure A.1: Feynman diagram of a Feynman propagator representing correlation function Eq. (A.2)
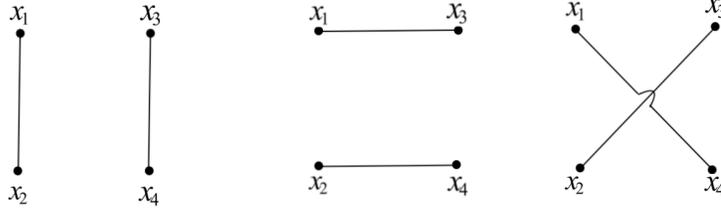
Figure A.2: Feynman diagrams representing correlation function Eq. (A.5)

Lets assume the weak perturbative lagrangian of Eq. (2.17) with the quartic self-interaction of Eq. (2.19). We can rewrite Eq. (2.18) as

$$
\begin{aligned}
Z[J] &= \exp\left[-\frac{i\lambda}{\hbar 4!}\int d^4x\left(\frac{\delta}{\delta J(x)}\right)^4\right]\int \mathcal{D}\phi \exp\left[i/\hbar\int d^4x(\mathcal{L}_0 + J\phi)\right] \\
&= \exp\left[-\frac{i\lambda}{\hbar 4!}\int d^4x\left(\frac{\delta}{\delta J(x)}\right)^4\right]Z_0[J] + ...
\end{aligned}
$$

(A.6)

with $Z_0[J]$ being the free scalar field functional of Eq. (2.9). If we assume λ is small and take $Z_0[J]$ equal to Eq. (2.10) then we can Taylor expand Eq. (A.6) i.e. create

a perturbation series as

$$Z[J] = Z_0[J] - \frac{i\lambda}{\hbar 4!} \int d^4x \left(\frac{\delta}{\delta J(x)} \right)^4 \Bigg|_{J=0} Z_0[J]$$
$$- \frac{1}{2} \left(\frac{i\lambda}{\hbar 4!} \right)^2 \int d^4x d^4y \left(\frac{\delta}{\delta J(x)} \right)^4 \left(\frac{\delta}{\delta J(y)} \right)^4 \Bigg|_{J=0} Z_0[J] + ...$$

(A.7)

with the λ- and λ^2-terms known as the *leading order(LO)* and *next-to-leading order(NLO)* terms. Lets recall the definition of the n-point correlation function.

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)...\phi(x_n)|0\rangle = \left(\frac{1}{i} \right)^n \frac{\delta}{\delta J(x_1)} \frac{\delta}{\delta J(x_2)} ... \frac{\delta}{\delta J(x_n)} Z[J] \Bigg|_{J=0}$$

(A.8)

Let's use Eq. (A.7) and A.8 to define a 4-point correlation function for the ϕ^4-interaction.

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)|0\rangle = \left(\frac{1}{i} \right)^4 \frac{\delta}{\delta J(x_1)} \frac{\delta}{\delta J(x_2)} \frac{\delta}{\delta J(x_3)} \frac{\delta}{\delta J(x_4)} Z_0[J] \Bigg|_{J=0}$$
$$- \frac{i\lambda}{\hbar 4!} \int d^4x \left(\frac{1}{i} \right)^4 \frac{\delta}{\delta J(x_1)} \frac{\delta}{\delta J(x_2)} \frac{\delta}{\delta J(x_3)} \frac{\delta}{\delta J(x_4)} \left(\frac{\delta}{\delta J(x)} \right)^4 Z_0[J] \Bigg|_{J=0}$$
$$- \frac{1}{2} \left(\frac{i\lambda}{\hbar 4!} \right)^2 \int d^4x d^4y \left(\frac{1}{i} \right)^4 \frac{\delta}{\delta J(x_1)} \frac{\delta}{\delta J(x_2)} \frac{\delta}{\delta J(x_3)} \frac{\delta}{\delta J(x_4)} \left(\frac{\delta}{\delta J(x)} \right)^4 \left(\frac{\delta}{\delta J(y)} \right)^4 Z_0[J] \Bigg|_{J=0} + ...$$

(A.9)

Now recall the Taylor expansion of $Z_0[J]$ of Eq. (A.3). This means that the 4-point correlation function of Eq. (A.9) is determined by the expansion orders of Eq. (A.3) and Eq. (A.7). Let's just look at the leading order term of Eq. (A.7) and up to 4-th order of Eq. (A.3). The 4-th order is needed to avoid the functional derivatives removing the interaction terms i.e. it is the minimum number of propagators $\Delta(x-y)$

needed for a 4-point interaction.

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)|0\rangle = \langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)|0\rangle_0$$
$$-\frac{i\lambda}{\hbar}\int d^4x \Delta(x_1-x)\Delta(x_2-x)\Delta(x_3-x)\Delta(x_4-x) + ... \tag{A.10}$$

with $\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)|0\rangle_0$ being the free scalar 4-point correlation function of Eq. (A.5). The new term of the correlation function can be represented by a Feynman diagram as well i.e. one with 4 propagators starting at x_1, x_2, x_3 and x_4 and 1 interaction vertex at x(Fig. A.3).
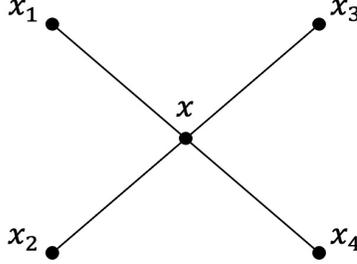

Figure A.3: A Feynman diagram representing the first interaction term of the 4-point correlation function in Eq. (A.10)

We can rewrite the interaction term of Eq. (A.10) by using Eq. (2.11).

$$\langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)|0\rangle = \langle 0|\mathcal{T}\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)|0\rangle_0$$
$$-\frac{i\lambda}{\hbar}\int d^4x \int \frac{d^4k_1}{(2\pi)^4}\frac{e^{ik_1(x_1-x)}}{k_1^2+m^2-i\epsilon}\int \frac{d^4k_2}{(2\pi)^4}\frac{e^{ik_2(x_2-x)}}{k_2^2+m^2-i\epsilon} \tag{A.11}$$
$$\int \frac{d^4k_3}{(2\pi)^4}\frac{e^{ik_3(x_3-x)}}{k_3^2+m^2-i\epsilon}\int \frac{d^4k_4}{(2\pi)^4}\frac{e^{ik_4(x_4-x)}}{k_4^2+m^2-i\epsilon} + ...$$

There are other additional leading order interaction terms corresponding to the different combinations of the space points x_1, x_2, x_3, x_4 and x. However, these all contain delta functions δ that require the incoming and outgoing momenta to be the same i.e. represent "no scattering" events like the free scalar amplitude $\langle p_1, p_2|p_3, p_4\rangle_0$. These

terms are all represented by disconnected vacuum diagrams like the one shown in Fig.
A.4.
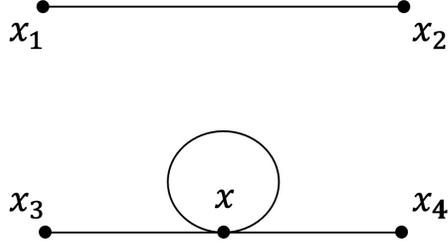
Figure A.4: A disconnected Feynman diagram representing a 4-point correlation function term that doesn't contribute the interaction amplitude.

A.3 $SU(2)_L$ Gauge Symmetry

Because the $SU(2)_L$ makes a distinction between right-handed ψ_R and left-handed ψ_L fermion fields, two different gauge transformations need to be introduced.

$$\psi_R \to e^{-i\beta(x)\frac{Y_W}{2}}\psi_R \tag{A.12}$$

$$\psi_L \to e^{-i\alpha_i(x)\frac{\tau_i}{2}-i\beta(x)\frac{Y_W}{2}}\psi_L \tag{A.13}$$

These transformation give rise to two different covariant derivatives defined as

$$D_\mu = \partial_\mu + ig_1\frac{Y_W}{2}B_\mu \tag{A.14}$$

$$D_\mu = \boldsymbol{I}(\partial_\mu + g_1\frac{Y_W}{2}B_\mu) + ig_2\frac{\boldsymbol{\tau}}{2}\boldsymbol{W_\mu}. \tag{A.15}$$

The gauge constants g_1 and g_2 define the coupling strength of the introduced interactions. We impose the following transformation on the new gauge fields

$$W_\mu^i \to W_\mu^i + \alpha^j(x)\epsilon^{ijk}W_\mu^k + \frac{1}{g_2}\partial_\mu\alpha^i(x) \tag{A.16}$$

$$B_\mu \to B_\mu + \frac{1}{g_1}\partial_\mu\beta(x) \tag{A.17}$$

where ϵ^{ijk} is the Levi-civita function and originates from the generator commutation rules $[\sigma_i, \sigma_j] = 2i\epsilon^{ijk}\sigma_k$. The field strength tensors for the new gauge fields $W^i_{\mu\nu}$ and $B_{\mu\nu}$ are defined as

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g_2\epsilon^{ijk}W^j_\mu W^k_\nu \tag{A.18}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{A.19}$$

A.4 $SU(3)_C$ Gauge Symmetry

The quarks are arranged in triplets in the $SU(3)_C$ representation and are transformed in the following way.

$$\psi \to \psi e^{-i\alpha_a(x)\frac{\lambda_a}{2}} \tag{A.20}$$

These transformations give rise to a new covariant derivative

$$D_\mu = \partial_\mu + ig_3\frac{\lambda_a}{2}G^a_\mu \tag{A.21}$$

where g_3 is the strong coupling constant and the 8 gauge fields G^i_μ correspond to the gluon gauge bosons. To reinstate the symmetry the following gauge field transformation is imposed.

$$G^a_\mu \to G^a_\mu + \alpha^b(x)f^{abc}G^c_\mu + \frac{1}{g_3}\partial_\mu\alpha^a(x) \tag{A.22}$$

Here f^{abc} are the structure constants of the $SU(3)_C$ group which originate from the generator commutation rules $[\lambda_a, \lambda_b] = 2if^{abc}\lambda_c$. We define the field strength tensors for the new gauge fields as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial^a_\nu - g_3 f^{abc} G^b_\mu G^c_\nu \tag{A.23}$$

A.5 Vector Boson Masses

The starting expression is the Higgs Lagrangian 2.49 with the potential 2.50.

$$\mathcal{L}_H = (\partial^\mu \Phi)^\dagger (\partial_\mu \Phi) - V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{A.24}$$

Lets impose the left-handed doublet covariant derivative of the $SU(2)_L \otimes U(1)_Y$ gauge symmetry stated in equation A.15 and include the complex scalar field in unitary gauge stated in equation 2.54. The kinetic terms become

$$D^\mu \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} = \left(\boldsymbol{I}(\partial_\mu + \frac{g_1}{2} B_\mu) + i g_2 \frac{\boldsymbol{\tau}}{2} \boldsymbol{W_\mu} \right)$$

$$= \begin{pmatrix} 0 \\ \partial_\mu h(x) \end{pmatrix} + i \frac{g_2}{2} (g_1 B_\mu - g_2 W^3_\mu) \begin{pmatrix} v + h(x) \\ 0 \end{pmatrix} + i \frac{g_2}{2} (g_1 B_\mu - g_2 W^3_\mu) \tag{A.25}$$

and

$$\left(D^\mu \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \right)^\dagger = (0, \partial_\mu h(x)) - i \frac{g_2}{2} (W^1_\mu + i W^2_\mu)(0, v + h(x))$$

$$- i \frac{g_2}{2\sqrt{2}} (W^1_\mu - i W^2_\mu)(0, v + h(x)). \tag{A.26}$$

Note that the following definition

$$\tau \cdot \boldsymbol{W}_\mu = \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & W_\mu^3 \end{pmatrix} = \begin{pmatrix} W_\mu^3 & W_\mu^+ \\ W_\mu^- & W_\mu^3 \end{pmatrix} \tag{A.27}$$

and

If equations A.25, A.26 and A.27 are plugged into equation A.24 the Higgs-Gauge section is redefined as

$$\mathcal{L}_{HG} = \frac{1}{2}\partial_\mu h(x)\partial^\mu h(x) - \frac{\mu^2}{2}(v + h(x))^2 - \frac{\lambda}{4}(v + h(x))^4$$
$$+ \frac{g_2^2}{2}(v + h(x))^2 W_\mu^+ W^{\mu -} + \frac{1}{8}(g_1 B^\mu - g_2 W^{3\mu})(g_1 B_\mu - g_2 W_\mu^3)(v + h(x))^2. \tag{A.28}$$

Additionally it is possible to redefine the fields as

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \tag{A.29}$$

with

$$\sin\theta_W = \frac{g_1}{(g_2^2 + g_1^2)^{1/2}}. \tag{A.30}$$

This gives the final expression with the mass terms for the W and Z bosons but leaving the photon massless.

$$\mathcal{L}_{HG} = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{g_2^2}{4}(v + h)^2 W_\mu^+ W^{-\mu} + \frac{1}{8}\frac{g_2^2}{\cos^2(\theta_W)}(v + h)^2 Z_\mu Z^\mu$$
$$+ \frac{\mu^2}{2}(v + h)^2 - \frac{\lambda}{4}(v + h)^4. \tag{A.31}$$

A.6 Toy Scalar Theory

Lets illustrate the construction and application of an effective lagrangian with a scalar theory involving a light scalar field ϕ with mass m_L and a heavy scalar field X with mass M. Consider the lagrangian

$$\mathcal{L}_{UV} = \frac{1}{2}[(\partial_\mu \phi)^2 - m^2\phi^2 + (\partial_\mu X)^2 - M^2 X^2] - \frac{\lambda_0}{4!}\phi^4 - \frac{\lambda_1}{2}M\phi^2 X. \tag{A.32}$$

Now lets consider an on-shell $2 \to 2$ ϕ-scattering process. The tree-level scattering amplitude given Eq. (A.32) consists out of a contribution from the ϕ^4-contact interaction and the s-, t- and u-channel exchange of X. See Fig. A.5 for the corresponding Feynman diagrams. The amplitude is given by

$$\mathcal{M}_{UV} = -\lambda_0 - \lambda_1 M^2 \left[\frac{1}{s + M^2} + \frac{1}{t - M^2} + \frac{1}{u - M^2} \right] \tag{A.33}$$

with the Mandelstam variables $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$ and $u = (p_1 - p_4)^2$. In case the mass of the heavy scalar M is much bigger than the transferred momenta, i.e. $M \gg s, t, u$, we can approximate the terms of Eq. (A.33) in the brackets with a Taylor expansion in M^{-2}.

$$\mathcal{M}_{UV} \approx -\lambda_0 - 3\lambda_1^2 - 4\lambda_1^2 \frac{m_L^2}{M^2} + \mathcal{O}(M^{-4}) \tag{A.34}$$

Note the use of the definition $4m_L^2 = s + t + u$. Now lets say we want to construct an effective lagrangian valid for energies much lower than the mass scale, i.e. for $E \ll M$, and that gives a similar result. According to Eq. (2.66) we can write down an effective lagrangian as a power expansion of operators in terms of the light field ϕ with a cut-off scale $\Lambda = M$.

$$\mathcal{L}_{EFT}(\phi) = \mathcal{L}^{D\leq 4}(\phi) + \frac{1}{M}\mathcal{L}^{D=5}(\phi) + \frac{1}{M^2}\mathcal{L}^{D=6}(\phi) + ... \tag{A.35}$$

Assuming the effective field theory inherits a parity symmetry $\phi \to -\phi$ of the UV theory odd operator dimensions are omitted eg. $D = 5$. A very straightforward choice for an effective lagrangian would then be

$$\mathcal{L}_{EFT} = \frac{1}{2}[(\partial_\mu \phi)^2 - m^2\phi^2] - \frac{c_1}{4!}\phi^4 - \frac{c_2}{6!M^2}\phi^6. \tag{A.36}$$

Note that we only included terms of the EFT expansion up until $D = 6$. Here we also introduced the two Wilson coefficients c_1 and c_2. If we would use this effective lagrangian to calculate the scattering amplitude again for the same $2 \to 2$ ϕ-scattering process we get

$$\mathcal{M}_{EFT} = -c_1. \tag{A.37}$$

We see that the EFT and UV theory give the same results under the condition

$$c_1 = \lambda_0 + 3\lambda_1 2 - 4\lambda_1^2 \frac{m_L^2}{M^2} \tag{A.38}$$

Constructing the effective lagrangian is in general not this straightforward and we only considered a very simple scalar theory for simple scattering process on tree-level. Also, starting with a UV theory in mind is also known as the *top-down* approach. It is more common to adopt a *bottom-up* approach, i.e. constructing the effective lagrangian first and matching it to UV theories later. However, this simple example shows that with one single effective field theory one can infer values of parameters of multiple higher energy UV theories.
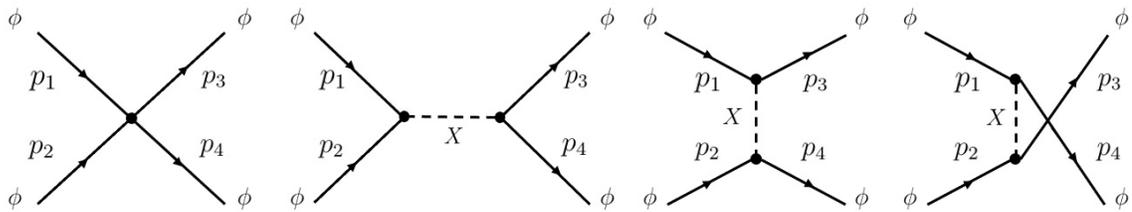
Figure A.5: Tree-level Feynman diagrams contributing to the $2 \rightarrow 2$ ϕ-scattering process.

Appendix B

Differential Cross Sections Appendix

This appendix contains additional fiducial differential cross section and covariance plots of Ch. 6.

B.1 List of nuisance parameters

This section includes all the nuisance parameters included in the analysis described in Ch. 6 and as defined by the ATLAS analysis team.

The following nuisance parameters are included in the $b$tagging category of systematic uncertainty:

- bTagSF_DL1r_extrapolation

- bTagSF_DL1r_extrapolation_from_charm

- bTagSF_DL1r_eigenvars_B1

- bTagSF_DL1r_eigenvars_B2

- bTagSF_DL1r_eigenvars_B3

- bTagSF_DL1r_eigenvars_B4

- bTagSF_DL1r_eigenvars_B5

- bTagSF_DL1r_eigenvars_B6

- bTagSF_DL1r_eigenvars_B7

- bTagSF_DL1r_eigenvars_B8

- bTagSF_DL1r_eigenvars_B9

- bTagSF_DL1r_eigenvars_C1

- bTagSF_DL1r_eigenvars_C2

- bTagSF_DL1r_eigenvars_C3

- bTagSF_DL1r_eigenvars_C4

- bTagSF_DL1r_eigenvars_Light1

- bTagSF_DL1r_eigenvars_Light2

- bTagSF_DL1r_eigenvars_Light3

- bTagSF_DL1r_eigenvars_Light4

The following nuisance parameters are included in the detector category of systematic uncertainty:

- CategoryReduction_JET_BJES_Response

- CategoryReduction_JET_EffectiveNP_Detector1

- CategoryReduction_JET_EffectiveNP_Detector2

- CategoryReduction_JET_EffectiveNP_Mixed1

- CategoryReduction_JET_EffectiveNP_Mixed2

- CategoryReduction_JET_EffectiveNP_Mixed3

- CategoryReduction_JET_EffectiveNP_Modelling1

- CategoryReduction_JET_EffectiveNP_Modelling2

- CategoryReduction_JET_EffectiveNP_Modelling3

- CategoryReduction_JET_EffectiveNP_Modelling4

- CategoryReduction_JET_EffectiveNP_Statistical1

- CategoryReduction_JET_EffectiveNP_Statistical2

- CategoryReduction_JET_EffectiveNP_Statistical3

- CategoryReduction_JET_EffectiveNP_Statistical4

- CategoryReduction_JET_EffectiveNP_Statistical5

- CategoryReduction_JET_EffectiveNP_Statistical6

- CategoryReduction_JET_EtaIntercalibration_Modelling

- CategoryReduction_JET_EtaIntercalibration_NonClosure_2018data

- CategoryReduction_JET_EtaIntercalibration_NonClosure_highE

- CategoryReduction_JET_EtaIntercalibration_NonClosure_negEta

- CategoryReduction_JET_EtaIntercalibration_NonClosure_posEta

- CategoryReduction_JET_EtaIntercalibration_TotalStat

- CategoryReduction_JET_Flavor_Composition

- CategoryReduction_JET_Flavor_Response

- CategoryReduction_JET_JER_DataVsMC_MC16

- CategoryReduction_JET_JER_EffectiveNP_1

- CategoryReduction_JET_JER_EffectiveNP_2

- CategoryReduction_JET_JER_EffectiveNP_3

- CategoryReduction_JET_JER_EffectiveNP_4

- CategoryReduction_JET_JER_EffectiveNP_5

- CategoryReduction_JET_JER_EffectiveNP_6

- CategoryReduction_JET_JER_EffectiveNP_7restTerm

- CategoryReduction_JET_Pileup_OffsetMu

- CategoryReduction_JET_Pileup_OffsetNPV

- CategoryReduction_JET_Pileup_PtTerm

- CategoryReduction_JET_Pileup_RhoTopology

- CategoryReduction_JET_PunchThrough_MC16

- CategoryReduction_JET_SingleParticle_HighPt

- EG_RESOLUTION_ALL

- EG_SCALE_AF2

- EG_SCALE_ALL

- MET_SoftTrk_ResoPara

- MET_SoftTrk_ResoPerp

- MET_SoftTrk_Scale

- MUON_ID

- MUON_MS

- MUON_SAGITTA_RESBIAS

- MUON_SAGITTA_RHO

- MUON_SCALE

The following nuisance parameters are included in the reconstruction category of systematic uncertainty:

- pile-up

- leptonSF_EL_SF_Trigger

- leptonSF_EL_SF_Reco

- leptonSF_EL_SF_ID

- leptonSF_EL_SF_Isol

- leptonSF_MU_SF_Trigger_STAT

- leptonSF_MU_SF_Trigger_SYST

- leptonSF_MU_SF_ID_STAT

- leptonSF_MU_SF_ID_SYST

- leptonSF_MU_SF_ID_STAT_LOWPT

- leptonSF_MU_SF_ID_SYST_LOWPT

- leptonSF_MU_SF_Isol_STAT

- leptonSF_MU_SF_Isol_SYST

- leptonSF_MU_SF_TTVA_STAT

- leptonSF_MU_SF_TTVA_SYST

- jvt

The following nuisance parameters model theory uncertainties from the top quark pair signal process and several background processes. These source of uncertainty have not been included in the analysis described in Ch. 6 but should be in any future analyses.

- Single top (DR vs. DS)

- W+jets + Z+jets normalisation

- Diboson

- ttV

- Fakes normalisation

- Parton-shower and hadronisation models (Pwg+Hw7)

- Matching between ME parton-shower model (aMC@NLO+Py8)

- Hard gluon radiation (Pwg+Py8 hdamp)

- top quark mass

- higher-order corrections

- initial-state radiation

- final-state radiation

- PDFs

B.2 Fiducial Differential Cross Sections

(a)

(b)

(c)

(d)

Figure B.1: Measured fiducial differential cross sections with the Full Run 2 dataset of 139 fb^{-1} unfolded to particle level with the Gaussian Processes method including all statistical and systematic uncertainties. The unfolded data is compared with the testing subsample of $t\bar{t}$ events from the POWHEG-BOX v2+PYTHIA8 signal sample described in Sec. 6.2 and includes a χ^2 with corresponding p-value.
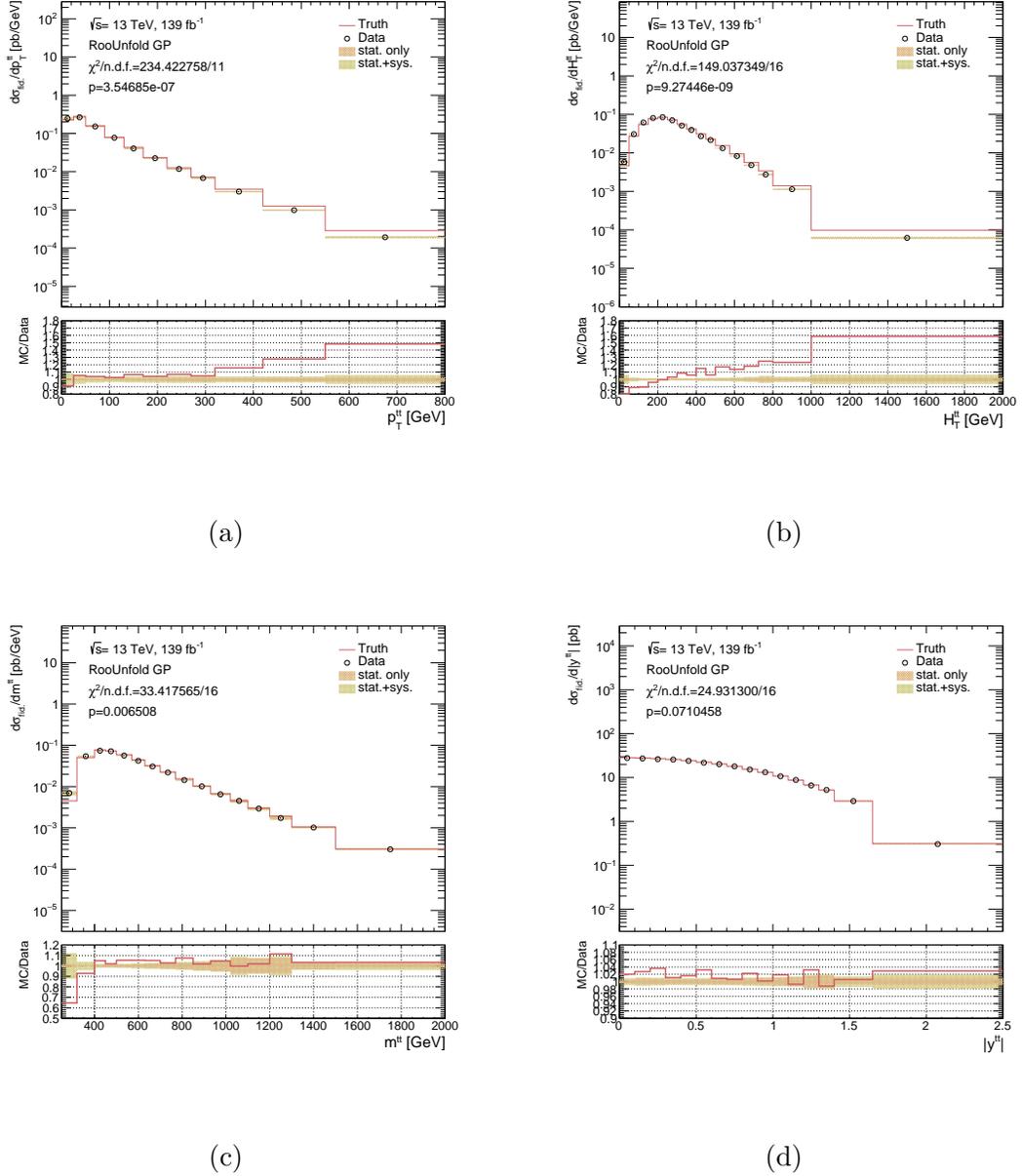
(a)

(b)

(c)
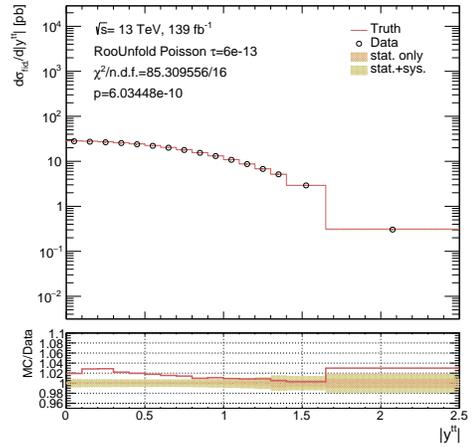
(d)

Figure B.2: Measured fiducial differential cross sections with the Full Run 2 dataset of 139 fb^{-1} unfolded to particle level with the Poisson method including all statistical and systematic uncertainties. The unfolded data is compared with the testing subsample of $t\bar{t}$ events from the POWHEG-BOX V2+PYTHIA8 signal sample described in Sec. 6.2 and includes a χ^2 with corresponding p-value.
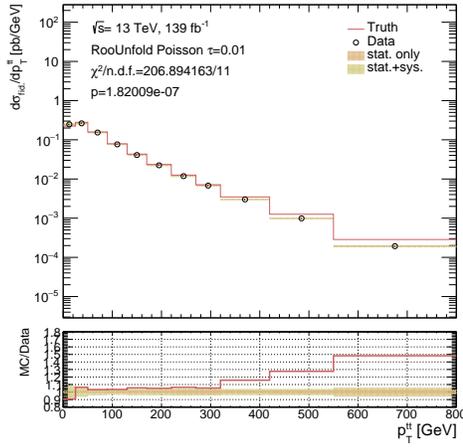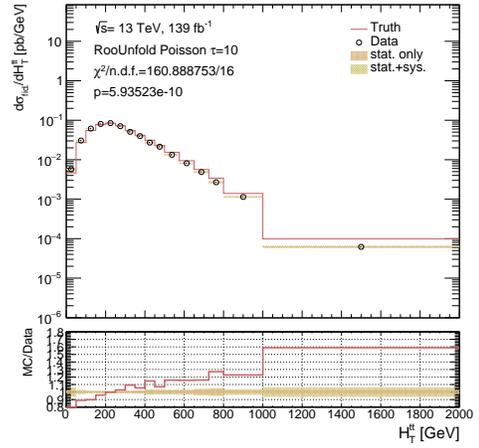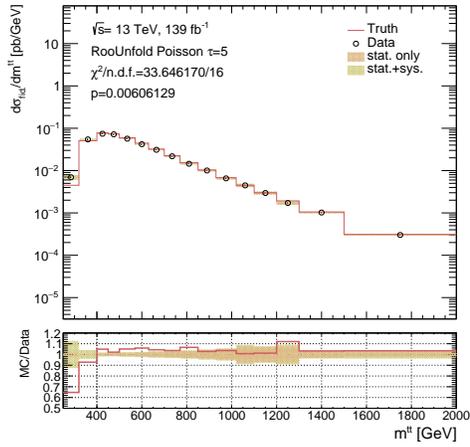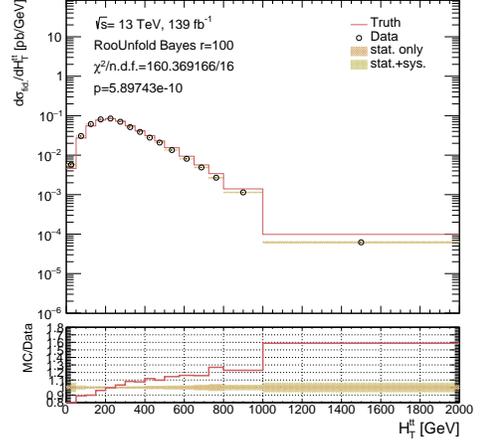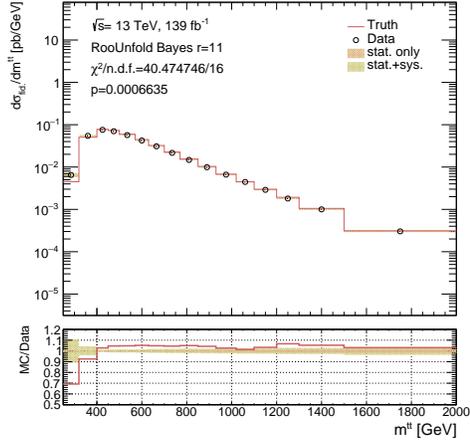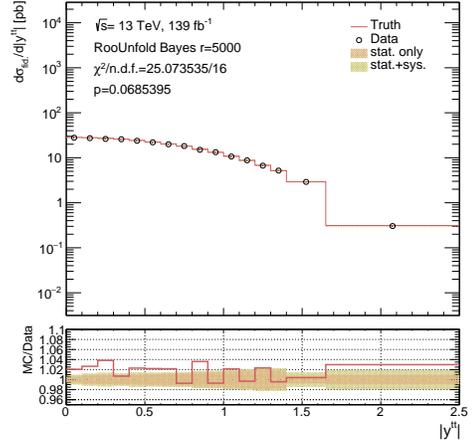
Figure B.3: Measured fiducial differential cross sections with the Full Run 2 dataset of 139 fb^{-1} unfolded to particle level with the Bayes method including all statistical and systematic uncertainties. The unfolded data is compared with the testing subsample of $t\bar{t}$ events from the Powheg-Box v2+Pythia8 signal sample described in Sec. 6.2 and includes a χ^2 with corresponding p-value.

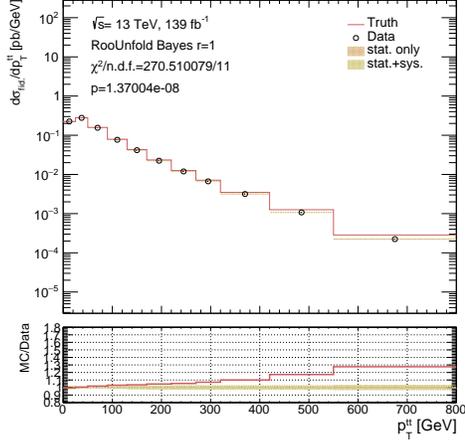## B.3   Covariances

(a)

(b)
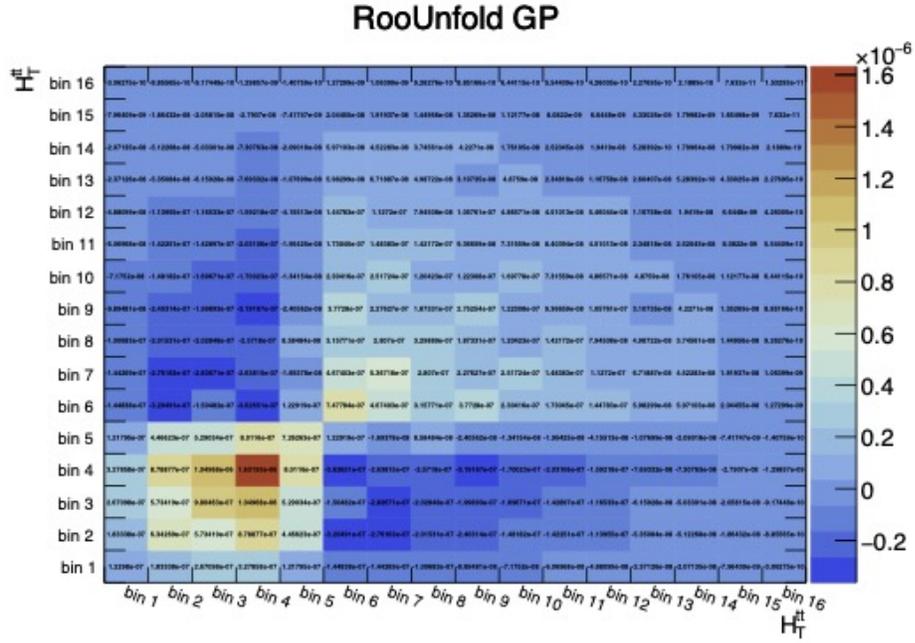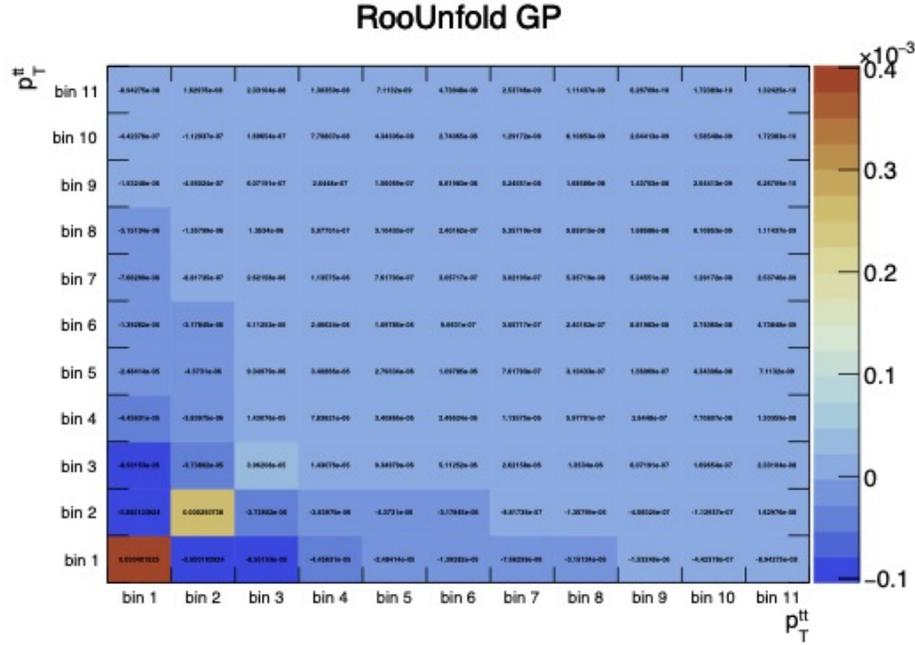
Figure B.4: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $p_T^{t\bar{t}}$ and $H_T^{t\bar{t}}$ unfolded with the Gaussian Processes method.

(a)

(b)
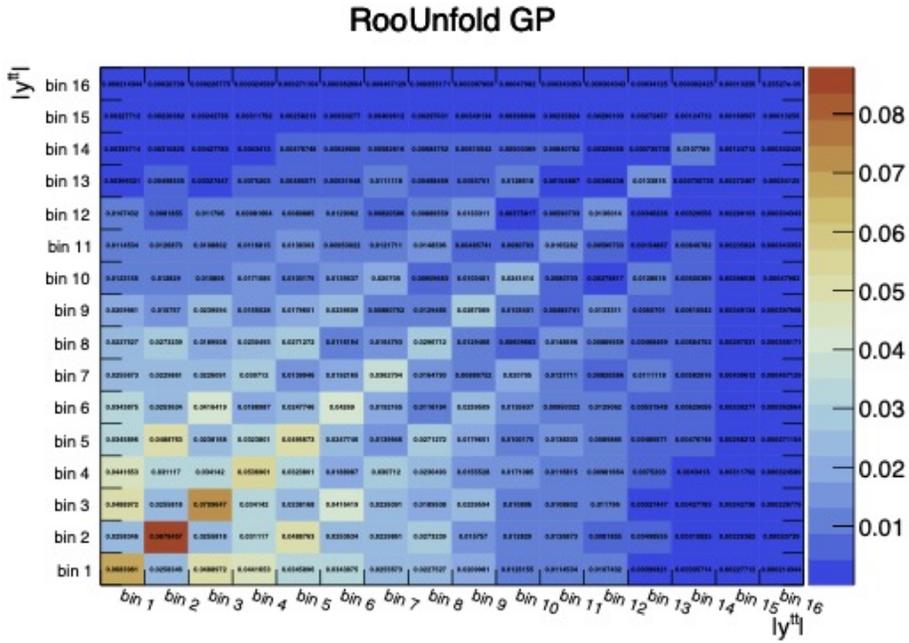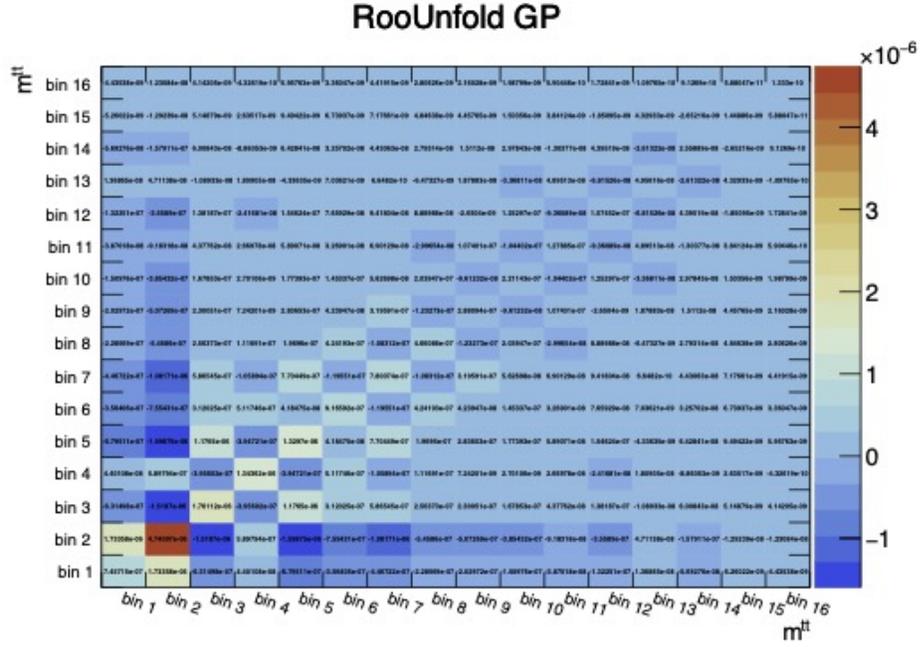
Figure B.5: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$ unfolded with the Gaussian Processes method.

(a)

(b)
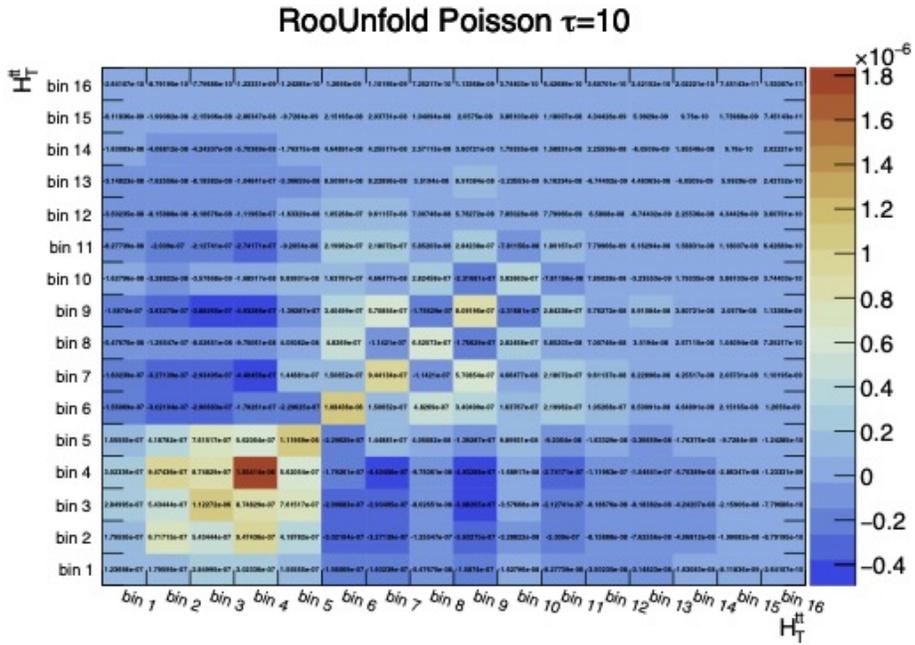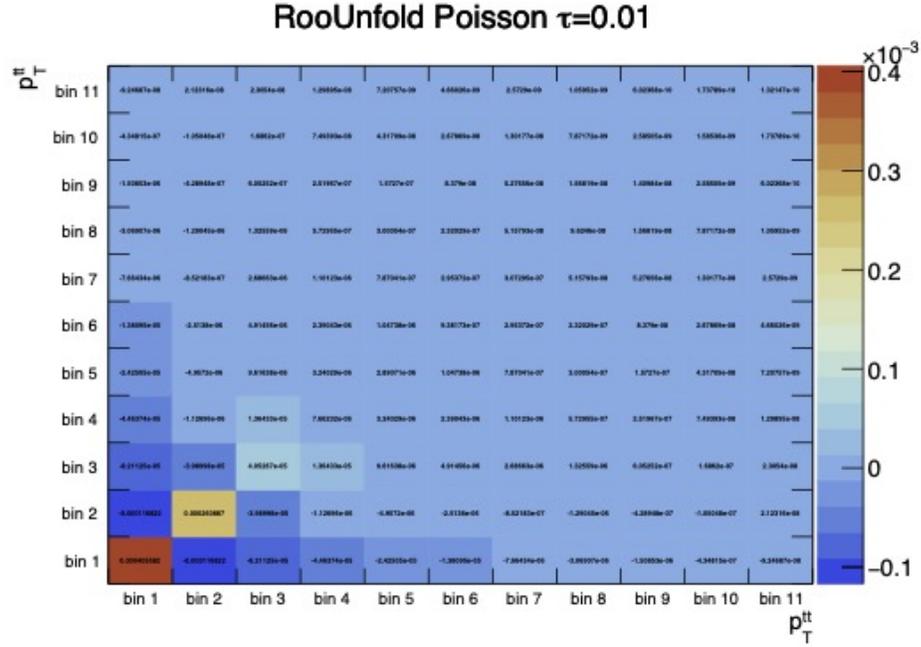
Figure B.6: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $p_T^{t\bar{t}}$ and $H_T^{t\bar{t}}$ unfolded with the Poisson method.

(a)

(b)
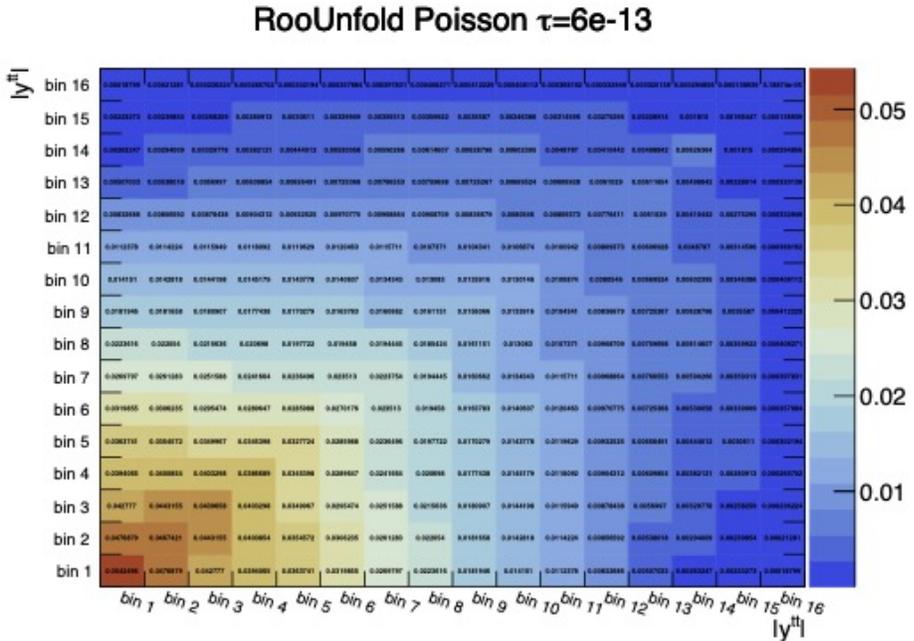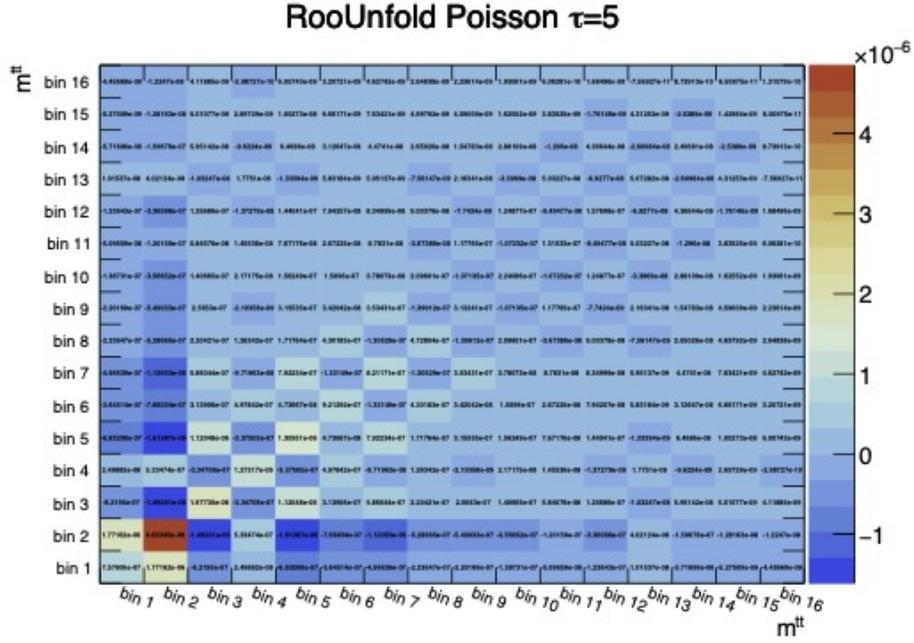
Figure B.7: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$ unfolded with the Poisson method.

(a)
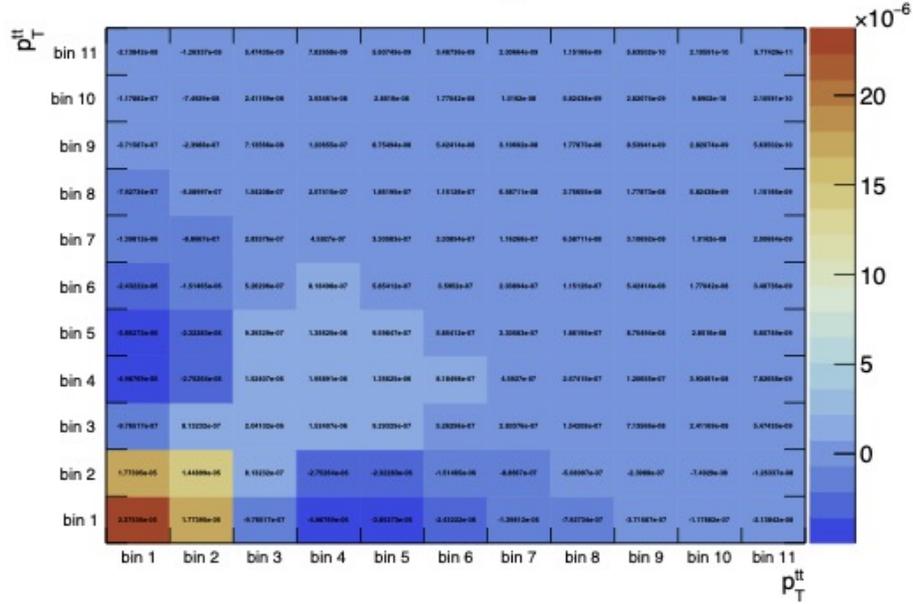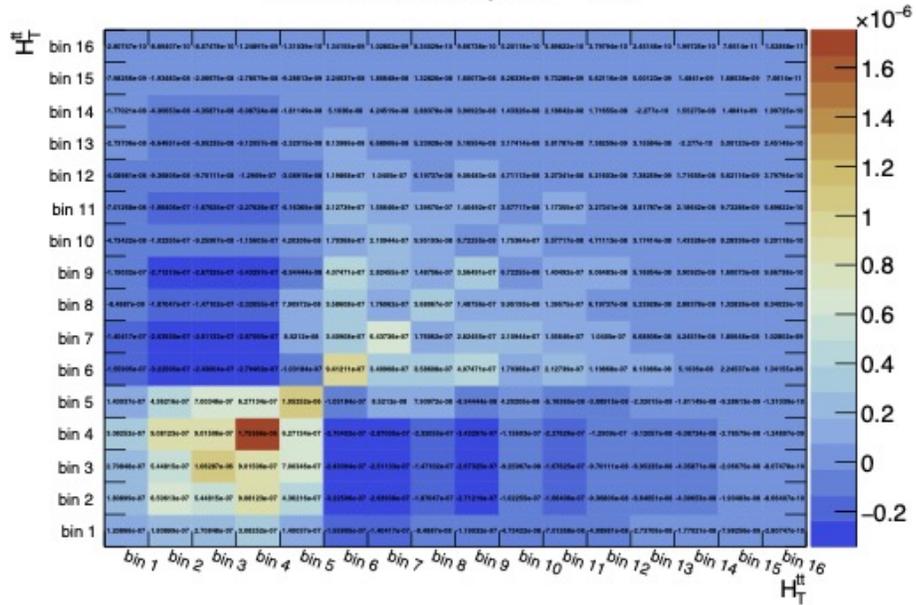
(b)

Figure B.8: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $p_T^{t\bar{t}}$ and $H_T^{t\bar{t}}$ unfolded with the Bayes method.

(a)

(b)
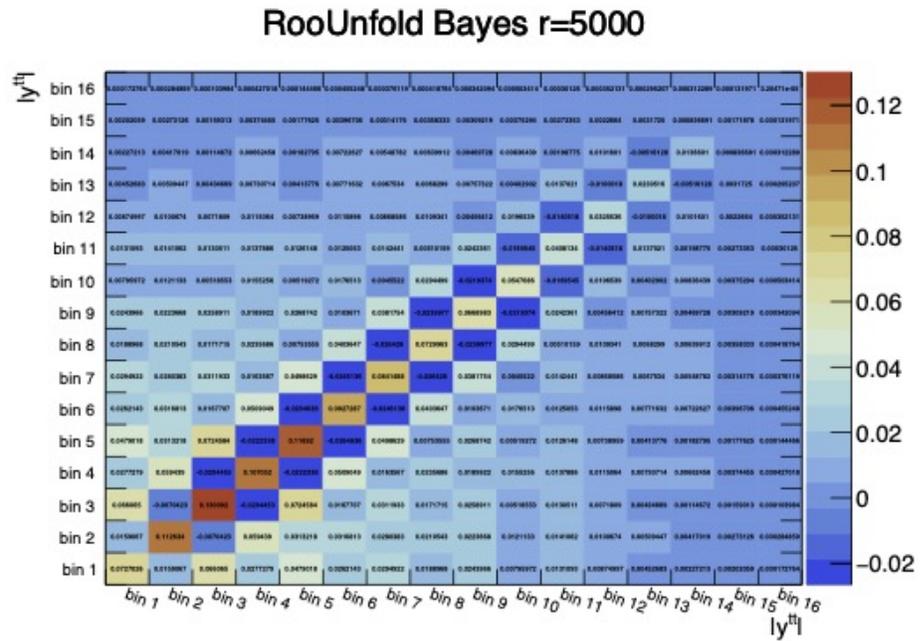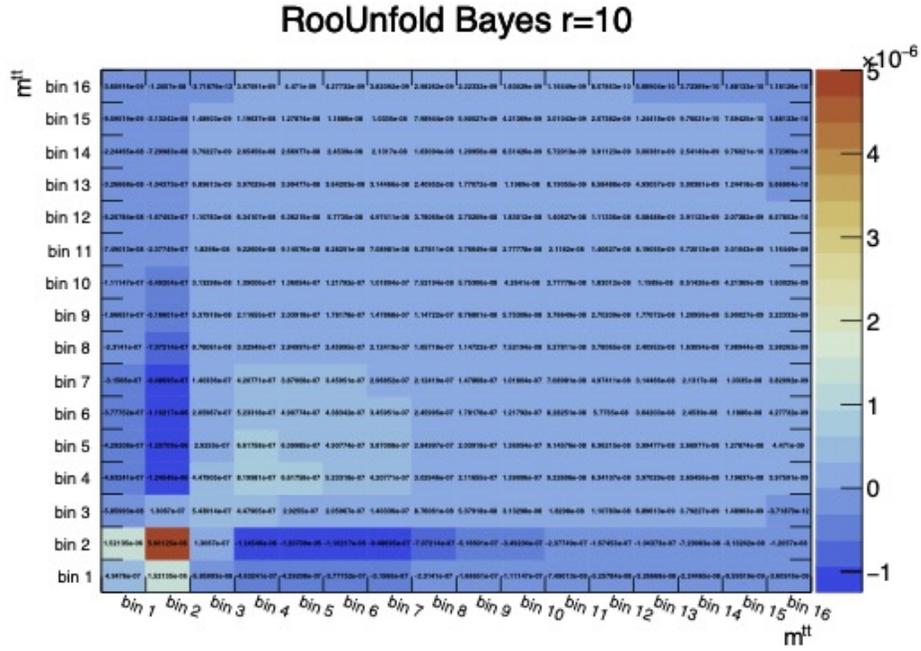
Figure B.9: Covariance matrix estimates corresponding to the measured fiducial differential cross sections for the $m^{t\bar{t}}$ and $|y^{t\bar{t}}|$ unfolded with the Bayes method.

Bibliography

[1] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory.* Reading, USA: Addison-Wesley, 1995.

[2] M. Srednicki, *Quantum field theory.* Cambridge University Press, 1 2007.

[3] B. R. Holstein, "Graviton Physics," *Am. J. Phys.*, vol. 74, pp. 1002–1011, 2006.

[4] B. Schulz, "Review on the quantization of gravity," 9 2014. arXiv:1409.7977.

[5] A. Ashtekar and E. Bianchi, "A short review of loop quantum gravity," *Rept. Prog. Phys.*, vol. 84, no. 4, p. 042001, 2021.

[6] CERN, "The standard model." https://home.cern/science/physics/standard-model.

[7] D. Galbraith and B. C., "The standard model of particle physics." http://davidgalbraith.org/portfolio/ux-standard-model-of-the-standard-model/.

[8] J. B. MARION, "Chapter 11 - kinematics of two-particle collisions," in *Classical Dynamics of Particles and Systems* (J. B. MARION, ed.), pp. 315–341, Academic Press, 1965.

[9] L. D. Landau and E. M. Lifshitz, *Mechanics, Third Edition: Volume 1 (Course of Theoretical Physics).* Butterworth-Heinemann, 3 ed., Jan. 1976.

[10] H. Goldstein, *Classical Mechanics*. Addison-Wesley, 1980.

[11] H. Kragh, *Quantization (First, Second)*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[12] J. Conlon, "Canonical quantization." http://www-thphys.physics.ox.ac.uk/courses/C6/C6_2017/C6Notes/canonicalJC.pdf.

[13] D. S. Lemons, *Perfect Form: Variational Principles, Methods, and Applications in Elementary Physics*. Princeton University Press, 1997.

[14] Wikipedia, "Principle of least action." https://en.wikipedia.org/wiki/Principle_of_least_action.

[15] R. P. Feynman, "Space-time approach to non-relativistic quantum mechanics," *Rev. Mod. Phys.*, vol. 20, pp. 367–387, Apr 1948.

[16] H. Lehmann, K. Symanzik, and W. Zimmermann, "Zur Formulierung quantisierter Feldtheorien," *Nuovo Cimento*, vol. 1, pp. 205–25, 1955.

[17] G. C. Wick, "The Evaluation of the Collision Matrix," *Phys. Rev.*, vol. 80, pp. 268–272, 1950.

[18] R. P. Feynman, "Space-time approach to quantum electrodynamics," *Phys. Rev.*, vol. 76, pp. 769–789, Sep 1949.

[19] K. Kumericki, "Feynman diagrams for beginners," 2016. arXiv:1602.04182.

[20] E. Noether, "Invariante variationsprobleme," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1918, pp. 235–257, 1918.

[21] P. A. M. Dirac, "Quantum theory of emission and absorption of radiation," *Proc. Roy. Soc. Lond. A*, vol. 114, p. 243, 1927.

[22] M. Thomson, *Modern Particle Physics*. Cambridge University Press, 2013.

[23] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett. 19*, vol. 19, pp. 1264–1266, Nov. 1967.

[24] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc. C*, vol. 680519, pp. 367–377, 1968.

[25] S. L. Glashow, "The renormalizability of vector meson interactions," *Nuclear Physics*, vol. 10, pp. 107–117, 1959.

[26] M. Gell-Mann, "Symmetries of baryons and mesons," *Phys. Rev.*, vol. 125, pp. 1067–1084, Feb 1962.

[27] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.

[28] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[29] A. Lahiri and D. Mukhopadhyay, "Unitarity in $WW \to WW$ elastic scattering without a Higgs boson," 7 2011.

[30] V. C. Rubin and J. Ford, W. Kent, "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions," *Astrophysical Journal*, vol. 159, p. 379, Feb. 1970.

[31] V. C. Rubin, D. Burstein, J. Ford, W. K., and N. Thonnard, "Rotation velocities of 16 SA galaxies and a comparison of Sa, SB and SC rotation properties.," *Astrophysical Journal*, vol. 289, pp. 81–104, Feb. 1985.

[32] R. Clausius, "Xvi. on a mechanical theorem applicable to heat," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 40, no. 265, pp. 122–127, 1870.

[33] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln," *Helv. Phys. Acta*, vol. 6, pp. 110–127, 1933.

[34] A. Coc, E. Vangioni-Flam, P. Descouvemont, A. Adahchour, and C. Angulo, "Updated big bang nucleosynthesis compared withwilkinson microwave anisotropy probeobservations and the abundance of light elements," *The Astrophysical Journal*, vol. 600, p. 544–552, Jan 2004.

[35] C. e. a. Alcock, "The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations," *Astrophysical Journal*, vol. 542, pp. 281–307, Oct. 2000.

[36] G. Jungman, M. Kamionkowski, and K. Griest, "Supersymmetric dark matter," *Physics Reports*, vol. 267, p. 195–373, Mar 1996.

[37] A. Birkedal, A. Noble, M. Perelstein, and A. Spray, "Little higgs dark matter," *Physical Review D*, vol. 74, Aug 2006.

[38] D. Hooper and S. Profumo, "Dark matter and collider phenomenology of universal extra dimensions," *Physics Reports*, vol. 453, p. 29–115, Dec 2007.

[39] F. Chadha-Day, J. Ellis, and D. J. E. Marsh, "Axion dark matter: What is it and why now?," 2021. arXiv:2105.01406.

[40] M. Schumann, "Direct detection of wimp dark matter: Concepts and status," 2019. arXiv:1903.03026.

[41] J. Conrad, "Indirect detection of wimp dark matter: a compact review," 2014. arXiv:1411.1925.

[42] M. Aaboud and et al., "Search for dark matter produced in association with bottom or top quarks in $\sqrt{s} = 13$ pp collisions with the atlas detector," *The European Physical Journal C*, vol. 78, Jan 2018.

[43] K. Sung, "Search for dark matter in final states with top quarks in CMS data," in *53rd Rencontres de Moriond on Electroweak Interactions and Unified Theories*, pp. 207–212, 2018.

[44] D. Casadei, "Searches for cosmic antimatter," 2004. arXiv:0405417.

[45] I. Galaktionov, "Antimatter in cosmic rays," *Rept. Prog. Phys.*, vol. 65, pp. 1243–1270, 2002.

[46] A. V. Grobov and S. G. Rubin, "Formation and search of large scale antimatter regions," *International Journal of Modern Physics D*, vol. 24, p. 1550093, Sep 2015.

[47] A. D. Sakharov, "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe," *Pisma Zh. Eksp. Teor. Fiz.*, vol. 5, pp. 32–35, 1967.

[48] K. S. Babu *et al.*, "Working Group Report: Baryon Number Violation," in *Community Summer Study 2013: Snowmass on the Mississippi*, 11 2013.

[49] D. Phillips and et al., "Neutron-antineutron oscillations: Theoretical status and experimental prospects," *Physics Reports*, vol. 612, pp. 1–45, 2016. Neutron-Antineutron Oscillations: Theoretical Status and Experimental Prospects.

[50] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π decay of the k_2^0 meson," *Phys. Rev. Lett.*, vol. 13, pp. 138–140, Jul 1964.

[51] A. Sanda, "The road to the discovery of large cp violation in b decays: —a miracle story—," *Progress of theoretical physics*, vol. 122, no. 1, pp. 103–123, 2009.

[52] B. Aubert, D. Boutigny, J.-M. Gaillard, A. Hicheur, Y. Karyotakis, J. Lees, P. Robbe, V. Tisserand, A. Palano, G. Chen, *et al.*, "Observation of cp violation in the b 0 meson system," *Physical review letters*, vol. 87, no. 9, p. 091801, 2001.

[53] G. Valencia, "Cp violation in top-quark pair production and decay," 2013. arXiv:1007.4765.

[54] S. Berge, M. Mühlleitner, A. Moreno Briceño, D. Wackeroth, and M. Wiebusch, "Cp-violation in top quark pair production in the complex mssm at hadron colliders," *Nuclear and Particle Physics Proceedings*, vol. 267-269, pp. 294–301, 2015. X Latin American Symposium of High Energy Physics.

[55] H. Georgi and S. L. Glashow, "Unity of all elementary-particle forces," *Phys. Rev. Lett.*, vol. 32, pp. 438–441, Feb 1974.

[56] J. C. Pati, "Discovery of proton decay: A must for theory, a challenge for experiment," *AIP Conference Proceedings*, 2000.

[57] S. Dimopoulos, S. Raby, and F. Wilczek, "Supersymmetry and the scale of unification," *Phys. Rev. D*, vol. 24, pp. 1681–1683, Sep 1981.

[58] S. Dimopoulos and H. Georgi, "Softly broken supersymmetry and su(5)," *Nuclear Physics B*, vol. 193, no. 1, pp. 150–162, 1981.

[59] "Gauge coupling running in minimal su(3) × su(2) × u(1) superstring unification," *Physics Letters B*, vol. 272, no. 3, pp. 251–260, 1991.

[60] F. Gursey, P. Ramond, and P. Sikivie, "A Universal Gauge Theory Model Based on E6," *Phys. Lett. B*, vol. 60, pp. 177–180, 1976.

[61] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, and et al., "Search for supersymmetry in proton-proton collisions at 13 tev using identified top quarks," *Physical Review D*, vol. 97, Jan 2018.

[62] A. M. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, and et al., "Search for natural

supersymmetry in events with top quark pairs and photons in pp collisions at $\sqrt{s} = 8$ tev," *Journal of High Energy Physics*, vol. 2018, Mar 2018.

[63] Lumen Physics, "GUTs: The Unification of Forces." https://courses.lumenlearning.com/physics/chapter/33-6-guts-the-unification-of-forces/.

[64] A. A. Coley and D. L. Wiltshire, "What is general relativity?," *Physica Scripta*, vol. 92, p. 053001, Apr 2017.

[65] R. Hedrich, "Quantum gravity: Motivations and alternatives," 2009. arXiv:0908.0355.

[66] P. A. M. Casares, "A review on loop quantum gravity," 2018. arXiv:1808.01252.

[67] D.-W. Chiou, "Loop quantum gravity," *International Journal of Modern Physics D*, vol. 24, p. 1530005, Dec 2014.

[68] M. Shifman, "Large extra dimensions: Becoming acquainted with an alternative paradigm," *International Journal of Modern Physics A*, vol. 25, pp. 199–225, 2010.

[69] X.-G. Wu and Z.-Y. Fang, "Revisiting the real graviton effects at cern lhc within the quantum gravity theory with large extra dimensions," *Physical Review D*, vol. 78, Nov 2008.

[70] T. V. Obikhod and I. A. Petrenko, "Searches for Massive Graviton Resonances at the LHC," *Adv. High Energy Phys.*, vol. 2018, p. 3471023, 2018.

[71] S. Dey, A. Bhat, D. Momeni, M. Faizal, A. F. Ali, T. K. Dey, and A. Rehman, "Probing noncommutative theories with quantum optical experiments," *Nuclear Physics B*, vol. 924, pp. 578–587, 2017.

[72] M. J. Duff, "String and m-theory: Answering the critics," *Foundations of Physics*, vol. 43, p. 182–200, Nov 2011.

[73] A. Miemiec and I. Schnakenburg, "Basics of m-theory," *Fortschritte der Physik*, vol. 54, p. 5–72, Jan 2006.

[74] R. Durrer and J. Hasenkamp, "Testing superstring theories with gravitational waves," *Phys. Rev. D*, vol. 84, p. 064027, Sep 2011.

[75] R. Kallosh and A. Linde, "Testing string theory with cosmic microwave background," *Journal of Cosmology and Astroparticle Physics*, vol. 2007, p. 017–017, Apr 2007.

[76] G. L. Kane, "String/m-theories about our world are testable in the traditional physics way," 2016. arXiv:1601.07511.

[77] J. Sonneveld, "Searches for physics beyond the standard model at the lhc," 2019. arXiv:1905.06239.

[78] S. Rappoccio, "The experimental status of direct searches for exotic physics beyond the standard model at the large hadron collider," *Reviews in Physics*, vol. 4, p. 100027, Nov 2019.

[79] B. Clerbaux, "Searches for physics beyond the standard model at atlas and cms," *Nuclear Physics B - Proceedings Supplements*, vol. 245, pp. 86–93, 2013. The Proceedings of the 7th Joint International Hadron Structure'13 Conference.

[80] A. V. Manohar, "Introduction to effective field theories," 2018. arXiv:1804.05863.

[81] H. Georgi, "Effective field theory." https://www.people.fas.harvard.edu/~hgeorgi/review.pdf.

[82] A. Helset and A. Kobach, "Baryon Number, Lepton Number, and Operator Dimension in the SMEFT with Flavor Symmetries," *Phys. Lett. B*, vol. 800, p. 135132, 2020.

[83] A. Kobach, "Baryon number, lepton number, and operator dimension in the standard model," *Physics Letters B*, vol. 758, p. 455–457, Jul 2016.

[84] I. Brivio and M. Trott, "The standard model as an effective field theory," *Physics Reports*, vol. 793, p. 1–98, Feb 2019.

[85] B. Grzadkowski, M. Iskrzyński, M. Misiak, and J. Rosiek, "Dimension-six terms in the standard model lagrangian," *Journal of High Energy Physics*, vol. 2010, Oct 2010.

[86] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, "The strongly-interacting light higgs," *Journal of High Energy Physics*, vol. 2007, p. 045–045, Jun 2007.

[87] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang, "Automated one-loop computations in the standard model effective field theory," *Physical Review D*, vol. 103, May 2021.

[88] J. A. A. Saavedra and et al., "Interpreting top-quark lhc measurements in the standard-model effective field theory," 2018. arXiv:1802.07237.

[89] G. Durieux, F. Maltoni, and C. Zhang, "Global approach to top-quark flavor-changing interactions," *Physical Review D*, vol. 91, Apr 2015.

[90] C. Degrande, F. Maltoni, J. Wang, and C. Zhang, "Automatic computations at next-to-leading order in qcd for top-quark flavor-changing neutral processes," *Physical Review D*, vol. 91, Feb 2015.

[91] C. Zhang and S. Willenbrock, "Effective-field-theory approach to top-quark production and decay," *Physical Review D*, vol. 83, Feb 2011.

[92] N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang, "A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector," *JHEP*, vol. 04, p. 100, 2019.

[93] L. Szymanowski, "QCD collinear factorization, its extensions and the partonic distributions," *PoS*, vol. QNP2012, p. 015, 2012.

[94] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*, vol. 8. Cambridge University Press, 2 2011.

[95] P. Barnreuther, "Top quark pair production at the lhc," 2012. PhD Thesis.

[96] H. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J. Owens, J. Pumplin, and W. Tung, "Global qcd analysis of parton structure of the nucleon: Cteq5 parton distributions," *The European Physical Journal C*, vol. 12, p. 375–392, Feb 2000.

[97] D. E. Soper, "Parton distribution functions," *Nuclear Physics B - Proceedings Supplements*, vol. 53, p. 69–80, Feb 1997.

[98] M. Guzzi, P. Nadolsky, E. Berger, H.-L. Lai, F. Olness, and C. P. Yuan, "Ct10 parton distributions and other developments in the global qcd analysis," 2011. arXiv:1101.0561.

[99] J. Rojo, "Progress in the NNPDF global analyses of proton structure," in *55th Rencontres de Moriond on QCD and High Energy Interactions*, 4 2021.

[100] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, "Lhapdf6: parton density access in the lhc precision era," *The European Physical Journal C*, vol. 75, Mar 2015.

[101] H. G. Katzgraber, "Introduction to monte carlo methods," 2011. arXiv:0905.1629.

[102] S. Weinzierl, "Introduction to monte carlo methods," 07 2000. arXiv:0006269.

[103] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box," *Journal of High Energy Physics*, vol. 2010, Jun 2010.

[104] S. Frixione, G. Ridolfi, and P. Nason, "A positive-weight next-to-leading-order monte carlo for heavy flavour hadroproduction," *Journal of High Energy Physics*, vol. 2007, p. 126–126, Sep 2007.

[105] P. Nason, "A new method for combining nlo qcd with shower monte carlo algorithms," *Journal of High Energy Physics*, vol. 2004, p. 040–040, Nov 2004.

[106] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *Journal of High Energy Physics*, vol. 2014, Jul 2014.

[107] E. Bothmann, G. Singh Chahal, S. Höche, J. Krause, F. Krauss, S. Kuttimalai, S. Liebschner, D. Napoletano, M. Schönherr, H. Schulz, and et al., "Event generation with sherpa 2.2," *SciPost Physics*, vol. 7, Sep 2019.

[108] T. Gleisberg and S. Höche, "Comix, a new matrix element generator," *Journal of High Energy Physics*, vol. 2008, p. 039–039, Dec 2008.

[109] F. Cascioli, P. Maierhöfer, and S. Pozzorini, "Scattering amplitudes with open loops," *Physical Review Letters*, vol. 108, Mar 2012.

[110] N. D. Christensen and C. Duhr, "Feynrules – feynman rules made easy," *Computer Physics Communications*, vol. 180, p. 1614–1641, Sep 2009.

[111] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, and E. Norrbin, "High-energy-physics event generation with pythia 6.1," *Computer Physics Communications*, vol. 135, no. 2, pp. 238–259, 2001.

[112] R. Corke, "Multiple interactions in pythia 8," 2009. arXiv:0901.2852.

[113] Z. M. and, "Simulation of pile-up in the ATLAS experiment," *Journal of Physics: Conference Series*, vol. 513, p. 022024, jun 2014.

[114] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An introduction to pythia 8.2," *Computer Physics Communications*, vol. 191, p. 159–177, Jun 2015.

[115] U. A. Wiedemann, "HIP and HEP," tech. rep., Jan 2021. arXiv:2101.01971.

[116] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*. Cambridge monographs on particle physics, nuclear physics, and cosmology, Cambridge: Cambridge University Press, 2003. Photography by S. Vascotto.

[117] M. Bähr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Plätzer, P. Richardson, M. H. Seymour, A. Sherstnev, and et al., "Herwig++ physics and manual," *The European Physical Journal C*, vol. 58, p. 639–707, Nov 2008.

[118] M. H. Seymour and A. Siódmok, "Constraining mpi models using eff and recent tevatron and lhc underlying event data," *Journal of High Energy Physics*, vol. 2013, Oct 2013.

[119] S. Schumann and F. Krauss, "A parton shower algorithm based on catani-seymour dipole factorisation," *Journal of High Energy Physics*, vol. 2008, p. 038–038, Mar 2008.

[120] "CERN Rapport annuel 2019," tech. rep., CERN, Geneva, 2020.

[121] L. Evans and P. Bryant, "LHC machine," *Journal of Instrumentation*, vol. 3, pp. S08001–S08001, aug 2008.

[122] K. Schindl, "The injector chain for the LHC," in *9th LEP Performance Workshop*, pp. 47–52, 3 1999.

[123] CERN, "Facts and figures on the lhc." https://home.cern/resources/faqs/facts-and-figures-about-lhc.

[124] D. Boussard and T. P. R. Linnecar, "The LHC Superconducting RF System," tech. rep., CERN, Geneva, Dec 1999.

[125] CERN, "Pulling together: Superconducting electromagnets." https://home.cern/science/engineering/pulling-together-superconducting-electromagnets.

[126] G. Aad *et al.*, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, p. S08003, 2008.

[127] S. Chatrchyan *et al.*, "The CMS Experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[128] K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC," *JINST*, vol. 3, p. S08002, 2008.

[129] A. A. Alves, Jr. *et al.*, "The LHCb Detector at the LHC," *JINST*, vol. 3, p. S08005, 2008.

[130] B. Goddard, M. Gyr, J. Uythoven, R. Veness, and W. Weterings, "Lhc beam dumping system: extraction channel layout and acceptance," in *Proceedings of the 2003 Particle Accelerator Conference*, pp. 1646 – 1648 vol.3, 06 2003.

[131] E. Mobs, "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018," Aug 2018. General Photo.

[132] J. Pequenao, "Computer generated image of the whole ATLAS detector." Mar 2008.

[133] A. Collaboration, "Luminosity public results run2." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[134] CERN, "Atlas." https://home.cern/science/experiments/atlas.

[135] *ATLAS inner detector: Technical Design Report, 1.* Technical design report. ATLAS, Geneva: CERN, 1997.

[136] C. N. Booth, "Measurement in a magnetic field." https://cbooth.staff.shef.ac.uk/phy6040det/magfield.html.

[137] Y. Takubo, "ATLAS IBL operational experience," tech. rep., CERN, Geneva, Nov 2016.

[138] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, "ATLAS Insertable B-Layer Technical Design Report," tech. rep., Sep 2010.

[139] N. Wermes and G. Hallewel, *ATLAS pixel detector: Technical Design Report.* Technical design report. ATLAS, Geneva: CERN, 1998.

[140] H. Pernegger, "The pixel detector of the ATLAS experiment for LHC run-2," *Journal of Instrumentation*, vol. 10, pp. C06012–C06012, jun 2015.

[141] J. N. Jackson, "The ATLAS semiconductor tracker (SCT)," *Nucl. Instrum. Meth. A*, vol. 541, pp. 89–95, 2005.

[142] M. Capéans-Garrido, "The transition radiation tracker of the ATLAS experiment," *IEEE Trans. Nucl. Sci.*, vol. 51, pp. 994–1000, 2004.

[143] E. Hines, "Performance of Particle Identification with the ATLAS Transition Radiation Tracker," in *Meeting of the APS Division of Particles and Fields*, 9 2011.

[144] B. Mindur, "ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider," *Nucl. Instrum. Meth. A*, vol. 845, pp. 257–261, 2017.

[145] *ATLAS liquid-argon calorimeter: Technical Design Report.* Technical design report. ATLAS, Geneva: CERN, 1996.

[146] N. Nikiforou, "Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade," in *3rd International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications*, 6 2013.

[147] D. E. Boumediene, "ATLAS calorimeters: Run-2 performances and Phase-II upgrades," tech. rep., CERN, Geneva, Oct 2017.

[148] P. Puzo, "ATLAS calorimetry," *Nucl. Instrum. Meth. A*, vol. 494, pp. 340–345, 2002.

[149] J. Abdallah, S. Angelidakis, G. Arabidze, N. Atanov, J. Bernhard, R. Bonnefoy, J. Bossio, R. Bouabid, F. Carrio, T. Davidek, and et al., "Study of energy response and resolution of the atlas tile calorimeter to hadrons of energies from 16 to 30 gev," *The European Physical Journal C*, vol. 81, Jun 2021.

[150] F. Ariztizabal and et al., "Construction and performance of an iron-scintillator hadron calorimeter with longitudinal tile configuration," *Nucl. Instrum. Methods Phys. Res., A*, vol. 349, pp. 384–397. 32 p, Mar 1994.

[151] O. Gildemeister, F. Nessi-Tedaldi, and M. Nessi, "An economic concept for a barrel hadron calorimeter with iron scintillation sampling and WLS-fiber readout," tech. rep., CERN, Geneva, Nov 1991.

[152] S. Palestini, "The muon spectrometer of the atlas experiment," *Nuclear Physics B - Proceedings Supplements*, vol. 125, pp. 337–345, 2003. Innovative Particle and Radiation Detectors.

[153] G. Aad *et al.*, "Readiness of the ATLAS Tile Calorimeter for LHC collisions," *Eur. Phys. J. C*, vol. 70, pp. 1193–1236, 2010.

[154] E. Diehl, "Calibration and Performance of the ATLAS Muon Spectrometer," in *Meeting of the APS Division of Particles and Fields*, 9 2011.

[155] C. Bini, "Study of the performance of the ATLAS Muon Spectrometer," pp. 1265–1268, 2011.

[156] J. Goodson, *Search for Supersymmetry in States with Large Missing Transverse Momentum and Three Leptons including a Z-Boson.* PhD thesis, Stony Brook University, May 2012. Presented 17 Apr 2012.

[157] J. P. Badiou, J. Beltramelli, J. M. Baze, and J. Belorgey, *ATLAS barrel toroid: Technical Design Report.* Technical design report. ATLAS, Geneva: CERN, 1997. Electronic version not available.

[158] *ATLAS end-cap toroids: Technical Design Report.* Technical design report. ATLAS, Geneva: CERN, 1997. Electronic version not available.

[159] *ATLAS central solenoid: Technical Design Report.* Technical design report. ATLAS, Geneva: CERN, 1997. Electronic version not available.

[160] CERN, "Trigger and data acquisition." `https://atlas.cern/discover/detector/trigger-daq`.

[161] *ATLAS level-1 trigger: Technical Design Report.* Technical design report. AT-LAS, Geneva: CERN, 1998.

[162] M. E. Pozo Astigarraga, "THE ATLAS DATA ACQUISITION SYSTEM IN LHC RUN 2," tech. rep., CERN, Geneva, Nov 2017.

[163] A. Borga, A. Kugel, M. Joos, J. Schumacher, L. Tremblet, W. Vandelli, J. Vermeulen, P. Werner, and F. Wickens, "Evolution of the ReadOut System of the ATLAS experiment," tech. rep., CERN, Geneva, Jun 2014.

[164] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report.* Technical design report. ATLAS, Geneva: CERN, 2003.

[165] K. Bos and et al., *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005).* Technical design report. LCG, Geneva: CERN, 2005.

[166] S. Agostinelli and et al., "Geant4—a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003.

[167] E. Richter-Was, D. Froidevaux, and L. Poggioli, "ATLFAST 2.0 a fast simulation package for ATLAS," tech. rep., CERN, Geneva, Nov 1998.

[168] F. Dias, "The new ATLAS Fast Calorimeter Simulation," tech. rep., CERN, Geneva, Oct 2016.

[169] G. Aad, B. Abbott, J. Abdallah, A. A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, and et al., "The atlas simulation infrastructure," *The European Physical Journal C*, vol. 70, p. 823–874, Sep 2010.

[170] J. Henning, "Spec cpu2000: measuring cpu performance in the new millennium," *Computer*, vol. 33, pp. 28 – 35, 08 2000.

[171] J. Catmore, J. Cranshaw, T. Gillam, E. Gramstad, P. Laycock, N. Ozturk, and G. A. Stewart, "A new petabyte-scale data derivation framework for ATLAS," *J. Phys. Conf. Ser.*, vol. 664, no. 7, p. 072007, 2015.

[172] R. Brun, F. Rademakers, P. Canal, A. Naumann, O. Couet, L. Moneta, V. Vassilev, S. Linev, D. Piparo, G. GANIS, B. Bellenot, E. Guiraud, G. Amadio, wverkerke, P. Mato, TimurP, M. Tadel, wlav, E. Tejedor, J. Blomer, A. Gheata, S. Hageboeck, S. Roiser, marsupial, S. Wunsch, O. Shadura, A. Bose, CristinaCristescu, X. Valls, and R. Isemann, "root-project/root: v6.18/02," Aug. 2019.

[173] ATLAS, "Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment," 3 2012. ATLAS-CONF-2012-042.

[174] V. Lacuesta, "Track and vertex reconstruction in the ATLAS experiment," *Journal of Instrumentation*, vol. 8, pp. C02035–C02035, feb 2013.

[175] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl. Instrum. Meth. A*, vol. 262, pp. 444–450, 1987.

[176] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, "Calorimeter clustering algorithms: Description and performance," 5 2008. ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003.

[177] M. Aaboud, G. Aad, B. Abbott, D. C. Abbott, O. Abdinov, B. Abeloos, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, and et al., "Electron reconstruction and identification in the atlas experiment using the 2015 and 2016 lhc proton–proton collision data at $\sqrt{s} = 13$ tev," *The European Physical Journal C*, vol. 79, Aug 2019.

[178] J. Illingworth and J. Kittler, "A survey of the hough transform," *Computer Vision, Graphics, and Image Processing*, vol. 44, no. 1, pp. 87–116, 1988.

[179] D. Primor, G. Mikenberg, and H. Messer, "Muon detection in the ATLAS CSC detector," *IEEE Trans. Nucl. Sci.*, vol. 54, pp. 635–642, 2007.

[180] G. Aad and et al., "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 *pp* collision data set at $\sqrt{s} = 13$ TeV," *Eur. Phys. J. C*, vol. 81, no. 7, p. 578, 2021.

[181] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, R. Aben, M. Abolins, O. S. AbouZeid, N. L. Abraham, and et al., "Muon reconstruction performance of the atlas detector in proton–proton collision data at $\sqrt{s} = 13$ tev," *The European Physical Journal C*, vol. 76, May 2016.

[182] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, and et al., "Topological cell clustering in the atlas calorimeters and its performance in lhc run 1," *The European Physical Journal C*, vol. 77, Jul 2017.

[183] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-ktjet clustering algorithm," *Journal of High Energy Physics*, vol. 2008, p. 063–063, Apr 2008.

[184] Y. Dokshitzer, G. Leder, S. Moretti, and B. Webber, "Better jet clustering algorithms," *Journal of High Energy Physics*, vol. 1997, p. 001–001, Aug 1997.

[185] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, and et al., "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s}$ s =8 tev using the atlas detector," *The European Physical Journal C*, vol. 76, Oct 2016.

[186] "Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run," tech. rep., CERN, Geneva, Jul 2017.

[187] G. Aad, B. Abbott, D. C. Abbott, A. A. Abud, K. Abeling, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, and et al., "Atlas b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ tev," *The European Physical Journal C*, vol. 79, Nov 2019.

[188] G. Aad, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O. S. AbouZeid, and et al., "Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using pp collisions at $\sqrt{s} = 7$ tev in the atlas detector," *Journal of High Energy Physics*, vol. 2015, Jun 2015.

[189] P. A. Zyla and et al., "Review of Particle Physics," *Progress of Theoretical and Experimental Physics*, vol. 2020, 08 2020.

[190] T. Adye, "Unfolding algorithms and tests using RooUnfold," in *PHYSTAT 2011*, (Geneva), pp. 313–318, CERN, 2011.

[191] E. De Micheli and G. A. Viano, "Fredholm integral equations of the first kind and topological information theory," *Integral Equations and Operator Theory*, vol. 73, p. 553–571, May 2012.

[192] G. Cowan, *Statistical data analysis.* Oxford University Press, USA, 1998.

[193] F. B. Diniz, "Condition number and matrices," 2017. arXiv:1703.04547.

[194] P. Pathak, *Breakthroughs in Statistics.* Springer New York, 1992.

[195] H. Cramer, *Mathematical methods of statistics.* Princeton University Press Princeton, 1946.

[196] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[197] F. Santosa and W. W. Symes, "Linear inversion of band-limited reflection seismograms," *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 4, pp. 1307–1330, 1986.

[198] D. E. Hilt and D. W. Seegrist, "Ridge: a computer program for calculating ridge regression estimates," in *Research Note NE-236*, 1977.

[199] M. Gruber, *Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimators.* Routledge, 11 2017.

[200] A. Tikhonov and V. Y. Arsenin, "Solutions of ill-posed problems," in *Dokl. Akad. Nauk SSSR*, 1977.

[201] S. Schmitt, "Tunfold, an algorithm for correcting migration effects in high energy physics," *Journal of Instrumentation*, vol. 7, p. T10003–T10003, Oct 2012.

[202] A. Höcker and V. Kartvelishvili, "Svd approach to data unfolding," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 372, p. 469–481, Apr 1996.

[203] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.*, vol. 79, pp. 745–754, 1974.

[204] W. H. Richardson, "Bayesian-based iterative method of image restoration∗," *J. Opt. Soc. Am.*, vol. 62, pp. 55–59, Jan 1972.

[205] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem," *Nucl. Instrum. Meth. A*, vol. 362, pp. 487–498, 1995.

[206] B. Malaescu, "An iterative, dynamically stabilized (ids) method of data unfolding," pp. 271–275, 2011. arXiv:1106.3107.

[207] A. Bozson, G. Cowan, and F. Spanò, "Unfolding with Gaussian Processes," 11 2018.

[208] M. Kuusela and V. M. Panaretos, "Statistical unfolding of elementary particle spectra: Empirical Bayes estimation and bias-corrected uncertainty quantification," *Ann. Appl. Stat.*, vol. 9, pp. 1671–1705, 2015.

[209] P. C. HansenDepartment, "The l-curve and its use in the numerical treatment of inverse problems," 2000.

[210] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems.* Society for Industrial and Applied Mathematics, 1998.

[211] P. C. Hansen and D. O'leary, "The use of the l-curve in the regularization of discrete ill-posed problems," *SIAM J. Sci. Comput.*, vol. 14, pp. 1487–1503, 11 1993.

[212] B. Bischl, O. Mersmann, H. Trautmann, and C. Weihs, "Resampling methods for meta-model validation with recommendations for evolutionary computation," *Evolutionary computation*, vol. 20, pp. 249–75, 02 2012.

[213] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, (San Francisco, CA, USA), p. 1137–1143, Morgan Kaufmann Publishers Inc., 1995.

[214] J. Rodríguez, A. Pérez, and J. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, pp. 569 – 575, 04 2010.

[215] L. Brenner, R. Balasubramanian, C. Burgard, W. Verkerke, G. Cowan, P. Verschuuren, and V. Croft, "Comparison of unfolding methods using RooFitUnfold," *Int. J. Mod. Phys. A*, vol. 35, no. 24, p. 2050145, 2020.

[216] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[217] G. Cowan, "Response matrix from mc." http://www.pp.rhul.ac.uk/~cowan/stat/notes/response_matrix_from_mc.pdf.

[218] J. L. Horowitz, "Bootstrap methods in econometrics," *Annual Review of Economics*, vol. 11, no. 1, pp. 193–224, 2019.

[219] J. Chimka, "Bootstrap methods: A practitioner's guide," *IIE Transactions*, vol. 35, no. 6, pp. 583–583, 2003.

[220] F. Dekking, C. Kraaikamp, H. Lopuhaä, and L. Meester, *A Modern Introduction to Probability and Statistics: Understanding Why and How*. Springer Texts in Statistics, Springer, 2005.

[221] R. D. Cousins and V. L. Highland, "Incorporating systematic uncertainties into an upper limit," *Nuclear Instruments and Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment*, vol. 320, pp. 331–335, 1992.

[222] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O. AbouZeid, and et al., "Measurements of

normalized differential cross sections for tt production in pp collisions at $\sqrt{s} = 7$ tev using the atlas detector," *Physical Review D*, vol. 90, Oct 2014.

[223] G. Aad and et al., "Measurements of top-quark pair differential and double-differential cross-sections in the l+jets channel with pp collisions at $\sqrt{s} = 13$ tev using the atlas detector," *The European Physical Journal C*, vol. 79, Dec 2019.

[224] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A. A. Abdelalim, O. Abdinov, R. Aben, B. Abi, and et al., "Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ tev using the atlas detector at the lhc," *The European Physical Journal C*, vol. 73, Aug 2013.

[225] ATLAS, "Summary of ATLAS Pythia 8 tunes," tech. rep., CERN, Geneva, Aug 2012. ATL-PHYS-PUB-2012-003.

[226] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and et al., "Parton distributions with lhc data," *Nuclear Physics B*, vol. 867, p. 244–289, Feb 2013.

[227] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, and et al., "Parton distributions for the lhc run ii," *Journal of High Energy Physics*, vol. 2015, Apr 2015.

[228] "ATLAS Pythia 8 tunes to 7 TeV data," tech. rep., CERN, Geneva, Nov 2014. ATL-PHYS-PUB-2014-021.

[229] J. Gao, C. S. Li, and H. X. Zhu, "Top-quark decay at next-to-next-to-leading order in qcd," *Physical Review Letters*, vol. 110, Jan 2013.

[230] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders," *Comput. Phys. Commun.*, vol. 185, p. 2930, 2014.

[231] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation," *Physics Letters B*, vol. 710, p. 612–622, Apr 2012.

[232] A. Giammanco, "Single top quark production at the lhc," *Reviews in Physics*, vol. 1, p. 1–12, Nov 2016.

[233] S. Frixione, E. Laenen, P. Motylinski, C. White, and B. R. Webber, "Single-top hadroproduction in association with awboson," *Journal of High Energy Physics*, vol. 2008, p. 029–029, Jul 2008.

[234] "Studies on top-quark Monte Carlo modelling for Top2016," tech. rep., CERN, Geneva, Sep 2016. ATL-PHYS-PUB-2016-020.

[235] E. Boos, V. Bunichev, L. Dudko, I. Myagkov, and M. Perfilov, "Modeling of associated tw single top production including the interference with diagrams," *EPJ Web of Conferences*, vol. 158, p. 04003, 01 2017.

[236] "ATLAS simulation of boson plus jets processes in Run 2," tech. rep., CERN, Geneva, May 2017. ATL-PHYS-PUB-2017-006.

[237] C. Anastasiou, L. Dixon, K. Melnikov, and F. Petriello, "High-precision qcd at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order," *Physical Review D*, vol. 69, May 2004.

[238] G. Aad, B. Abbott, J. Abdallah, A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, and et al., "Measurement of the

top quark pair production cross-section with atlas in the single lepton channel," *Physics Letters B*, vol. 711, p. 244–263, May 2012.

[239] "Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," tech. rep., CERN, Geneva, Oct 2014. ATLAS-CONF-2014-058.

[240] G. Aad, B. Abbott, D. C. Abbott, A. A. Abud, K. Abeling, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al., "Measurements of top-quark pair differential and double-differential cross-sections in the l+jets channel with pp collisions at $\sqrt{s} = 13$ tev using the atlas detector," *The European Physical Journal C*, vol. 79, Dec 2019.

[241] C. Glasman, M. A. Principe Martin, J. Terron Cuadrado, M. Faucci Giannelli, D. Malito, E. Tassi, S. Palazzo, M. Romano, P. J. Verschuuren, L. Bellagamba, V. Boisvert, G. Cowan, F. Curcio, M. Negrini, and F. Spano, "Internal Document: Measurements of top-quark pair differential cross sections in the ℓ+jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 139 fb^{-1} of ATLAS data (Restricted)," tech. rep., CERN, Geneva, Apr 2021. ATL-COM-PHYS-2021-187.

[242] A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, and et al., "Measurements of differential cross sections of top quark pair production as a function of kinematic event variables in proton-proton collisions at $\sqrt{s} = 13$ tev," *Journal of High Energy Physics*, vol. 2018, Jun 2018.

[243] J. Mazzitelli, P. F. Monni, P. Nason, E. Re, M. Wiesemann, and G. Zanderighi, "Next-to-Next-to-Leading Order Event Generation for Top-Quark Pair Production," *Phys. Rev. Lett.*, vol. 127, no. 6, p. 062001, 2021.

[244] S. Höche, S. Kuttimalai, and Y. Li, "Hadronic final states in dis at nnlo qcd with parton showers," *Physical Review D*, vol. 98, Dec 2018.

[245] E. Re, "MiNNLO$_{PS}$: a new method to match NNLO QCD with parton showers," in *55th Rencontres de Moriond on QCD and High Energy Interactions*, 7 2021.

[246] "Study of top-quark pair modelling and uncertainties using ATLAS measurements at $\sqrt{s}$=13 TeV," tech. rep., CERN, Geneva, Sep 2020. ATL-PHYS-PUB-2020-023.

[247] "Constraints on Higgs boson properties using $WW^*(\to e\nu\mu\nu)\ jj$ production in 36.1fb^{-1} of $\sqrt{s} = 13TeV$ pp collisions with the ATLAS detector," tech. rep., CERN, Geneva, Nov 2020. ATLAS-CONF-2020-055.

[248] G. A. et al., "Measurement of the associated production of a higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at s=13 tev with the atlas detector," *Physics Letters B*, vol. 816, p. 136204, 2021.

[249] G. Aad *et al.*, "Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector," *Eur. Phys. J. C*, vol. 81, no. 2, p. 178, 2021.

[250] G. Aad *et al.*, "Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Eur. Phys. J. C*, vol. 80, p. 957. 76 p, Apr 2020.

[251] G. Aad *et al.*, "Measurements and interpretations of Higgs-boson fiducial cross sections in the diphoton decay channel using 139fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector," tech. rep., CERN, Geneva, Jul 2019. ATLAS-CONF-2019-029.

[252] G. Aad *et al.*, "Measurements of $W^+W^- + \geq 1$ jet production cross-sections in *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP*, vol. 2106, p. 003. 58 p, Mar 2021.

[253] G. Aad *et al.*, "Measurements of differential cross-sections in four-lepton events in 13 TeV proton-proton collisions with the ATLAS detector," *JHEP*, vol. 07, p. 005, 2021.

[254] G. Aad *et al.*, "Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS," *Eur. Phys. J. C*, vol. 81, no. 2, p. 163, 2021.

[255] M. Aaboud *et al.*, "Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Eur. Phys. J. C*, vol. 79, no. 10, p. 884, 2019.

[256] M. J. Kareem, "Combination of the W boson polarization measurements in top quark decays using ATLAS and CMS data at $\sqrt{s} = 8$ TeV (Restricted)," Jul 2020.

[257] "Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector," tech. rep., CERN, Geneva, Jul 2019. ATLAS-CONF-2019-026.

[258] M. e. a. Aaboud, "Measurement of the $t\bar{t}z$ and $t\bar{t}w$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the atlas detector," *Phys. Rev. D*, vol. 99, p. 072009, Apr 2019.

[259] G. Aad *et al.*, "Search for flavour-changing neutral currents in processes with one top quark and a photon using 81 fb^{-1} of *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS experiment," *Phys. Lett. B*, vol. 800, p. 135082, 2020.

[260] "Search for flavour-changing neutral current top quark decays $t \to qZ$ in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector," tech. rep., CERN, Geneva, Sep 2017. ATLAS-CONF-2017-070.

[261] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, "Ufo – the universal feynrules output," *Computer Physics Communications*, vol. 183, p. 1201–1214, Jun 2012.

[262] I. Brivio, Y. Jiang, and M. Trott, "The smeftsim package, theory and tools," *Journal of High Energy Physics*, vol. 2017, Dec 2017.

[263] I. Brivio, "Smeftsim 3.0 — a practical guide," *Journal of High Energy Physics*, vol. 2021, Apr 2021.

[264] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang, "Operator and wilson coefficient definitions of smeftatnlo." http://feynrules.irmp.ucl.ac.be/attachment/wiki/SMEFTatNLO/definitions.pdf.

[265] J. J. Ethier, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang, "Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC," arXiv:2105.00006.

[266] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff, and C. Zhang, "O new physics, where art thou? a global search in the top sector," *Journal of High Energy Physics*, vol. 2020, Feb 2020.

[267] E. Conte, B. Fuks, and G. Serret, "Madanalysis 5, a user-friendly framework for collider phenomenology," *Computer Physics Communications*, vol. 184, p. 222–256, Jan 2013.

[268] J. Y. Araz, B. Fuks, and G. Polykratis, "Simplified fast detector simulation in madanalysis 5," *The European Physical Journal C*, vol. 81, Apr 2021.

[269] J. Y. Araz, M. Frank, and B. Fuks, "Reinterpreting the results of the lhc with madanalysis 5: uncertainties and higher-luminosity estimates," *The European Physical Journal C*, vol. 80, Jun 2020.

[270] E. Conte and B. Fuks, "Confronting new physics theories to lhc data with madanalysis 5," *International Journal of Modern Physics A*, vol. 33, p. 1830027, Oct 2018.

[271] B. Dumont, B. Fuks, S. Kraml, S. Bein, G. Chalons, E. Conte, S. Kulkarni, D. Sengupta, and C. Wymant, "Toward a public analysis database for lhc new physics searches using madanalysis 5," *The European Physical Journal C*, vol. 75, Feb 2015.

[272] E. Conte, B. Dumont, B. Fuks, and C. Wymant, "Designing and recasting lhc analyses with madanalysis 5," *The European Physical Journal C*, vol. 74, Oct 2014.

[273] O. Mattelaer, "On the maximal use of monte carlo samples: re-weighting events at nlo accuracy," *The European Physical Journal C*, vol. 76, Dec 2016.

[274] C. Bierlich, A. Buckley, J. Butterworth, C. H. Christensen, L. Corpe, D. Grellscheid, J. F. Grosse-Oetringhaus, C. Gutschow, P. Karczmarczyk, J. Klein, and et al., "Robust independent validation of experiment and theory: Rivet version 3," *SciPost Physics*, vol. 8, Feb 2020.

[275] Rivet, "Rivet Routine ATLAS_2019_I1750330." https://rivet.hepforge.org/analyses/ATLAS_2019_I1750330.html.

[276] G. Aad, B. Abbott, D. C. Abbott, A. A. Abud, K. Abeling, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al., "Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s} = 13$ tev using the atlas detector," *The European Physical Journal C*, vol. 79, Dec 2019.

[277] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang.

[278] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff, and C. Zhang, "O new physics, where art thou? A global search in the top sector," *JHEP*, vol. 02, p. 131, 2020.

[279] A. Buckley, P. Ilten, D. Konstantinov, L. Lönnblad, J. Monk, W. Pokorski, T. Przedzinski, and A. Verbytskyi, "The hepmc3 event record library for monte carlo event generators," *Computer Physics Communications*, vol. 260, p. 107310, 2021.

[280] V. Gherardi, D. Marzocca, and E. Venturini, "Matching scalar leptoquarks to the smeft at one loop," *Journal of High Energy Physics*, vol. 2020, Jul 2020.

[281] J. Ellis, M. Madigan, K. Mimasu, V. Sanz, and T. You, "Top, higgs, diboson and electroweak fit to the standard model effective field theory," *Journal of High Energy Physics*, vol. 2021, Apr 2021.

[282] J. C. Criado, "Matchingtools: A python library for symbolic effective field theory calculations," *Computer Physics Communications*, vol. 227, p. 42–50, Jun 2018.

[283] S. D. Bakshi, J. Chakrabortty, and S. K. Patra, "Codex: Wilson coefficient calculator connecting smeft to uv theory," *The European Physical Journal C*, vol. 79, Jan 2019.

[284] J. Cresser, "Macquarie university, lecture notes phys301: Observables and measurements in quantum mechanics." `http://physics.mq.edu.au/~jcresser/Phys301/Chapters/Chapter13.pdf`.

[285] L. D. Debbio, "University of edinburgh, lecture notes qm: Observables." `https://www2.ph.ed.ac.uk/~ldeldebb/docs/QM/lect2.pdf`.

[286] P. Denteneer, "Leiden university, lecture notes for quantum theory: Second quantization." `http://wwwhome.lorentz.leidenuniv.nl/~pjhdent/SecQuant08.pdf`.

[287] F. Essler, "Oxford university, lecture notes for advanced quantum theory." `http://www-thphys.physics.ox.ac.uk/courses/C6/C6_2017/C6Notes/lecturenotes2017AQT.pdf`.

[288] P. Stamp, "University of california, berkeley: Lecture notes for quantum field theory ii." `https://phas.ubc.ca/~stamp/TEACHING/PHYS508/NOTES/B1.pdf`.