Dissertation

Automated optimization of sensitivity in a search for boosted VBF Higgs pair production in the $b\bar{b}b\bar{b}$ quark final state with the ATLAS detector

For the attainment of the academic degree doctor rerum naturalium

(Dr. rer. nat.) in the subject: Physics

Frederic Renner

Berlin, 01.12.2023

Faculty of Mathematics and Natural Sciences of the Humboldt University of Berlin

1st Supervisor: Dr. Clara Elisabeth Leitgeb

2nd Supervisor: Prof. Dr. Cigdem Issever

(Only after the disputation for publication in the university library according to § 15 of the doctoral regulations enter the names and the date):

Reviewers:

1st:

2nd:

3rd:

Date of the oral examination:

Abstract

I am an abstract.

Contents

Chapter 1

The HH→4b analysis

The search for the exact shape of the Higgs potential is an interesting endeavor as its not only directly related to Electroweak Symmetry Breaking (EWSB) but also could solve some fundamental questions about the nature of the universe as described in section **??**. Since the Higgs interacts via Yukawa couplings from equation **??** the coupling strengths for fermions are directly proportional to their mass and thus the Higgs couples most strongly to heavy particles. The main production modes at the Large Hadron Collider (LHC) are shown in figure 1.1. All couplings in the following are scaled with respect to their Standard Model (SM) values and are denoted with $\kappa_c = c/c_{sm}$ so that $\kappa_c = 1$ represents the SM value of some coupling c.

The dominant Higgs pair production processes are shown in figure 1.1. The first two gluon-gluon fusion (GGF) diagrams (a) and (b) have a cross-section of $\sigma^{SM}_{vbf\,HH} = 31.05\,\text{fb}$ calculated at a center of mass energy of 13 TeV at next-to-next-to-leading order (NNLO) [1] while the vector-boson fusion (VBF) processes (c), (d) and (e) of figure 1.1 have a production cross-section of $\sigma^{SM}_{vbf\,HH} = 1.73\,\text{fb}$ at next-to-next-to-next-to-leading order (N^3LO) [2]. A characteristic of the VBF processes is that the Higgs pair products are accompanied by two additional quarks. The VBF cross section is about 3×10^4 times smaller than the production cross section for single Higgs $\sigma^{SM}_{H} = 48.58\,\text{pb}$ at the LHC [3] which already illustrates the challenge of discovering Higgs pairs in these final states.

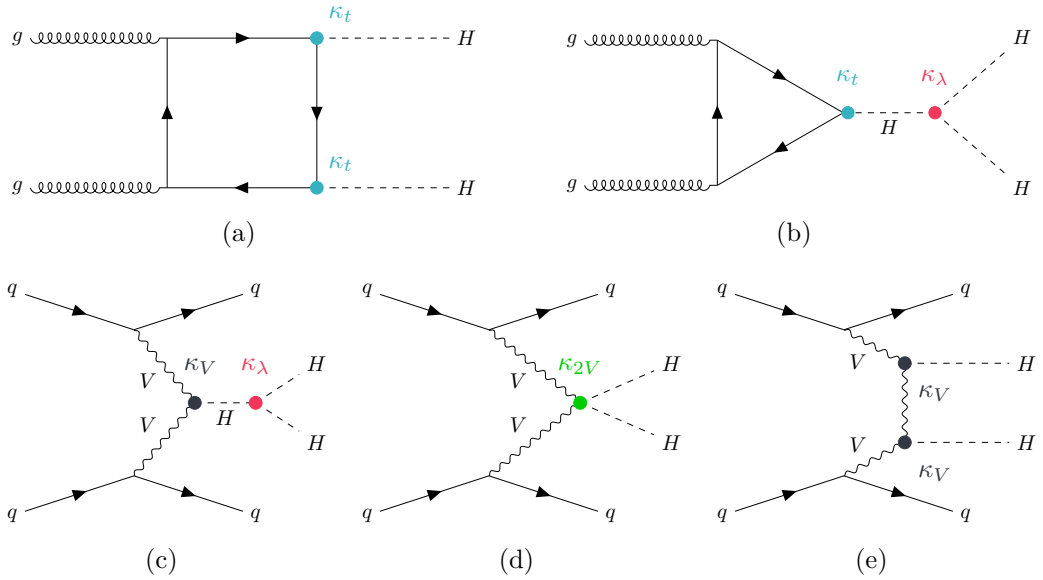


Figure 1.1: Leading Higgs Pair production processes at the LHC. (a), (b) shows GGF and (c), (d), (e) VBF processes. Adopted from [4].

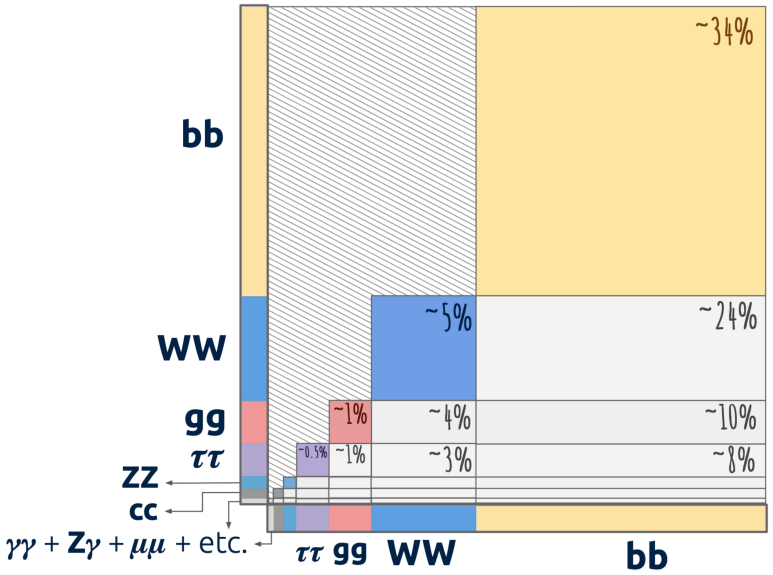


Figure 1.2: Contributions of final states represented by area for a pair of Higgs. Adopted from [5].

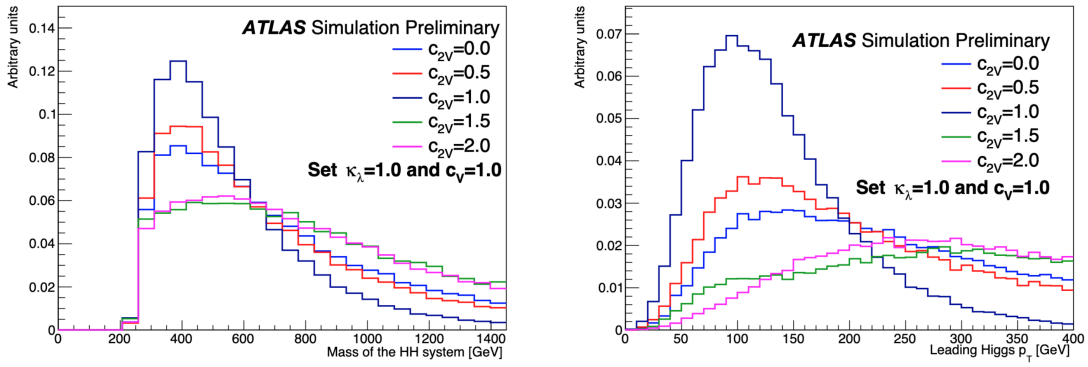


Figure 1.3: Invariant mass of the Higgs pair system and the leading Higgs candidate jet p_{T} reconstructed from simulation for different $\kappa_{2\mathrm{V}}$. Adopted from [7].

Figure 1.2 reveals that an interesting channel is the final state with the largest branching fraction of about 34 percent consisting of four b-quarks. However as this a fully hadronic final state it comes with the challenge of large Quantum Chromodynamics (QCD) backgrounds.

This work focuses on the boosted topology of highly energetic jets which do not allow reconstruction of b-jets individually but rather of final states consisting of two collimated b-jets inside a larger jet. The advantage of this selection is that it reduces greatly the QCD backgrounds since highly energetic jets aare more likely to come from heavy particles such as b quarks. Furthermore it is easy to trigger on events containing jets with large p_{T}. Although they represent a comparatively clean signal such events are rare and therefore have limited statistical power. For this reason other search strategies are better suited for the discovery of the Higgs pair production process.

A reason for the low cross-section is that diagrams (d) and (e) of figure 1.1 cancel each other destructively for SM values. In turn if $\kappa_{2\mathrm{V}}$ is moved to non-SM values the production cross-section increases significantly $\sigma_{\kappa_{2\mathrm{V}}=0} \approx 20\sigma_{\kappa_{2\mathrm{V}}=1}$ [6] and the decay products have much larger transverse momentum as shown in figure 1.3. Because of this behavior the power of this analysis lies in constraining and proving the existence of the $\kappa_{2\mathrm{V}}$ couplings within the SM to which it is directly sensitive.

1.1 Data and Monte Carlo Simulation

This analysis uses the full run 2 data taken by A Toroidal LHC Apparatus (ATLAS) between 2015 and 2018. The dataset contains 140.1 fb^{-1} of data good for physics at a center of mass energy of 13 TeV [8].

Monte Carlo (MC) generation in ATLAS is done in three steps. At first at parton level events are generated with MadGraph (v.2.7.3p3.atlas6) [9]. The output of this tool are stochastically sampled four vectors of the final states of the process of interest.

Afterwards the hadronization process is simulated with Pythia8 [10].For this the NNPDF3.0nlo Parton Density Function (PDF) is used. The SM VBF cross-section is rendered to the 4b branching ratio by multiplying it with $\mathcal{B}(4b) = 0.3392$.

add samples linear combination

1.2 Analysis strategy

This section describes the event selection and analysis strategy. A detailed description of reconstructed physical objects used is described in chapter **??**.

1.2.1 Trigger

As outlined in section **??** events need to be preselected/triggered. The high level trigger (HLT) applied in this analysis selects events with a large transverse energy E_{T} large R jet. The definition slightly changed over the data taking years as can be seen in table 1.1. Previous studies have shown that they become fully efficient

Table 1.1: Trigger selections per data taking year and minimum requirements on transverse energy E_{T} and mass m on the large R jet.

Year	E_{T}	m
2015	> 360	0
2016	> 420	0
2017	> 420	> 35
2018	> 420	> 40

at about $p_{\text{T}} > 420\,\text{GeV}$ [5, 11].

1.2.2 Large Radius Jets

To fully capture the boosted Higgs pair topology two large $R = 1.0$ jets clustered with the Anti-k_t algorithm from Track CaloClusters (TCCs) are used as described in section **??**. These enclose the two boosted collimated b-jets in each of them to form the Higgs candidates. If there are several large-R jets the two with the highest p_{T} are chosen. To be fully efficient on the trigger the leading large-R jet is required to have $p_{\mathrm{T}} > 450\,\mathrm{GeV}$. In general for decay products to be inside a jet approximately holds $R \approx 2m/p_{\mathrm{T}}$ with the mass m and tranverse momentum of the parent particle [5]. For a Higgs mass of 125 GeV to be contained inside a large-R jet the Higgs candidate therefore must have $p_{\mathrm{T}} \gtrsim 250\,\mathrm{GeV}$ and is thus chosen as the p_{T} requirement on the sub-leading Higgs candidate. Additionally both Higgs candidates have a mass requirement $m > 50\,\mathrm{GeV}$ to reduce QCD background.

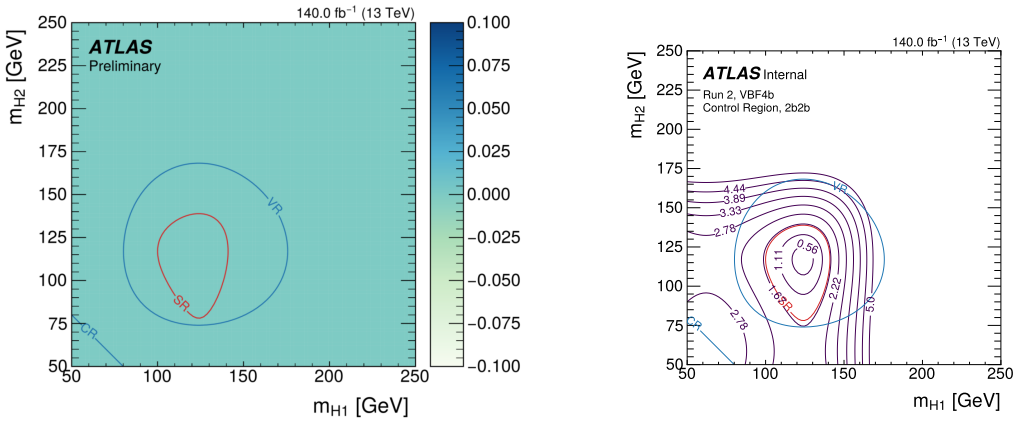
The identification of b-jets inside the selected large-R jets is done with the $X \to bb$ tagger explained in section **??**. The top fraction f_{top} is set to 0.25 and the 60 % Higgs efficiency working point (WP) is required. Studies with the more inclusive 60 % WP displayed slightly worse limit results [11].

1.2.3 Small Radius Jets

Two small radius $R = 0.4$ jets are required for the VBF signature and are referred to as VBF jets in the following. They are also reconstructed with the anti-k_t algorithm and as Particle Flow Objects (PFOs) as described in **??**. The tight WP for the jet vertex tagger (JVT) ad the LooseBad WPfor the event cleaning are applied both described in **??**. Small-R jets j are selected for $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 4.5$ and are required to be outside of the Higgs candidate large-R jets J by imposing $\Delta R(J, j) > 1.4$. Further cuts applied on the jet system optimized on significance are $|\Delta\eta(j, j)| > 3$ and $m_{jj} > 1\,\mathrm{TeV}$.

1.2.4 Kinematic Regions

Signal Region (SR),variable radius (VR) and Control Region (CR) are explored and optimized in previous analyses [4, 11] in the m_{H1}, m_{H2} plane and are defined



(a) (b)

Figure 1.4: REDO

as

$$SR = X_{hh} = \sqrt{\left(\frac{m_{H1} - 124\,\text{GeV}}{1500/m_{H1}}\right)^2 + \left(\frac{m_{H2} - 117\,\text{GeV}}{1900/m_{H2}}\right)^2} < 1.6, \tag{1.2.1}$$

$$VR = \sqrt{\left(\frac{m_{H1} - 124\,\text{GeV}}{0.1\ln(m_{H1})}\right)^2 + \left(\frac{m_{H2} - 117\,\text{GeV}}{0.1\ln(m_{H2})}\right)^2} < 100, \tag{1.2.2}$$

and

$$CR = \sqrt{\left(\frac{m_{H1} - 124\,\text{GeV}}{0.1\ln(m_{H1})}\right)^2 + \left(\frac{m_{H2} - 117\,\text{GeV}}{0.1\ln(m_{H2})}\right)^2} > 100\ \&\ < 170. \tag{1.2.3}$$

Figure depicts the regions in the m_{H1}, m_{H2} plane on the SM signal sample

1.2.5 Cutflow

TODO, also fine like that?

1.2.6 Analysis Optimization

present here the used NN, retrieved from neos from chapter blah

Selection	Event	Fraction [%]	Total Fraction [%]
Initial	16854036422.000		
Preselections (MNT + Jet Cleaning)	670573995.000	100.000	100.000
PassTrigBoosted	63944638.000	9.536	9.536
PassTwoFatJets	57510800.000	89.938	8.576
PassTwoHbbJets	12875.000	0.0223	<0.001
PassVBFJets	5762.000	44.753	<0.001
PassFatJetPt	3902.000	67.720	<0.001
PassVBFCut	314.000	8.047	<0.001

Table 1.2: Cut-flow table for data before signal region cut

Selection	Event	Fraction [%]	Total Fraction [%]
Initial	1475.226		
Preselections (MNT + Jet Cleaning)	547.960	100.000	100.000
PassTrigBoosted	20.926	3.819	3.819
PassTwoFatJets	14.141	67.576	2.581
PassTwoHbbJets	5.353	37.852	0.977
PassVBFJets	2.243	41.903	0.409
PassFatJetPt	1.408	62.793	0.257
PassVBFCut	0.148	10.539	0.027
PassSR	0.097	65.484	0.018
OverlapRemoval	0.059	61.200	0.011

Table 1.3: Cut-flow table for DSID = 600463

1.2.7 Background Estimation

Since the final state of this analysis is hadronic it remains an infeasible task to estimate the contributions from the plethora of QCD processes that contribute to the background. Therefore the well established ABCD method is employed to derive a data-driven background estimate [12, 13]. Its is based on the idea to use two independent variables e.g. f and g to define four orthogonal regions A,B,C and D as illustrated in figure 1.5 so that for some combination of the ratio of the event yields in the regions hold

$$\frac{N_A}{N_B} = \frac{N_C}{N_D}. \tag{1.2.4}$$

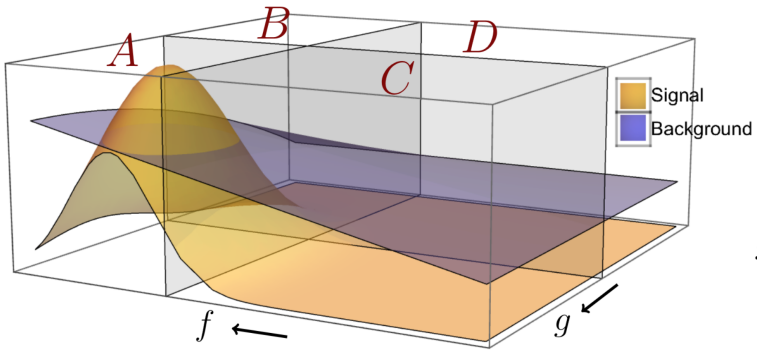


Figure 1.5: Illustration of four orthogonal regions A,B,C and D defined by two variables f and g in the horizontal and signal and background yields in the vertical dimension. Adopted from [13].

By rearranging the equation for the unknown N_A an estimate for the background of the signal region can be derived from the other known quantities that lie in the regions dominated by the background. This approach only works well if the change in shape of the background in Figure 1.5 does not vary greatly between C to D and A to B. Therefore the method should always be tested in another region to proof its validity.

In this analysis the two orthogonal variables are defined via the amount of $X \to bb$ Higgs tagged large-R jets denoted as Xbb and the kinematic regions of the SR and CR defined in 1.2.4. This gives the four orthogonal regions shown in

Table 1.4: Four orthogonal region definitions for the ABCD method

2 Xbb in CR	2 Xbb in SR
1 Xbb in CR	1 Xbb in SR

table 1.4 so that the background in the SR is estimated with a weight extracted from the CR

$$N_{\mathrm{SR}}^{\mathrm{2Higgs}} = \frac{N_{\mathrm{CR}}^{\mathrm{2Higgs}}}{N_{\mathrm{CR}}^{\mathrm{1Higgs}}} N_{\mathrm{SR}}^{\mathrm{1Higgs}} = w_{\mathrm{CR}} N_{\mathrm{SR}}^{\mathrm{1Higgs}}. \tag{1.2.5}$$

question is if we put the validation already here, conflict with neos description, beacuse would need to explain the nn used before to understand validation

studies with a bin-wise trained Neural Network (NN) didnt gain anything (like in the resolved analysis), but lets put this as well in the results part

Chapter 2

Analysis Optimization

After a thoughtful design of physical objects that correspond to the final state of the process of interest the main task remains to construct some quantity to compare the prediction with the observation. Usually this is done with some kinematic variable e.g. in this analysis it could be the invariant mass of the Higgs pair system m_{HH} designed to differentiate signal from background. Events are then counted with a histogram for this variable to be then tested on the statistical significance of an hypothesis.

Since the purpose of this quantity is to classify it is actually better suited to be treated with a Machine Learning (ML) model as already applied widely in particles physics [14–17]. However a key issue with any optimization or ML model used in a typical particle physics analysis is that it is not optimized for the actual final quantity of interest like a discovery significance or the level of confidence if a tested hypothesis should be accepted or rejected. Usually optimization is done with some ratio of signal to background or in the case of ML models by optimizing on the best separation of signal and background(s).

Even though the signal to background ratio or the separation power of the ML model is correlated in some way to the expressivity of the statistical test it is not optimal and in particular does not account for uncertainties that can crucially influence the result of the statistical test. This means if the statistical test performs very well on the nominal values it is trained or tested on its behavior can change

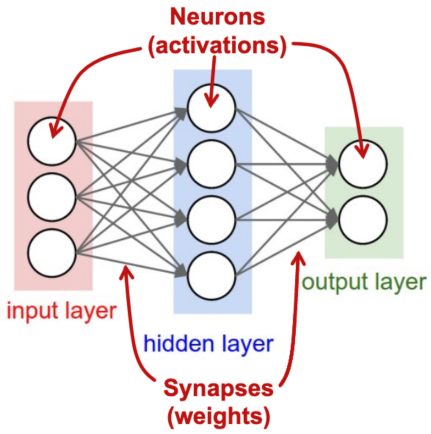


Figure 2.1: Structure of artificial feed-forward neural networks. Adopted from [20].

dramatically if given inputs that deviate from the nominal values (uncertainties) and the power of the statistical test degrades accordingly.

This issue can be addressed with: *neural end-to-end-optimized summary statistics (NEOS)* [18]. It is based on the observation that the classifier training can be done with respect to the actual final quantity of interest by including the statistcal model into the training.

2.1 Machine Learning

ML refers to algorithms that enable computers to learn from data to make predictions for some specific task without being explicitly programmed for [19]. One particular subset of ML are Artificial Neural Networks (ANNs) inspired by the human brain. Their fundamental unit are nodes, the neurons, that are interconnected to several other neurons organized in consecutive layers with an initial input and a desired output as shown in figure 2.1. The signals between neurons are transferred weighted and each neuron has an activation function that converts the received input stimulus into an output strength. Thus learning occurs through adjusting the weights between the neurons. ANNs can be designed in various ways with many hidden or deep layers depending on the specific task.

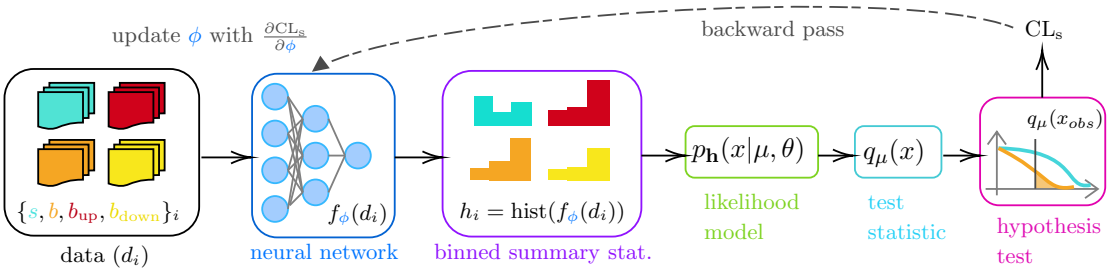


Figure 2.2: Typical particle physics analysis chain. For NEOS the CL_s value is back-propagated to train the neural network parameters $\boldsymbol{\varphi}$. Adopted from [18].

The training of NNs is usually done with the back-propagation algorithm which seeks to minimize a cost function $C(\boldsymbol{\varphi})$ that is designed to measure the deviance of a presented input to a desired output dependent on the model parameters $\boldsymbol{\varphi}$. For a feed-forward NN these model parameters are the mentioned weights of the neurons. Via gradient descent the minimum of the cost function can be found stepwise with the first derivative and a chosen learning rate parameter γ

$$\boldsymbol{\varphi}_{n+1} = \boldsymbol{\varphi}_n - \gamma \nabla C(\boldsymbol{\varphi}_n). \tag{2.1.1}$$

In three dimensions this corresponds to a mountaineer searching for the valley by going in the direction of the steepest crescent with a stepsize proportional to γ.

2.2 NEOS

The key idea of NEOS is now to choose the cost function as the quantity which is to be minimized. In this case the CL_s value is a reasonable choice as discussed section ??. For a typical analysis chain as shown in figure 2.2 this corresponds mathematically to a functional concatenation of the individual steps in the chain, as each step depends on the previous one. Hence the cost function CL_s is a function of the Dataset $\mathcal{D}$ and the ML model parameters $\boldsymbol{\varphi}$

$$\mathrm{CL}_s = f(\mathcal{D}, \boldsymbol{\varphi}) = (f_{\text{sensitivity}} \circ f_{\text{test stat}} \circ f_{\text{likelihood}} \circ f_{\text{histogram}} \circ f_{\text{observable}})(\mathcal{D}, \boldsymbol{\varphi}). \tag{2.2.1}$$

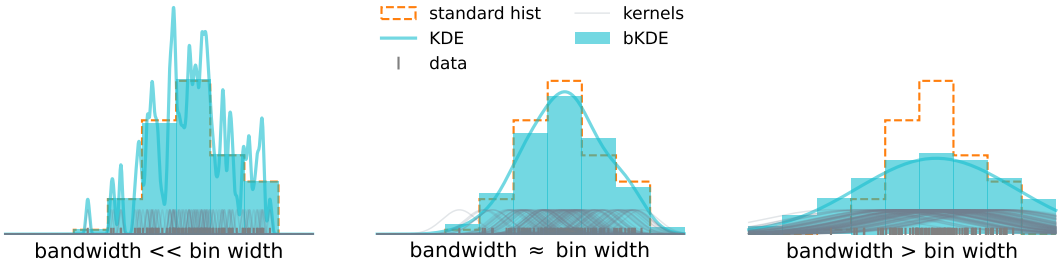


Figure 2.3: Dependence of the histogram approximation with normal distributions of different standard deviations called bandwidth in the context of KDE. Depicted in grey small bars on the x-axis are the data points and their kernel estimates above them, the standard histogram binned from data, the KDE and the binned Kernel Density Estimation (bKDE) as a histogram resulting from the KDE. Adopted from [18].

In order to find the minimum via the gradient descent of equation 2.1.1 CL_s needs to be differentiable with respect to $\boldsymbol{\varphi}$. By applying the chain rule this reads for one model parameter φ_i

$$\frac{\partial \mathrm{CL}_s}{\partial \varphi_i} = \frac{\partial f_{\text{sensitivity}}}{\partial f_{\text{test stat}}} \frac{\partial f_{\text{test stat}}}{\partial f_{\text{likelihood}}} \frac{\partial f_{\text{likelihood}}}{\partial f_{\text{histogram}}} \frac{\partial f_{\text{histogram}}}{\partial f_{\text{observable}}} \frac{\partial f_{\text{observable}}}{\partial \varphi_i}. \tag{2.2.2}$$

Apart from histogramming which is inherently discrete all steps are differentiable. One way to make histograms differentiable is Kernel Density Estimation (KDE) [21]. Instead of counting a quantity in bins, a histogram can be approximated by representing each data point by a normal distribution, the kernel, with the mean of the data point and a chosen value for the standard deviation, also called the bandwidth in this context. Since the area under the Gaussian is equal to one, the summation of all Gaussian's can produce a smoothed estimate of a histogram that is inherently differentiable. However the bandwidth should be chosen around the desired binning width as the estimate depends crucially on it, as exemplified in Figure 2.3. Although this can be a source of uncertainty one can always revert any differentiable step or block in figure 2.2 and go back to the exact calculation.

2.2.1 Implementation

To trace the differentiation through software is again achieved with a simple observation that in fact any software is only a concatenation of elementary arithmetic operations. Since their analytic solutions are known all there is to do is the concatenate of all these operations and differtiate this function to calculate. This is also known as *Automatic Differentiation* and is achieved here with the help of the software package JAX [22].

Although the idea seems simple and one of the reasons it has not been attempted previously is that it can be conceivably difficult to differentiate through multiple individual software frameworks. That this became feasible in a reasonable time-frame builds on the efforts to port the individual steps to the PYTHON programming language away from C++ ROOT-based software [23]. While ROOT was indispensable at its time of emergence it is currently suboptimal for contemporary scientific analysis. This is because it is not only difficult to integrate within other software but also comparatively user-unfriendly compared to readily available PYTHON solutions maintained by a scientific community doing data-analysis. Using PYTHON in this context presents a notable advantage since many problems have already been addressed and instead of relying on a small group for maintenance there is support from a large community. This enables out-of-the-box applications as for example the use of JAX which a priori has no connection to High Energy Physics (HEP).

The original proposers Simpson and Heinrich [18] of NEOS developed differentiable versions of common HEP tasks like the upper mentioned hypothesis test, histogramming and the optimization of a cut in a packaged called RELAXED [24]. Using this they tested the whole pipeline successfully in a toy model [25]. These efforts were transferred and further developed to be applicable to this analysis with [26]. how about a better repo name: sensitivo, sensitivity optimizer?

Chapter 3

Systematics Uncertainties

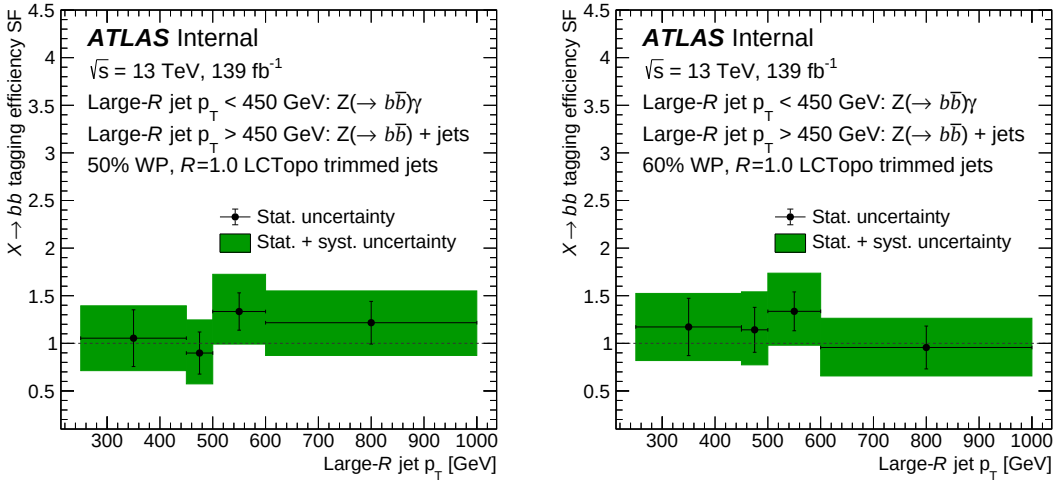
Any measurement needs to consider uncertainties in order to determine its validity. In this analysis they can be divided into systematic errors for the reconstructed objects, uncertainties from theoretical calculations, methodological errors and statistical uncertainties and are described in the following

3.1 Jet Uncertainties

Jets are calibrated using well known reference objects as described in section **??**. These corrections are themselves subject to uncertainties related to detector effects, modeling and statistics leading to corrections of the jet energy and are collectively referred to as Jet Energy Scale (JES) [27, 28]. Since simulations of jets have a higher accuracy than observed jets the uncertainties of the simulated jets are broadened to be consistent with the jets observed in the data. These uncertainties are known as Jet Energy Resolution (JER). Furthermore large-R jets are additionally corrected for their mass. The uncer tainties related to this procedure are called Jet Mass Resolution (JMR) [29].

3.2 $X \to bb$ Tagger Uncertainties

The NN of the $X \to bb$ tagger was trained using simulations leading to potential discrepancies in selection efficiencies between observed data and simulation. Cal-



(a) (b)

Figure 3.1: Derived scale factors in large-R jet p_{T} for the **(a)** 50 % and **(b)** 60 % WP from the calibration of the $X \to bb$ tagger.

ibration is conducted with $Z(\to b\bar{b})+$ jets and $Z(\to b\bar{b}) + \gamma$ applying the same methodology as in [30]. However as of this analysis the b-tagging algorithm for the VR track jets has been updated to the DL1d algorithm described in section **??**. The differences between MC and data are measured in large-R jet p_{T} and the extracted scale factors and their corresponding combined systematic and statistical uncertainties are shown in figure 3.1.

3.3 Theory Uncertainties

The cross-section calculation for some process initiated by a proton proton collision calculated at n-th order has a functional form [31]

$$\sigma^{(n)} = PDF(x_1, \mu_F)PDF(x_2, \mu_F)\hat{\sigma}^{(n)}(x_1, x_2, \mu_R), \tag{3.3.1}$$

with the Parton Density Functions (PDFs) carrying momentum fraction $x_{1,2}$ of the partons and the factorization scale μ_F. This scale is named after the assumption that the cross-sections of the initial particle can be calculated by factorizing it in

its parton contributions [32]. The term $\hat{\sigma}^{(n)}$ in equation 3.3.1 is the calculable part of the cross-section at renormalization scale μ_R as described in section **??** and is expanded to a desired order n in the strong coupling constant α_s with the usual Quantum Field Theory (QFT) ansatz outlined in section **??**

$$\hat{\sigma}^{(n)} = \alpha_s\hat{\sigma}^{(0)} + \alpha_s^2\hat{\sigma}^{(1)} + \ldots + \alpha_s^n\hat{\sigma}^{(n)} + \mathcal{O}(\alpha_s^{n+1}). \tag{3.3.2}$$

Modelling cross-sections via PDFs is necessary since the approximation of the perturbation ansatz of section **??** breaks down for low energy scales Q^2 as described in section **??** which is the energy scale for which the approximation would need to hold to describe the partons inside a proton. However similar to renormalization a scaling behavior can be derived which allows to deduce an estimate of the PDFs by measuring it at a some energy scale Q^2 to extrapolate it to another. The equations enabling this are also expanded in α_s to a desired order and are knwon as DGLAP equations [32]. Three main sources of uncertainty arise in this calculation described in the following.

Scale Variations

α_s is expanded to some order n in the cross-section calculation and as well in estimating the PDFs. To account for missing higher orders corrections of theses expansions scale variations of the renormalization and factorization scales are performed pairwise $\{\mu_\mathrm{r}, \mu_\mathrm{f}\} \times \{0.5, 0.5\}, \{1, 0.5\}, \{0.5, 1\}, \{1, 1\}, \{2, 1\}, \{1, 2\}, \{2, 2\}$. For the cross-section calculation this accounts essentially for the term $\mathcal{O}(\alpha_s^{n+1})$ in equation 3.3.2. The envelope that gives the largest variation is taken as the scale uncertainty.

PDF Uncertainties

PDFs need to be deduced from experiment and thus come by themselves with experimental uncertainties. Further uncertainties arise from the functional forms assumed for the PDFs.

α_s Uncertainties

α_s is also experimentally deduced at the scale of the Z mass which is subject to uncertainties. In all perturbative calculations it is truncated at some order that needs to be accounted for.

The uncertainties on α_s and the PDFs are both estimated by varying α_s. Even though there correlation is not strong they are usually applied combined [31].

3.3.1 Uncertainty on HH cross section

The cross-section calculation for the VBF Higgs pair production process has associated uncertainties for the scale variations $^{+0.03\%}_{-0.04\%}$ and the combined PDF+α_s uncertainty is $\pm 2.1\%$ [3].

3.3.2 Uncertainty on Acceptance

Theoretical uncertainties on the final acceptance are evaluated on MC simulations for scale variations and PDFs + α_s TODO, although shouldnt matter...

3.3.3 Parton Shower

Pythia 8 and Herwig 7 are compared. TODO

3.3.4 Branching Ratio Uncertainty

The error estimate for the branching ratio takes into account theoretical uncertainties (THU) and parametric uncertainties (PU) that are included in the SM calculations. The theoretical uncertainties mainly considers missing higher orders while for the parameters p the four leading non-negligible contributions of $p = \alpha_s, m_c, m_b, m_t$ are considered.

Parametric uncertainties are Gaussian errors and are added in quadrature which ensures unity in the Branching Ratio calculation. Theoretical uncertainties in turn are not Gaussian and would lead to underestimated errors and are therefore are added linearly [3]. By assuming a Higgs mas of 125 GeV since and considering that

there are two Higgs decaying to two b-quarks the error on the branching is

$$\Delta\mathrm{BR} = 2 \times \left(\Delta\mathrm{BR}(\mathrm{THU}) + \sqrt{\sum_p \Delta\mathrm{BR}(\mathrm{PU}_p)^2}\right) = ^{+3.4\%}_{-3.5\%} . \tag{3.3.3}$$

3.3.5 Reweighting Uncertainties

3.4 Background Modelling Uncertainties

3.5 Statistical Uncertainties

As discussed in the chapter on statistics **??** the bin content for histograms in this work follows a Poisson distribution thus the standard error for N events is the square root of the variance $\sigma = \sqrt{\mathrm{Var}} = \sqrt{N}$. Since histograms are filled weighted this needs to be taken into account. The variance of a bin filled with weights w_i is $\sigma^2 = \mathrm{Var}_{\mathrm{bin}}(w_i)$. Since the variance is additive and invariant with respect to constants it follows for the variance of the bin decomposed per event

$$\sigma^2_{\mathrm{stat}} = \mathrm{Var}_{\mathrm{bin}}\left(\sum_i w_i\right) = \underbrace{\sum_i \mathrm{Var}(w_i \times 1 \text{ event})}_{\mathrm{Var}(i+j)=\mathrm{Var}(i)+\mathrm{Var}(j)} = \underbrace{\sum_i w_i^2 \mathrm{Var}(1 \text{ event})}_{\mathrm{Var}(aX)=a^2\mathrm{Var}(X)} \tag{3.5.1}$$

$$= \sum_i w_i^2 \sqrt{(1 \text{ event})}, \tag{3.5.2}$$

so that the statistical error of a bin is

$$\sigma^{\mathrm{bin}}_{\mathrm{stat}} = \sqrt{\sum_i w_i^2} \tag{3.5.3}$$

methodische errors, bkg unc
modelling errors
luminosity

Appendix A

Acronyms

CERN Organisation européenne pour la recherche nucléaire

ATLAS A Toroidal LHC Apparatus

SM Standard Model

QFT Quantum Field Theory

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

EW Electroweak

EWSB Electroweak Symmetry Breaking

VEV Vacuum Expectation Value

CKM Cabibbo-Kobayashi-Maskawa

EM electromagnetic

IP impact parameter of tracks

ML Machine Learning

neos neural end-to-end-optimized summary statistics

HEP High Energy Physics

LHC Large Hadron Collider

HL-LHC High Luminosity LHC

ID Inner Detector

SCT semiconductor tracker

TRT transition radiation tracker

IBL insertable *b*-layer

HLT high level trigger

L1 Level-1

PFO Particle Flow Object

TCC Track CaloCluster

UFO Unified Flow Object

JES Jet Energy Scale

JER Jet Energy Resolution

JMR Jet Mass Resolution

GGF gluon-gluon fusion

VBF vector-boson fusion

NNLO next-to-next-to-leading order

N^{3}LO next-to-next-to-next-to-leading order

SR Signal Region

VR Validation Region

CR Control Region

KDE Kernel Density Estimation

bKDE binned Kernel Density Estimation

MC Monte Carlo

PDF Parton Density Function

PV primary vertex

JVT jet vertex tagger

NN Neural Network

ANN Artificial Neural Network

WP working point

VR variable radius

Bibliography

[1] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J. M. Lindert, and J. Mazzitelli. Higgs boson pair production at NNLO with top quark mass effects. *Journal of High Energy Physics*, 2018(5), may 2018. doi:10.1007/jhep05(2018)059. URL `https://doi.org/10.1007%2Fjhep05%282018%29059`.

[2] Frédéric A. Dreyer and Alexander Karlberg. Vector-boson fusion higgs pair production at n^3LO. *Phys. Rev. D*, 98:114016, Dec 2018. doi:10.1103/PhysRevD.98.114016.

[3] Daniel de Florian, D Fontes, J Quevillon, M Schumacher, FJ Llanes-Estrada, AV Gritsan, E Vryonidou, A Signer, P de Castro Manzano, D Pagani, et al. *arXiv: Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. Number arXiv: 1610.07922. Cern, 2016.

[4] ATLAS Collaboration. Search for nonresonant pair production of higgs bosons in the b b^- b b^- final state in p p collisions at s= 13 tev with the atlas detector. *Phys. Rev. D*, 108(5):052003, 2023. doi:10.1103/PhysRevD.108.052003.

[5] Dale Charles Abbott, William Keaton Balunas, Lucas Santiago Borgna, Alexander Emerman, James Frost, Sean Joseph Gasiorowski, James Cameron Grundy, Nicole Michelle Hartman, Shota Hayashida, Todd Brian Huffman, Cigdem Issever, Michael Kagan, Yu Nakahama, Santiago Rafael Paredes Saenz, Attilio Picazio, Jana Schaarschmidt, Todd Seiss, Mel Shochet, Beojan Stanislaus, Maximilian J Swiatlowski, Rafael Teixeira De Lima, Stephane Willocq, Anna Goussiou, Nikolaos Konstantinidis, Sau Lan Wu, Chen-Hsun Chan, Chen Zhou, Rui Zhang, Christopher Gubbels, Marta Maja Czurylo, Raif Rafideen Bin Norisam, Teng Jian Khoo, Arely Cortes-Gonzalez, Daniel Guest, Liaoshan Shi, Iza Veliscek, Marco Valente, Alessandra Betti, Christopher Don Milke, and Katharine Leney. Supporting Document: The Search for Resonant HH Production Decaying to the 4b Final State Using the Full Run-2 Data and the Boosted Analysis Channel. Technical report, CERN, Geneva, 2020. URL `https://cds.cern.ch/record/2708599`.

[6] Fady Bishara, Roberto Contino, and Juan Rojo. Higgs pair production in vector-boson fusion at the lhc and beyond. *The European Physical Journal C*, 77:1–26, 2017. doi:10.1140/epjc/s10052-017-5037-9.

[7] Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector. Technical report, CERN, Geneva, 2019. URL `https://cds.cern.ch/record/2665057`. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2019-007.

[8] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. 2022. doi:10.48550/ARXIV.2212.09379.

[9] Johan Alwall, R Frederix, S Frixione, V Hirschi, Fabio Maltoni, Olivier 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*, 2014(7):1–157, 2014.

[10] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015. doi:10.1016/j.cpc.2015.01.024.

[11] Sau Lan Wu, Ashutosh Kotwal, Arely Cortes Gonzalez, Michael Kagan, Shu Li, Maximilian J Swiatlowski, Liaoshan Shi, Janna Katharina Behr, Valentina Cairo, Thomas Andrew Schwarz, Sebastien Rettie, Yanlin Liu, Rui Zhang, Rachel Jordan Hyneman, Sanmay Ganguly, Dilia Maria Portillo Quintero, Kunlin Ran, Marco Valente, Mohamed Belfkir, Rafael Teixeira De Lima, Zhen Wang, Daariimaa Battulga, Jem Aizen Mendiola Guhit, Yuwen Ebony Zhang, Russell Bate, Karl Ver Hage Falb, Salah Nasri, Hemza Azri, and Marcus Vinicius Gonzalez Rodrigues. Search for resonant and non-resonant boosted Higgs boson pair production in *bbbb* final state via vector-boson-fusion (VBF) production using the full Run 2 data with ATLAS detector. Technical report, CERN, Geneva, 2023. URL `https://cds.cern.ch/record/2848140`.

[12] Will Buttinger. Background estimation with the abcd method. URL `https://twiki.cern.ch/twiki/pub/Main/ABCDMethod/ABCDGuide_draft18Oct18.pdf`. Last accessed: 2023-11-21.

[13] Gregor Kasieczka, Benjamin Nachman, Matthew D. Schwartz, and David Shih. Automating the abcd method with machine learning. *Phys. Rev. D*, 103:035021, Feb 2021. doi:10.1103/PhysRevD.103.035021. URL `https://link.aps.org/doi/10.1103/PhysRevD.103.035021`.

[14] Kim Albertsson, Piero Altoe, Dustin Anderson, John Anderson, Michael Andrews, Juan Pedro Araque Espinosa, Adam Aurisano, Laurent Basara, Adrian Bevan, Wahid Bhimji, Daniele Bonacorsi, Bjorn Burkle, Paolo Calafiura, Mario Campanelli, Louis Capps, Federico Carminati, Stefano Carrazza, Yi fan Chen, Taylor Childers, Yann Coadou, Elias Coniavitis, Kyle Cranmer, Claire David, Douglas Davis, Andrea De Simone, Javier Duarte, Martin Erdmann, Jonas Eschle, Amir Farbin, Matthew Feickert, Nuno Filipe Castro, Conor Fitzpatrick, Michele Floris, Alessandra Forti, Jordi Garra-Tico, Jochen Gemmler, Maria Girone, Paul Glaysher, Sergei Gleyzer, Vladimir Gligorov, Tobias Golling, Jonas Graw, Lindsey Gray, Dick Greenwood, Thomas Hacker, John Harvey, Benedikt Hegner, Lukas Heinrich, Ulrich Heintz, Ben Hooberman, Johannes Junggeburth, Michael Kagan, Meghan Kane, Konstantin Kanishchev, Przemysław Karpiński, Zahari Kassabov, Gaurav Kaul, Dorian Kcira, Thomas Keck, Alexei Klimentov, Jim Kowalkowski, Luke Kreczko, Alexander Kurepin, Rob Kutschke, Valentin Kuznetsov, Nicolas Köhler, Igor Lakomov, Kevin Lannon, Mario Lassnig, Antonio Limosani, Gilles Louppe, Aashrita Mangu, Pere Mato, Narain Meenakshi, Helge Meinhard, Dario Menasce, Lorenzo Moneta, Seth Moortgat, Mark Neubauer, Harvey Newman, Sydney Otten, Hans Pabst, Michela Paganini, Manfred Paulini, Gabriel Perdue, Uzziel Perez, Attilio Picazio, Jim Pivarski, Harrison Prosper, Fernanda Psihas, Alexander Radovic, Ryan Reece, Aurelius Rinkevicius, Eduardo Rodrigues, Jamal Rorie, David Rousseau, Aaron Sauers, Steven Schramm, Ariel Schwartzman, Horst Severini, Paul Seyfert, Filip Siroky, Konstantin Skazytkin, Mike Sokoloff, Graeme Stewart, Bob Stienen, Ian Stockdale, Giles Strong, Wei Sun, Savannah Thais, Karen Tomko, Eli Upfal, Emanuele Usai, Andrey Ustyuzhanin, Martin Vala, Justin Vasel, Sofia Vallecorsa, Mauro Verzetti, Xavier Vilasís-Cardona, Jean-Roch Vlimant, Ilija Vukotic, Sean-Jiun Wang, Gordon Watts, Michael Williams, Wenjing Wu, Stefan Wunsch, Kun Yang, and Omar Zapata. Machine learning in high energy physics community white paper, 2019.

[15] Jonathan Shlomi, Peter Battaglia, and Jean-Roch Vlimant. Graph neural networks in particle physics. *Machine Learning: Science and Technology*, 2(2): 021001, 2020.

[16] Matthew Feickert and Benjamin Nachman. A living review of machine learning for particle physics, 2021.

[17] Matthew D. Schwartz. Modern Machine Learning and Particle Physics. *Harvard Data Science Review*, 3(2), may 13 2021. https://hdsr.mitpress.mit.edu/pub/xqle7lat.

[18] Nathan Simpson and Lukas Heinrich. neos: End-to-end-optimised summary statistics for high energy physics. *Journal of Physics: Conference Series*, 2438 (1):012105, feb 2023. doi:10.1088/1742-6596/2438/1/012105. URL `https://dx.doi.org/10.1088/1742-6596/2438/1/012105`.

[19] Miroslav Kubat. *An introduction to machine learning*, volume 3. Springer, 2021. doi:https://doi.org/10.1007/978-3-030-81935-4.

[20] Vivienne Sze, Yu-Hsin Chen, Tien-Ju Yang, and Joel S. Emer. Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12):2295–2329, 2017. doi:10.1109/JPROC.2017.2761740.

[21] Kyle Cranmer. Kernel estimation in high-energy physics. *Computer Physics Communications*, 136(3):198–207, 2001. ISSN 0010-4655. doi:https://doi.org/10.1016/S0010-4655(00)00243-5. URL `https://www.sciencedirect.com/science/article/pii/S0010465500002435`.

[22] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018. URL `http://github.com/google/jax`.

[23] I. Antcheva, M. Ballintijn, B. Bellenot, M. Biskup, R. Brun, N. Buncic, Ph. Canal, D. Casadei, O. Couet, V. Fine, L. Franco, G. Ganis, A. Gheata, D. Gonzalez Maline, M. Goto, J. Iwaszkiewicz, A. Kreshuk, D. Marcos Segura, R. Maunder, L. Moneta, A. Naumann, E. Offermann, V. Onuchin, S. Panacek, F. Rademakers, P. Russo, and M. Tadel. Root — a c++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 180(12):2499–2512, 2009. ISSN 0010-4655. doi:https://doi.org/10.1016/j.cpc.2009.08.005. 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures.

[24] Nathan Simpson. relaxed: version 0.3.0, July 2023. URL `https://github.com/gradhep/relaxed`.

[25] Nathan Simpson and Lukas Heinrich. neos: version 0.2.0, January 2021. URL `https://github.com/gradhep/neos`.

[26] Frederic Renner. hh_neos. `https://gitlab.cern.ch/frenner/hh_neos`, 2023. Accessed: November 27, 2023.

[27] ATLAS Collaboration atlas. publications@ cern. ch, Georges Aad, B Abbott, DC Abbott, A Abed Abud, K Abeling, DK Abhayasinghe, SH Abidi, OS AbouZeid, NL Abraham, et al. Jet energy scale and resolution measured in proton–proton collisions at s= 13 tev with the atlas detector. *The European Physical Journal C*, 81(8):689, 2021. doi:10.48550/arXiv.2007.02645.

[28] The ATLAS collaboration. In situ calibration of large-radius jet energy and mass in 13 tev proton–proton collisions with the atlas detector. *The European Physical Journal C*, 79(2):135, 2019. doi:10.1140/epjc/s10052-019-6632-8. URL `https://doi.org/10.1140/epjc/s10052-019-6632-8`.

[29] ATLAS Collaboration. Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with $80\,\mathrm{fb}^{-1}$ of $\sqrt{s} = 13\,\mathrm{TeV}$ pp data. ATLAS-CONF-2020-022, 2020. URL `https://cds.cern.ch/record/2724442`.

[30] ATLAS Collaboration. Efficiency corrections for a tagger for boosted $H \to b\bar{b}$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical report, CERN, Geneva, 2021. URL `https://cds.cern.ch/record/2777811`.

[31] The ATLAS collaboration. Particle Modeling Group systematic uncertainty recipes. URL `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PmgSystematicUncertaintyRecipes`. Last accessed: 2023-11-29.

[32] Francis Halzen, A Martin, and Leptons Quarks. An introductory course in modern particle physics. *John and Wiley*, 1984.

Statutory Declaration - Eidesstattliche Erklärung

I declare that I have authored this thesis independently, that I have not used other than the declared sources/ resources and that I have explicitly marked all materials which has been quoted either literally or by content form the used sources.

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Berlin, 01.12.2023

Frederic Renner