Dissertation

End to end optimization in a search for boosted Higgs boson pair production in the $bbbb$ final state via vector-boson-fusion (VBF) production using the run 2 dataset with the ATLAS detector

For the attainment of the academic degree doctor rerum naturalium

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

M.Sc. Frederic Renner

Berlin, 10.07.2023

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

1 Theory **1**

1.1 Feynman rules from field theory 1

1.2 Probability of a process 2

1.3 The Standard Model 3

1.4 Statistics 4

1.4.1 Building the likelihood 4

1.4.2 From test statistic to p-value 5

1.4.3 The CL_s value 8

1.4.4 The HistFactory model 8

1.4.5 The Modifiers 11

1.4.6 The constraint terms 12

Bibliography **15**

List of Figures

List of abbreviations

ATLAS	The Atlas Experiment
pdf	probability density function

Chapter 1

Theory

Where to start? Sometimes it semms like that particle physics is bringing it all together, as it tries to give a comprehensive picture of the world by describing the structure of matter from quantum mechanics to cosmology. So I would say to start shallow (since we are experimentalists) at the very beginning, and then dive a bit deeper into Standard Model to have a plausible thread how the need and the development of the Higgs mechanism came about. The following is mainly based on [1, 2] and intended to make the calculation of cross sections plausible.

1.1 Feynman rules from field theory

The fact that elementary particles can seemingly be born out of nothing and die again led to the development of their currently most successful description through quantum field theories. Heuristically it can be understood by the uncertainty principle, which states that energy can vary greatly on short time scales, and by special relativity, which allows the property energy to be converted into the property mass. This marriage between quantum mechanics and special relativity is what drove the development of quantum field theory.

can be deduced through perturbation theory, just on term, The conventional strategy is perturbation theory with the free fields as starting point, treating the interaction as a small perturbation

To make a field one assigns a quantity to some region in spacetime, e.g. $\phi(\boldsymbol{x}, t)$. A Lagrangian $L(\phi(\boldsymbol{x}, t))$ then governs the dynamics, like excitations or interactions of this field, which can e.g. represent the birth and death of particles or interactions by the exchange of a particle between them. One formulation of quantum field theory is by use of the path integral formulation. It then basically boils down to integrals of the form $\int D\phi e^{i\int d^4x L(\phi(\boldsymbol{x},t))}$. Where $\int D\phi$ is the integral over all possible paths/ways a particle could take. Through back and forth expansions of the e functions the integral can be solved and the result is a probability - the amplitude $\mathcal{M}$ of e.g. an interaction between two particles, like scattering, usually depicted in the form of Feynman Diagrams. As this follows a pattern the formalism can be contracted into the infamous Feynman rules (for details see [2]).

1.2 Probability of a process

Probes of elementary particle interactions are accessible via bound states, decays and scattering. The first can be studied within classical quantum mechanics whereas the latter uses the preceding. Since this work deals with a collider experiment I think its at least useful to see how one can calculate in principle a cross section σ. It is a measure of how possible an interaction is when shooting something at each other. Calculating reaction rates in quantum mechanics is done by Fermi's golden rule. Here the relativistic version for a scattering process like $1+2 \to 3+4+\cdots+n$ is given [2]

$$\begin{aligned}\sigma =& \frac{S\hbar^2}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2 c^2)^2}} \int |\mathcal{M}|^2 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 \cdots - p_n) \\ & \times \prod_{j=3}^{n} 2\pi\delta(p_j^2 - m_j^2 c^2)\Theta(p_j^0) \frac{\mathrm{d}^4 p_j}{(2\pi)^4}.\end{aligned} \tag{1.2.1}$$

S is a statistical factor accounting for identical particles (e.g. $a \to b+b+c+c+c$, then $S = (1/2!)(1/3!)$), p_i are four momenta of particle i over which one integrates, $\mathcal{M}(p_1, \ldots, p_n)$ is the amplitude of the process calculable with the Feynman rules, the δ^4 ensures energy and momentum conservation, the last δ ensures that particles are on their mass shell ($E_j^2/c^2 - \boldsymbol{p}_j^2 = m_j^2 c^2$) and the Heaveside Θ makes sure that outgoing energies are positive $p_j^0 = E_j/c > 0$. With this, a particle physicist can calculate the probability of any process at a collider experiment.

1.3 The Standard Model

1.1

dirac, require local gauge invariance -> qed Lagrangian

gauge field blah only about Lagrangians, no field solutions needed

1.4 Statistics

Every scientific investigation starts with a hypothesis that is to be tested empirically. The main objective is to evaluate if the proposed hypothesis agrees or disagrees with observed data, to either accept or reject it against the null-hypothesis. The metric at hand to do so is the p-value that arises within hypothesis testing.

In the field of high energy physics, a framework based on likelihood statistics has been developed specifically for this task. This section begins to lay out the mathematical fundamentals of the approach and explains its implementation in PYHF [3, 4]. The following is based on [3, 5, 6].

1.4.1 Building the likelihood

The statistical model must take into account how compatible are the theoretical predictions with the observed collision events. This can be described by a likelihood $L(\boldsymbol{x}|\boldsymbol{\phi})$ which is in in general a probability for an observation $\boldsymbol{x}$ under a given set of parameters $\boldsymbol{\phi}$. Through $\boldsymbol{\psi}$ the theoretical predictions are incorporated into the probability calculation. Given that this is a counting experiment, the preferred tool of analysis are bins of a histogram $\boldsymbol{h} = (h_1, ..., h_N)$.

It is useful to subdivide a measurement $\boldsymbol{x} = (\boldsymbol{n}, \boldsymbol{a})$ further into observable histograms $\boldsymbol{n}$, (e.g. the invariant mass of a particle) and auxiliary measurement histograms $\boldsymbol{a}$ that help to constrain the model. Additionally, in the context of hypothesis testing it is useful to split the set of parameters $\boldsymbol{\phi} = (\boldsymbol{\psi}, \boldsymbol{\Theta})$ into so-called parameters of interest $\boldsymbol{\psi}$ and nuisance parameters $\boldsymbol{\Theta}$. For this subsection it is instructive to consider only one parameter of interest μ, the signal strength μ.

The bin contents can then be expressed in terms of the amount of signal $s_i(\boldsymbol{\Theta})$ and background $b_i(\boldsymbol{\Theta})$ in bin i that depend on the nuisance parameters. The prediction (expectation value) of the histogram bins of the observable n_i can then be expressed as

$$\langle n_i(\mu, \boldsymbol{\Theta})\rangle = \mu s_i(\boldsymbol{\Theta}) + b_i(\boldsymbol{\Theta}). \tag{1.4.1}$$

Similarly, for auxiliary measurement bins a_i their expectation value can be calculated from functions $u_i(\boldsymbol{\Theta})$ that also depend on the nuisance parameters and help to constrain the model

$$\langle a_i(\mathbf{\Theta})\rangle = u_i(\mathbf{\Theta}). \tag{1.4.2}$$

Since this is a counting experiment in which events occur at a constant mean rate and independently of time, each bin follows a Poisson distribution

$$\frac{r^k e^{-r}}{k!}. \tag{1.4.3}$$

r is the expected rate of occurrences, which translates as our prediction, whereas k are the actual measured occurrences. Therefore the likelihood is a product of Poisson probabilities

$$L(\mu, \mathbf{\Theta}) = \prod_{j=1}^{N} \frac{(\mu s_j(\mathbf{\Theta}) + b_j(\mathbf{\Theta}))^{n_j}}{n_j!} e^{-(\mu s_j(\mathbf{\Theta}) + b_j(\mathbf{\Theta}))} \prod_{k=1}^{M} \frac{u_k(\mathbf{\Theta})^{a_k}}{a_k!} e^{-u_k(\mathbf{\Theta})}. \tag{1.4.4}$$

The last product can also be thought of penalizing the likelihood if e.g. an auxiliary measurement displays a very improbable value for a quantity. To test for a hypothesized value of μ, the best choice according to the Neyman-Pearson lemma [6], is the profile likelihood ratio that reduces the dependence to the parameter(s) of interest

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\mathbf{\Theta}}})}{L(\hat{\mu}, \hat{\mathbf{\Theta}})} \tag{1.4.5}$$

The denominator is the unconditional maximum likelihood estimation so that $\hat{\mu}$ and $\hat{\mathbf{\Theta}}$ both are free to vary to maximize L, whereas the numerator is the found maximum likelihood conditioned on some chosen μ and the set of nuisance parameters $\hat{\hat{\mathbf{\Theta}}}$ that maximize the likelihood for that particular μ. This definition gives $0 \leq \lambda \leq 1$. For a $\lambda \approx 1$ the hypothesized value of μ shows good agreement to the Poissonian model.

1.4.2 From test statistic to p-value

For this subsection μ can be a set of parameters of interest (defined above as $\boldsymbol{\psi}$). This is standard in the literature and I will stick with it to avoid confusion. To test for alternative hypotheses it is useful to transform the profile likelihood into a test statistic t_μ

$$t_\mu = -2 \log \lambda(\mu). \tag{1.4.6}$$

This translates to $t_\mu \to 0$ as good agreement and $t_\mu \to \infty$ as bad agreement to the model. A right-tail p-value can then be calculated from the probability density function of t_μ: $\mathrm{pdf}(t_\mu) = f(t_\mu \mid \mu)$

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu \mid \mu) \mathrm{d}t_\mu \tag{1.4.7}$$

$t_{\mu,obs}$ is the test statistic t_μ evaluated with the observed data. This is like plugging the same values for r and k into the Poisson distributions in eq. 1.4.3. Just like a probability density function for a standard normal distribution, intuitively the pdf is how probable a particular value of the test statistic t_μ is under a fixed value of the signal strength (how often it occurs compared to all other values t_μ can have).

This particular form is handy because there exist approximations for $f(t_\mu \mid \mu)$ [5]. Wald [7] proved that for the null hypothesis in the large sample limit, the test statistic follows a normalized sum of squared distances between the tested parameters of interest μ and its maximum likelihood estimate $\hat{\mu}$. The result was extended by Wilk [8] for any hypothesis, so the test statistic becomes

$$t_\mu = \sum_i \frac{(\mu_i - \hat{\mu}_i^2)}{\sigma_i^2} + \mathcal{O}(1/\sqrt{N}). \tag{1.4.8}$$

The $\hat{\mu}_i$ are in the large sample limit normally distributed with mean μ' (true values) and standard deviation σ_i. This is the definition of a non-central χ-squared distribution with degrees of freedom equal to the number of parameters of interest (see section 3.1 in [5]). For one parameter of interest the distribution reads

$$f(t_\mu \mid \mu) = \frac{1}{2\sqrt{t_\mu}} \frac{1}{\sqrt{2\pi}} \left[\exp\left(-\frac{1}{2} \left(\sqrt{t_\mu} + \sqrt{\Lambda_\mu} \right) \right) + \exp\left(-\frac{1}{2} \left(\sqrt{t_\mu} - \sqrt{\Lambda_\mu} \right) \right) \right], \tag{1.4.9}$$

with non-centrality parameter

$$\Lambda_\mu = \frac{(\mu - \mu')^2}{\sigma^2}. \tag{1.4.10}$$

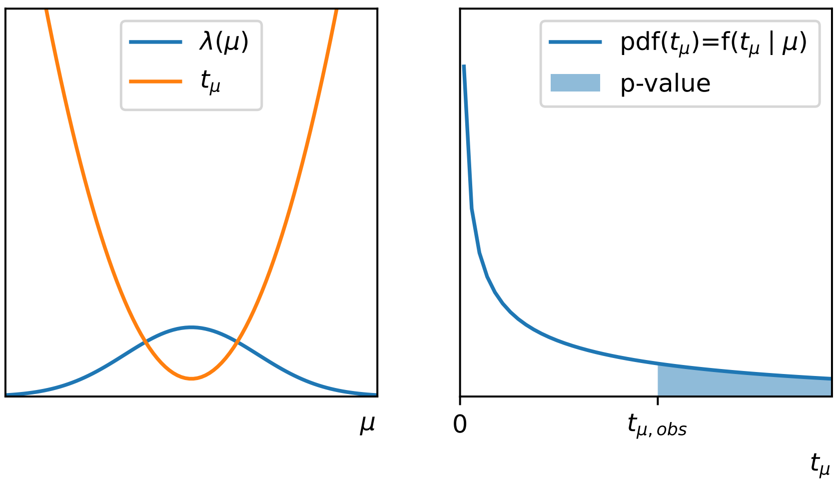


Figure 1.1: A sketch to follow the steps to calculate p-values. (**left**) The profile likelihood (■) has essentially some hill-like form with a maximum at $\lambda(\hat{\mu}, \hat{\boldsymbol{\Theta}})$, t_μ (■) is $-2\ln(\lambda)$. (**right**) For one parameter of interest in the large sample limit $f(t_\mu \mid \mu)$ follows a non-central chi-squared distribution with one degree of freedom, equation 1.4.9. The blue shaded area under the pdf is a right hand sided p-value.

Figure 1.1 illustrates the different steps. Being able to calculate p-values allows now to state how likely it is that the proposed hypothesis is reflected by the observed data. In other words, when the experiment is to be repeated, the p-value represents the probability that the result favors the alternative hypothesis over the null hypothesis.

In the scientific community a widely accepted threshold for this is a p-value of 0.05. Though particle physicists only claim discovery of a new phenomenon for $p < 2.87 \times 10^{-7}$ corresponding to 5 standard deviations of the standard normal distribution. Further particle physicists exclude hypotheses if the p-value is not below 2 standard deviations of the standard normal distribution $p \lesssim 0.05$. One caveat here is that this particular form of t_μ assumes μ can also be negative, which can be non-physical depending on the impact of a new process. Test statistics and their pdf approximations considering the different cases are covered in [5].

1.4.3 The CL_s value

Particle physicists are usually interested in two things when making statistical tests for the discovery of new phenomena: how well is the modeling of backgrounds (things we know) and whether there is evidence in the observations for a new phenomenon. This means one needs to test two hypotheses: a background only (b) and a signal plus background ($s+b$) hypothesis. Each will result in a p-value of their own. For example, $p_b = 0$ would mean that the backgrounds are perfectly reflected by the observations and a $p_{s+b} < 0.05$ could be a sign of e.g. new physics. To combine these two metrics into a single score, particle physicists came up with the pseudo Confidence Level/p-value called CL_s incorporating also the goodness of the modeling of the backgrounds

$$\mathrm{CL}_s = \frac{p_{s+b}}{1-p_b} = \frac{\int_{t_{\mu,obs}}^{\infty} f(t_\mu \mid \mu)\mathrm{d}t_\mu}{1-\int_{t_{\mu,obs}}^{\infty} f(t_\mu \mid \mu)\mathrm{d}t_\mu}. \tag{1.4.11}$$

Intuitively the numerator is again just the value for the alternative hypothesis whereas the denominator penalizes CL_s if the modeling of the backgrounds is not reflected in the observations. This can also be understood visually from the first figure of the paper that introduced the CL_s quantity [9] (see description of fig. 1.2).

1.4.4 The HistFactory model

A model used widely to build a likelihood as described in section 1.4.1 is called HistFactory [10] and is implemented within PYHF [3]. This section is based on the introduction to the model within the documentation of PYHF. HistFactory reduces the building of a likelihood into a small number of basic components. In order to do that it is again useful to think of another splitting of the model parameters $\boldsymbol{\phi}$ into

$$L(\boldsymbol{x}|\boldsymbol{\phi}) = L(\boldsymbol{x}|\overbrace{\boldsymbol{\psi}}^{\text{parameters of interest}}, \underbrace{\boldsymbol{\theta}}_{\text{nuisance parameters}}) = L(\boldsymbol{x}|\overbrace{\boldsymbol{\eta}}^{\text{free}}, \underbrace{\boldsymbol{\chi}}_{\text{constrained}}) \tag{1.4.12}$$

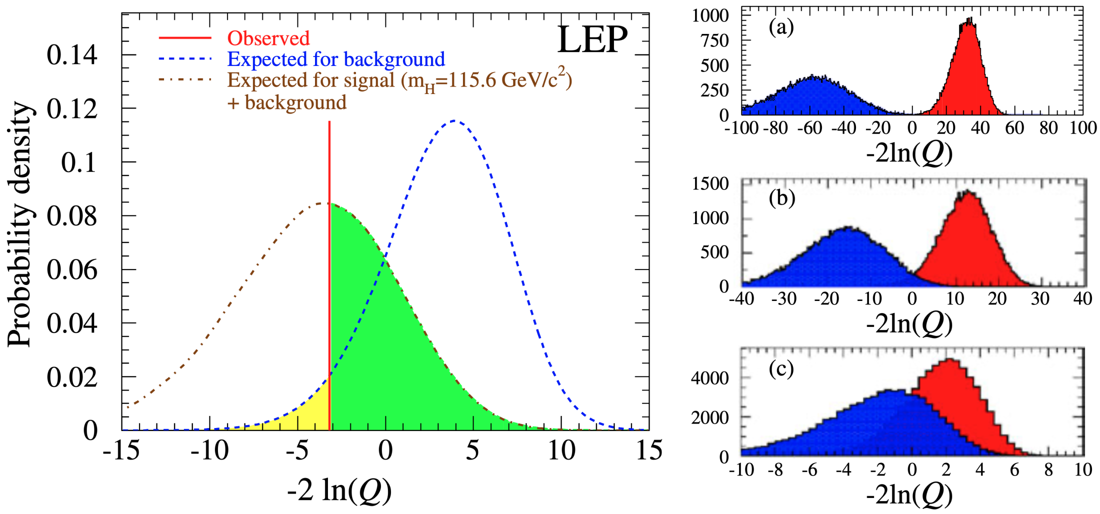


Figure 1.2: Probability density functions of test statistics from a Higgs search at LEP illustrating the calculation of p-values (λ becomes Q). (**left**) The pdf's of the test statistic $f(t_\mu \mid \mu)$ of the signal + background (/) and background (/) only hypotheses. The p-value is calculated by integration from $t_{\mu,obs}$ (the red observed line (/)) to infinity (see eq. 1.4.7). The green shaded area (■) corresponds to p_{s+b} whereas the yellow area (■) corresponds to $1 - p_b$ since the integral over one whole pdf is 1. (**right**) Degradation of search sensitivity from (a) to (c). Note that the colors of the pdf's change here to signal + background (■) and background only (■). For example putting the observation ($t_{\mu,obs}$) on the x-axis at 0 in these plots, one would get for plot (a) $p_b \approx 1$ and $p_{s+b} \approx 0$ resulting in a $\mathrm{CL}_s \approx 0$, whereas with increasing overlap the CL_s value increases and the sensitivity decreases. Taken from [9].

free parameters $\boldsymbol{\eta}$ and constrained parameters $\boldsymbol{\chi}$. Free parameters are free to choose in the model and can be for example a cross-section of a process. Constrained parameters are used to incorporate uncertainties into the likelihood to constrain it. Further there might be several histograms of an observable, for example measured in orthogonal kinematic regions, that are called channels c. Bins have the index b here and constraint terms are denoted c_χ. With that the likelihood can be described by

$$L(\boldsymbol{n}, \boldsymbol{a} \,|\, \boldsymbol{\eta}, \boldsymbol{\chi}) = \underbrace{\prod_{c \in \text{channels}} \prod_{b \in \text{bins}_c} \text{Pois}\left(n_{cb} \,|\, \nu_{cb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\substack{\text{Simultaneous measurement} \\ \text{of multiple channels}}} \underbrace{\prod_{\chi \in \boldsymbol{\chi}} c_\chi(a_\chi | \, \chi)}_{\substack{\text{constraint terms} \\ \text{for auxiliary measurements}}} . \tag{1.4.13}$$

The n_{cb} is the observation and $\nu_{cb}(\boldsymbol{\eta}, \boldsymbol{\chi})$ the prediction. The $c_\chi(a_\chi | \, \chi)$ are calculated from auxiliary measurements a_χ (the uncertainties) to constrain the parameter χ and can be any function (e.g. Gaussian, Poissonian,...) the parameter/uncertainty is believed to be distributed.

The prediction is a sum of nominal bin counts[1] ν^0_{scb} over all samples s (e.g. $t\bar{t}$, multijet-background, etc.). These nominal bin counts are subject to uncertainties. Therefore the bin counts can be varied within the bounds of these uncertainties. However the effect of this modification to the likelihood must be taken into account which is through the constraint terms. These penalize the likelihood the larger the modification to a nominal value becomes. The ν^0_{scb} are varied with multiplicative κ_{scb} and additive modifiers Δ_{scb}

$$\nu_{cb}\left(\boldsymbol{\phi}\right) = \sum_{s \in \text{samples}} \nu_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right) \tag{1.4.14}$$

$$= \sum_{s \in \text{samples}} \underbrace{\left(\prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\text{multiplicative modifiers}} \left(\nu^0_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right) + \underbrace{\sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)}_{\text{additive modifiers}}\right). \tag{1.4.15}$$

The different types of modifiers are explained in section 1.4.5 and the constraint terms c_χ in section 1.4.6.

[1] also called rates, like in the definition of a Poisson distribution

Why this is useful can be seen by considering one uncertainty to a nominal bin count estimate ν^0_{scb}. By modifying ν^0_{scb} with a factor κ in a way that increases the Poisson probability while the corresponding constraint term $c_\kappa(\kappa)$ stays around 1, it can be beneficial for the goal of maximizing the likelihood. This means the most likely/compatible value to the observed data within the modeling of the uncertainties can be found.

1.4.5 The Modifiers

In HistFactory there are by convention four types $\{\lambda, \mu, \gamma, \alpha\}$ of such multiplicative rate modifiers that are explained in this section. There are **free rate modifiers** λ **and** μ that affect all bins equally, like the cross-section of a process or the luminosity

$$\nu_{scb}(\mu) = \mu\nu^0_{scb}. \tag{1.4.16}$$

These are bin-independent normalization factors and preserve the shape of the histogram. Further there are **bin-wise modifiers** γ_b (uncorrelated shape)

$$\nu_{scb}(\gamma_b) = \gamma_b\nu^0_{scb}. \tag{1.4.17}$$

These are useful for example to include uncertainties of a per bin data-driven background estimate. This type without a constraint term is not of much use as if there is only one sample or channel, the fit would always match the data perfectly. In addition there exist **interpolation parameters** α (shape factors) that enter the modeling through an interpolation function η instead of being the factor itself. They exist in multiplicative versions

$$\nu_{scb}(\alpha) = \eta(\alpha)\nu^0_{scb}, \tag{1.4.18}$$

and additive versions

$$\nu_{scb}(\alpha) = \nu^0_{scb} + \eta(\alpha). \tag{1.4.19}$$

This is useful to include systematic uncertainties. In a typical ATLAS analysis usually one knows the one standard deviation of a bin count $\eta_{-1} = \nu^{\text{1down}}_{scb}$ and $\eta_1 = \nu^{\text{1up}}_{scb}$ to the nominal value ν^0_{scb} of an uncertainty. These are used to construct interpolation functions that modify the nominal value with a nuisance parameter that is also used to apply a penalization c_α according to the modeling of the uncertainty.

In HistFactory there exists four of such interpolation functions. For those exist an identity operator

$$\eta_0 = \eta(\alpha = 0) = \begin{cases} 1, & \text{multiplicative modifier, } (\kappa) \\ 0, & \text{additive modifier, } (\lambda). \end{cases} \tag{1.4.20}$$

One example of these interpolation functions that scales the bin count linearly over the known deviations $\eta_{-1} = \nu_{scb}^{1\text{down}}$ and $\eta_1 = \nu_{scb}^{1\text{up}}$ is

$$\eta_{\text{linear}}(\alpha) = \begin{cases} \alpha(\eta_0 - \eta_1), & \alpha > 0 \\ \alpha(\eta_0 - \eta_{-1}), & \alpha < 0 \end{cases} \tag{1.4.21}$$

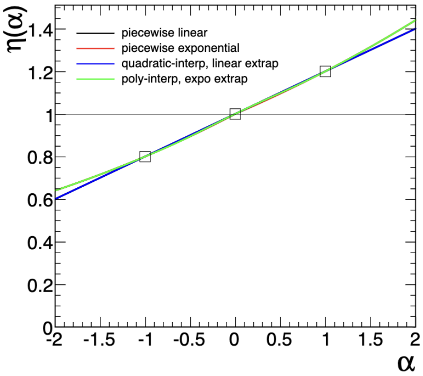
This is illustrated in fig. 1.3(a). For the other ones see e.g. [11]. It is noted that α is the nuisance parameter and not the function $\eta(\alpha)$ and there is an associated constraint term c_α to each α.

1.4.6 The constraint terms

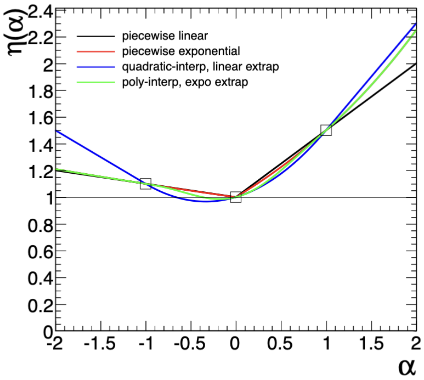
Uncertainties are modeled either Gaussian or Poissonian. The Gaussian implementation is straightforward as the uncertainty appears squared in the definition. For the interpolation function the nuisance parameter is scaled to the standard deviation values as described before Gauss($\alpha \mid a, \sigma = 1$).

For a Poissonian constraint to a multiplicative modifier γ_b, with a nominal (most probable) value $\gamma_0 = 1$, the Poisson distribution must be scaled with a factor f, so it reflects the original bin-count uncertainty σ. To find the corresponding Poisson distribution all parameters are multiplied by a factor f and is then solved for the one with the desired uncertainty. Since the variance of a Poissonian like eq. 1.4.3 is the rate parameter λ it follows

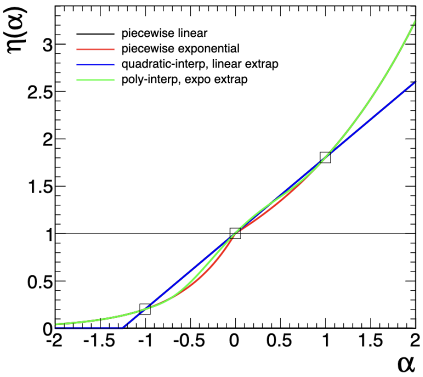
$$\text{Var}\left[\text{Pois}(k = f\gamma_0, \lambda = f\gamma)\right] = \lambda \overset{\gamma=\gamma_0}{=} f\gamma_0 = (f\sigma)^2 \quad \rightarrow \quad f = (1/\sigma^2). \tag{1.4.22}$$



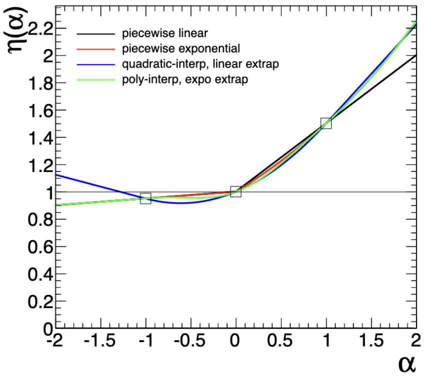
(a)



(b)



(c)



(d)

Figure 1.3: The four interpolation functions $\eta(\alpha)$ for different up and down standard deviation values. For example in (a) the bin count will be scaled with a factor of 0.8 for an $\alpha = -1$ (1.2 for an $\alpha = 1$). From [10].

Table 1.1: Modifiers and constraint terms used in HistFactory implemented by PYHF. Note that the interpolation functions are called f_p and g_p here instead of η as chosen in the full text. Taken from [3]

Description	Modification	Constraint Term c_χ	c_χ input
Uncorrelated Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \mathrm{Pois}\left(r_b = \sigma_b^{-2} \middle| \rho_b = \sigma_b^{-2}\gamma_b\right)$	σ_b
Correlated Shape	$\Delta_{scb}(\alpha) = f_p\left(\alpha \middle| \Delta_{scb,\alpha=-1}, \Delta_{scb,\alpha=1}\right)$	$\mathrm{Gaus}\left(a = 0 \middle| \alpha, \sigma = 1\right)$	$\Delta_{scb,\alpha=\pm1}$
Normalisation Unc.	$\kappa_{scb}(\alpha) = g_p\left(\alpha \middle| \kappa_{scb,\alpha=-1}, \kappa_{scb,\alpha=1}\right)$	$\mathrm{Gaus}\left(a = 0 \middle| \alpha, \sigma = 1\right)$	$\kappa_{scb,\alpha=\pm1}$
MC Stat. Uncertainty	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \mathrm{Gaus}\left(a_{\gamma_b} = 1 \middle| \gamma_b, \delta_b\right)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
Luminosity	$\kappa_{scb}(\lambda) = \lambda$	$\mathrm{Gaus}\left(l = \lambda_0 \middle| \lambda, \sigma_\lambda\right)$	$\lambda_0, \sigma_\lambda$
Normalisation	$\kappa_{scb}(\mu_b) = \mu_b$		
Data-driven Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$		

This completes all the requirements needed for the creation of HistFactory models. The different types of modifiers and their constraint terms are summarized in table 1.1.

Bibliography

[1] A. Zee, *Quantum field theory in a nutshell*, Vol. 7 (Princeton university press, 2010).

[2] D. Griffiths, *Introduction to elementary particles* (John Wiley & Sons, 2020).

[3] L. Heinrich, M. Feickert, and G. Stark, "pyhf: v0.7.2," Https://github.com/scikit-hep/pyhf/releases/tag/v0.7.2.

[4] L. Heinrich, M. Feickert, G. Stark, and K. Cranmer, Journal of Open Source Software **6**, 2823 (2021).

[5] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, The European Physical Journal C **71**, 1 (2011).

[6] O. Behnke, K. Kröninger, G. Schott, and T. Schörner-Sadenius, *Data analysis in high energy physics: a practical guide to statistical methods* (John Wiley & Sons, 2013).

[7] A. Wald, Transactions of the American Mathematical society **54**, 426 (1943).

[8] S. S. Wilks, The annals of mathematical statistics **9**, 60 (1938).

[9] A. L. Read, Journal of Physics G: Nuclear and Particle Physics **28**, 2693 (2002).

[10] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke (ROOT), "HistFactory: A tool for creating statistical models for use with RooFit and RooStats," Tech. Rep. (New York U., New York, 2012).

[11] L. Heinrich, *Searches for Supersymmetry, RECAST, and Contributions to Computational High Energy Physics*, Ph.D. thesis, New York University (2019).

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, 10.07.2023

Frederic Renner