
Determination of non-perturbative corrections in diphoton production at the LHC using iterative methods

Bachelor-Arbeit
zur Erlangung des Hochschulgrades
Bachelor of Science
im Bachelor-Studiengang Physik

vorgelegt von
Jakob Beyer
geboren 20.07.1994 in Dresden
Mat.-Nr.: 3991017

Institut für Kern- und Teilchenphysik
Fachrichtung Physik
Fakultät Mathematik und Naturwissenschaften
Technische Universität Dresden

Emmy-Noether junior research group
"Precision Monte Carlo event generators and LHC measurements"
2016

Eingereicht am 14. Juni 2016

1. Gutachter: Dr. Frank Siegert
2. Gutachter: Prof. Dr. Arno Straessner

Abstract

This thesis investigates non-perturbative corrections for diphoton events from pp collisions at the LHC from parton level to hadron level. Iterative unfolding based on Bayes' theorem, as proposed by D'Agostini [11], is used on fake diphoton parton level samples. The results are compared to bin-by-bin unfolding. Influences of the diphoton identification cut parameters on the results are tested.

Diphoton events are simulated using Sherpa 2.2.0 [13] and analyzed using Rivet 2.4.1 [7]. Unfolding was performed using the ROOT framework RooUnfold [5].

Non-perturbative corrections are primarily observed as efficiencies. Smearing effects are low for the investigated observables. The unfolding results show improved accuracy of iterative unfolding for one or two iterations over bin-by-bin corrections. This is not influenced when varying the cut parameters in the diphoton identification analysis. However, limitations of comparing the unfolded results to the hadron level distributions of independent test samples show, that further investigation is necessary.

Zusammenfassung

In der vorliegenden Arbeit werden nicht-perturbative Korrekturen auf Diphoton Ereignisse am LHC, für den Übergang von Parton-Level zu Hadron-Level, untersucht. Das von D'Agostini vorgeschlagene iterative Entfaltungsverfahren, basierend auf Bayes' Theorem [11], wird auf künstliche Diphoton Parton-Level Verteilungen angewendet. Die Ergebnisse werden mit Bin-by-bin-Entfalten verglichen. Zusätzlich wird der Einfluss der Diphoton-Identifikations-Parameter auf die Ergebnisse getestet.

Diphoton-Ereignisse werden mit Sherpa 2.2.0 [13] generiert und mit Rivet 2.4.1 [7] analysiert. Für das Entfalten wird das ROOT Framework RooUnfold [5] genutzt.

Nicht-perturbative Korrekturen werden in erster Linie als Effizienz-Faktoren beobachtet. Migrationen zwischen den Bins in den Histogrammen der untersuchten Observablen sind gering. Die Resultate des Entfaltens zeigen, dass die Genauigkeit durch iteratives Entfalten mit ein oder zwei Iterationen gegenüber Bin-by-bin-Entfalten verbessert werden kann. Das Variieren der Diphoton-Identifikations-Parameter hat auf dieses Ergebnis keinen Einfluss. Eine Untersuchung der Gültigkeit des Vergleiches des Entfaltungs-Ergebnisses mit den Hadron-Level Verteilungen der Test-Proben zeigt, dass weitere Untersuchungen notwendig sind.

Contents

1	Introduction	5
2	Theoretical Framework	7
2.1	Non-perturbative effects and event levels	7
2.1.1	Non-perturbative effects	7
2.1.2	Hadron collision event generators and event levels	7
2.2	Unfolding	9
2.2.1	Bin-by-bin unfolding	9
2.2.2	Iterative method	10
2.3	Diphoton events	12
2.3.1	Diphoton identification criteria	12
2.3.2	Isolation conditions	12
2.3.3	Diphoton production in pp collisions	13
3	Implementation	15
3.1	Simulations using SHERPA	15
3.2	Event selection and analysis	15
3.3	Fake data for testing of unfolding	16
3.4	Unfolding	17
3.4.1	Error estimation for iterative unfolding	17
4	Results and discussion	19
4.1	Comparison using ATLAS analysis	19
4.2	Modified analysis	21
4.2.1	Additional observables	21
4.2.2	Response matrices	21
4.2.3	Influence of analysis cut parameters	22
4.3	Unfolding of fake diphoton samples	23
4.3.1	Original cut parameters	23
4.3.2	Variation of cut parameters	25
4.3.3	Validity of results	27
5	Conclusion	31
A	Appendix	33
A.1	Figures	33
A.2	Tables	49
A.3	Code	50

1 Introduction

Particle physics aims to find and understand the most fundamental parts of this universe. On the theoretical side, the Standard Model already describes most of what is relevant to us in everyday life, with the notable exception of gravity. It also provides predictions about particles additional to the well-known electrons and protons (which turned out to not be fundamental but composed). 12 elementary matter particles, 6 similar to the electron (leptons) and 6 similar to the composites of the proton (quarks), are included in the model and have been observed. These are Spin 1/2 particles and called fermions. Additionally, the Standard Model describes the interactions between those fermions using corresponding fields. In doing so, it introduces particles (with integer spin) carrying the force of each interaction, named bosons. The most prominent example must be the photon, the force-carrier of the electromagnetic field, since they are quanta of light. Furthermore, the strong and weak interactions with their bosons are predicted and described.

In 2012, particle physics suddenly gained unknown popularity, when one of the last big predictions of the Standard Model was observed at CERN: the Higgs boson [3], [8]. The Higgs boson is the carrier of the Higgs field, which is mathematically needed to give mass to the boson of the weak interaction [15]. At the Large Hadron Collider at CERN protons are collided with energies high enough to produce these heavy and short-lived bosons. However, since they decay rapidly, only the decay products of the Higgs boson can be measured. One of these decays, that was important in the discovery of the Higgs, yields two photons [1]. Such collision events, having two high energy photons in the final state, are called diphoton events. At hadron colliders such as the LHC, these events do not only occur in Higgs production, but have also background from strong and electroweak interactions [18]. Furthermore, in December 2015 the ATLAS and CMS collaborations (detectors at the LHC) announced the detection of a yet unexplained statistically significant deviation in the diphoton invariant mass spectrum around 750 GeV [10], [9].

In order to investigate these small peaks and deviations in the background diphoton spectra, the background must be well understood. This can be achieved to a large extent by Monte Carlo simulations of diphoton events in hadronic collisions. There, large numbers of these events are randomly generated using theoretical calculations and phenomenological models about the underlying processes.

However, other theoretical simulations, e.g. those performed by diphoton generators such as DIPHOX [6], may not necessarily include all parts of such an event as it would occur at a hadronic collider. For example, the simulation may only look at the one scattering process that produces the two photons of a diphoton event. Additional scatterings of other constituents of the colliding protons ("partons"), than the two involved in that process may be completely neglected. Simulations may also only calculate the event up to a certain decay stage of the produced particles.

If the results of the simulations are supposed to be compared to measurements, these non-perturbative effects must be accounted for again. This is achieved by generating the events one time accounting for all effects and comparing to the simulation that ignores some of the effects. From this, corrections on the incomplete simulations can be obtained. These can then be applied to all further simulations, potentially rendering further complete simulations unnecessary.

Several approaches to derive these corrections are available, collectively referred to as unfolding. So far, the typically used method applies simple correction factors in each distribution bin for the

observable. This is known as bin-by-bin unfolding.

In this thesis non-perturbative effects in diphoton events are investigated and corrected with iterative unfolding using Bayes' theorem. This more sophisticated approach was proposed by D'Agostini [11]. It is able to account for effects such as offsets and migration of events between bins. A comparison to the results obtained by using bin-by-bin unfolding is made.

The aim is to investigate, whether the corrections can be improved using this more advanced approach. In the first chapter 2, a theoretical overview of event generation, the methods of unfolding and diphoton events is given. Chapter 3 treats the event simulation using Sherpa [13] and the implementation of the event analyses and unfolding methods. The results of this thesis are presented and discussed in chapter 4. Finally, the conclusions of the works is drawn and an outlook is given in chapter 5.

2 Theoretical Framework

2.1 Non-perturbative effects and event levels

2.1.1 Non-perturbative effects

In order to calculate the probability of a particle scattering process, the interaction matrix element is needed. When calculating interaction matrix elements, particle physics usually relies on perturbation theory. More specific, Feynman diagrams are used up to a maximum order in the coupling constant. This approximation requires, that the coupling constant be small compared to one. However, for the strong coupling constant α_S , this is not always the case. As described by Wilczek and Gross [14], α_S is inversely proportional to the logarithm of the energy scale μ involved in the strong interaction

$$\alpha_S \sim \frac{1}{\ln \frac{\mu}{\Lambda_{\text{QCD}}}}, \text{ for } \mu \gg \Lambda_{\text{QCD}}. \quad (2.1)$$

Λ_{QCD} is the QCD scale and itself energy-dependent. For examples, $\Lambda_{\text{QCD}} \approx 210 \text{ MeV}$ is found [16] at energies close to the Z boson mass. Consequently, the coupling constant can become large for interactions happening at small energy scales μ . Such interactions cannot be described using perturbation theory in α_S and are called *non-perturbative effects*.

2.1.2 Hadron collision event generators and event levels

Fig. 2.1 shows a pp collision event as produced by a Monte Carlo hadron collision event generator [13]. The event generation is divided into different stages, or levels, distinguished by their energy scale.

First, from the two incoming protons (big green incoming ellipses) one parton each is randomly selected for the hard collision (red gluons interacting at big red dot). The hard collision is the interaction between two partons of the protons that occurs at the highest energy. Heavy particles created in the hard interaction (red lines springing from big red dot) will decay rapidly (smaller red dots). Depending on their lifetime, they may be simulated to decay directly in the hard interaction vertex or at a small displacement. Due to the high involved energies, those (red) interactions can be calculated in perturbation theory.

Additionally, lower energy parton showers (blue) are originating from the initial protons or the hard process. Perturbation theory can no longer be applied on their energy scales. The process must therefore be approximated by the generator.

The parton showers and remaining hard collision decay products are called final state partons and leptons. Together these particles will be referred to as the *parton level* in the following.

What follows are relatively low energy interactions. These include *secondary parton interactions*, where additional partons from the incoming protons are interacting with each other (purple ellipse). Predominantly, low energy QCD radiation is produced by the secondary scattering (purple radiation from ellipse). However, heavy particles are rarely created in these interactions and are normally neglected.

Further, all final state partons are simulated to decay into hadrons at low energy scales (light green

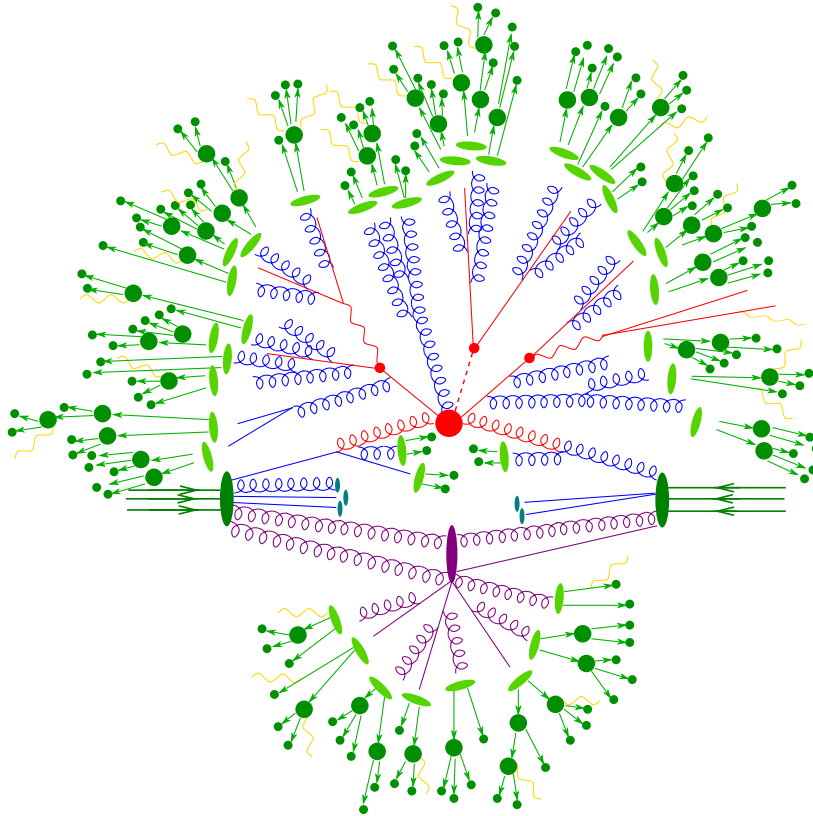


Figure 2.1: $t\bar{t}h$ event as generated by event generators artistically represented. The different colors distinguish the interaction or energy stages in the way the generator would divide them. [13]

ellipses). The subsequent unstable hadrons can again decay (dark green dots) to stable hadrons. These two processes are called *hadronization*.

All particles remaining as the result of the described simulations are together called *hadron level*.

In all described steps, QED radiation (bremsstrahlung) (yellow) may be generated by charged particles and assigned to all subsequent levels.

Hadronization consists of low energy scale interactions. It can therefore not be approximated by perturbation theory and presents a non-perturbative effect. Different models can be used to simulate the effects in Monte Carlo generators. Those are called *fragmentation theories*, an overview is found in [17].

Simulating the hadronization effects must be done for each event. This is time and resource consuming. Additionally, hadronization cannot always be simulated for all orders that the parton level is simulated in.

It would be beneficial, if the influences of hadronization on the parton level could be corrected without applying simulation for each study of the collision process.

Using unfolding could allow the determination and application of such corrections. The following section will describe the methods of unfolding. Thereafter, a specific process is chosen on which the methods will be tested.

2.2 Unfolding

In physics, measuring the "true" distribution of what one sets out to investigate often proves basically impossible. Setups and detectors are never completely ideal, leading to inefficiencies, biases and smearing in the measurement.

Methods accounting for these effects and trying to trace them back to get the true distributions are collectively known as unfolding. For clarity, this section will discuss unfolding specifically from distributions after the measurement (detector level) to those before (particle level). However, later in this thesis unfolding will be applied to correct several other effects, such as hadronization.

Two example of unfolding methods are described below: bin-by-bin unfolding and iterative unfolding based on Bayes' theorem. They focus on unfolding binned histograms to find the true underlying histograms.

To understand those methods it is helpful to think of the histograms in a vector representation (fig. 2.2), where the vector elements are the values for each bin.

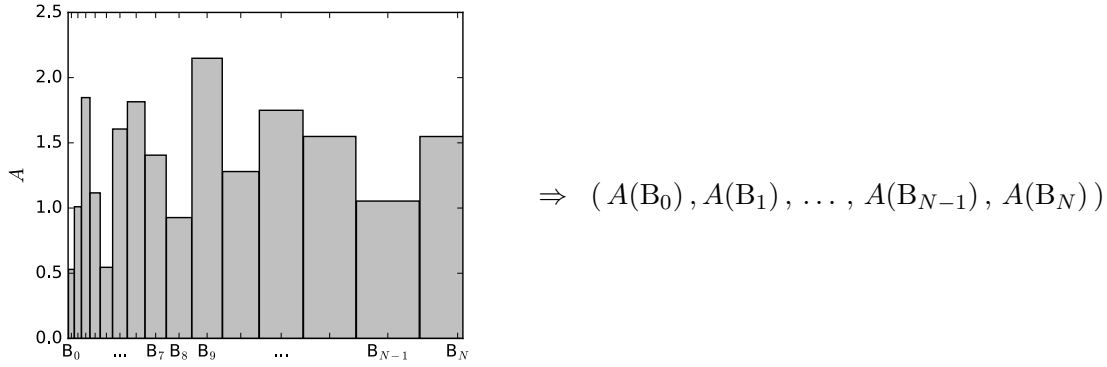


Figure 2.2: A binned histogram of a variable A with the bins B_i can be represented as a vector. The values at each bin are then the vector components.

2.2.1 Bin-by-bin unfolding

Bin-by-bin unfolding starts with the assumptions, that measured and true histogram have the same binning and that their difference is only a bin-dependent efficiency. The true histogram vector $\{A_t\}$ is therefore distorted to the measured histogram vector $\{A_m\}$ by a diagonal matrix with elements that correspond to the efficiency ϵ_i in each bin.

$$\begin{pmatrix} A_m(B_0) \\ A_m(B_1) \\ \vdots \\ A_m(B_N) \end{pmatrix} = \begin{pmatrix} \epsilon_0 & 0 & \cdots & 0 \\ 0 & \epsilon_1 & & \\ \vdots & & \ddots & 0 \\ 0 & & 0 & \epsilon_N \end{pmatrix} \begin{pmatrix} A_t(B_0) \\ A_t(B_1) \\ \vdots \\ A_t(B_N) \end{pmatrix} \quad (2.2)$$

Inverting this matrix is trivial. As a result, the true bin content is found as

$$A_t(B_i) = \frac{1}{\epsilon_i} \cdot A_m(B_i) =: C_i \cdot A_m(B_i). \quad (2.3)$$

Here the *Correction factors* $C_i := 1/\epsilon_i$ are defined. They can be determined by Monte Carlo simulating events on the true and the measured level. The correction factors are then calculated by dividing the binned histograms. Subsequently, they can be used for measurements to calculate the true histograms.

Bin-by-bin is the simplest unfolding method. However it also comes with a number of disadvantages. Most prominently, any kind of correlation between the bins is neglected. This includes, for example, smearing between the bins, i.e. when an event in a truth level bin can cause a measurement in the neighbouring bins. The simulated correction factors only account for smearing in the specific simulated distributions. In a different simulated distribution, the effect of smearing on the factors would differ. Additionally, the method does not treat biases in the measurement, i.e. an event in bin B_i always being measured in bin B_{i+k} .

The assumption that both histograms have the same binning also poses a big constraint on the application. Especially in cases of strong smearing or offsets, one might wish to choose different binnings for truth and measurement.

2.2.2 Iterative method

Iterative unfolding based on Bayes' theorem, developed by D'Agostini [11], aims to be more generally applicable approach than bin-by-bin. Since the method allows for different binnings in the true and measured histograms, the true bins will now be written as T_j ($j = 0, 1, \dots, N$) and the measured bins as M_i ($i = 0, 1, \dots, L$). Those may now also include bins for events, that do not fall inside the range of the histogram. Bayes' theorem describes the probability, that a certain effect E_i was caused by a certain source S_j , when the probability that any source causes any effect $P(E_k|S_l)$ is known for all effects and sources.

$$P(S_j|E_i) = \frac{P(E_i|S_j) \cdot P(S_j)}{\sum_l P(E_i|S_l) \cdot P(S_l)} \quad (2.4)$$

Going back to the example of detector measurement these probabilities mean the following. $P(S_j|E_i) \equiv P(T_j|M_i)$ is the probability, that a measurement in bin M_i was caused by an event originating from true bin T_j . Analogous, $P(E_i|S_j) \equiv P(M_i|T_j)$ is the probability, that an event in bin T_j causes a measurement in bin M_i . Finally, $P(S_j) \equiv P(T_j)$ is the probability, that an event occurs in bin T_j .

The value in true bin T_j can as a result be calculated as:

$$A_t(T_j) = \frac{1}{\epsilon_j} \cdot \sum_i P(T_j|M_i) \cdot A_m(M_i) =: \sum_i U_{ji} A_m(M_i) , \quad (2.5)$$

with

$$P(T_j|M_i) = \frac{P(M_i|T_j) \cdot P(T_j)}{\sum_l P(M_i|T_l) \cdot P(T_l)} \quad (2.6)$$

from Bayes' theorem, and the efficiency

$$\epsilon_j = \sum_i P(M_i|T_j) \quad (2.7)$$

with which an event originating in bin T_j is detected at all. This can be written in vector form:

$$\begin{pmatrix} A_t(T_0) \\ A_t(T_1) \\ \vdots \\ A_t(T_N) \end{pmatrix} = \begin{pmatrix} U_{00} & U_{01} & \cdots & U_{0L} \\ U_{10} & U_{11} & & \vdots \\ \vdots & & \ddots & \\ U_{N0} & \cdots & & U_{NL} \end{pmatrix} \begin{pmatrix} A_m(M_0) \\ A_m(M_1) \\ \vdots \\ A_m(M_L) \end{pmatrix}. \quad (2.8)$$

The matrix $\{P(M_i|T_j)\}$ determined by what event can cause measurements in which bin is called *Smearing matrix* or *Response matrix*. The matrix $\{U_{ji}\}$ in eq. (2.8) with the elements

$$U_{ji} = \frac{P(M_i|T_j) \cdot P(T_j)}{(\sum_k P(M_k|T_j)) \cdot (\sum_l P(M_i|T_l) \cdot P(T_l))} \quad (2.9)$$

is used to calculate the true histogram and consequently called *Unfolding matrix*. In order to use Bayes' theorem to determine the true histogram from the measured histogram, the probabilities that an event occurred in any bin $P(T_j)$ must be known. However, this is exactly what has to be determined.

To solve this problem, D'Agostini [11] developed an iterative method. It starts out in the simplest case with assuming a uniform distribution of the events among the true bins ($P(T_j) = 1/\text{Number of } T_i \forall j$). The iteration then repeats through the following steps:

- Using the given true distribution $\{P(T_j)\}$ and the known smearing matrix $\{P(M_i|T_j)\}$ the unfolding matrix $\{U_{ji}\}$ is calculated.
- Using eq. (2.8) the unfolding matrix is applied to the measured histogram $\{A_m(M_i)\}$ to get an estimate $\{A_t(T_j)\}$ of the true histogram.
- The new estimate for the true distribution $\{P(T_j)\}$ is determined by normalizing the found histogram distribution:

$$P(T_j) = \frac{A_t(T_j)}{\sum_l A_t(T_l)} \quad (2.10)$$

- Using a χ^2 -comparison between the new truth estimate $\{P(T_j)\}$ and the one from the previous step it is determined whether further iteration is necessary. If so, the process is repeated from the first point with the new estimate for the true distribution.
- If χ^2 is low enough, the last calculated $\{A_t(T_j)\}$ gives the final estimate for the true histogram.

D'Agostini observed, that the method reaches a minimum in χ^2 and starts to fluctuate strongly, if more iterations are used [11]. Therefore, an optimal number of iterations must be found. Increasing that number will then not yield better results.

2.3 Diphoton events

In this thesis, unfolding will be applied to correct from the parton level to the hadron level of diphoton events.

Diphoton events at particle colliders are defined as those collision that produce two isolated high- p_T photons. Recently they played an important role in the discovery of the Higgs boson at CERN [3] [8], as the Higgs can decay as $H \rightarrow \gamma\gamma$. Additionally, in December 2015 it was reported that the ATLAS and CMS detectors at CERN observed a yet unexplained peak in the diphoton production around 750GeV [10]. An overview of diphoton production processes at hadronic colliders is given in section 2.3.3. It is therefore important to test whether improvements in predictions about this process can be achieved. For example, higher order parton level simulations together with iterative unfolding may yield more accurate results.

2.3.1 Diphoton identification criteria

When deciding to analyse a specific production process such as the diphoton channel, the process must first of all be well defined to ensure comparable results. For diphoton production, the identification criteria are:

- The final state has to contain two photons.
- Each photon must have transverse momentum above a threshold. The threshold is varied depending on what is to be analysed.
- The photons must each be sufficiently isolated from hadronic activity. This is discussed in further detail below.
- To prevent significant energy deposits of the photons in each others isolation cone, they must have a minimal angular separation $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. $\Delta\eta$ is the pseudorapidity difference between the photon and a particle and $\Delta\phi$ their polar angle difference.

2.3.2 Isolation conditions

In hadron collisions the primary interest is in the hard collision, not so much the hadronization. Accordingly, one should ensure that the photons do not originate from fragmentation and are not part of a jet. This is done by requiring them to be isolated. Two approaches to define such a condition will be used in this thesis. Both set a requirement for the energy inside a cone around the photon. The cone is defined by all particles within a certain ΔR around the photon. Consequently, the cone energy $E_{\text{Cone}}(\Delta R)$ is defined as the sum over the energies of all particles inside the ΔR -Cone, except the photon. The difference between the approaches is the condition to the particles inside this cone:

1. Fixed maximum energy cone:

The simplest condition to the cone energy is a fixed maximum for a fixed cone. With the maximum allowed energy in the cone $E_{\text{max,Cone}} > 0$ and the fixed cone radius ΔR this approach can be written as:

$$E_{\text{Cone}} = \sum_{\substack{\text{particles p in} \\ \Delta R \text{ around photon}}} E_p \stackrel{!}{<} E_{\text{max,Cone}} . \quad (2.11)$$

The simplicity of this condition however comes with the disadvantage that collinear particles to the photon are allowed. Since fragmentation is a purely collinear process, this fixed energy cut does not reliably exclude photons originating from fragmentation [12].

2. Dynamic maximum energy cone:

A modified approach that by definition excludes collinear particles was developed by Frixione [12]. He proposed a condition that depends on the cone radius:

$$E_{\text{max,Cone}}(\Delta R) = \tilde{E} \cdot f(\Delta R) , \quad (2.12)$$

where $\tilde{E}$ is an arbitrary energy scale and $f(\Delta R)$ the cone-dependency. Since no collinear particles should be allowed the latter must fulfill

$$f(\Delta R) \rightarrow 0 \text{ for } \Delta R \rightarrow 0 . \quad (2.13)$$

In his paper, Frixione proposes

$$f(\Delta R) = \left(\frac{1 - \cos(\Delta R)}{1 - \cos(\Delta R_0)} \right)^n \quad (2.14)$$

$$\tilde{E} = E_\gamma \cdot \epsilon_\gamma \quad (2.15)$$

with the photon energy E_γ , the maximum cone radius ΔR_0 and two constant parameters ϵ_γ and n , as isolation function. This choice is mostly arbitrary, its main property being that it meets condition eq. (2.13).

When analysing data from hadronic collisions, the energies E_i in the calculations must be replaced by the transverse energies $E_{T,i}$.

2.3.3 Diphoton production in pp collisions

The probability of a scattering process can be assessed by the number of interactions within the process. Each of those interactions contributes an order of magnitude in the coupling constant α_{coup} of the involved interaction. If α_{coup} is smaller than one, higher orders in α_{coup} mean a stronger suppression. In this case, processes with matrix elements proportional to α_{coup}^0 are called leading order (LO) in the involved interaction. Those proportional to α_{coup}^1 are named next to leading order (NLO) in the involved interaction, and so on.

For hadronic scattering, the dominant interaction is the strong interaction with the coupling constant α_S .

Since photons do not carry an electric charge, the initial state particles of a diphoton production process must have a total charge 0. At hadronic colliders the initial state is made up of partons. Therefore diphoton production occurs as $\text{parton} + \text{parton} \rightarrow \gamma\gamma + X$. If the initial state contains quarks or antiquarks, a diphoton final state can be produced through LO (in QCD) processes. For gluon fusion ($gg \rightarrow \gamma\gamma + X$) this is not possible. Here the lowest possible order in QCD is α_S^2 and contains loop diagrams (see fig. 2.3b). However, they still contribute significantly to the diphoton events [18] and must therefore be included in simulations.

A table of the processes included in the Monte Carlo simulation used in this paper alongside their Feynman diagrams is given in A.1.

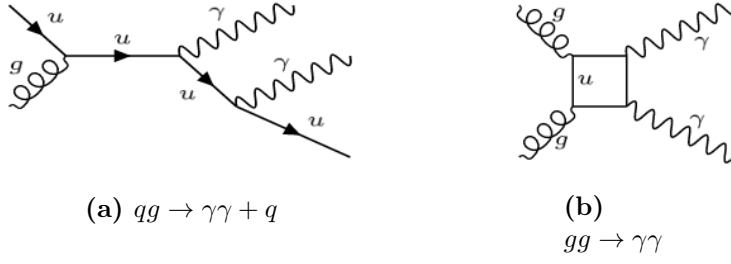


Figure 2.3: Examples for diphoton production processes in pp collision.

3 Implementation

To investigate the effect of iterative unfolding, collision events had to be simulated, analysed and finally unfolded. This chapter will describe how the individual steps were carried out and which programs were used.

3.1 Simulations using SHERPA

In order to generate diphoton events the Monte Carlo event generator SHERPA 2.2.0 is employed [13]. Only the most important settings chosen for the generation of diphoton events are discussed here. A version of the settings file ("Run card") including commentary is given in the appendix A.1. Two proton beams with an energy of 3.5 TeV each are chosen as colliding beams. This reflects the 7 TeV run of the LHC. The energy scale, at which the QED coupling constant α_{QED} is evaluated, is set to zero, corresponding to the photon mass.

Generated events must contain two final state photons and up to two additional jets are allowed in the final state. They are only simulated at leading order. Only for gluon-gluon fusion higher order interactions through loop diagrams are allowed, since they contribute significantly [18].

The two final state photons (not originating from fragmentation) must fulfill a dynamic isolation cone cut as described in section 2.3, to ensure convergence of the matrix elements. Additionally, loose conditions such as a low p_T minimum for the two photons are imposed to decrease the number of simulated events, that are going to be rejected in the analysis.

For the generation at parton level, three more settings are made, corresponding to the definition of the levels in section 2.1.

1. The hadronization of the final state partons is turned off.
2. No multi-parton interactions are allowed (= no secondary hard interactions in the same pp collision).
3. The transverse momentum of the initial partons within the proton is fixed to zero. This is a necessary approximation for perturbative calculation and should therefore be included into the parton level simulation.

To use the generated events to create response matrices for unfolding, the events on parton and hadron level must be correlated. This means, that the event data must be recoverable on parton and hadron level for each event. Therefore, a common seed for the Sherpa random number generator is used.

The correlated outputs for hadron and parton level can then be put into the analysis.

For the main diphoton simulation, that is used in this thesis to create the cross-section distributions and response-matrices, one million diphoton events were simulated using Sherpa.

3.2 Event selection and analysis

For the analysis of the simulated data the toolkit Rivet 2.4.1 is used [7]. It allows the user to write an analysis that exerts cuts on the given input simulation and then calculate and plot observables of

the events that made it through the cuts. Several experimental analyses have been implemented as standard into Rivet.

This work is based on the standard implemented analysis used in an ATLAS measurement of isolated diphoton pair production [2]. This analysis is modified to include variations of the cut parameters and to differentiate between two types of input particles: the parton level and the hadron level.

The cuts correspond to the definition of a diphoton event, as described in section 2.3.1. First, at least two photons must be present. A fixed maximum isolation cut, as discussed in section 2.3.2, is then applied to all photons in the event. For the isolation cut, the cone radius is varied ($\Delta R_{\text{Cone}} \in \{0.1, 0.2, 0.4\}$) as well as the maximum allowed cone energy ($E_{\text{max,Cone}} \in \{4, 6, 10\}$ GeV). The event is required to have at least two isolated photons. The two highest- p_T isolated photons must then each pass a minimal- p_T -cut, which too is varied ($p_{T,\text{min}} \in \{15, 20, 25\}$ GeV). Finally, the two photons, called leading photon and sub-leading photon, must have an angular separation of at least $\Delta R_{\gamma\gamma} = 0.4$. In the original analysis [2], that is modified here, the cut parameters were fixed to $\Delta R_{\text{Cone}} = 0.4$, $E_{\text{max,Cone}} = 4$ GeV and $p_{T,\text{min}} = 25$ GeV.

All cuts are applied to hadron and parton level each. Two separate sets of 1D-histograms with the same observables are created for the two levels. If one of the levels fulfilled all requirements, the corresponding diagrams were filled. Additionally, a set of 2D-histograms for the same observables with the x -axis as the parton level value and the y -axis as the hadron level value was initialized. Those histograms only were filled if both levels made it through the cuts and serve as response matrices for the unfolding.

The analysis from [2] investigated four observables: the diphoton invariant mass $M_{\gamma\gamma}$, the diphoton transverse momentum $p_{T,\gamma\gamma}$, the angle between the diphoton pair $\Delta\phi_{\gamma\gamma}$ and the cosine of the polar scattering angle in the diphoton Collins-Soper frame $\cos\theta_{\gamma\gamma}^*$ [4]. Two new observables were added to the analysis to test the performance of iterative unfolding: The single photon transverse momentum for the leading isolated photon $p_{T,\gamma L}$ and the sub-leading isolated photon $p_{T,\gamma SL}$ in a diphoton event.

3.3 Fake data for testing of unfolding

In order to compare the iterative unfolding to the bin-by-bin method, two test samples of random fake data are created. This aims to generate a distributions of the diphoton observables on hadron and parton level, which are different to the one used in the response matrix generation. This is achieved by changing the simulation as well as the analysis. In the simulation setting, the weighting of the events is enabled.

For the first test sample the analysis then randomly fluctuates the weight of each event according to a Gaussian. The mean of the Gaussian is set to the input weight and its the standard deviation to one tenth of the input weight.

A second fake sample is created by setting all weights to 1.0. This is equivalent to ignoring the weights.

For both samples, Sherpa simulation with 500,000 events each was used.

3.4 Unfolding

The ROOT framework RooUnfold [5] is used to implement bin-by-bin unfolding as well as Bayes iterative unfolding in a Python script. It uses the 2D-histograms from the Rivet analysis and the corresponding 1D-histograms for both parton and hadron level to create a response matrix. The parton level 1D histogram is used to take into account events that are identified as diphoton only on parton level. With the hadron level 1D histogram the inefficiency of each hadron level bin can be calculated.

With that, it unfolds another 1D-histogram parton level set that was produced using the same Rivet analysis on a different set of simulated data.

Additionally to the comparison of bin-by-bin and iterative unfolding, the number of iterations for the Bayes' method is varied between $N_{\text{iter}} = \{1, 2, 3, 4, 7, 11\}$.

The resulting hadron level histograms are then compared to the known true hadron level. For this thesis, the previously described normally weighted and analysed simulation results are used to unfold the fake data.

3.4.1 Error estimation for iterative unfolding

Three sources of uncertainties were taken into account for the estimation of the error in the iterative unfolding. The first two are the statistical uncertainties on the parton level 1D-histogram, that has to be unfolded, and on the response matrix. Those uncertainties arise from a finite Monte Carlo sample assuming Poisson distribution. An estimation of their influence on the unfolded result is achieved by testing the effects of fluctuating them within their uncertainty. For each histogram, 2000 toy experiments are created by using a Gaussian distribution for each bin, with the original value as expectation value and the original uncertainty as standard deviation, to get a new random bin value. The error on the observable A in bin B_i by fluctuating the histogram can then be estimated as:

$$\Delta A(B_i) = \sqrt{\langle (A(B_i) - \langle A(B_i) \rangle_{\text{Toys}})^2 \rangle_{\text{Toys}}} = \sqrt{\langle A(B_i)^2 \rangle_{\text{Toys}} - \langle A(B_i) \rangle_{\text{Toys}}^2} \quad (3.1)$$

where $\langle \cdot \rangle_{\text{Toys}}$ is the mean value over all toy histograms. This is done independently for the parton level histograms and the response matrices. A third source of uncertainty is the influence of the number of iterations of the result. To approximate this error for the unfolding with N_{iter} iterations, the unfolding with $N_{\text{iter}} + 3$ iterations is performed and the absolute difference is taken as estimate. For each bin, the three uncertainties are added quadratically:

$$\Delta A_{\text{total}}(B_i) = \sqrt{(\Delta A_{\text{parton level}}(B_i))^2 + (\Delta A_{\text{response}}(B_i))^2 + (\Delta A_{\text{iteration}}(B_i))^2}. \quad (3.2)$$

4 Results and discussion

4.1 Comparison using ATLAS analysis

The diphoton data from the Sherpa MC simulations are evaluated using the Rivet analysis from [2] for hadron and parton level, respectively. In the simulations, no NLO effects (except gg loops) are considered. This yields a smaller total hadron level cross section from the simulations compared to the 4.9 fb^{-1} ATLAS data from 2011, which are used as comparison by the analysis.

$$\begin{aligned}\sigma_{ATLAS} &= 44.0^{+3.2}_{-4.2} \text{ pb} , \quad \sigma_{\text{MC,had}} \approx 37 \text{ pb} \\ \Rightarrow \frac{\sigma_{ATLAS}}{\sigma_{\text{MC,had}}} &\approx 1.2\end{aligned}\tag{4.1}$$

To account for the left out NLO effects, the MC simulations are multiplied by the factor 1.2. The resulting comparison can be seen in fig. 4.1.

A comparison of the MC simulation results created for this thesis to the Sherpa MC results from [2] shows to a large extent agreement within the errors. Some deviation is expected, since a different renormalization scale is chosen in this work. This difference is observed mainly in the low $M_{\gamma\gamma}$ region, where the simulation of this are closer to the ATLAS data, and the low $p_{T,\gamma\gamma}$ region, in which the simulations from [2] match the ATLAS data better.

Additionally, the here simulated hadron level data largely agrees with the ATLAS data within the margins of error, when corrected for the missing higher order effects. However, the higher order effects also influence the shape of the differential cross-section. Since those effects cannot be accounted for by a global factor, they are still noticeable in the distributions. This can be observed mainly in the $p_{T,\gamma\gamma}$ distribution, where the differential cross-section for $p_{T,\gamma\gamma} \lesssim 5 \text{ GeV}$ is underestimated and for $p_{T,\gamma\gamma} \gtrsim 10 \text{ GeV}$ is partially overestimated. It is also visible in underestimations by the simulation at the edges of the $\cos\theta_{\gamma\gamma}^*$ distribution.

For the comparison of hadron and parton level, it is again important to look at the calculated total cross-sections of the simulations.

$$\begin{aligned}\sigma_{\text{MC,par}} &= 43.0^{+3.2}_{-4.2} \text{ pb} , \quad \sigma_{\text{MC,had}} \approx 37 \text{ pb} \\ \Rightarrow \frac{\sigma_{\text{MC,par}}}{\sigma_{\text{MC,had}}} &\approx 1.2\end{aligned}\tag{4.2}$$

It should be noted, that the similarity of this number compared to the data/hadron-level ratio is a coincidence, since the ratios are dominated by two different effects.

This factor of 1.2 is the main difference between the two levels in the investigated observables. Due to the correlation between the simulations, the shape of both levels is similar even to the statistical fluctuations. A deviation is still observable: In the $M_{\gamma\gamma}$ distribution below 40 GeV , in the $p_{T,\gamma\gamma}$ distribution near zero and in the $\Delta\phi_{\gamma\gamma}$ distribution below 1.0 , the parton level increases or stays constant towards smaller values, while the hadron level decreases. This may be contributed to fragmentation photons in the hadron level, that are misidentified as isolated photons.

The overall similar shape of hadron and parton level in this analysis suggest, that smearing does strongly effect the transition from one level to the other.

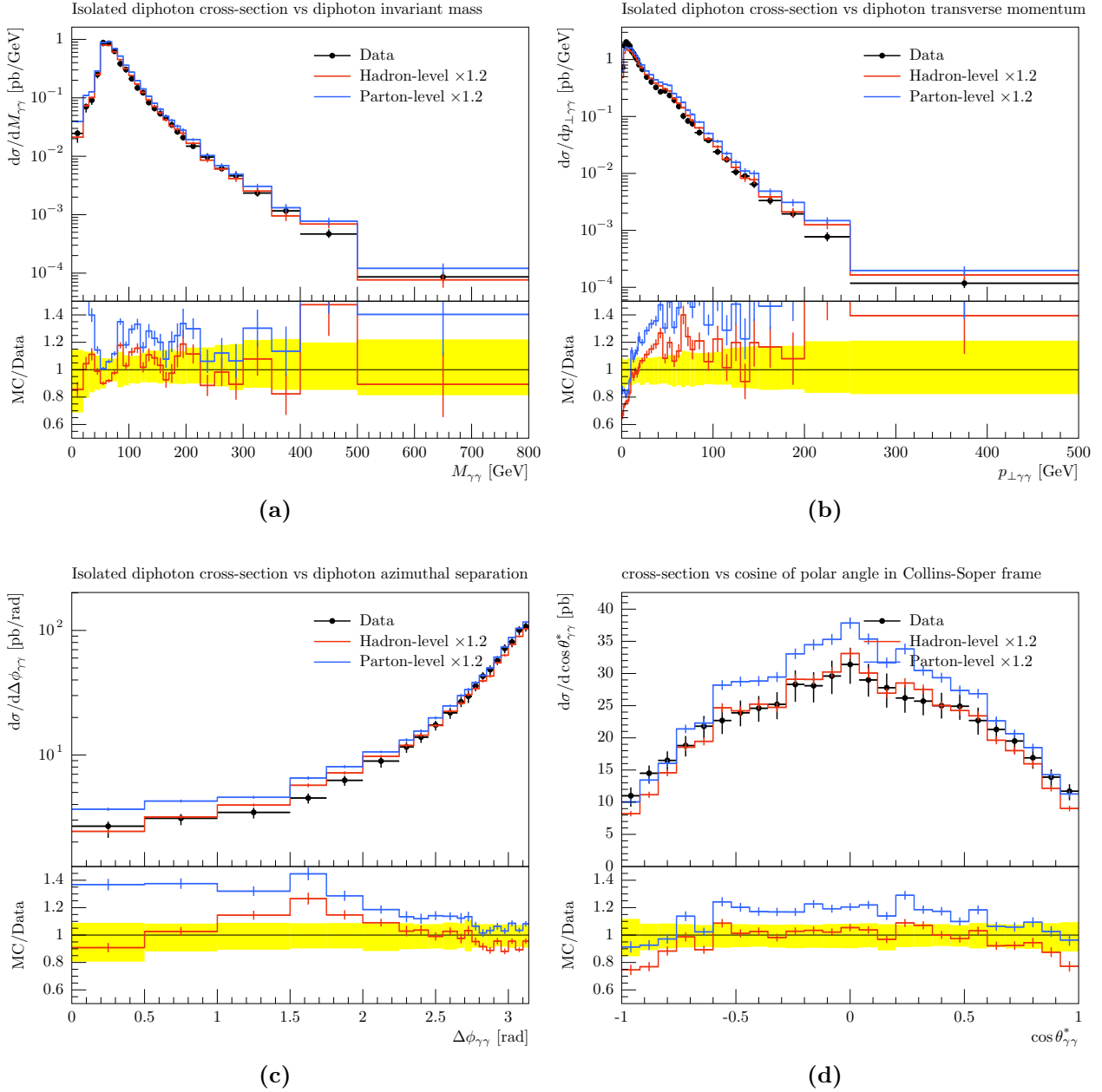


Figure 4.1: Comparison of differential cross-sections from Monte Carlo simulations using Sherpa to the 4.9fb^{-1} ATLAS data from 2011. (a) diphoton invariant mass. (b) diphoton transverse momentum. (c) polar angle between leading and subleading isolated photon. (d) cosine of Collins-Soper diphoton rest frame polar angle. The simulations (red, blue) are scaled by 1.2, which is roughly the ratio of data total cross section to MC total cross section.

4.2 Modified analysis

4.2.1 Additional observables

To investigate the different unfolding methods on additional variables, the transverse momentum of leading and sub-leading photon of the diphoton event are added as observables in the analysis. Their differential cross-section distributions are shown in Fig. 4.2.

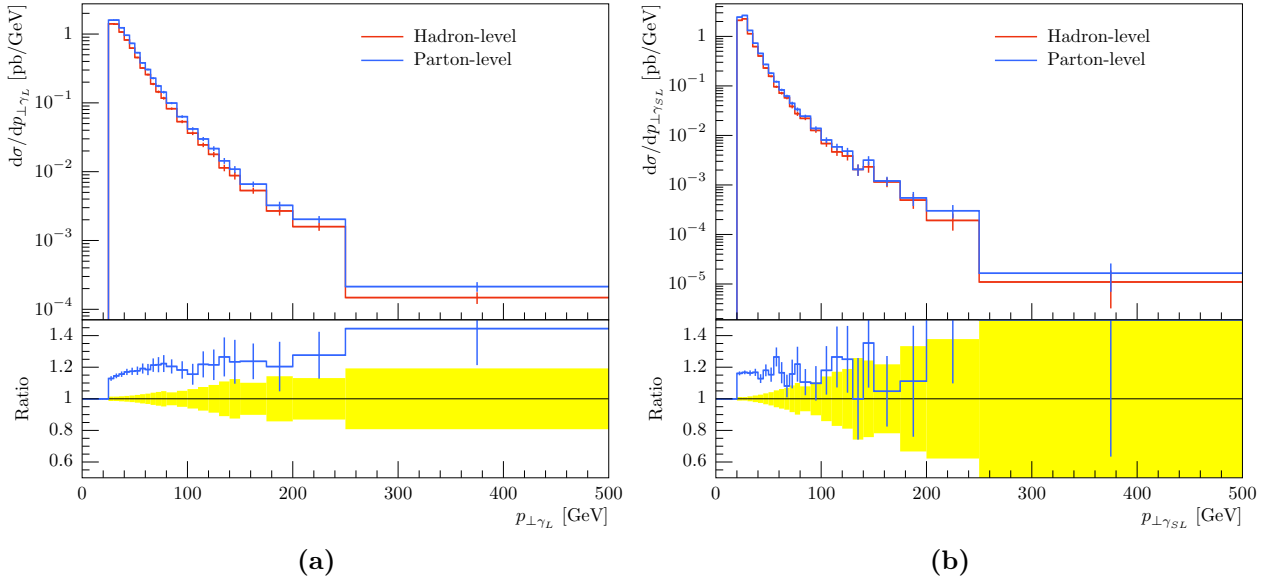


Figure 4.2: Newly added observables: single isolated photon transverse momentum of leading (a) and sub-leading (b) photon of diphoton event. Since no data is available, the ratio plots are ratio between parton level and hadron level. Cut parameters are the same as in [2].

Both hadron and parton level distributions are zero below $p_{T,\gamma} = 25$ GeV due to the p_T cut on the two highest p_T photons. The shape is, as in the other observables, largely similar between the both levels. A main difference is the factor of around 1.2 from hadron to parton level, as explained in the previous section. For the leading photon this factor decreases towards low $p_{T,\gamma}$, but stays above 1.1.

4.2.2 Response matrices

In the modified Rivet analysis, 2D-histograms are filled with events that fit all requirements to be identified as diphoton event on hadron level as well as parton level. The axes are the value of the observable on parton level and hadron level, respectively. Those 2D-histograms can therefore be identified as the response-matrices from hadron to parton level. However, they do not include events, that fall outside the range of the histogram, and are not normalized. Later, this is corrected using RooUnfold, which includes the 1D-histograms for each level in the response-matrix calculation.

Here, the matrices for the cut parameters used in [2] are discussed. The generated matrices depicted in fig. 4.3 and in the appendix in fig. A.1.

It is observed, that all matrices are primarily diagonal (same values at hadron and parton level). This means, that smearing or biases between hadron and parton level are small or that the binning is not fine enough to capture them. The notable exception is the $\cos\theta_{\gamma\gamma}^*$ 2D-histogram, where an additional perpendicular diagonal is found. Those are events, where the observable changes its sign

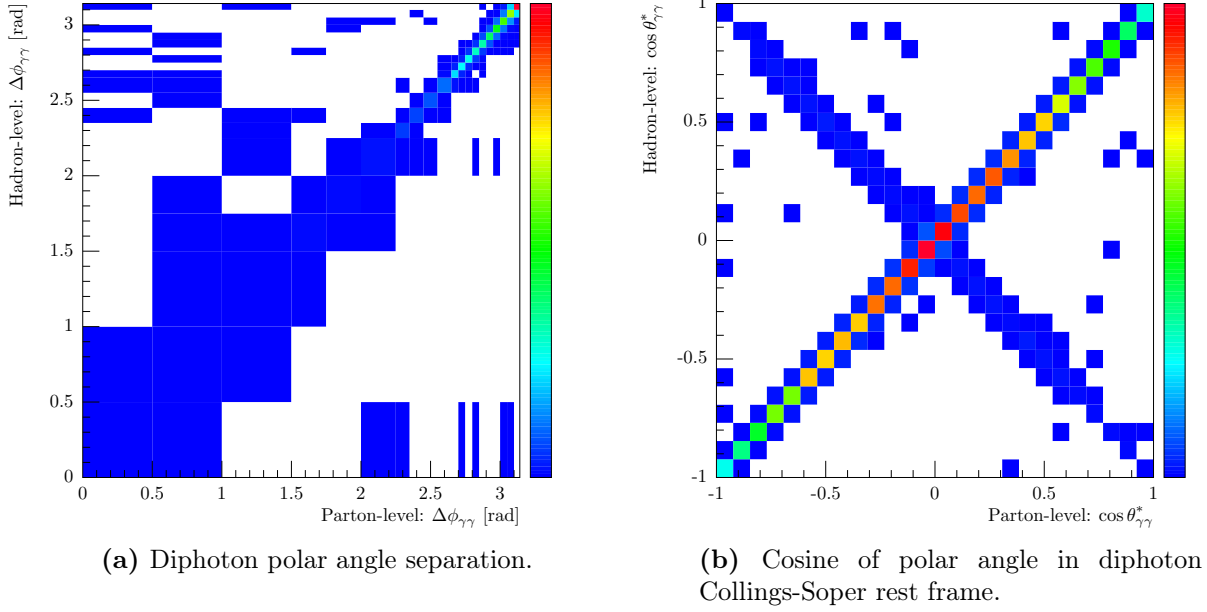


Figure 4.3: Response matrices for $\Delta\phi_{\gamma\gamma}$ and $\cos\theta_{\gamma\gamma}^*$ obtained from a 1 million event Sherpa diphoton simulation analyzed in Rivet. The cut parameters are as in [2]. The color-scale is linear and in arbitrary units with dimension $[\text{pb}/(\text{GeV})^2]$.

from one level to the other. A possible explanation is found in the definition of $\cos\theta_{\gamma\gamma}^*$, where a dependence $\cos\theta_{\gamma\gamma}^* \propto \sinh(\eta_{\gamma_L} - \eta_{\gamma_{SL}})$ occurs. Should the order of the two leading photons change, i.e. the leading photon loses enough energy to become the sub-leading photon, then due to the $\sinh$ the sign of $\cos\theta_{\gamma\gamma}^*$ changes.

4.2.3 Influence of analysis cut parameters

The cut parameters $p_{T\gamma\min}$ (minimal transverse momentum of each of the leading isolated photons), ΔR_{Iso} (photon isolation cone radius) and $E_{\text{max,Cone}}$ (maximum allowed energy in photon isolation cone) in the Rivet analysis are varied. This aims to investigate their influence on the differential cross-section distributions and response matrices. If changes are only in the total cross-section and not in the shape of the distributions, the cut parameters could be chosen more inclusive in future research.

As an example the diphoton invariant mass histograms for the variations are given in the appendix in figures A.2, A.3, A.4 and A.5.

First, the deviations of the 1D parton or hadron level distributions are discussed.

When varying the minimum transverse momentum that the leading photons need to have $p_{T\gamma,\min}$, changes are only observed below $M_{\gamma\gamma} \approx 100 \text{ GeV}$ on both hadron and parton level. Differences are only observed below $M_{\gamma\gamma} \approx 60 \text{ GeV}$ for $p_{T\gamma,\min} = 20 \text{ GeV}$. There, the distribution increases due to the allowed lower p_T diphoton pairs. This behaviour is observed for both levels.

Varying the radius of the photon isolation cone ΔR_{iso} yields a difference over the complete $M_{\gamma\gamma}$ range and influences the shape of the distribution. With decreasing cone radius, the differential cross-section increases, due to more photons passing the isolation cut. The effect is stronger on hadron level than on parton level. This could be caused by photons that emerged from fragmentation being misidentified as isolated due to the small cone. A smaller cone size should consequently be chosen with caution to

this effect. Within the scope of this thesis, they can however be useful to demonstrate the unfolding effect on non-perturbative effect that include smearing.

Variations in the maximum allowed energy in the photon isolation cone $E_{\text{max,Cone}}$ show only for $M_{\gamma\gamma} < 60 \text{ GeV}$ changes in the distribution shape. Above that value, the deviations for different $E_{\text{max,Cone}}$ are cut-dependent constant factor within the margins of error. For the parton level, this effect is again about half as strong, because of the hadronization to the hadron level, which causes more particles to be in the cone around any photon.

When considering unfolding, the variations of the response-matrices (fig. A.5) are important, since they describe smearing and offsets between the levels. As above, for the variation of $p_{T\gamma,\text{min}}$ changes are only observed for low $M_{\gamma\gamma}$. No new high value off-diagonal elements appear when changing this parameter. Changing the isolation cone radius ΔR_{iso} adds new off-diagonal elements to the response-matrix. However, the diagonal elements to a large extent exceed the off-diagonal elements. The variation of $E_{\text{max,Cone}}$ too causes new off-diagonal elements to appear, but fewer than the ΔR_{iso} variation and the diagonal is again dominant.

It is concluded, that looser diphoton identification cuts, while producing off-diagonal elements, do not lead to dominant smearing. Accordingly, varying the cuts is expected to change the comparison of the unfolding methods only slightly.

4.3 Unfolding of fake diphoton samples

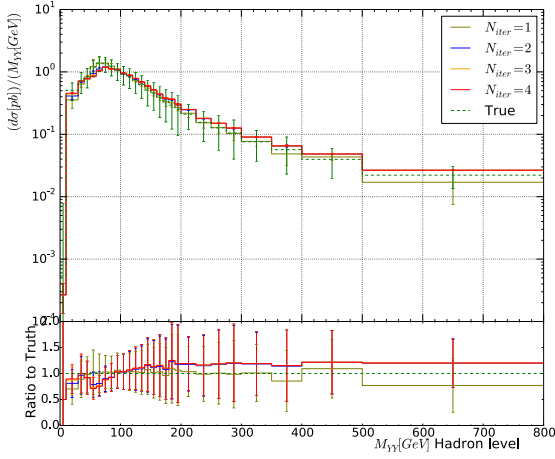
In this chapter the results of unfolding the test samples are presented. First, the results for the analysis cut parameters as in [2] are discussed. Thereafter, the influence of varying the analysis cut parameters on the unfolding result is investigated. Concluding the chapter is an examination of the influence, that the reweighting in the analysis had on the response matrices of the fake diphoton samples.

Both fake diphoton samples for testing the unfolding methods were produced from a weighted simulation done with Sherpa. The first fake sample, in which all event weights are set to unity in the analysis, is called the unweighted sample. The second fake sample, where each event weight is randomly varied, is dubbed random weight sample. The sample using the actual weights given by Sherpa is referred to as normal weight sample.

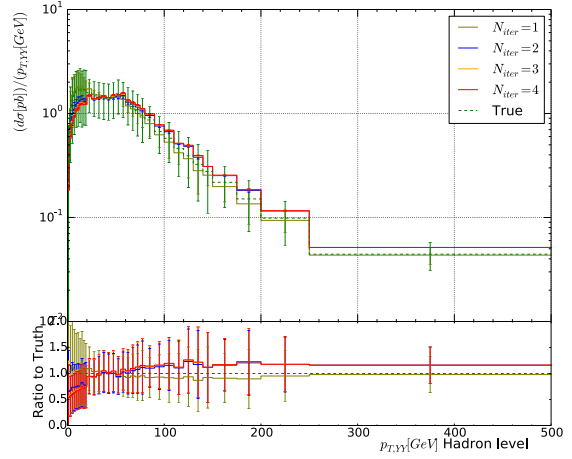
4.3.1 Original cut parameters

The results of unfolding the unweighted sample parton level distributions with the response matrix from the normal weight sample are presented in figures 4.4 and 4.5 and in the appendix in figures A.6 to A.9. For comparison, the true hadron level distributions taken from the same analysis are drawn as well.

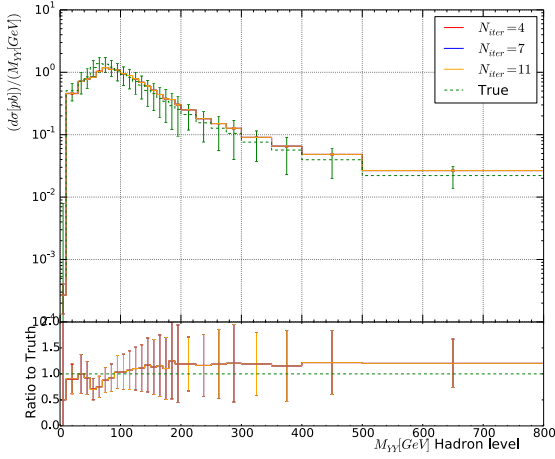
The distributions closest to the known true hadron level were achieved with iterative unfolding for either one iteration ($M_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, $p_{T\gamma_L}$ and $p_{T\gamma_{SL}}$) or two iterations ($\Delta R_{\gamma\gamma}$). For the $\cos\theta_{\gamma\gamma}^*$ distribution, one and two iterations are similarly close to the true distribution. More iterations in the iterative unfolding method do not yield better results. In the cases of the $p_{T\gamma_L}$ and $p_{T\gamma_{SL}}$ distributions, more iterations exhibit a divergence from the true distribution. This is expected for the iterative method after reaching the optimal number of iterations. The comparison with bin-by-bin



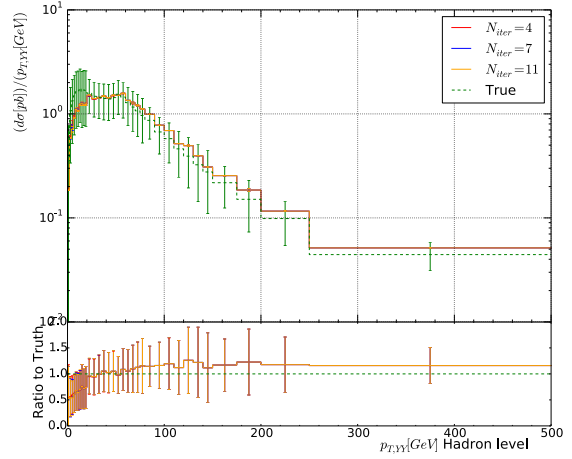
(a) Bayes iterative unfolding for low numbers of iterations.



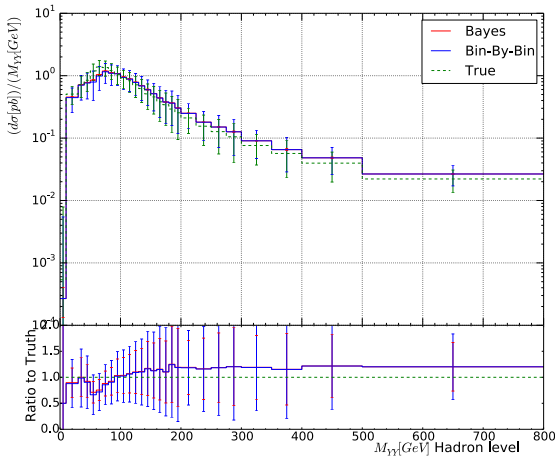
(a) Bayes iterative unfolding for low numbers of iterations.



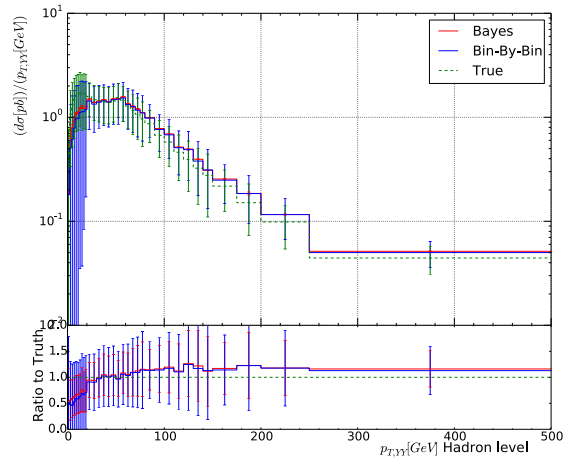
(b) Bayes iterative unfolding for higher numbers of iterations.



(b) Bayes iterative unfolding for higher numbers of iterations.



(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.



(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.

Figure 4.4: Result of unfolding fake unweighted parton level $M_{\gamma\gamma}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

Figure 4.5: Result of unfolding fake unweighted parton level $p_{T,\gamma\gamma}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

unfolding shows that the iterative unfolding for low numbers of iterations (one or two) delivers results closer to the true distributions. For the $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$, $\Delta R_{\gamma\gamma}$ and $\cos\theta_{\gamma\gamma}^*$ distributions, the outcome of bin-by-bin unfolding includes the truth within the margin of error. At values of the single photon p_T distributions close to the cut value, both bin-by-bin and iterative unfolding for four iterations yield results that differ from the true distribution and the differences are not covered by the statistical uncertainty.

In all distributions, the iterative method with only few iterations results in smaller uncertainties than the bin-by-bin method.

The unfolded distributions for the fake diphoton sample with randomly varied weights are shown in figures 4.6 and 4.7, and further in the appendix in the figures A.10 to A.13.

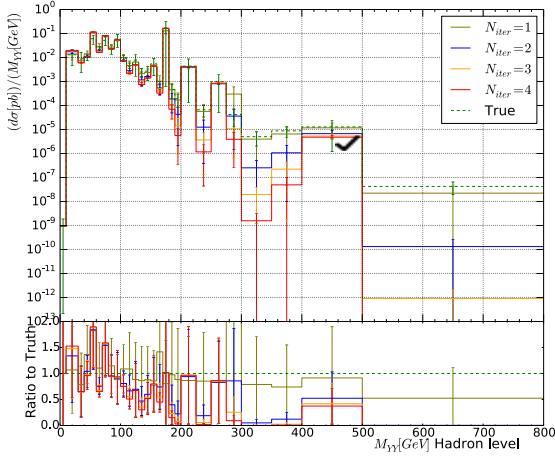
For this sample, due to the random distributions, it is not straightforward to determine the result closest to the true hadron level distributions. However, for all distributions iterative unfolding is shown to be able to achieve better results than bin-by-bin unfolding. Only one or two iterations are necessary. Additional iterations show a behaviour diverging from the truth, as expected for the method. Again, the iterative unfolding shows to a large extent smaller uncertainties than the bin-by-bin method.

The shown unfolded results indicate that using the iterative method developed by D’Agostini with only one or two iterations tends to yield more accurate results than the bin-by-bin method. This should be further investigated using other test samples, e.g. some created with another diphoton parton level tool. The diverging behaviour for higher numbers of iterations is expected for the iterative method [11]. In this context it should be noted that the distributions of the fake samples may be unrealistic. This further supports the need for other tests using physically meaningful distributions. Additionally, a χ^2 -comparison between the unfolding results and the known hadron level should be implemented to provide a quantitative measure of the accuracy.

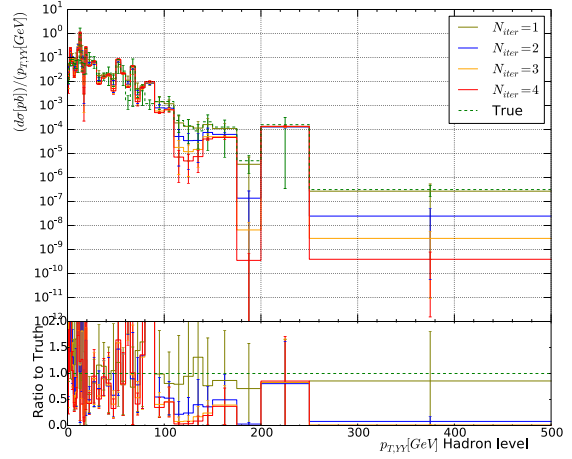
4.3.2 Variation of cut parameters

The influence of varying the cut parameters $p_{T\gamma,\min}$, ΔR_{Iso} and $E_{\text{max,Cone}}$ on the precision of the unfolding methods is investigated for the example of the $p_{T,\gamma\gamma}$ distributions. This is done for both fake diphoton samples. The results are shown in figures 4.6 and 4.7 and further in the appendix in figures A.10 to A.19.

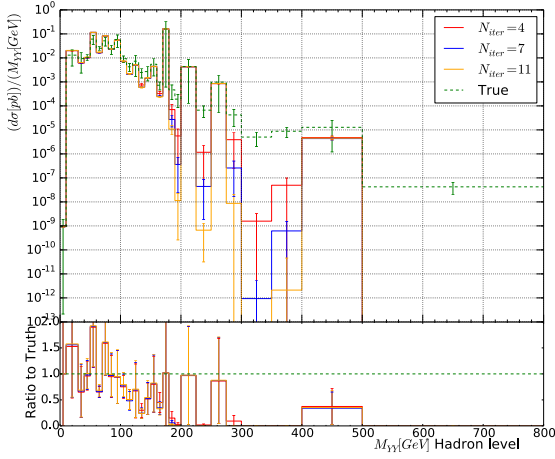
Varying the cut parameters, while changing the distributions themselves, does not seem to impact the accuracy of the unfolding methods. For both samples, the results closest to the true hadron level distributions are achieved using iterative unfolding with one iteration. More iterations in the iterative method lead to the expected divergence of the result away from the true distribution. Bin-by-bin unfolding is observed to yield results further away from the true distribution than the result of iterative unfolding. However, for the unweighted sample the true distribution is within the statistical uncertainty of the bin-by-bin unfolding results. For the random weight sample, the bin-by-bin method does not capture the true distribution within the uncertainties. This is expected, since looser cuts are observed to result in additional migration.



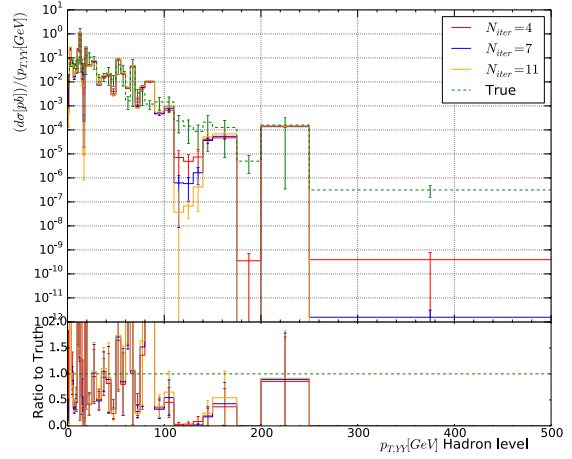
(a) Bayes iterative unfolding for low numbers of iterations.



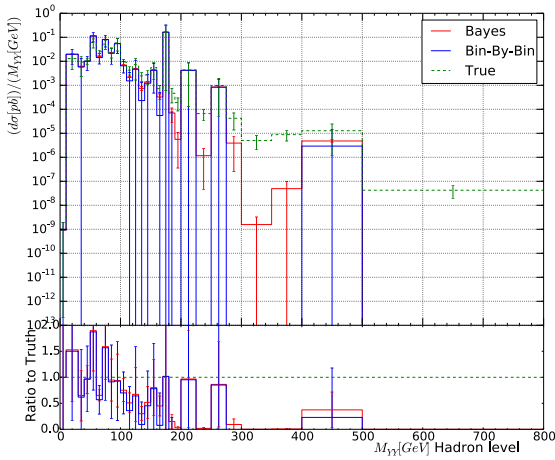
(a) Bayes iterative unfolding for low numbers of iterations.



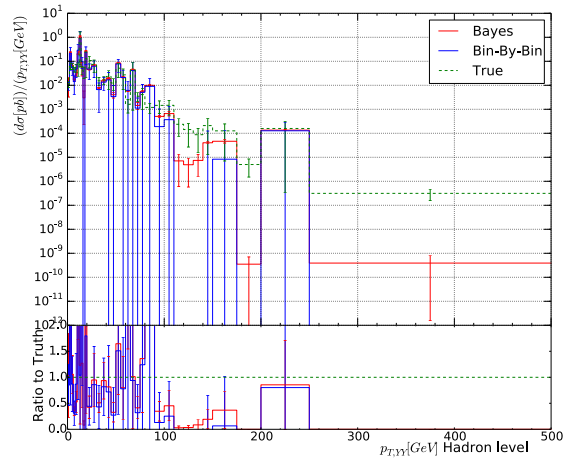
(b) Bayes iterative unfolding for higher numbers of iterations.



(b) Bayes iterative unfolding for higher numbers of iterations.



(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.



(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.

Figure 4.6: Result of unfolding fake random weight parton level $M_{\gamma\gamma}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

Figure 4.7: Result of unfolding fake random weight parton level $M_{\gamma\gamma}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

4.3.3 Validity of results

In order to assess the accuracy of the unfolding methods, the above results are compared to the hadron level distributions of the same fake sample. This section investigates if this comparison should be made.

The response matrices gained from a normal weighted sample are applied to the different distributions from the fake samples. As the ideal unfolding result one expects the hadron level from the same analysis as the fake parton level. For this to be true, the same response matrix must be valid for the normal weighted and the fake distribution. This means, that the probability for an event in a specific truth (hadron) level bin to cause an event in a specific effect (parton) level bin, must be equal for both samples used. In other words, the physical effects between both levels must be the same. Moreover, the distribution of events, on which each physical effect may act, is changed.

For the fake diphoton samples, the alternation of the weights assigned by Sherpa may affect the discussed probabilities. It should therefore be validated, that the same response matrices are obtained for the normally weighted sample and the reweighted samples.

In order to compare the 2D response matrices, an effective ratio r_{eff} is defined. It is supposed to show deviations, while weighting them according to their "importance". This includes how big the uncertainty on the bin value is and the percentage the bin contributes to the column of the matrix. The latter is chosen so that changes in bins, that may not strongly affect the outcome of the unfolding, are perceived as less significant. For the comparison, the values v in each bin are first normalized over each column, as this is done when using the response matrix for unfolding.

$$r_{\text{eff}}(B_{ij}) = \left(\frac{v^{\text{I}}(B_{ij})/N^{\text{I}}(j)}{v^{\text{II}}(B_{ij})/N^{\text{II}}(j)} - 1 \right) \cdot F(B_{ij}) + 1 \quad (4.3)$$

$$N(j) = \sum_i v(B_{ij}). \quad (4.4)$$

Here, $v^{\text{I}}(B_{ij})$ is the value in bin B_{ij} of matrix I and $F(B_{ij})$ the weight factor on the deviation. If only one of the values $v^{\text{I,II}}(B_{ij})$ is zero, $r_{\text{eff}}(B_{ij})$ is defined as zero. If both are zero, $r_{\text{eff}}(B_{ij})$ is set to unity. The weighting factor is separated into two parts for the two discussed significance measures.

$$F = F_{\text{column}} \cdot F_{\text{error}} \quad (4.5)$$

$$F_{\text{column}} = \frac{1}{2} \cdot \left(\frac{v^{\text{I}}(B_{ij})}{N^{\text{I}}(j)} + \frac{v^{\text{II}}(B_{ij})}{N^{\text{II}}(j)} \right) \quad (4.6)$$

$$F_{\text{error}} = \left(1 - \sqrt{\frac{\Delta v^{\text{I}}(B_{ij})}{v^{\text{I}}(B_{ij})} \cdot \frac{\Delta v^{\text{II}}(B_{ij})}{v^{\text{II}}(B_{ij})}} \right)^2. \quad (4.7)$$

At this point, it is important to stress that this choice is purely arbitrary and has no physical meaning. It is chosen only for the sake of visualization. Furthermore, the weighting factor is always smaller or equal 1.0, so the weighting always leads to a reduction of the deviations.

Examples for results obtained by this effective ratio to compare the fake diphoton samples (I) to the normal weighted sample (II) are presented in the figures 4.8 (unweighted sample) and 4.9 (random weight sample) and in the appendix in figures A.20 and A.21.

In both samples matrices with elements not equal one are found. The response matrices for $M_{\gamma\gamma}$ and $p_{\text{T},\gamma\gamma}$ are least affected by the reweighting of the test samples. For both samples, the $M_{\gamma\gamma}$ response

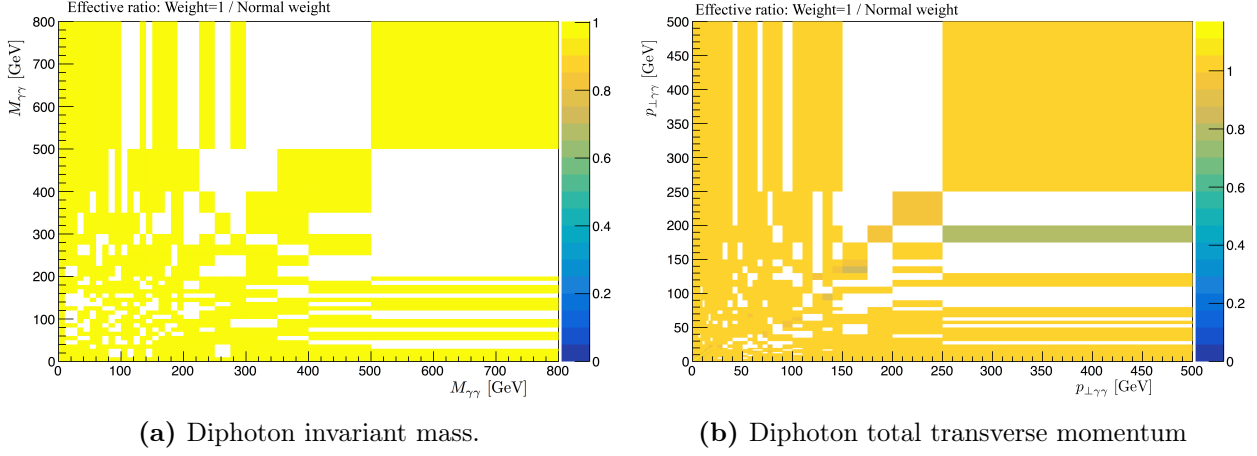


Figure 4.8: Effective ratio of fake unweighted diphoton sample response matrices to those from correctly weighted diphoton simulation. Cut parameter are as in [2].

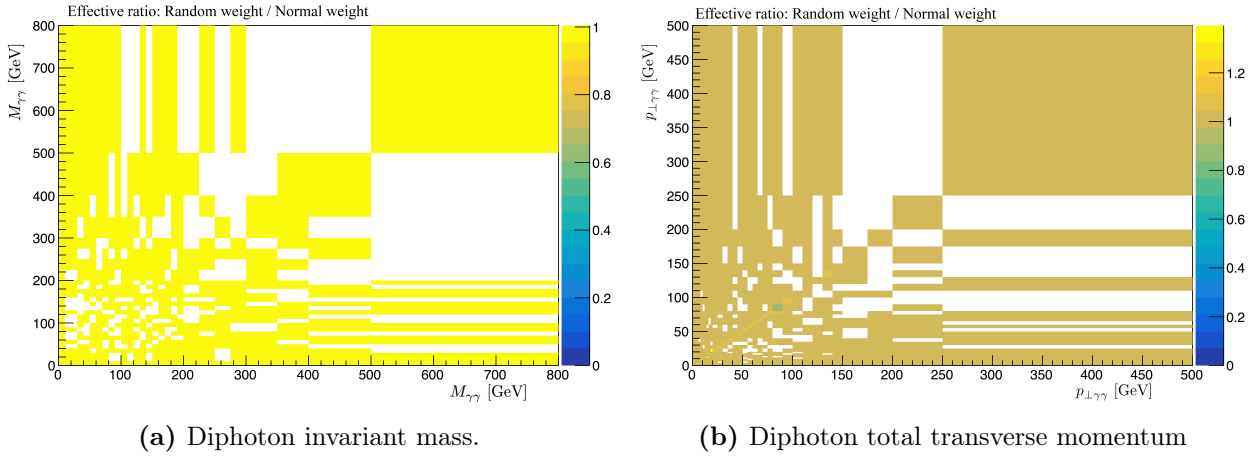


Figure 4.9: Effective ratio of fake random weight diphoton sample response matrices to those from correctly weighted diphoton simulation. Cut parameter are as in [2].

matrix shows $r_{\text{eff}}(B_{ij}) \approx 1$ for all occupied bins. The $p_{T,\gamma\gamma}$ response only yields a few off-diagonal deviations in the unweighted sample, while the random weight sample also yields deviations in the diagonal elements. For the single photon transverse momenta of leading and subleading photon both test samples show only few deviations from $r_{\text{eff}}(B_{ij}) \approx 1$. The strongest deviations are found in the $\Delta R_{\gamma\gamma}$ and $\cos\theta_{\gamma\gamma}^*$ response matrices. While for the unweighted sample the effective ratio never exceeds 1.2, the random weight sample shows values of $r_{\text{eff}}(B_{ij}) \approx 1$ up to 2 and in a single case of the $\cos\theta_{\gamma\gamma}^*$ response matrix even a value around 7.

These deviations from the in principle necessary equality of the response matrices suggest, that the comparison of the unfolding results with hadron level from the sample analyses may not be valid.

The unfolding should consequently be further tested using samples that have by construction the same response matrix as the normal weighted simulation results. This can for example be achieved by reweighting all events that fall into a specific bin of the hadron level with a bin-dependent constant factor. This would affect the cross-section distributions over the observables, while keeping the probability distribution for an event of a specific hadron level bin to fall into any parton level bin constant.

5 Conclusion

The goal of this thesis was the determination and application of non-perturbative corrections from hadron level to parton level for diphoton events. Specifically, iterative unfolding based on Bayes' theorem was to be compared to the standard unfolding method of using simple bin-by-bin correction factors.

Using Sherpa 2.2.0 a diphoton event sample from pp collisions at $\sqrt{s} = 7\text{ TeV}$ is generated. The sample is analyzed using Rivet 2.4.1 and the implemented analysis to the ATLAS analysis [2], the results from the same paper can be reconstructed with few deviations due to improvements in Sherpa and updated simulation settings. Variations of the cut parameters of the analysis to more inclusive cuts show that a variation of the photon isolation cone radius and the minimal required single photon momentum lead to deviations in the shape of the cross-section distribution. The choice of those parameters should include an examination of the efficiency and misidentification rate. Varying the maximally allowed energy in the photon isolation cone changes the total cross-section. Deviations in the shape of the distribution are within the statistical uncertainties. Two new observables are added to the diphoton analysis, namely the single photon transverse momentum of the leading and the subleading photon. The response matrices of the investigated observables are shown to be mainly diagonal, illustrating that the non-perturbative effects do not cause major smearing or biases in the cross-section distributions. When unfolding test diphoton samples from parton level to hadron level, iterative unfolding [11] for low numbers of iterations (one or two) is observed to yield more accurate results than bin-by-bin unfolding. If the number of iterations is increased further, a slowly diverging behaviour of the iteratively unfolded results is observed. The influence of the iterative method may increase, if bin sizes are chosen smaller.

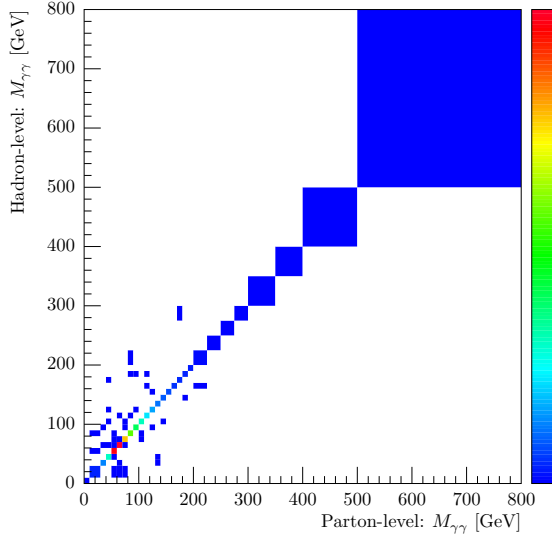
Limitations of comparing the unfolded results to the hadron level distributions of independent test samples are demonstrated for cases where the response matrices may be different between the nominal and the test samples. Further testing of the unfolding methods is recommended. This could include using samples that by default have the same response matrices.

For the investigation of the accuracy of iterative unfolding, the influence of the prior distribution on the unfolding result should be examined as well.

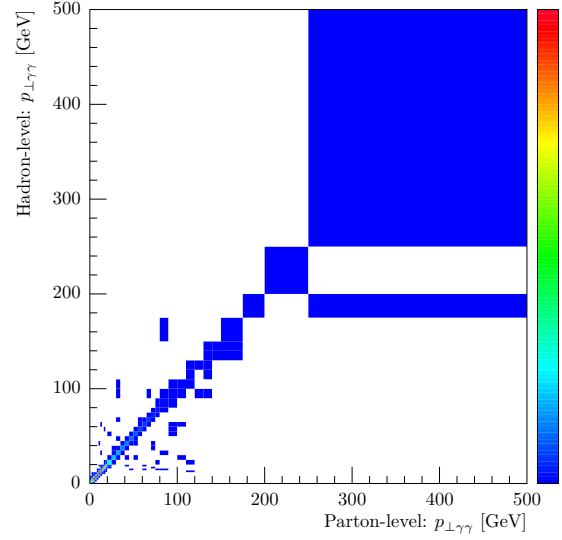
Since iterative unfolding is observed to improve the accuracy of non-perturbative corrections, further testing the application on other processes is recommended.

A Appendix

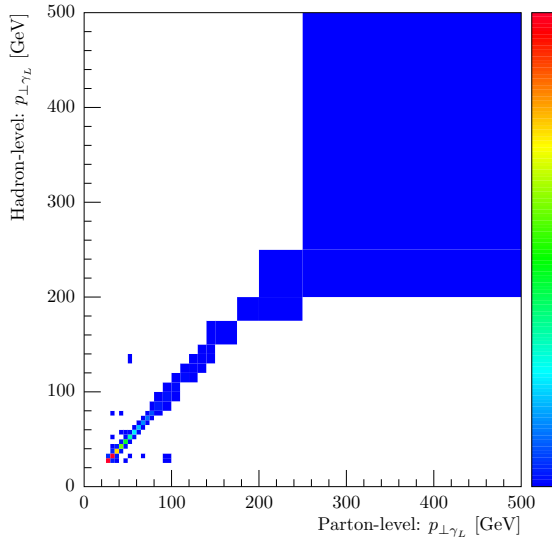
A.1 Figures



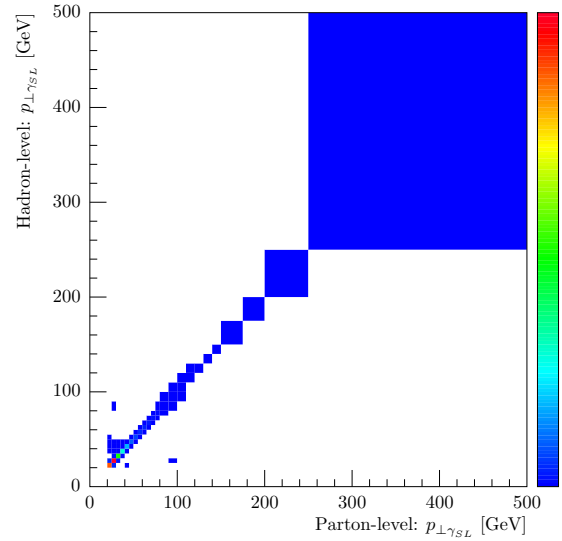
(a) Diphoton invariant mass.



(b) Diphoton total transverse momentum.



(c) Leading photon transverse momentum.



(d) Sub-leading photon transverse momentum.

Figure A.1: Response matrices obtained from a 1 million event Sherpa diphoton simulation analyzed in Rivet. The cut parameters are as in [2]. The color-scale is linear and in arbitrary units with dimension $[\text{pb}/(\text{GeV})^2]$.

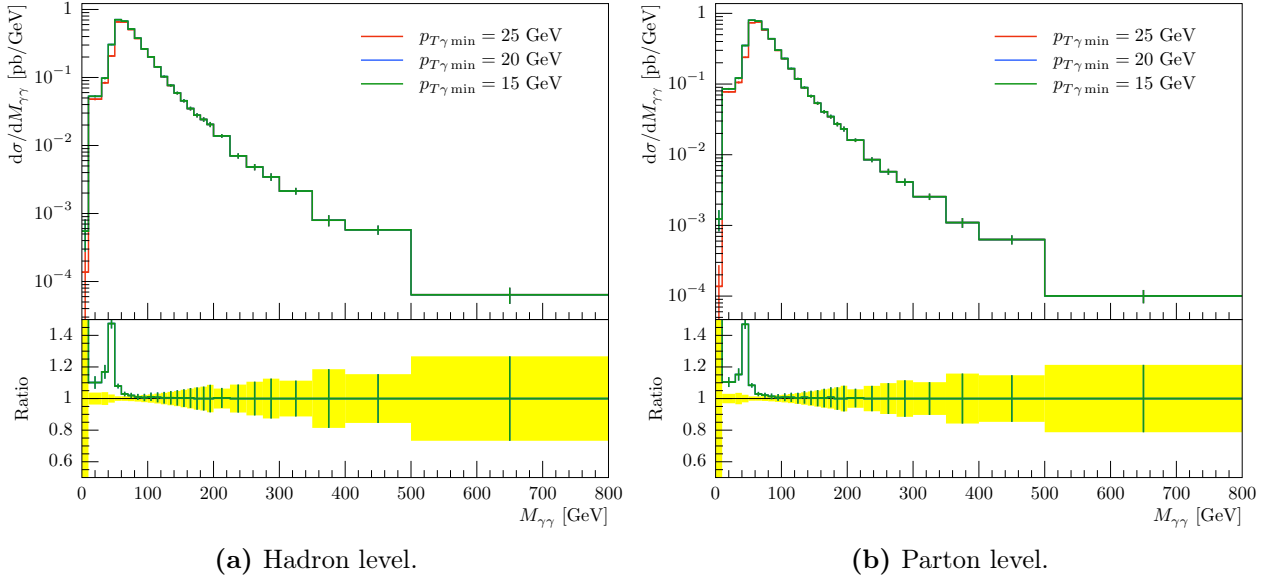


Figure A.2: Differential cross-section vs. invariant diphoton mass for varied cut parameter $p_{T\gamma, \min}$ (minimal transverse momentum of each of the leading isolated photons). The other cut parameters are as in [2]: $\Delta R_{\text{Iso}} = 0.4$, $E_{\text{max, Cone}} = 4$ GeV.

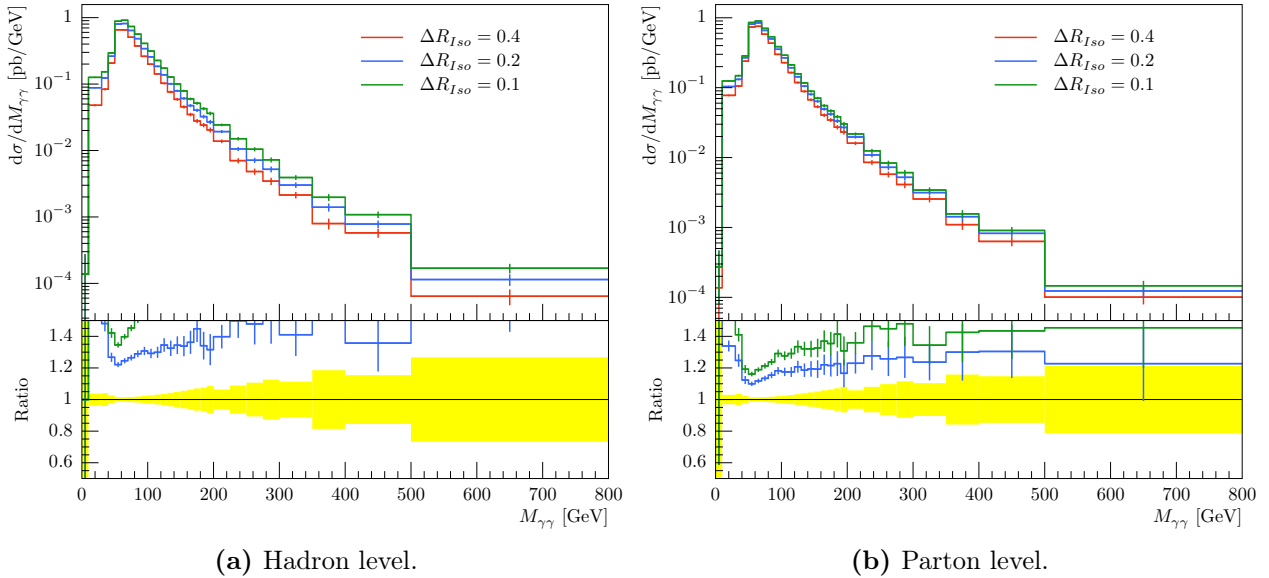


Figure A.3: Differential cross-section vs. invariant diphoton mass for varied cut parameter ΔR_{Iso} (photon isolation cone radius). The other cut parameters are as in [2]: $p_{T\gamma, \min} = 25$ GeV, $E_{\text{max, Cone}} = 4$ GeV.

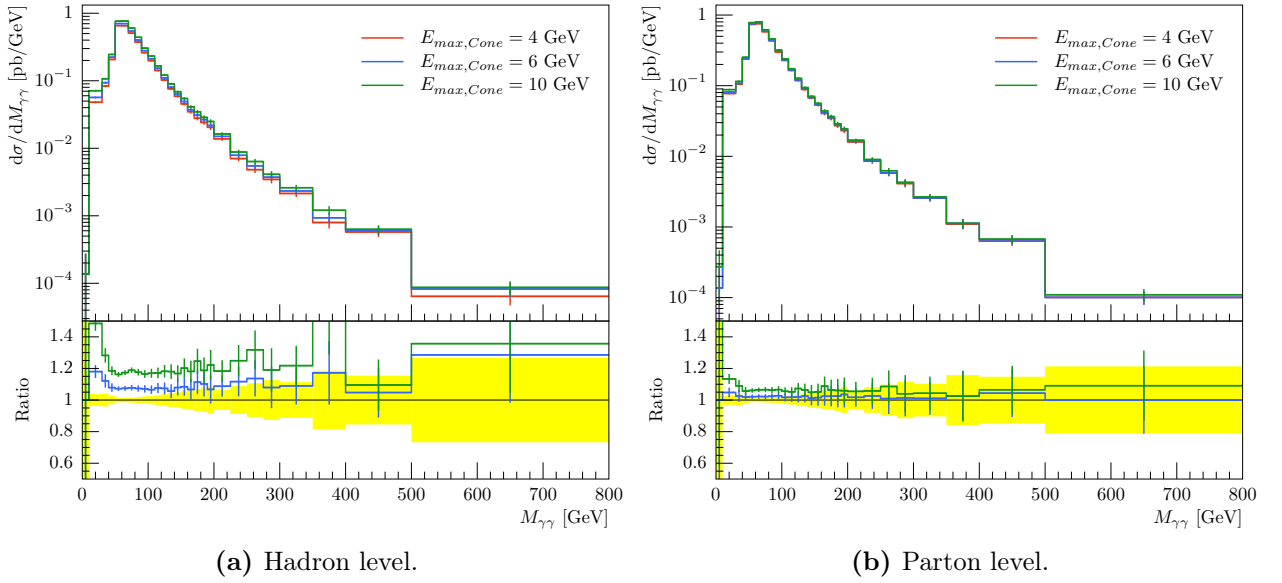
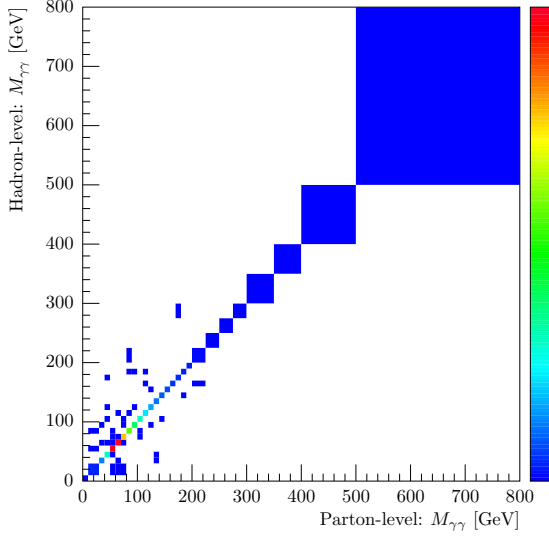


Figure A.4: Differential cross-section vs. invariant diphoton mass for varied cut parameter $E_{max,Cone}$ (maximum allowed energy in photon isolation cone). The other cut parameters are as in [2]: $p_{T\gamma,min} = 25$ GeV, $\Delta R_{Iso} = 0.4$.



(a) Original cut parameters.

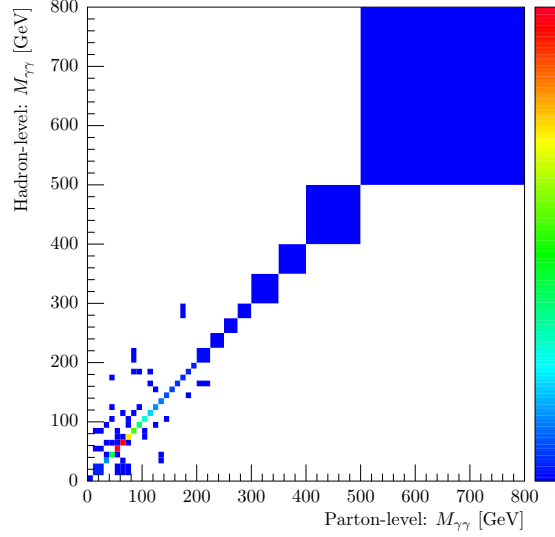
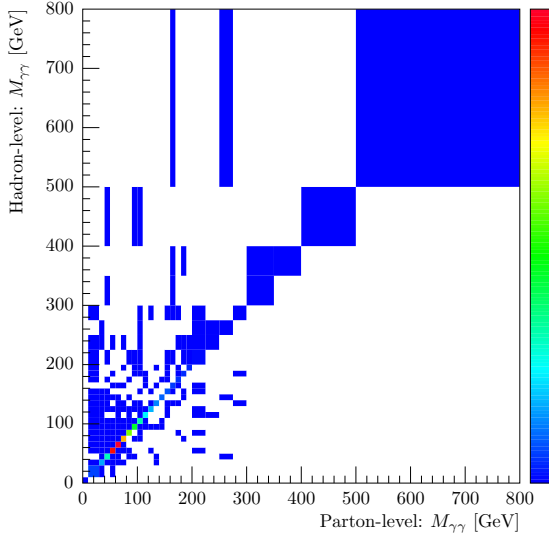
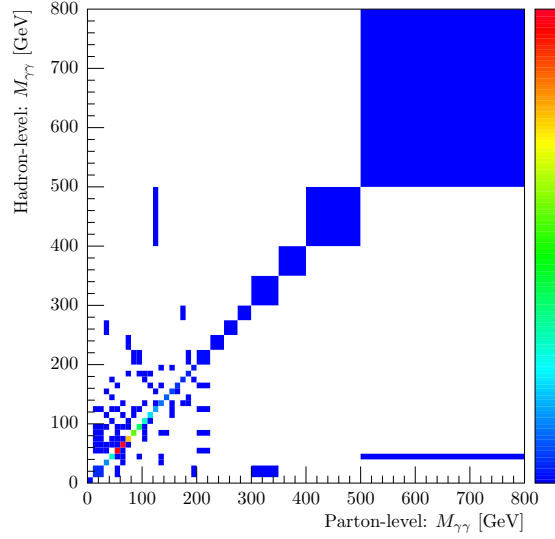
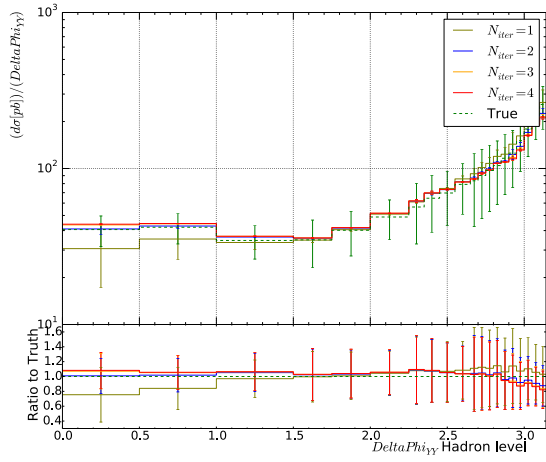
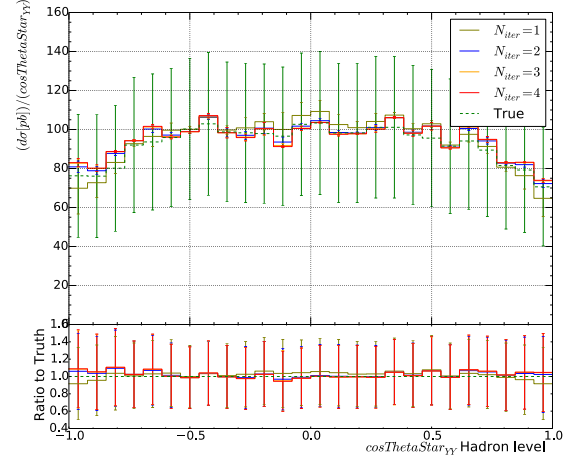
(b) Varied parameter: $p_{T\gamma} = 15$ GeV(c) Varied parameter: $\Delta R_{iso} = 0.1$ (d) Varied parameter: $E_{max,Cone} = 10$ GeV

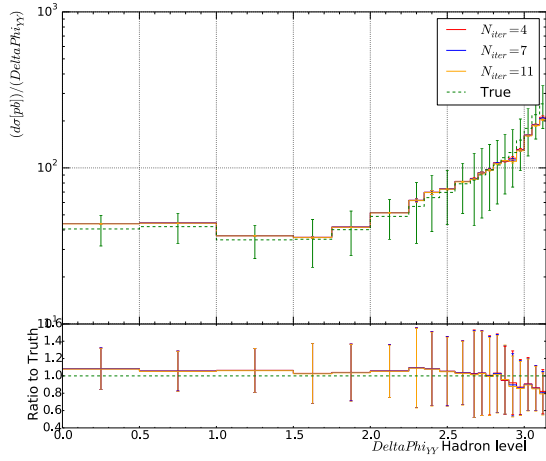
Figure A.5: Response matrices obtained from a 1 million event Sherpa diphoton simulation analyzed in Rivet. The cut parameters are varied, the original parameters from [2] being: $p_{T\gamma} = 25$ GeV, $\Delta R_{iso} = 0.4$ and $E_{max,Cone} = 4$ GeV. The color-scale is linear and in arbitrary units with dimension $[\text{pb}/(\text{GeV})^2]$.



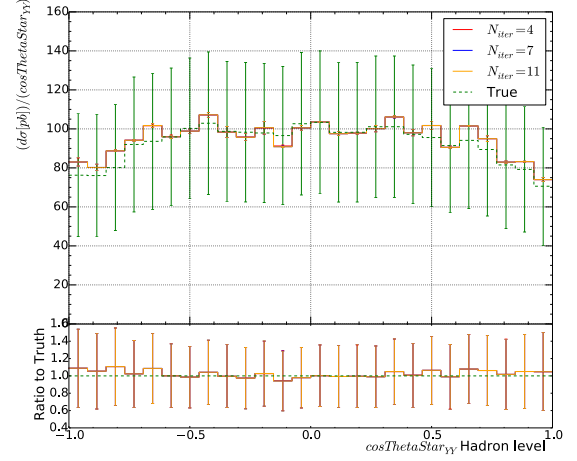
(a) Bayes iterative unfolding for low numbers of iterations.



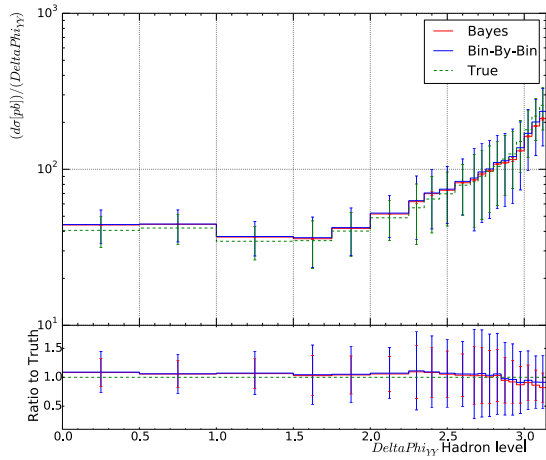
(a) Bayes iterative unfolding for low numbers of iterations.



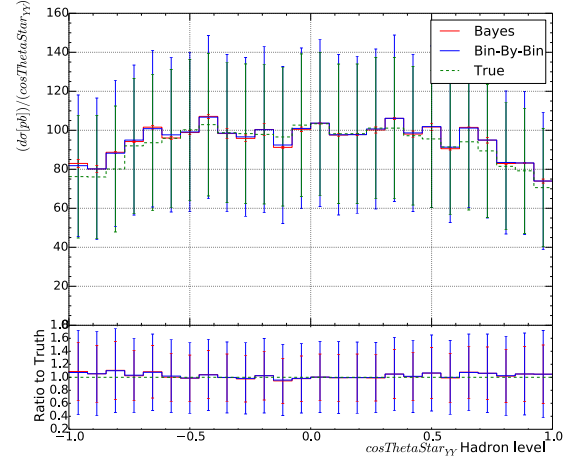
(b) Bayes iterative unfolding for higher numbers of iterations.



(b) Bayes iterative unfolding for higher numbers of iterations.



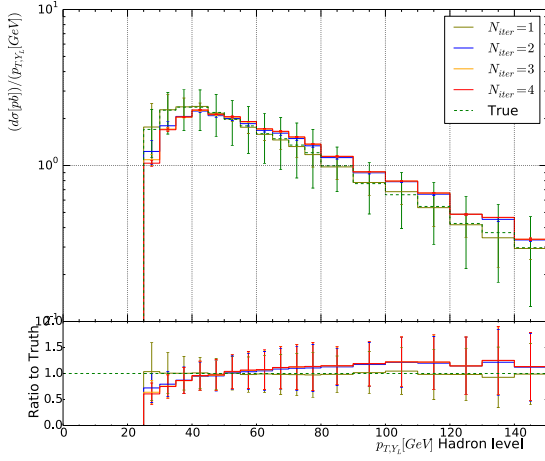
(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.



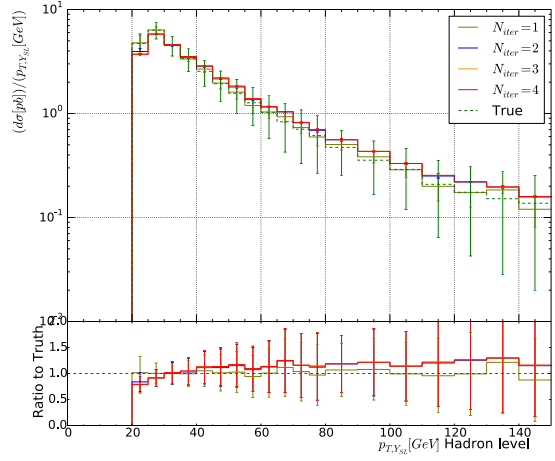
(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.

Figure A.6: Result of unfolding fake unweighted parton level $\Delta\phi_{\gamma\gamma}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

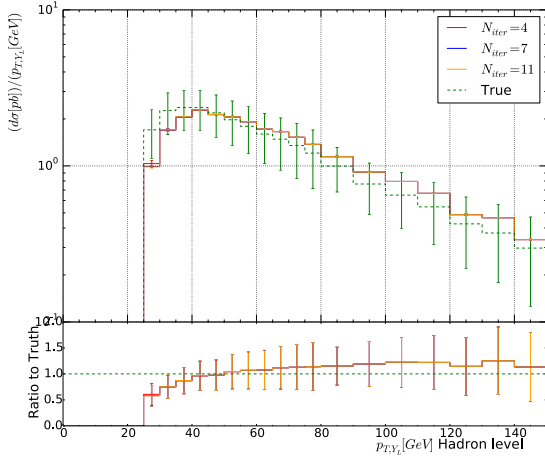
Figure A.7: Result of unfolding fake unweighted parton level $\cos\theta_{\gamma\gamma}^*$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].



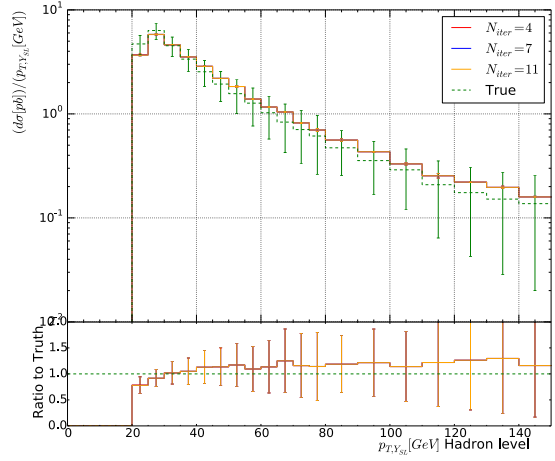
(a) Bayes iterative unfolding for low numbers of iterations.



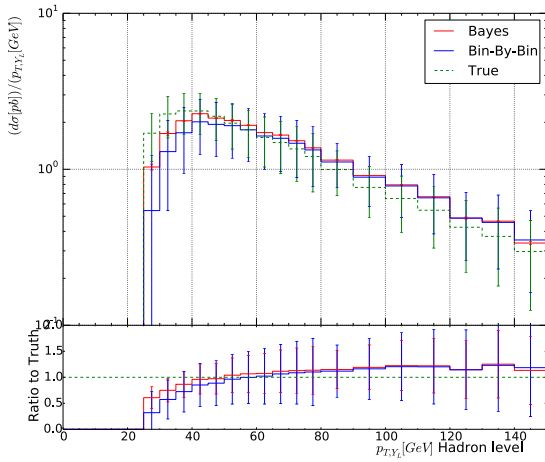
(a) Bayes iterative unfolding for low numbers of iterations.



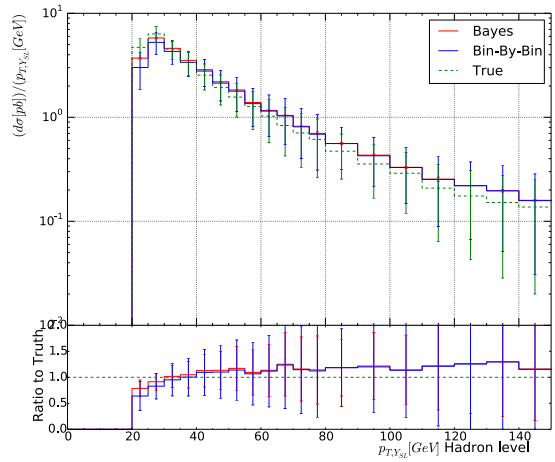
(b) Bayes iterative unfolding for higher numbers of iterations.



(b) Bayes iterative unfolding for higher numbers of iterations.



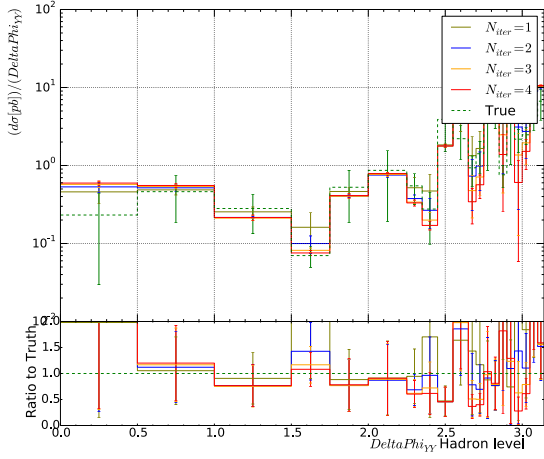
(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.



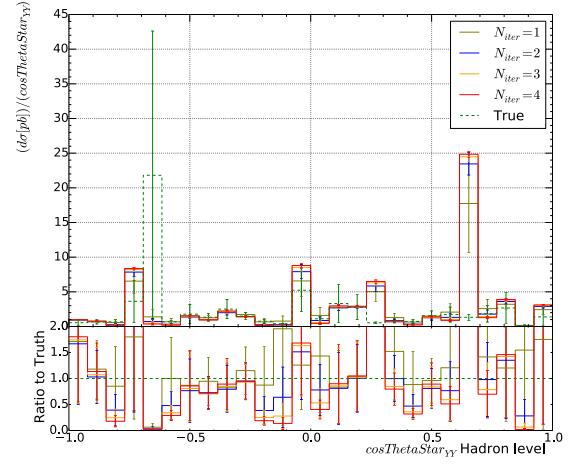
(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.

Figure A.8: Result of unfolding fake unweighted parton level p_{T,γ_L} distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

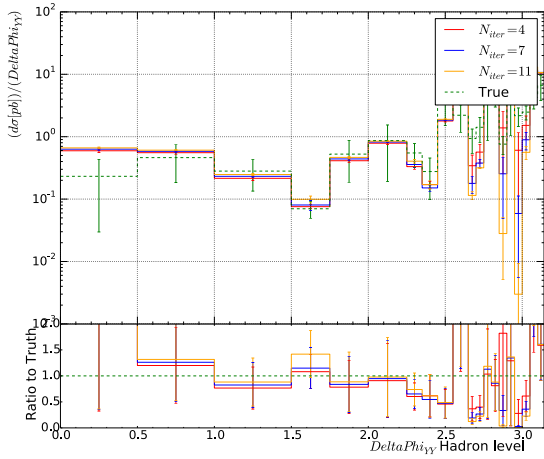
Figure A.9: Result of unfolding fake unweighted parton level $p_{T,\gamma_{SL}}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].



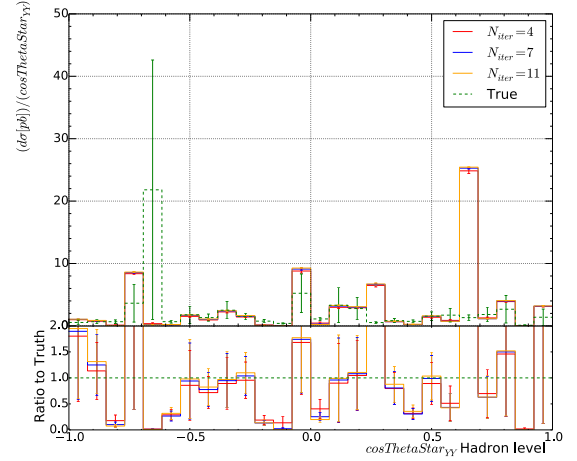
(a) Bayes iterative unfolding for low numbers of iterations.



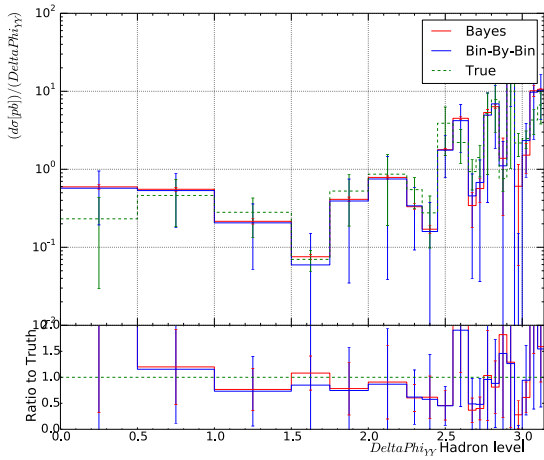
(a) Bayes iterative unfolding for low numbers of iterations.



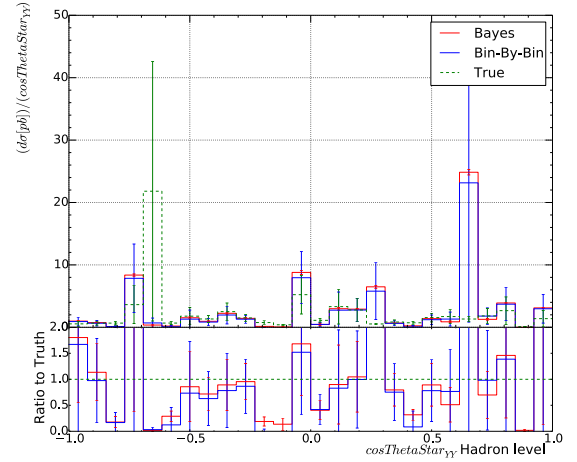
(b) Bayes iterative unfolding for higher numbers of iterations.



(b) Bayes iterative unfolding for higher numbers of iterations.



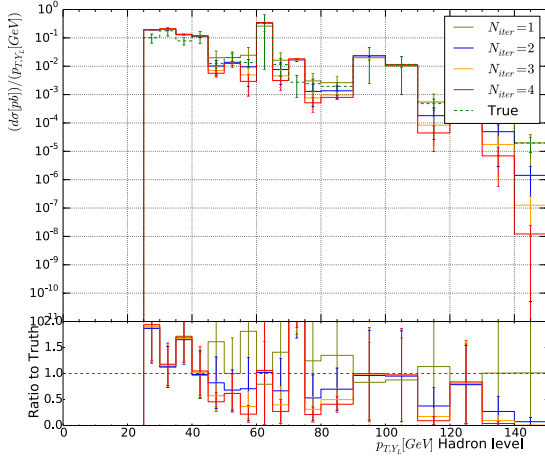
(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.



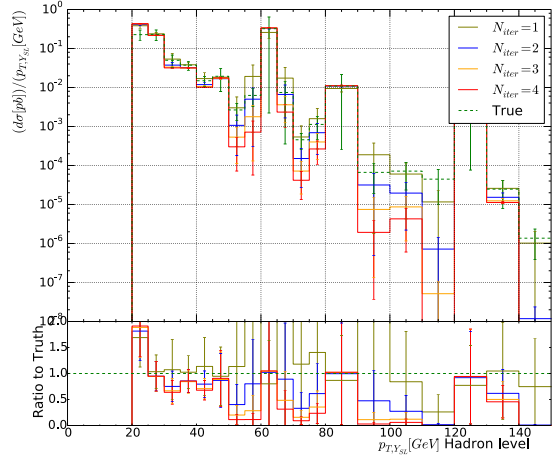
(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.

Figure A.10: Result of unfolding fake random weight parton level $\Delta\phi_{\gamma\gamma}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

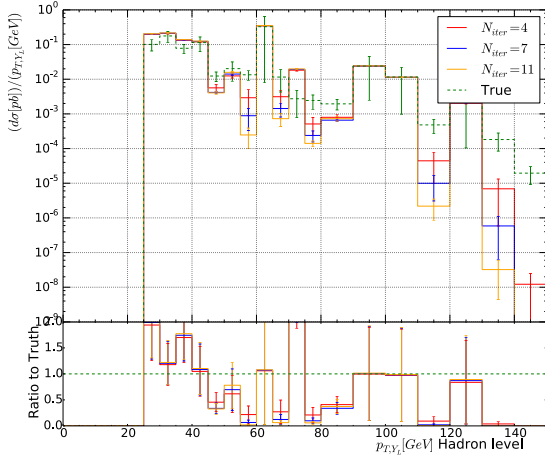
Figure A.11: Result of unfolding fake random weight parton level $\cos\theta_{\gamma\gamma}^*$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].



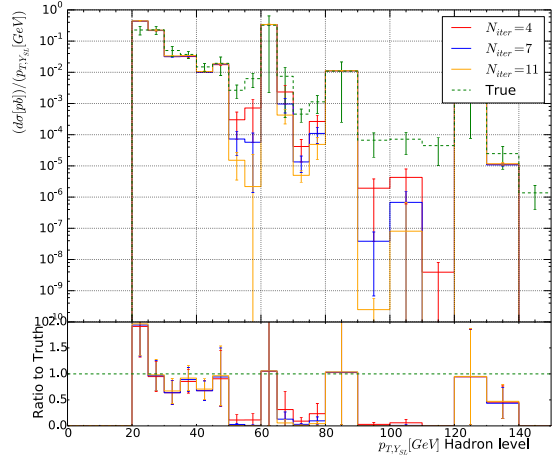
(a) Bayes iterative unfolding for low numbers of iterations.



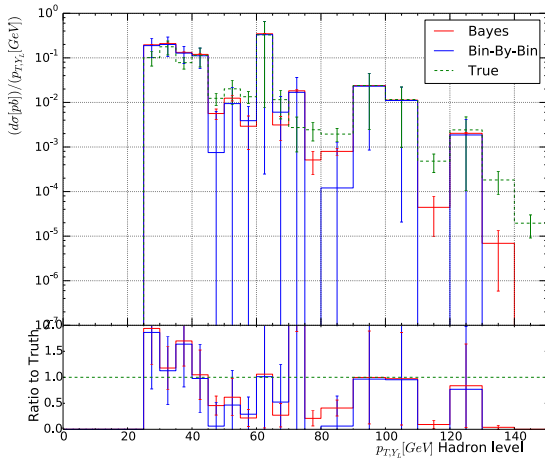
(a) Bayes iterative unfolding for low numbers of iterations.



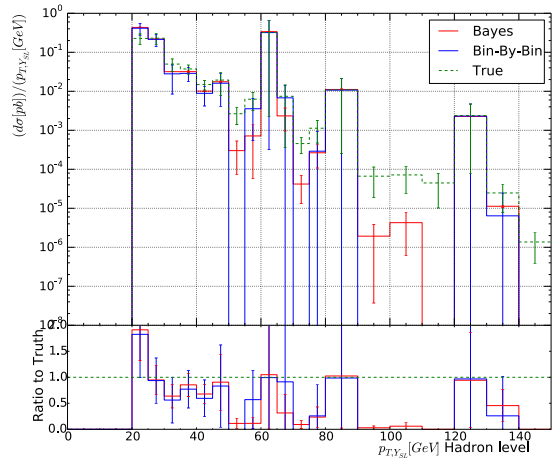
(b) Bayes iterative unfolding for higher numbers of iterations.



(b) Bayes iterative unfolding for higher numbers of iterations.



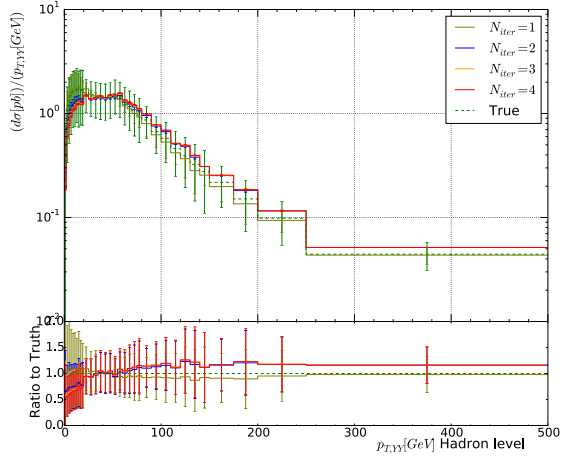
(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.



(c) Bayes iterative unfolding (4 iterations) vs. Bin-by-bin unfolding.

Figure A.12: Result of unfolding fake random weight parton level $p_{T\gamma_L}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].

Figure A.13: Result of unfolding fake random weight parton level $p_{T\gamma_{SL}}$ distribution to hadron level, compared with known true hadron level. Cut parameters as in [2].



(a) Original cut parameters.

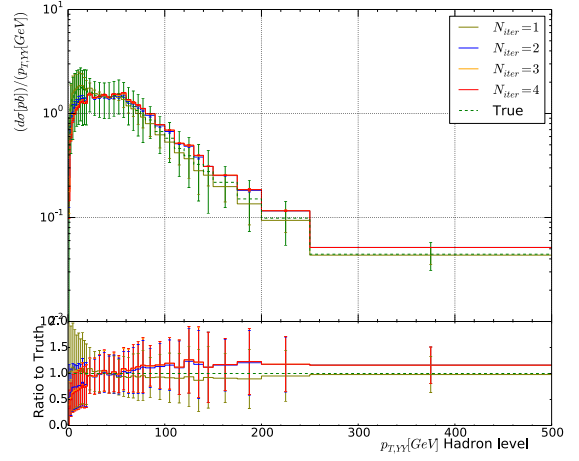
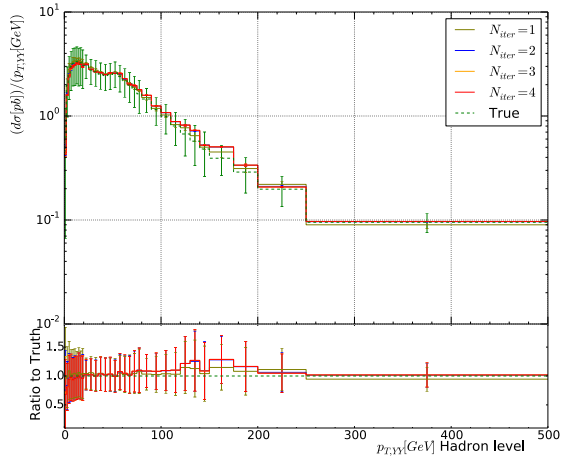
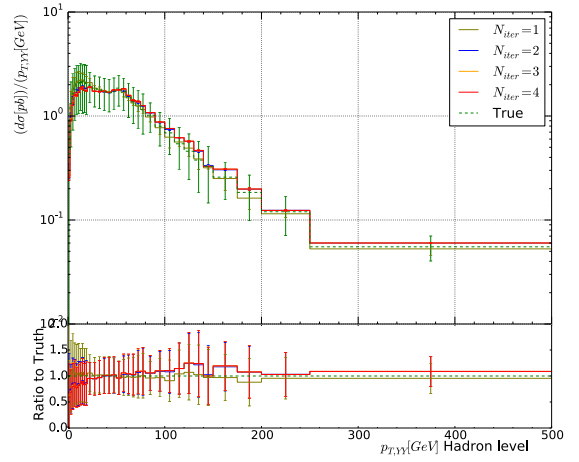
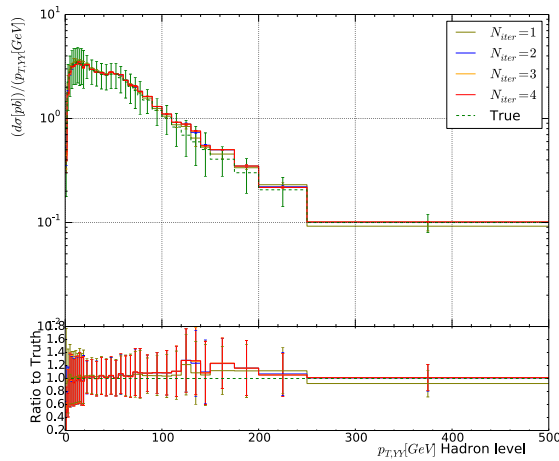
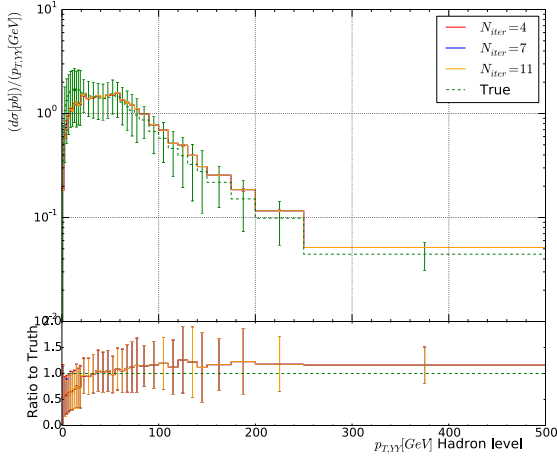
(b) Varied parameter: $p_{T\gamma} = 15$ GeV.(c) Varied parameter: $\Delta R_{iso} = 0.1$.(d) Varied parameter: $E_{max,Cone} = 10$ GeV.(e) Varied parameters: $p_{T\gamma} = 15$ GeV,
 $\Delta R_{iso} = 0.1$, $E_{max,Cone} = 10$ GeV.

Figure A.14: Results of unfolding fake unweighted parton level $p_{T\gamma\gamma}$ distribution to hadron level with Bayes iterative unfolding for low numbers of iterations. Compared with known true hadron level. Cut parameters are varied compared to original from [2]: $p_{T\gamma} = 25$ GeV, $\Delta R_{iso} = 0.4$ and $E_{max,Cone} = 4$ GeV.



(a) Original cut parameters.

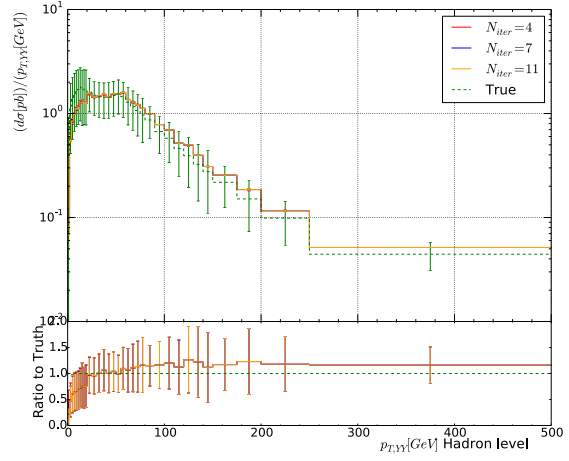
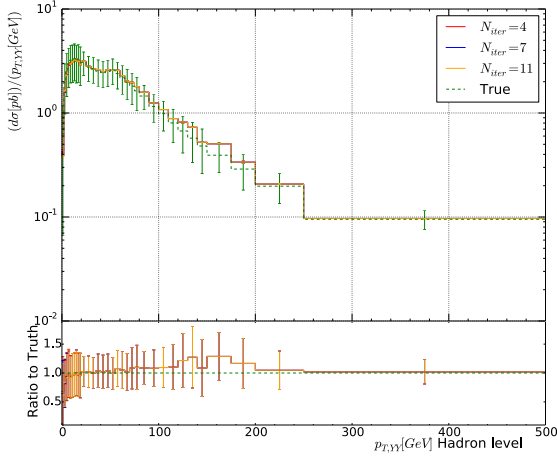
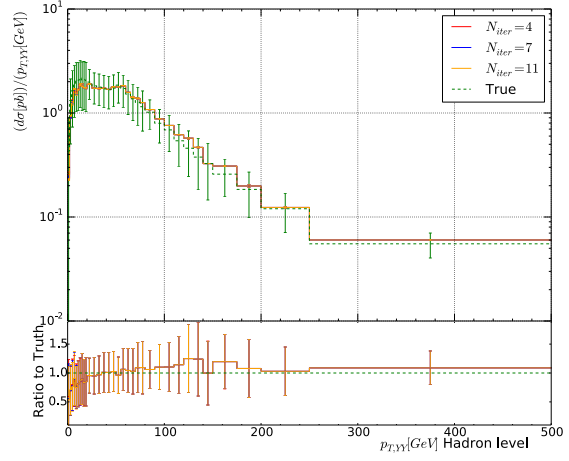
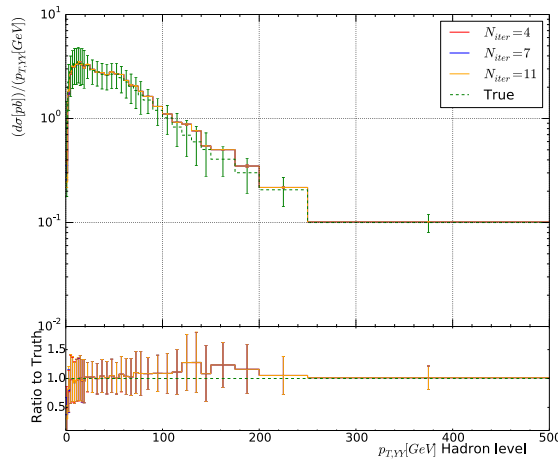
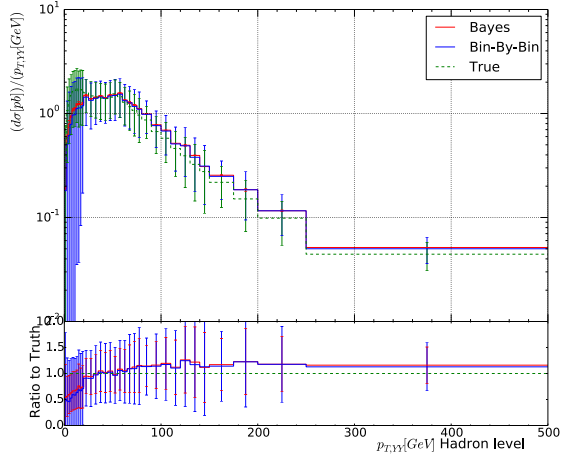
(b) Varied parameter: $p_{T\gamma} = 15$ GeV.(c) Varied parameter: $\Delta R_{\text{iso}} = 0.1$.(d) Varied parameter: $E_{\text{max,Cone}} = 10$ GeV.(e) Varied parameters: $p_{T\gamma} = 15$ GeV,
 $\Delta R_{\text{iso}} = 0.1$, $E_{\text{max,Cone}} = 10$ GeV.

Figure A.15: Results of unfolding fake unweighted parton level $p_{T\gamma\gamma}$ distribution to hadron level with Bayes iterative unfolding for higher numbers of iterations. Compared with known true hadron level. Cut parameters are varied compared to original from [2]: $p_{T\gamma} = 25$ GeV, $\Delta R_{\text{iso}} = 0.4$ and $E_{\text{max,Cone}} = 4$ GeV.



(a) Original cut parameters.

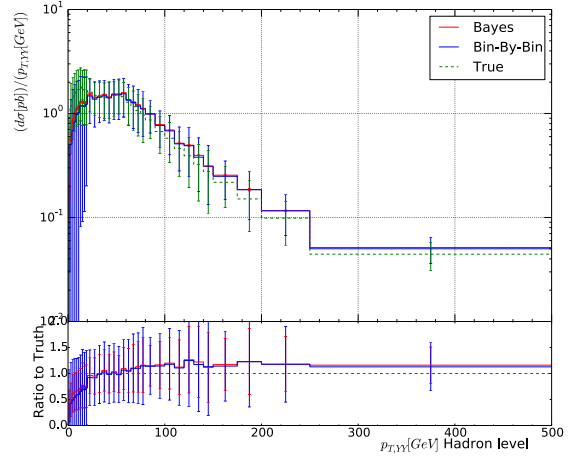
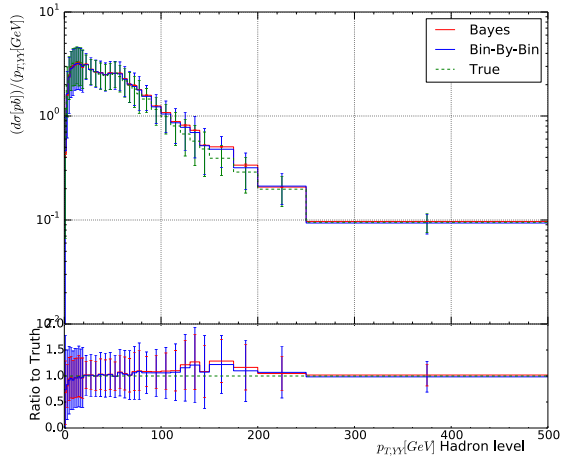
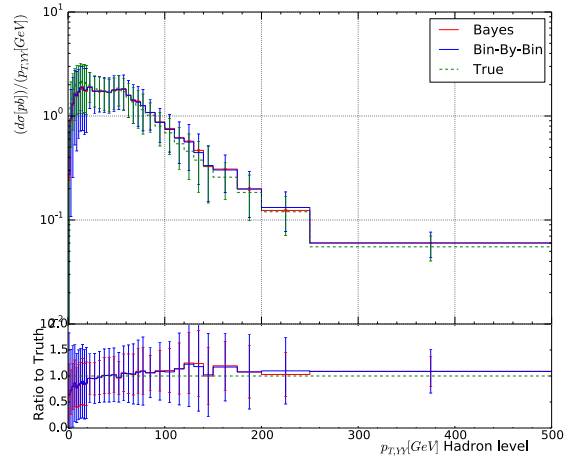
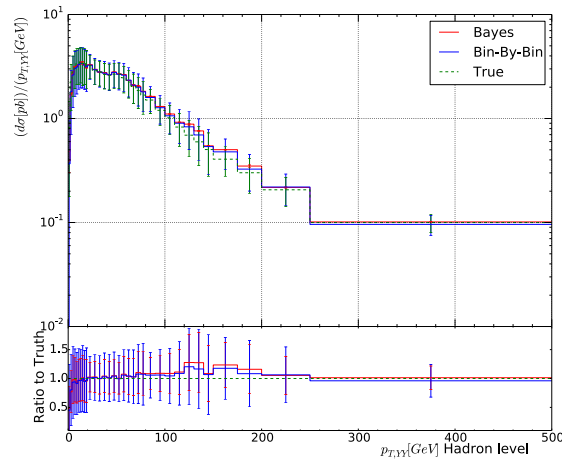
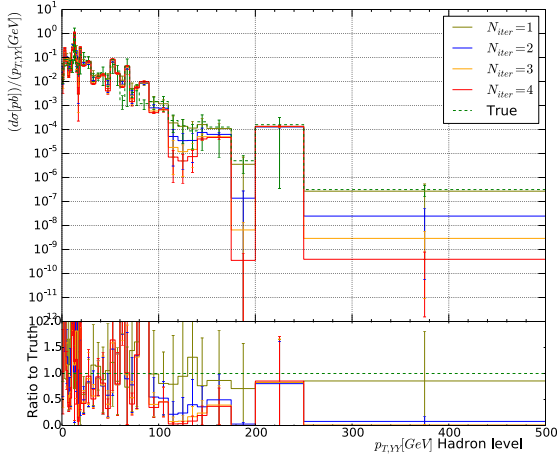
(b) Varied parameter: $p_{T\gamma} = 15$ GeV.(c) Varied parameter: $\Delta R_{\text{iso}} = 0.1$.(d) Varied parameter: $E_{\text{max,Cone}} = 10$ GeV.(e) Varied parameters: $p_{T\gamma} = 15$ GeV,
 $\Delta R_{\text{iso}} = 0.1$, $E_{\text{max,Cone}} = 10$ GeV.

Figure A.16: Results of unfolding fake unweighted parton level $p_{T\gamma\gamma}$ distribution to hadron level with Bayes iterative unfolding for four iterations and with Bin-by-bin unfolding. Compared with known true hadron level. Cut parameters are varied compared to original from [2]: $p_{T\gamma} = 25$ GeV, $\Delta R_{\text{iso}} = 0.4$ and $E_{\text{max,Cone}} = 4$ GeV.



(a) Original cut parameters.

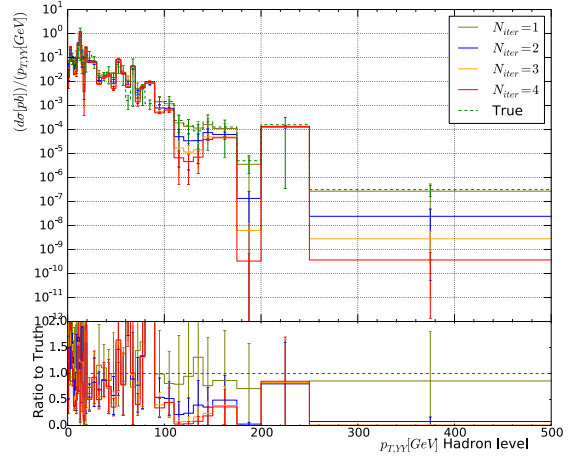
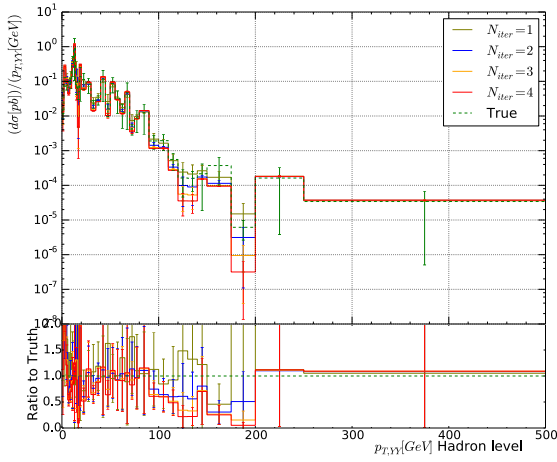
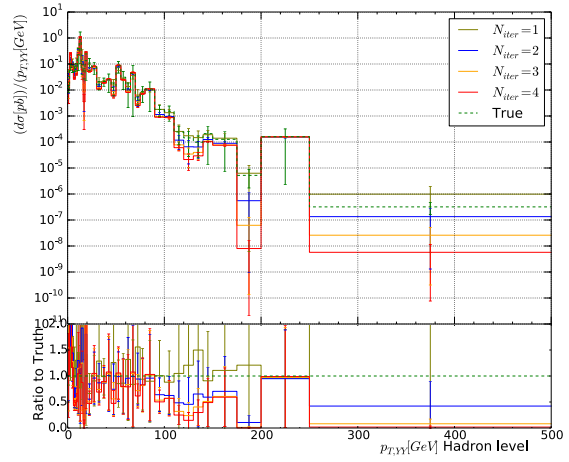
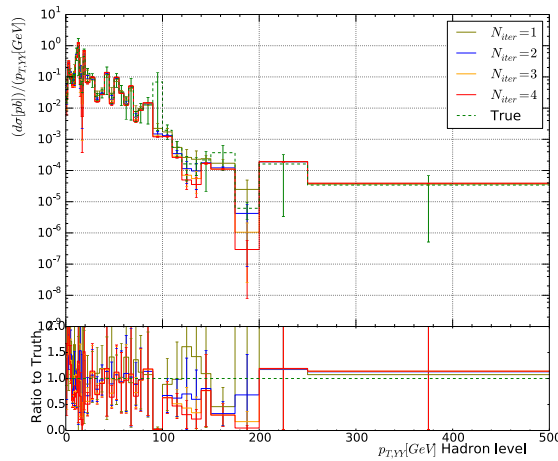
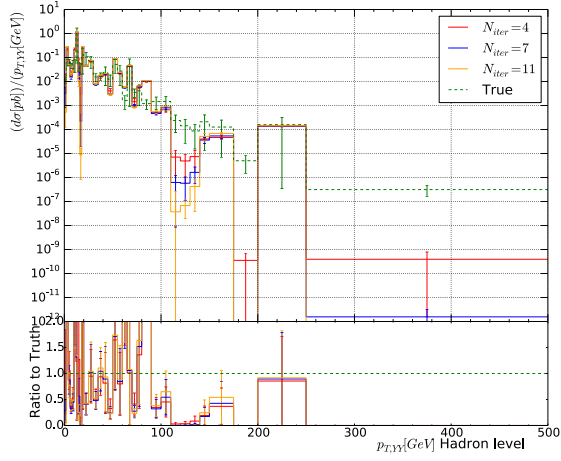
(b) Varied parameter: $p_{T\gamma} = 15$ GeV.(c) Varied parameter: $\Delta R_{iso} = 0.1$.(d) Varied parameter: $E_{max,Cone} = 10$ GeV.(e) Varied parameters: $p_{T\gamma} = 15$ GeV,
 $\Delta R_{iso} = 0.1$, $E_{max,Cone} = 10$ GeV.

Figure A.17: Results of unfolding fake random weight parton level $p_{T\gamma\gamma}$ distribution to hadron level with Bayes iterative unfolding for low numbers of iterations. Compared with known true hadron level. Cut parameters are varied compared to original from [2]: $p_{T\gamma} = 25$ GeV, $\Delta R_{iso} = 0.4$ and $E_{max,Cone} = 4$ GeV.



(a) Original cut parameters.

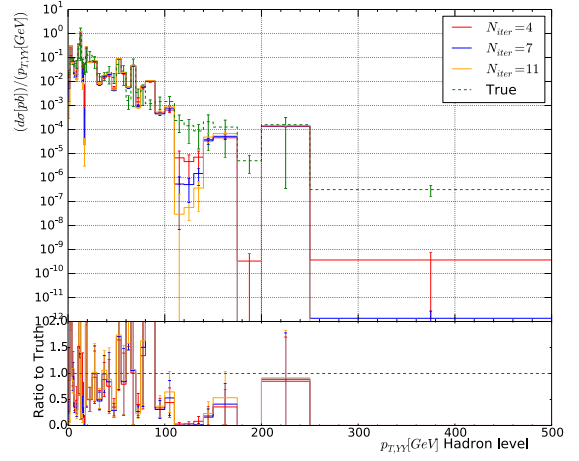
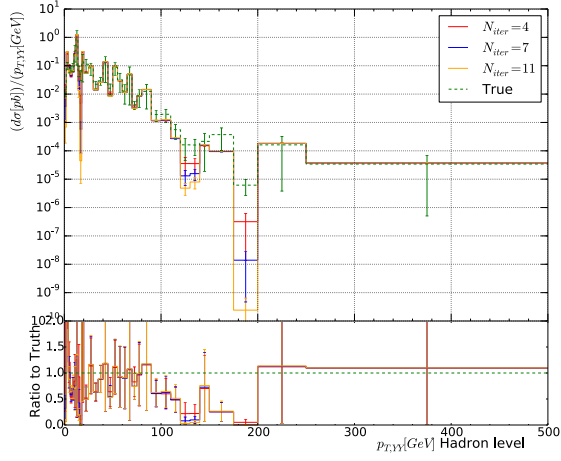
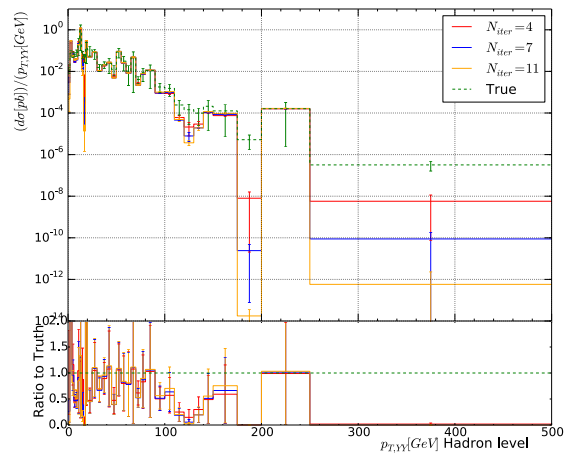
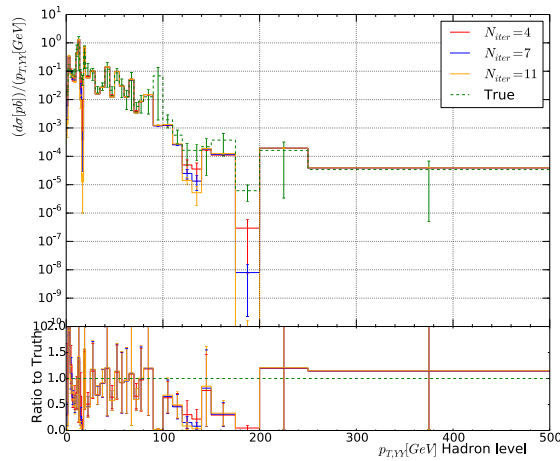
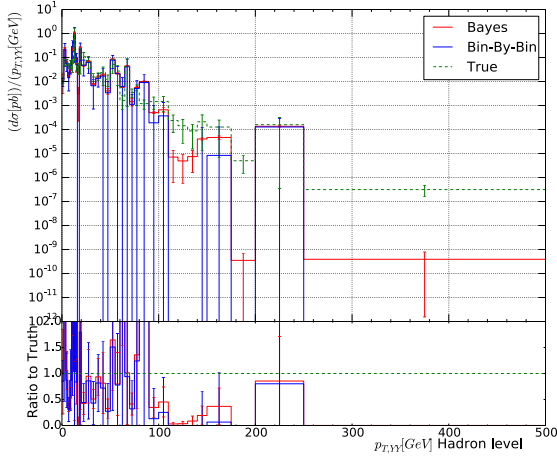
(b) Varied parameter: $p_{T\gamma} = 15$ GeV.(c) Varied parameter: $\Delta R_{iso} = 0.1$.(d) Varied parameter: $E_{max,Cone} = 10$ GeV.(e) Varied parameters: $p_{T\gamma} = 15$ GeV, $\Delta R_{iso} = 0.1$, $E_{max,Cone} = 10$ GeV.

Figure A.18: Results of unfolding fake random weight parton level $p_{T\gamma\gamma}$ distribution to hadron level with Bayes iterative unfolding for higher numbers of iterations. Compared with known true hadron level. Cut parameters are varied compared to original from [2]: $p_{T\gamma} = 25$ GeV, $\Delta R_{iso} = 0.4$ and $E_{max,Cone} = 4$ GeV.



(a) Original cut parameters.

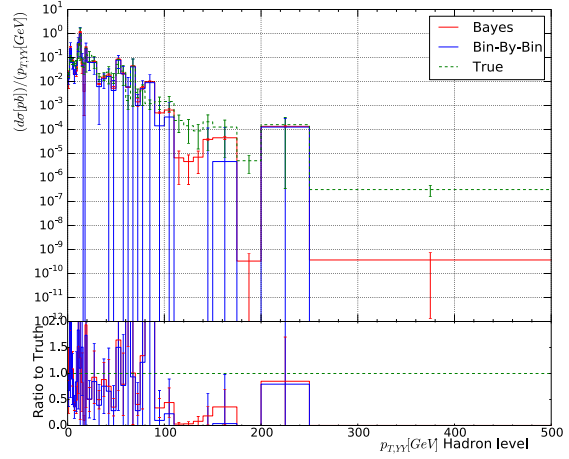
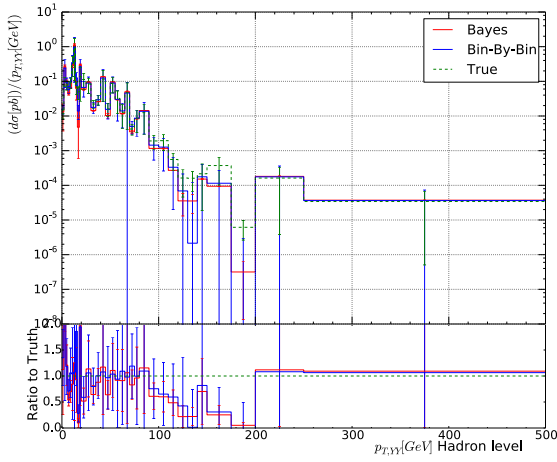
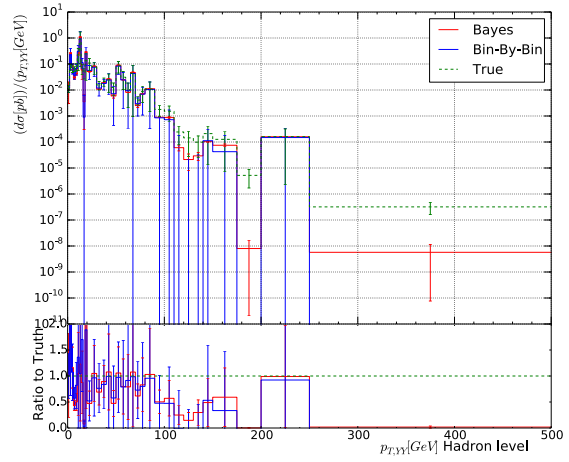
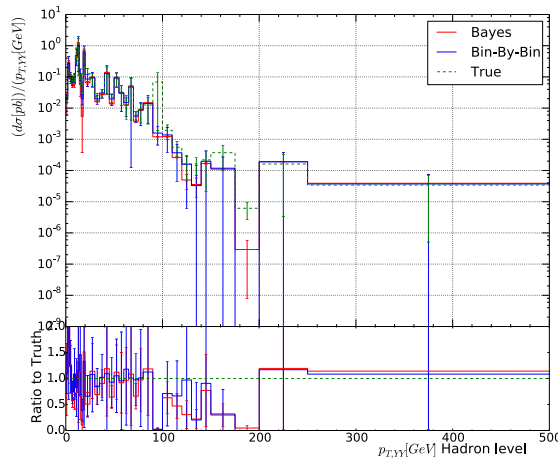
(b) Varied parameter: $p_{T\gamma} = 15$ GeV.(c) Varied parameter: $\Delta R_{\text{iso}} = 0.1$.(d) Varied parameter: $E_{\text{max,Cone}} = 10$ GeV.(e) Varied parameters: $p_{T\gamma} = 15$ GeV,
 $\Delta R_{\text{iso}} = 0.1$, $E_{\text{max,Cone}} = 10$ GeV.

Figure A.19: Results of unfolding fake random weight parton level $p_{T\gamma\gamma}$ distribution to hadron level with Bayes iterative unfolding for four iterations and with Bin-by-bin unfolding. Compared with known true hadron level. Cut parameters are varied compared to original from [2]: $p_{T\gamma} = 25$ GeV, $\Delta R_{\text{iso}} = 0.4$ and $E_{\text{max,Cone}} = 4$ GeV.

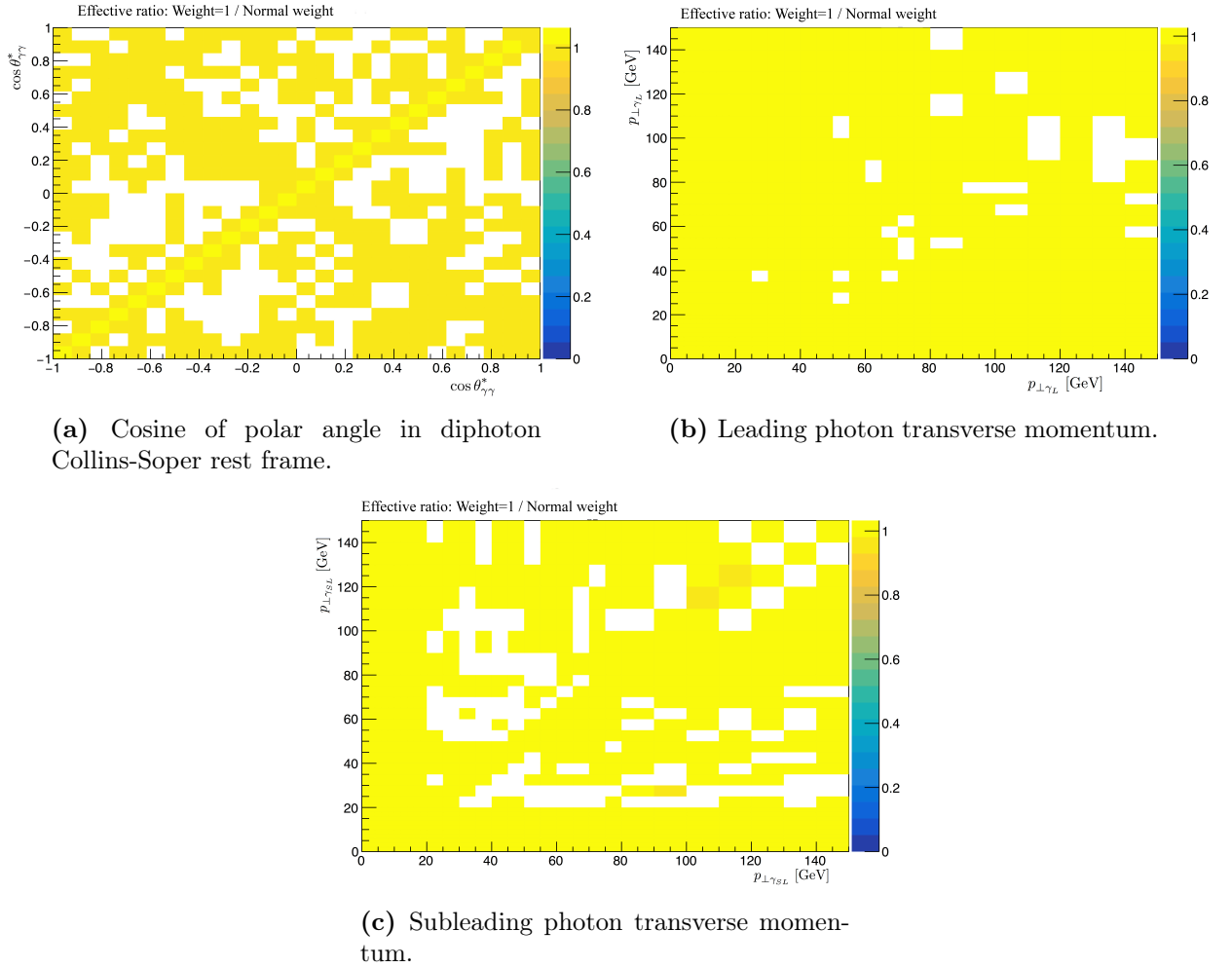
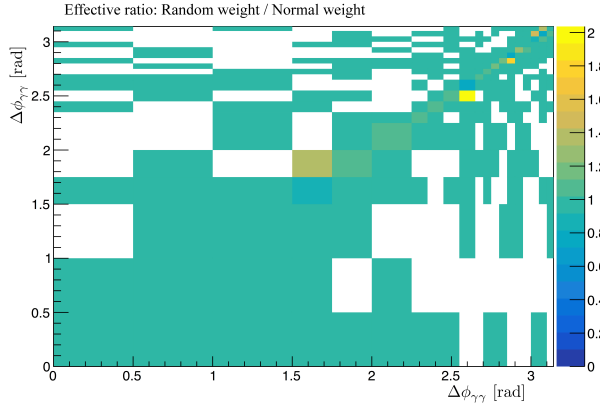
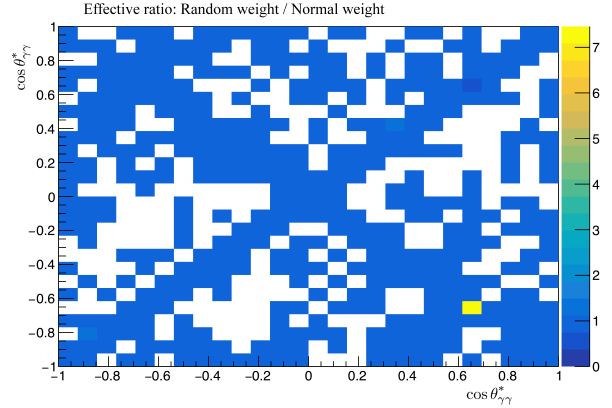


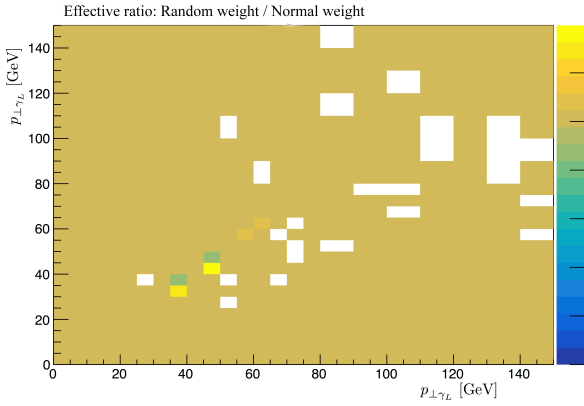
Figure A.20: Effective ratio (definition: section 4.3.3) of fake unweighted diphoton sample response matrices to those from correctly weighted diphoton simulation. Cut parameter are as in [2].



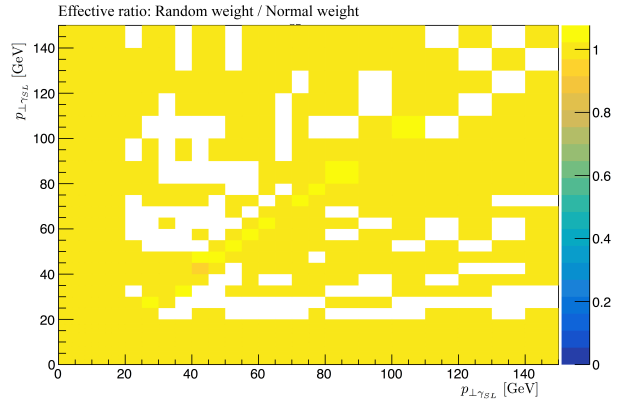
(a) Diphoton angular separation.



(b) Cosine of polar angle in diphoton Collins-Soper rest frame.



(c) Leading photon transverse momentum.

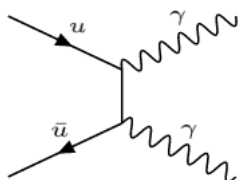
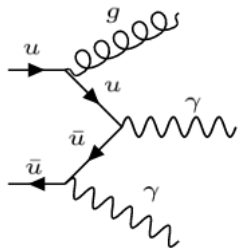
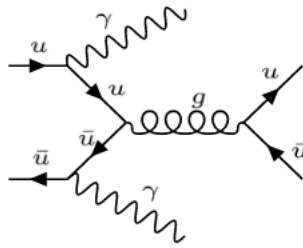
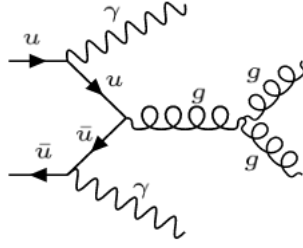
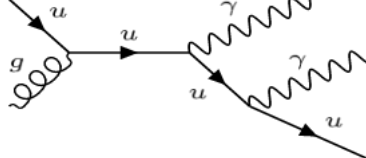


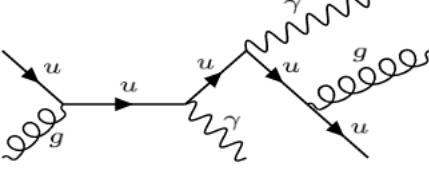
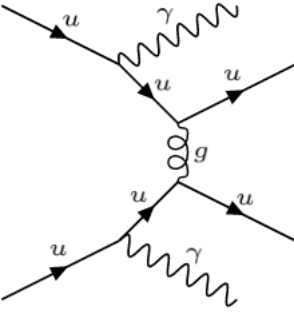
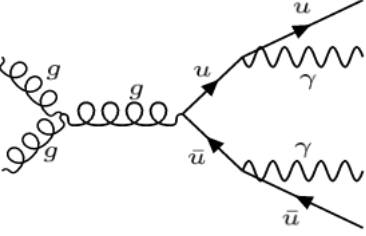
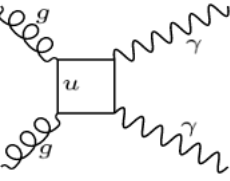
(d) Subleading photon transverse momentum.

Figure A.21: Effective ratio (definition: section 4.3.3) of fake random weight diphoton sample response matrices to those from correctly weighted diphoton simulation. Cut parameter are as in [2].

A.2 Tables

Table A.1: Feynman diagrams for diphoton production processes at pp colliders ($pp \rightarrow \gamma\gamma + X$) at leading order in α_S of the process. Here, X is allowed to be composed of up to two partons. Processes that are the same as the given with all antiquarks switched to quarks and all quarks switched to antiquarks do not yield new diagram structures and are left out.

Process	Example for lowest order Feynman diagram
$q\bar{q} \rightarrow \gamma\gamma$	
$q\bar{q} \rightarrow \gamma\gamma + g$	
$q\bar{q} \rightarrow \gamma\gamma + q\bar{q}$	
$q\bar{q} \rightarrow \gamma\gamma + gg$	
$qg \rightarrow \gamma\gamma + q$	

$qg \rightarrow \gamma\gamma + qg$	
$qq \rightarrow \gamma\gamma + qq$	
$gg \rightarrow \gamma\gamma + q\bar{q}$	
$gg \rightarrow \gamma\gamma$	

A.3 Code

Listing A.1: Settings file ('Run card') for the 1 million events parton level diphoton Sherpa simulation.

```
# General simulation settings and tags
(run){
  MULTIPROPAR:=3234;      # Tag to allow multiprocessing -> standard value
  RANDOM_SEED=MULTIPROPAR; # Multiprocessing varies this tag -> different seeds

  # For correlated results -> Data output into hepmc file with seed tag
  EVENT_OUTPUT=HepMC_Short[CorrelatedParton_MULTIPROPAR];

  ISOLATIONDR:=0.05;      # Photon isolation cone radius for selector
  ISOLATIONEXP:=2;        # Exponent in dynamic photon isolation cut
  ISOLATIONEPS:=0.10;     # Epsilon parameter in dynamic photon isolation cut
  DELTARYY:=0.2;          # Minimal angular separation between main photons
  PTYTAG:=10;             # Minimal transverse momentum of main photons
```

```

EVENTS 1M;                # One million events

ALPHAQED_DEFAULT_SCALE=0.0; # QED coupling set on photon mass scale
MASS[6]=172.5;             # Fixing masses and decay widths of heavy particles
WIDTH[6]=1.5;              # Number in [] brackets = PDG ID
MASS[15]=1.777;
WIDTH[15]=2.26735e-12;
MASS[23]=91.1876;
WIDTH[23]=2.4952;
MASS[24]=80.399;
WIDTH[24]=2.085;

# Tags for process setup
NJET:=2;    # Maximum number of additional jets in hard process
QCUT:=10;   # Tag for transition from matrix element to parton shower

# Matrix element generator settings
ME_SIGNAL_GENERATOR Comix Amegic LOOPGEN Internal;
EVENT_GENERATION_MODE=P;    # For smaller files
EVENT_SEED_MODE=3;         # For internal bookkeeping of random seed
LOOPGEN:=OpenLoops;        # Loop generator

# collider setup
BEAM_1 2212; BEAM_ENERGY_1 = 3500.; # Two proton beams with total sqrt(s)=7TeV
BEAM_2 2212; BEAM_ENERGY_2 = 3500.; # -> 7TeV LHC run

# For Simulations on Parton-Level
FRAGMENTATION=Off; # Deactivates hadronization
MI_HANDLER=None;   # Deactivates Multiparton-Interactions
K_PERP_MEAN_1=0.0; # Fix transverse momenta of partons in protons to zero
K_PERP_MEAN_2=0.0;
K_PERP_SIGMA_1=0.0; # Fixes width of gauss distribution of pt of
K_PERP_SIGMA_2=0.0; # partons on in protons to zero
}(run)

# Allowed processes
(processes){
  Process 21 21 -> 22 22;    # gg -> PP
  ME_Generator Internal;    # Internal matrix element generator
  Loop_Generator gg_yy;     # Specify loop generator
  Scales VAR{Abs2(p[2]+p[3])}; # Scale to invariant diphoton mass
  End process;

  Process 93 93 -> 22 22 93{NJET}; # pp -> PP + up to NJET partons
  Order (*,2);                  # No fixed number of QCD vertices, only 2 QED
  # Remove double counting for jet emissions between ME and parton shower
  CKKW sqr(QCUT/E_CMS)/(1.0+sqr(QCUT/0.6)/Abs2(p[2]+p[3]));
  End process;
}(processes)

# Cuts that specify what processes should be generated
(selector){

```

```
PT 22 PTYTAG E_CMS;          # PT cut on hard process photons
# Dynamic isolation cut on hard process photons -> section Diphoton-events
IsolationCut 22 ISOLATIONDR ISOLATIONEXP ISOLATIONEPS;
DeltaR 22 22 DELTARYY 1000.0; # Minimal angular distance between these photons
}(selector)
```

Bibliography

- [1] G. Aad et al. “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”. In: *Phys. Rev. Lett.* 114 (19 2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.114.191803>.
- [2] G. Aad et al. “Measurement of isolated-photon pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”. In: *Journal of High Energy Physics* 2013.1 (2013), pp. 1–42. ISSN: 1029-8479. DOI: 10.1007/JHEP01(2013)086. URL: [http://dx.doi.org/10.1007/JHEP01\(2013\)086](http://dx.doi.org/10.1007/JHEP01(2013)086).
- [3] G. Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: <http://dx.doi.org/10.1016/j.physletb.2012.08.020>. URL: <http://www.sciencedirect.com/science/article/pii/S037026931200857X>.
- [4] V.M. Abazov et al. “Measurement of the differential cross sections for isolated direct photon pair production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”. In: *Physics Letters B* 725.1–3 (2013), pp. 6–14. ISSN: 0370-2693. DOI: <http://dx.doi.org/10.1016/j.physletb.2013.06.036>. URL: <http://www.sciencedirect.com/science/article/pii/S037026931300511X>.
- [5] T. Adye. “Unfolding algorithms and tests using RooUnfold”. In: *ArXiv e-prints* (May 2011). arXiv: 1105.1160 [physics.data-an].
- [6] T. Binoth, J.Ph. Guillet, E. Pilon, and M. Werlen. “A full next-to-leading order study of direct photon pair production in hadronic collisions”. In: *The European Physical Journal C - Particles and Fields* 16.2 (2000), pp. 311–330. ISSN: 1434-6052. DOI: 10.1007/s100520050024. URL: <http://dx.doi.org/10.1007/s100520050024>.
- [7] A. Buckley, J. Butterworth, D. Grellscheid, H. Hoeth, L. Lonnblad, J. Monk, H. Schulz, and F. Siegert. “Rivet user manual”. In: *ArXiv e-prints* (Mar. 2010). arXiv: 1003.0694 [hep-ph].
- [8] S. Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: <http://dx.doi.org/10.1016/j.physletb.2012.08.021>. URL: <http://www.sciencedirect.com/science/article/pii/S0370269312008581>.
- [9] CMS Collaboration. “Search for new physics in high mass diphoton events in proton-proton collisions at 13TeV”. In: (2015).
- [10] The ATLAS collaboration. *Search for resonances decaying to photon pairs in 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2015. URL: <http://cds.cern.ch/record/2114853>.
- [11] G. D’Agostini. “A multidimensional unfolding method based on Bayes’ theorem”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2–3 (Aug. 1995), pp. 487–498. DOI: [http://dx.doi.org/10.1016/0168-9002\(95\)00274-X](http://dx.doi.org/10.1016/0168-9002(95)00274-X). URL: <http://www.sciencedirect.com/science/article/pii/016890029500274X>.

- [12] Stefano Frixione. “Isolated photons in perturbative QCD”. In: *Physics Letters B* 429.3–4 (June 1998), pp. 369–374. DOI: [http://dx.doi.org/10.1016/S0370-2693\(98\)00454-7](http://dx.doi.org/10.1016/S0370-2693(98)00454-7). URL: <http://www.sciencedirect.com/science/article/pii/S0370269398004547>.
- [13] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter. “Event generation with SHERPA 1.1”. In: *Journal of High Energy Physics* 2009.02 (2009), p. 007. URL: <http://stacks.iop.org/1126-6708/2009/i=02/a=007>.
- [14] David J. Gross and Frank Wilczek. “Ultraviolet Behavior of Non-Abelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [15] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [16] K.A. Olive and Particle Data Group. “Review of Particle Physics”. In: *Chinese Physics C* 38.9 (2014), p. 090001. URL: <http://stacks.iop.org/1674-1137/38/i=9/a=090001>.
- [17] BRYAN R. WEBBER. “FRAGMENTATION AND HADRONIZATION”. In: *International Journal of Modern Physics A* 15.suppl01b (2000), pp. 577–606. URL: <http://www.worldscientific.com/doi/abs/10.1142/S0217751X00005334>.
- [18] Carl Schmidt Zvi Bern Lance Dixon. “Di-photon production at the LHC”. In: (2003). URL: [arXiv:hep-ph/0306130](http://arxiv.org/abs/hep-ph/0306130).

Acknowledgements

I would like to thank Dr. Frank Siegert for the interesting theme, the insight into the workings of event generators and the helpful discussions.

I would like to express my gratitude towards Johannes Krause for answering whatever kind of question came up during the long debugging sessions and for helping in giving this thesis a proper form. Furthermore, I would like to thank Dr. Christian Gütschow for providing quick and efficient solutions where I got stuck and improving the readability of this thesis.

Finally, I want to thank my co-workers and fellow students, especially Paul Moder and David Tucholski, for all the second opinions, cookies and foosball matches.

Erklärung

Hiermit erkläre ich, dass ich diese Arbeit im Rahmen der Betreuung am Institut für Kern- und Teilchenphysik ohne unzulässige Hilfe Dritter verfasst und alle Quellen als solche gekennzeichnet habe.

Jakob Beyer
Dresden, Juni 2016