SPARKLING BEAUTY

A measurement of the $b\bar{b}\gamma\gamma$ cross-section at 8 TeV with ATLAS

WOUTER VAN DEN WOLLENBERG

Sparkling Beauty

A measurement of the $b\bar{b}\gamma\gamma$ cross-section at 8 TeV with ATLAS

Wouter van den Wollenberg

Graduation committee:

Prof.dr.ir. B. van Eijk	SUPERVISOR	*University of Twente*
Dr. M. Sławińska	CO-SUPERVISOR	*Polish Academy of Sciences*
Prof.dr.ir. J.W.M. Hilgenkamp	CHAIRMAN/SECRETARY	*University of Twente*
Prof.dr. A.P. Colijn	MEMBER	*University of Utrecht*
Prof.dr. R.H.P. Kleiss	MEMBER	*Radboud University Nijmegen*
Prof.dr. R.J.M. Snellings	MEMBER	*University of Utrecht*
Prof.dr. W. Verkerke	MEMBER	*University of Amsterdam*
Prof.dr.ir. H.J.W. Zandvliet	MEMBER	*University of Twente*
Dr. M.M.J. Dhalle	MEMBER	*University of Twente*
Dr. F. Filthaut	MEMBER	*Radboud University Nijmegen*
Dr. M. Kagan	MEMBER	*Stanford Linear Accelerator Center*

The research presented in this thesis was carried out at the *Nationaal Instituut voor Subatomaire Fysica (Nikhef)* in Amsterdam and at the *European Organization for Nuclear Research (CERN)* near Geneva and was part of the research program of the *Stichting voor Fundamenteel Onderzoek der Materie (FOM)*, which is financially supported by the *Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO)*.

The author was financially supported by the *University of Twente* and *FOM*.

SPARKLING BEAUTY
ISBN: 978-94-6233-552-3

Printed in the Netherlands by Gildeprint.

SPARKLING BEAUTY

A MEASUREMENT OF THE $b\bar{b}\gamma\gamma$ CROSS-SECTION AT 8 TeV WITH ATLAS

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr. T.T.M. Palstra,
on account of the decision of the graduation committee
to be publicly defended
on Wednesday the 29th of March 2017 at 16:45

by

WOUTER VAN DEN WOLLENBERG

born on July 9th, 1987
in Nijmegen, the Netherlands

This dissertation is approved by:

Supervisor: Prof.dr.ir. B. van Eijk

Co-supervisor: Dr. M. Sławińska

Voor mijn moeder en ook voor een heks van een andere planeet,

met dank, bewondering, en liefde...

"Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilization, every king and peasant, every young couple in love, every mother and father, hopeful child, inventor and explorer, every teacher of morals, every corrupt politician, every "superstar", every "supreme leader", every saint and sinner in the history of our species lived there. On a mote of dust suspended in a sunbeam.

The Earth is a very small stage in a vast cosmic arena. Think of the rivers of blood spilled by all those generals and emperors so that, in glory and triumph, they could become the momentary masters of a fraction of a dot. Think of the endless cruelties visited by the inhabitants of one corner of this pixel on the scarcely distinguishable inhabitants of some other corner, how frequent their misunderstandings, how eager they are to kill one another, how fervent their hatreds.

Our posturings, our imagined self-importance, the delusion that we have some privileged position in the Universe, are challenged by this point of pale light. Our planet is a lonely speck in the great enveloping cosmic dark. In our obscurity, in all this vastness, there is no hint that help will come from elsewhere to save us from ourselves.

The Earth is the only world known so far to harbor life. There is nowhere else, at least in the near future, to which our species could migrate. Visit, yes. Settle, not yet. Like it or not, for the moment the Earth is where we make our stand.

It has been said that astronomy is a humbling and character-building experience. There is perhaps no better demonstration of the folly of human conceits than this distant image of our tiny world. To me, it underscores our responsibility to deal more kindly with one another, and to preserve and cherish the pale blue dot, the only home we've ever known."

Carl Sagan, Pale Blue Dot: A Vision of the Human Future in Space

Contents

CHAPTER 1

Introduction

"Only entropy comes easy."

- Anton Pavlovich Chekhov

OUR Universe was just a tiny fledgling coming to terms with its own existence when the electroweak symmetry of the Laws of Nature was spontaneously broken. This ushered in a phase transition transforming the vacuum into one where most particles have a mass. Such is the Universe in its first few moments after the Big Bang according to the Higgs mechanism. This is the principal explanation for the mass of elementary particles in modern particle physics. It is an explanation that for years was shrouded in uncertainty until the discovery of the Higgs boson in 2012. Now we are quite certain the Higgs field exists, but it remains yet unclear whether it is capable of performing its fated duty, generating mass.

To illuminate this we need to measure a certain aspect of the Higgs field, its potential, and more specifically the illusive trilinear coupling, λ_{3h}. This coupling governs how strongly the quanta of the Higgs field, the Higgs bosons, interact with themselves. In this thesis we attempted to investigate this λ_{3h}, unfortunately we shall see this was a bit too optimistic as not enough data has been collected by the ATLAS detector yet to perform such a measurement. We then decided to move on to study a background that such a future measurement will surely encounter. This background has not yet been examined by ATLAS, except as a discovery channel for possible double Higgs production.

Our channel of choice, our background of choice we could say, is $pp \to b\bar{b}\gamma\gamma$, or the prompt creation of two b-quarks and two photons. Its interest from a Higgs perspective lies in the fact that $h \to b\bar{b}$ has a large branching ratio, and $h \to \gamma\gamma$ is very clean. Thus, one might be able to glimpse a $h \to hh$ splitting, governed by λ_{3h}, in this channel. And the value of that parameter should tell us much about the Higgs potential and electroweak symmetry breaking in the early Universe. In this work however, we are interested in $pp \to b\bar{b}\gamma\gamma$ in its entirety, the only real background to our analysis are physics objects

that are wrongly reconstructed or detected by the ATLAS software and detector[†]. We thus set out to measure the inclusive $pp \to b\bar{b}\gamma\gamma$ production cross-section at $\sqrt{8}$ TeV. Technically this will include the $h \to hh$ process, but its contribution to the overall cross-section is predicted to be far too tiny to play any role.

We will start the thesis with chapter 2 which gives a theoretical overview of the Standard Model and the Higgs mechanism we want to investigate, after which we will motivate our desire to analyze the Higgs mechanism. We then move on to see how to best measure the Higgs trilinear coupling, and then, having decided upon $pp \to b\bar{b}\gamma\gamma$, we will investigate what the Standard Model tells us about this channel. After this we venture on to a description of the ATLAS detector in chapter 3 and the reconstruction software in chapter 4. We will then prepare a data sample pure in $jj\gamma\gamma$ to base our main analysis on in chapter 5. The actual analysis will happen in chapter 6 before closing with a final conclusion and outlook.

Also included in this work, located in appendix A, is a treatise of half of the author's ATLAS qualification task, which involved calibrating a method for identifying particle cascades that find their origin in b-hadrons. The other half of this qualification work involved software development work in the ATLAS b-tagging group, which, though very interesting to the author and certainly important to ATLAS, makes for dry reading and was therefore left out.

[†]There is also a completely negligible background from cosmic rays. But these are exceedingly unlikely to produce the kind of signal we are examining.

CHAPTER 2

Theory and motivation

"Mathematics is the art of giving the same name to different things."

- Henri Poincaré

THIS chapter will focus on the theoretical aspects of our analysis but will also provide some of the motivation for it. We start with a treatise of the so-called Standard Model of elementary particle physics. This model has been developed over many years, but was molded into its current form around the 1970s. It unifies our understanding of three of the four fundamental forces of nature into a single mathematical framework whose predictions have held up to spectacular scrutiny ever since. The discovery of the Higgs boson in 2012 [1, 2] was a crowning achievement. We will not delve too much into the underlying theories upon which the Standard Model depends, namely special relativity and quantum mechanics. Instead we will briefly discuss the basics before moving on to an in-depth look at the Higgs mechanism. We shall then motivate our desire to investigate the Higgs mechanism. After which we analyze how difficult it is to measure the Higgs potential directly and will note that $pp \to b\bar{b}\gamma\gamma$ seems to be the most promising avenue. We will then close with an examination of what the theory actually predicts for our chosen channel of $pp \to b\bar{b}\gamma\gamma$.

2.1 The Standard Model

THE Standard Model (of elementary particle physics) has been the cornerstone of modern particle physics for over forty years. It is arguably the best-tested model in all of physics, and its predictions are precise and very accurate[†]. The model describes

[†]Note, for example, the impressive accuracy with which the electron magnetic moment [3] and the W and Z boson mass ratio [4] are predicted and experimentally verified.

the physics measured at colliders around the world very well, and was built up piece by piece from various earlier discoveries and models. At the moment it is the best description of elementary particle physics we have, and in this section we will very coarsely list the various particles described by the Standard Model. We will not dip into the mathematical underpinnings, except for those underlying the Higgs mechanism, which are discussed in section 2.2.

The particles contained in the Standard Model can be classified into two groups: matter particles and force carriers. The force carriers mirror three fundamental[†] forces of Nature: electromagnetism, and the weak and strong nuclear forces. In the Standard Model each such force has an associated symmetry group under which the Standard Model's mathematical formulation is invariant. The number of generators of these groups corresponds to the number of force carriers that mediate the force. These force carriers are often called gauge bosons due to the fact that they are all bosons, and because these symmetries are gauge symmetries. The matter particles, on the other hand, are particles that do not interact with each other directly but do so by exchanging force carriers. Another difference is that their properties can not be directly derived from symmetries.

2.1.1 Force carriers

Electromagnetism acts on particles with an electric charge and is mediated by the massless spin-1 photon. The associated symmetry group is $U(1)$. The strong nuclear force acts on particles that have a color charge and is responsible for the stability of an atomic nucleus. It is also important for the formation of mesons and baryons, providing the glue that keeps the quarks, making up these particles, together. The symmetry group, $SU(3)$, is non-Abelian, which translates into the fact that the mediating particles, the massless spin-1 gluons, have a color charge themselves. This means that the force interacts with itself, which is a situation unlike that of electromagnetism.

The strong nuclear force is described by quantum chromodynamics (QCD), which also shows that this force exhibits a unique feature called confinement. Confinement means that the force actually gets stronger over larger distances, like a stretching elastic band. Meaning that when colored particles move apart they will either be pulled back together, or the force will grow to be so powerful that the energy contained in the gluon field between the two particles is enough to create a matter-anti-matter pair. Thus, metaphorically snapping the band into two smaller, less stretched, ones. The result is the inability to experimentally "observe" a free colored particle; such particles will always be contained into a color neutral combination called a hadron. This process is known as hadronization and typically occurs on a timescale of around 10^{-23} s [5]. The sole exception is the top quark which is so unstable that it decays before it can hadronize. Its average lifetime is only 10^{-25} s [6]. Both photons and gluons form their own anti-particles.

The weak nuclear force had $SU(2)_L$ as its symmetry group, the L denoting that it only acts on particles with a left-handed chirality, but this symmetry is no longer realized in Nature. The symmetry has been broken, and in the Standard Model this is described

[†]One could argue about the word 'fundamental' after reading the section on the Higgs mechanism; clearly the electromagnetic and weak nuclear force may not be regarded as fundamental if they merge into the electroweak force at higher energies.

by the Higgs mechanism which is discussed in section 2.2. Due to its broken nature its gauge bosons have varied masses. A relationship with electromagnetism is evidenced by the electric charge of two of them. These are the two W bosons, W^+ and W^-, both with a mass of approximately 80.4 GeV. The third boson is called the Z boson, or Z^0 and has a mass of about 91.2 GeV. As with the other two forces described previously these bosons are spin-1 particles. Like gluons the gauge bosons of the weak nuclear force experience self-interaction because $SU(2)$ is also a non-Abelian group. The weak nuclear force acts on all left-handed matter particles, but there is no real analogue of the electric or colored charges that we had for electromagnetism and the strong nuclear force. This is because the symmetry is broken; there is no conserved charge. The weak nuclear force can change one type of particle into another. For example a proton may turn into a neutron. The weak nuclear force is also involved in β-decay and plays a role in nuclear fusion in the Sun. The Z boson is its own anti-particle, and the W^+ and W^- form each others' anti-particles.

2.1.2 Matter particles

The elementary matter particles have a spin of $\frac{1}{2}$ and are therefore fermions. They come into two different classes: leptons and quarks. Each class comes in three generations with each generation containing two particles. The masses of the particles in each subsequent generation are larger than those in the previous one, with the exception of the neutrinos; these are all taken to be massless. The exact reason for these three generations remains a mystery[†].

The quarks are defined by their ability to feel the effects of the strong nuclear force as they have a color charge. The first generation contains the up quark, and the slightly heavier down quark. The up quark (u) has a fractional electric charge[‡] of $\frac{2}{3}$, and the down quark (d) has an electric charge of $-\frac{1}{3}$. The other quarks follow this pattern, with in the second generation the charm quark (c, charge $\frac{2}{3}$) and the strange quark (s, charge $-\frac{1}{3}$), and in the third generation the top quark (t, charge $\frac{2}{3}$) and beauty quark (b, charge $-\frac{1}{3}$). Especially the beauty quark, often called b-quark, will play an important role in this thesis.

The leptons follow the same model as the quarks, except that they do not feel the effects of the strong nuclear force. The two particles in each generation in this case have an electric charge of -1 or 0. Those without any electric charge are called neutrinos and are massless in the Standard Model. Those with an electric charge are the electron (e^-), muon (μ^-), and tauon (τ^-). The corresponding neutrinos are then simply called the electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ).

The masses of all these particles can be found in table 2.1.1. Virtually all matter we see around us is made up of regular atoms built from up quarks, down quarks, and electrons, along with the relevant force carriers; gluons and photons. It can also be said that by far most of the mass that these atoms have comes from the binding energy of quarks inside the atomic nucleus. Thus, a very large portion of the mass around us comes from gluons, not the Higgs field.

[†]It can be said however that at least three generations are necessary for the Standard Model to predict CP violation at the rates that are experimentally shown to exist in Nature. When there are less than three generations CP violation can only happen via non-perturbative effects which have only a very small prob-

Generation	Quark	Mass	Lepton	Mass
1	u	2.2 ± 0.6 MeV	e^-	$0.511 \pm 3.1 \times 10^{-9}$ MeV
	d	4.7 ± 0.5 MeV	ν_e	0
2	c	1.27 ± 0.03 GeV	μ^-	$105.7 \pm 2.4 \times 10^{-7}$ MeV
	s	96 ± 8 MeV	ν_μ	0
3	t	1.73 ± 1.22 GeV	τ^-	1776.9 ± 0.12 MeV
	b	4.66 ± 0.04 GeV	ν_τ	0

Table 2.1.1: *List of matter particles in the Standard Model and their masses. Data taken from [8].*

2.1.3 Higgs boson

The Higgs boson is a bit of an oddity and cannot be classified as a force carrier or a matter particle. The boson is not associated to any gauge symmetry and so it is not a force carrier, and unlike the matter particles it is not a spin-$\frac{1}{2}$ fermion. It can therefore best be regarded as forming a category on its own. Its a scalar boson meaning it has spin-0, which is unique in the Standard Model and indeed in Nature. The Higgs boson is the first seemingly elementary scalar particle ever discovered, and therefore raises the question whether more such particles might exist, and whether it is truly elementary or consists of other particles. It is described in much more detail in the next section where we discuss the Higgs mechanism.

2.1.4 Issues with the Standard Model

For all its strengths it should, however, be noted that the Standard Model is known to be incomplete. There are phenomena not described by the model like the force of gravity. Astronomical observations also point to the existence of dark matter [10] and dark energy [11], both of which are not included in the Standard Model. And on top of this the neutrinos in the model are massless, whilst we know from experiments that at least two neutrinos have a mass [12]. There are also possible issues with the internal consistency of the model in the Higgs sector: the boson might be metastable [13] and its mass has the potential to diverge or require fine-tuning [14].

Because of these problems many physicists are working on extending the Standard Model or overhauling it. Popular models include supersymmetry (SUSY) and various models predicting new forces of Nature. Many of these also severely influence the Higgs

ability of occurring (see for instance [7]). As for the possibility of having more than three generations; this option is still on the table, but there are measurements that show that if these generations exist they should be quite heavy and experimentally out-of-reach for now. Bounds exists, for example, from measurements of the Z boson decay peak width (see a list of references mentioned under 'Number of Neutrino Types' in [8]), and from nucleosynthesis in the early universe [9].

‡All the electric charges mentioned here are in units of the positive elementary charge, which is about 1.6×10^{-19} C.

sector; either by adding more Higgs-like bosons, or by making the Higgs boson a composite particle. In the last forty years no experimental evidence has been found for any of these, making it all the more important to go over all the predictions the Standard Model makes and see if they can be experimentally verified. Any deviation could provide a handle on determining which extensions remain viable and which need to be redirected to the scrapheap. As such the importance of measuring the properties of the Higgs boson is paramount.

2.2 The Higgs mechanism

THE Higgs mechanism forms a crucial part of the motivation for the main analysis presented in this thesis. In this section we will provide a description of the Higgs mechanism as a model for electroweak symmetry breaking. This model can be regarded without much difficulty as the canonical treatise, focusing on symmetry principles that underpin much of the Standard Model. It should be mentioned that there is a more phenomenological focused treatise as well [15]. It is the author's opinion that this latter method is more intuitive and easier to understand for those without a lot of mathematical training, but will defer from treating it in this thesis for brevity.

The Standard Model's mathematical description is often presented in the form of the Lagrangian density, $\mathscr{L}$. As in classical mechanics the Lagrangian, L, describes the difference between the kinetic and the potential energy in a system. The equations of motion can then be retrieved using the Euler-Lagrange equations[†]. In classical mechanics one deals often with discrete particles, an approach that does not work in the context of the Standard Model since its elements are not particles but continuous fields. Instead one uses the Lagrangian density, $\mathscr{L}$:

$$L(t) = \int \mathscr{L} d^3x \tag{2.2.1}$$

It is important to note, and to avoid confusion, that basically everyone calls $\mathscr{L}$ the Lagrangian, instead of the more correct Lagrangian *density*. In this thesis we can do no other than bow down to peer-pressure and acquiesce to this. It is an observation in Nature that all Lagrangians corresponding to physical systems be local. This means that fields at different locations will not talk to each other, and the closest one can get to this is through a gradient term between two infinitesimally close locations. The Euler-Lagrange equations can be stated in terms of $\mathscr{L}$. For a scalar field, ϕ, we get:

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \right) - \frac{\partial \mathscr{L}}{\partial \phi} = 0 \tag{2.2.2}$$

This equation can be written down for any field, but the results will differ depending on the spacetime properties of the field, specifically its spin and mass. For a spin-0 particle of any mass, a scalar like the Higgs boson, the Euler-Lagrange equation leads to the

[†]Which can be derived from the principle of least action. This action is defined as $S = \int_{t_1}^{t_2} L\, dt$, and requiring $\delta S = 0$ whilst varying how the system changes over time will then lead to the Euler-Lagrange equations.

Klein-Gordon equation. For spin-$\frac{1}{2}$ fields, one usually obtains the Dirac equation if the field is massive as for an electron, or the Weyl equation if the field is massless as for a neutrino[†]. For spin-1 fields one obtains the Maxwell equations for massless vector bosons, such as the photon, or the Proca equation if they do have a mass, such as the Z boson.

In the Standard Model one typically starts with a Lagrangian with all the spin-$\frac{1}{2}$ matter fields that have been observed in Nature. All these matter fields are taken to be massless ab initio. One then invokes a principle known as local gauge invariance to generate new terms for the Lagrangian that describe all the known forces of Nature with the exception of gravity. This principle can be thought of as an extension to Noether's theorem, which states that if the Lagrangian is left unchanged due to some global operation then there will be a conserved quantity. Since the Lagrangian describes the laws of Nature this then means that conservation laws can be extracted from symmetries found in these laws. For example the Standard Model Lagrangian can be shown to be invariant under spatial coordinate translations, which then via Noether's theorem leads to the law of linear momentum conservation.

Local gauge invariance[‡] extends this by saying that the invariance should also hold if the variation is allowed to be different for every point in spacetime. So in the case of translation this means one translates each point in spacetime differently, and the Lagrangian should still stay the same. As an example consider the the following Lagrangian for the electron (the situation for other particles is analogous):

$$\mathscr{L} = i\bar{\psi}_e \gamma^\mu \partial_\mu \psi_e - m\bar{\psi}_e \psi_e \tag{2.2.3}$$

where ψ_e represents the electron field, m the electron mass, and γ is a Dirac gamma-matrix. This Lagrangian is invariant under the following phase transformation:

$$\psi_e(x) \rightarrow e^{i\zeta}\psi_e(x) \tag{2.2.4}$$
$$\bar{\psi}_e(x) \rightarrow e^{-i\zeta}\bar{\psi}_e(x) \tag{2.2.5}$$

This can be easily checked, and holds because the phase is not acted on by the partial derivative in the term $\partial_\mu \psi_e$. By Noether's theorem it can be shown that this fact can be related to the conservation of a quantity that can be identified as an electric charge. Moving on to local gauge invariance the transformation changes to make the ζ parameter coordinate dependent:

$$\psi_e(x) \rightarrow e^{i\zeta(x)}\psi_e(x) \tag{2.2.6}$$
$$\bar{\psi}_e(x) \rightarrow e^{-i\zeta(x)}\bar{\psi}_e(x) \tag{2.2.7}$$

Thus, at every point in spacetime the phase could be chosen differently. It is customary, and illustrative, to replace the phase variable, $\zeta(x)$, by two different quantities, one constant, and the other one depending on x:

$$\zeta(x) \rightarrow q\theta(x) \tag{2.2.8}$$

[†]One can also obtain the Majorana equation in the special case that the fermion is its own anti-particle. In the Standard Model this is never the case, although there is some speculation that neutrinos might be Majorana particles in this sense, which could provide a convenient way to grant them a mass.

[‡]The invariance previously considered is known as a global gauge invariance.

The reason for doing this is that the constant q will end up playing the role of the electric charge of the electron or more correctly, the coupling constant between the electron and the photon[†]. As it stands our Lagrangian (2.2.3) is not invariant under this transformation. The trouble arises in the first term due to the partial derivative there:

$$i\bar{\psi}_e\gamma^\mu\partial_\mu\psi_e \quad \rightarrow \quad ie^{-iq\theta(x)}\bar{\psi}_e\gamma^\mu\partial_\mu\left(e^{iq\theta(x)}\psi_e\right) \tag{2.2.9}$$

Which leads to:

$$\begin{aligned} & ie^{-iq\theta(x)}\bar{\psi}_e\gamma^\mu\partial_\mu\left(e^{iq\theta(x)}\psi_e\right) \\ & = ie^{-iq\theta(x)}\bar{\psi}_e\gamma^\mu\left[e^{iq\theta(x)}\partial_\mu\psi_e + \left(\partial_\mu e^{iq\theta(x)}\right)\psi_e\right] && (2.2.10) \\ & = i\bar{\psi}_e\gamma^\mu\partial_\mu\psi_e + ie^{-iq\theta(x)}\bar{\psi}_e\gamma^\mu\left(\partial_\mu e^{iq\theta(x)}\right)\psi_e && (2.2.11) \end{aligned}$$

The first term here is the sought-after result, but the second term spoils the invariance unless it is zero, but it is not:

$$\begin{aligned} ie^{-iq\theta(x)}\bar{\psi}_e\gamma^\mu\left(\partial_\mu e^{iq\theta(x)}\right)\psi_e &= ie^{-iq\theta(x)}\bar{\psi}_e\gamma^\mu\left(iq\,e^{iq\theta(x)}\partial_\mu\theta(x)\right)\psi_e && (2.2.12) \\ &= -q\bar{\psi}_e\gamma^\mu\left(\partial_\mu\theta(x)\right)\psi_e && (2.2.13) \end{aligned}$$

This is solved by introducing a new field with a certain transformation behavior to cancel this term. In this case one introduces a new spin-1 field[‡], the photon field, indicated by A_μ whose transformation character is as follows:

$$A_\mu \rightarrow A_\mu + \partial_\mu\theta \tag{2.2.14}$$

In addition one adds a term describing the interaction between the electron and the photon fields:

$$q\bar{\psi}_e\gamma^\mu A_\mu\psi_e \tag{2.2.15}$$

So that our Lagrangian now reads:

$$\mathscr{L} = i\bar{\psi}_e\gamma^\mu\partial_\mu\psi_e - m\bar{\psi}_e\psi_e + q\bar{\psi}_e\gamma^\mu A_\mu\psi_e \tag{2.2.16}$$

Which is completely invariant under the given local gauge transformation and where q is the electric charge of the electron. Adding a kinetic term for the photon field then yields the quantum electrodynamics (QED) Lagrangian for electrons:

$$\mathscr{L} = i\bar{\psi}_e\gamma^\mu\partial_\mu\psi_e - m\bar{\psi}_e\psi_e + q\bar{\psi}_e\gamma^\mu A_\mu\psi_e - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.2.17}$$

Here $F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$. This same prescription can be used to derive the other forces by using various other invariances. But a problem arises when trying to include the weak nuclear force. This force and its gauge fields, the W and Z bosons, can be found

[†] One can of course do the full derivation using $\zeta(x)$ and the result will be the same. However, the q will be invisible (being equal to unity) and so the resulting Lagrangian is less intuitive to interpret.

[‡] The spin of the new field can be derived from the exact form of the local gauge transformation one wishes to cancel out. In the Standard Model this will always end up being spin-1, but a naive (and ultimately doomed) approach to including gravity by requiring invariance under local coordinate transformations will require one to introduce a spin-2 field, often coined the graviton.

by requiring local $SU(2)_L$ invariance, however these bosons will be massless, and we shall see that it is impossible to add a mass term for them. And to make matters worse, the inclusion of a $SU(2)_L$ symmetry also makes it impossible to retain the mass terms for the fermions, such as the electron. It should be stressed that these problems do not exist for the electromagnetic force ($U(1)$ symmetry) or the strong nuclear force ($SU(3)$ symmetry), which is why the Higgs boson and its mechanism is so closely tied to the weak sector.

The $SU(2)_L$ invariance creates three new spin-1 fields typically enumerated as W_1^μ, W_2^μ, and W_3^μ, collectively written down as $\vec{W}^\mu$. The reason for this is that the $SU(2)$ group has three generators. The transformation of ψ_e then is:

$$\psi_e \to e^{ig\vec{\sigma}\cdot\vec{\zeta}}\psi_e \tag{2.2.18}$$

Where $\vec{\sigma}$ is a vector of the three Pauli matrices that generate the $SU(2)$ group. $\vec{\zeta}$ is a vector of parameters which can be chosen differently at every point in spacetime. The role it plays is analogous to the role of ζ before. Analogous to q we introduce a coupling constant g which is not dependent on time or space. The gauge transformation of $\vec{W}$ is then:

$$\vec{W}_\mu \to \vec{W}_\mu + \partial_\mu\vec{\zeta} - g\vec{\zeta}\times\vec{W}_\mu \tag{2.2.19}$$

However, a mass term for a gauge boson has the following form:

$$\frac{1}{2}m_W W_1^\mu W_{1\mu} \tag{2.2.20}$$

Already the partial derivative of $\vec{\zeta}$ in the transformation rule makes this break the symmetry, because there is no other term to absorb it into. For example, the combination $m_W W_1^\mu W_{1\mu}$ transforms to something that includes a term proportional to $m_W\partial^\mu\zeta_1\partial_\mu\zeta_1$, which cannot cancel against anything else in $\mathscr{L}$.

For the fermionic mass terms the situation is also problematic. This is because we are talking about a $SU(2)_L$ symmetry here, not only a $SU(2)$ symmetry. The electron mass term has the following form:

$$\frac{1}{2}m_e\bar{\psi}_e\psi_e \tag{2.2.21}$$

The complication arises because of the chirality of the imposed symmetry. It is introduced because it is observed that the weak nuclear force does not interact at all with right-handed particles. Any fermionic wave function can be decomposed[†] into a right-handed part, ψ_R, and left-handed part, ψ_L, and doing so puts the mass term in the following form:

$$\frac{1}{2}m_e\left(\bar{\psi}_{eR}\psi_{eL} + \bar{\psi}_{eL}\psi_{eR}\right) \tag{2.2.22}$$

But this cannot be invariant under $SU(2)_L$ as only the ψ_{eL} is affected by an $SU(2)_L$ transformation, and there is no other field to compensate[‡].

[†]This can be done via the use of the projection operator $\frac{1}{2}(1+\gamma^5)$ for right-handed fermions, and $\frac{1}{2}(1-\gamma^5)$ for left-handed fermions. To be exact these operators can also be used to define what handedness actually means; a right-handed wave function is an eigenstate of $\frac{1}{2}(1+\gamma^5)$ and a left-handed wave function is an eigenstate of $\frac{1}{2}(1-\gamma^5)$. In this case γ^5 is a special Dirac matrix[15].

[‡]Note that in the case of the phase transformation we discussed before, the phase of $\bar{\psi}_e$ will rotate in the opposite way of the rotation of the phase of ψ_e and so the two rotations cancel. The same does not happen here because the $\bar{\psi}_{eR}$ is a singlet under $SU(2)_L$.

The Higgs mechanism solves this by providing a gauge invariant Lagrangian whose groundstate, however, is not invariant. As such the invariance will be spontaneously broken once the Universe drops to its groundstate. The mechanism leaves the strong nuclear sector completely unchanged and so we can focus only on the electromagnetic and weak nuclear forces. Before the spontaneous symmetry breaking these forces are combined into one superforce known as the electroweak force. It has the associated symmetry group of $U(1)_Y \times SU(2)_L$. The subscript Y is to denote that this $U(1)$ symmetry is not the usual phase invariance which leads to the electromagnetic force. Instead this $U(1)_Y$ symmetry refers to a different version of electromagnetism whose analogue of an electric charge is known as a hypercharge.

The reason for this bricolage is that in the process of the symmetry breaking the $SU(2)_L$ invariance will be lost and along the way the $U(1)_Y$ symmetry will combine with the remnants of the $SU(2)_L$ symmetry to form the $U(1)$ electromagnetic symmetry along with massive vector bosons for the broken $SU(2)_L$ symmetry. This connection between $U(1)$ and $SU(2)_L$ is already hinted at by the fact that the W bosons have an electric charge and so interact with the photon. We are thus in the situation that the force carriers of two distinct forces of Nature interact.

The Higgs model introduces an $SU(2)_L$ doublet of the following form:

$$\phi \quad = \quad \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.2.23}$$

Where ϕ^+ is a complex scalar with a hypercharge[†], and ϕ^0 is a complex scalar without any (hyper)charge. The Lagrangian for this doublet is given by:

$$\mathscr{L} = (\partial^\mu \phi)^\dagger (\partial_\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda \left(\phi^\dagger \phi\right)^2 \tag{2.2.24}$$

Where the first term is the kinetic term and the latter two terms combined form the Higgs potential. This is the simplest expression that does the trick, and the one currently under experimental scrutiny. This simple model has two free parameters: μ and λ. The μ parameter describes the mass of this doublet and together with λ determines the shape of the potential of ϕ.

The Higgs mechanism works its magic by saying that $\mu^2 < 0$. This means that the mass of ϕ is imaginary and thus that ϕ is a tachyonic field. Tachyonic fields possess the feature of being extremely unstable, which is exactly what happens here. To see this we need to inspect the minima of the potential, depicted in figure 2.2.1. The potential has a groundstate that is not symmetric since the argument of the complex ϕ is not fixed. The potential can be said to have a degenerate groundstate. The reigning idea is that this happened within the first fraction of a second after the Big Bang when the temperature dropped enough. The minima of the potential lie on a circle defined by:

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda} \tag{2.2.25}$$

Because the vacuum of the Universe is not electrically charged it should only be ϕ^0 that obtains a non-zero vacuum expectation value. This might seem contrived but this is

[†]The conjugate of ϕ of course carries a field ϕ^-.

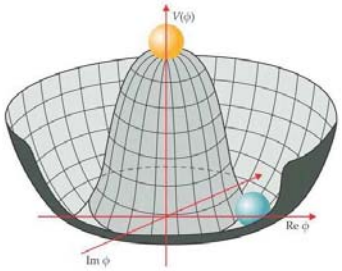

Figure 2.2.1: *Schematic representation of the potential, $V(\phi)$ of the ϕ scalar field. Overall the potential obeys $SU(2)_L \times U(1)$ symmetry but in the degenerate groundstate a definite choice has to be made which breaks it. Picture taken from [16].*

actually the way we defined ϕ^0 and ϕ^+ earlier[†]. One can then rewrite ϕ and expand it around the vacuum expectation value using the unitary gauge as:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \tag{2.2.26}$$

where v is the vacuum expectation value given by $v^2 \equiv -\frac{\mu^2}{2\lambda}$, and h is known as the Higgs field. We thus pass from using ϕ to using a real scalar field h.

The required bosonic mass terms are now generated by the following terms[‡], which are invariant under $U(1)_Y \times SU(2)_L$ as intended:

$$\frac{1}{4}\phi^\dagger \left(g^2 \vec{\sigma} \cdot \vec{W}_\mu \vec{\sigma} \cdot \vec{W}^\mu + q_y^2 B^\mu B_\mu + 2g\, q_Y \vec{\sigma} \cdot \vec{W}^\mu B_\mu \right) \phi \tag{2.2.27}$$

Here B^μ is the gauge field associated with the $U(1)$ hypercharge symmetry, and q_Y its coupling constant, analogous to the charge q for electromagnetism. These terms manage to be invariant under $SU(2)$ specifically because the ϕ and $\phi^\dagger$ help turn each term into an $U(1) \times SU(2)$ singlet, which was stopping a mass term of the form $\frac{1}{2}m^2 W_\mu W^\mu$ before. Note that the above expression does not expressly contain a mass term at all. The closest one gets are terms of the form $\phi^\dagger W_\mu W^\mu \phi$. However, with the spontaneous symmetry breaking we can generate such a mass term anyway. We can substitute in expression 2.2.26:

$$\frac{1}{8}\,(0, v + h) \left(g^2 \vec{\sigma} \cdot \vec{W}_\mu \vec{\sigma} \cdot \vec{W}^\mu + q_y^2 B^\mu B_\mu + 2g\, q_Y \vec{\sigma} \cdot \vec{W}^\mu B_\mu \right) \begin{pmatrix} 0 \\ v + h \end{pmatrix} \tag{2.2.28}$$

We are able to isolate the term containing the $B^\mu B_\mu$ combination:

$$\frac{1}{8}\,(0, v + h) \left(g^2 \vec{\sigma} \cdot \vec{W}_\mu \vec{\sigma} \cdot \vec{W}^\mu + 2g\, q_Y \vec{\sigma} \cdot \vec{W}^\mu B_\mu \right) \begin{pmatrix} 0 \\ v + h \end{pmatrix} + \frac{q_Y^2}{8} B_\mu B^\mu \left(v^2 + h^2 + v\, h \right) \tag{2.2.29}$$

[†]With respect to the potential these can be viewed as using polar coordinates. The ϕ^0 corresponds to the radius, and ϕ^+ to the angle, that can be taken to be zero without any loss of generality since the potential is cylindrically symmetric.

[‡]For completeness we should note that there are other terms as well, such as the kinetic term, but these are not relevant here.

Here we can already see the appearance of a mass term for the B^μ field of the form:

$$\frac{q_Y^2}{8} v^2 B^\mu B_\mu \tag{2.2.30}$$

This suggests the B^μ field has obtained a mass of:

$$m_B = \frac{1}{2} q_Y v \tag{2.2.31}$$

Let us reflect on how this term originated from the non-zero vacuum expectation value of the ϕ field. This essentially creates a lot of ϕ-independent terms featuring v. And because the v features only in the combination $v + h$ this ensures that there will also be an interaction terms with the Higgs field whose coupling strength is proportional to v and q_Y which both also feature in the mass. Thus, the Higgs field can be said to couple to particles with a strength proportional to their mass.

Expanding the $W^\mu W_\mu$ term leads to:

$$\begin{aligned}
&\frac{g^2}{8} (0, v+h)\, \vec{\sigma} \cdot \vec{W}_\mu \vec{\sigma} \cdot \vec{W}^\mu \begin{pmatrix} 0 \\ v+h \end{pmatrix} \\
&= \frac{g^2}{8} (0, v+h) \begin{pmatrix} W_{3\mu} & W_{1\mu} - iW_{2\mu} \\ W_{1\mu} + iW_{2\mu} & -W_{3\mu} \end{pmatrix} \vec{\sigma} \cdot \vec{W}^\mu \begin{pmatrix} 0 \\ v+h \end{pmatrix} && (2.2.32) \\
&= \frac{g^2}{8} (0, v+h) \begin{pmatrix} \vec{W}^\mu \cdot \vec{W}_\mu & 0 \\ 0 & \vec{W}^\mu \cdot \vec{W}_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v+h \end{pmatrix} && (2.2.33) \\
&= \frac{g^2}{8} (v+h)^2 \vec{W}^\mu \cdot \vec{W}_\mu && (2.2.34)
\end{aligned}$$

Out of this last line one may distill mass terms for W_1, W_2, and W_3 of the form:

$$\frac{g^2 v^2}{8} W_x^\mu W_{x\mu},\ x \in \{1, 2, 3\} \tag{2.2.35}$$

All three W bosons then have a mass of:

$$m_W = \frac{1}{2} g\, v \tag{2.2.36}$$

An interesting thing however happens with W_3^μ and B^μ. In the $W^\mu B_\mu$ term we can expand the Pauli matrices to obtain:

$$\frac{g\, q_Y}{4} (0, v+h) \left(\sigma_1 \cdot W_1^\mu + \sigma_2 \cdot W_2^\mu + \sigma_3 \cdot W_3^\mu\right) B_\mu \begin{pmatrix} 0 \\ v+h \end{pmatrix} \tag{2.2.37}$$

Which leads to:

$$\frac{g\, q_Y}{4} (0, v+h) \begin{pmatrix} W_3^\mu & W_1^\mu - iW_2^\mu \\ W_1^\mu + iW_2^\mu & -W_3^\mu \end{pmatrix} B_\mu \begin{pmatrix} 0 \\ v+h \end{pmatrix} \tag{2.2.38}$$

And eventually to:

$$-\frac{g\, q_Y}{4} (v+h)^2\, W_3^\mu B_\mu \tag{2.2.39}$$

Let us examine this term together with the mass terms pertaining to W_3^μ and B^μ obtained earlier. We get:

$$-\frac{g\,q_Y}{4}\,(v+h)^2\,W_3^\mu B_\mu + \frac{g^2}{8}(v+h)^2 W_3^\mu W_{3\mu} + \frac{q_Y^2}{8}B_\mu B^\mu\left(v^2+h^2+v\,h\right) \tag{2.2.40}$$

Focusing only on the terms with the Higgs field, h, in them we find:

$$-\frac{g\,q_Y v^2}{4}W_3^\mu B_\mu + \frac{g^2 v^2}{8}W_3^\mu W_{3\mu} + \frac{q_Y^2 v^2}{8}B_\mu B^\mu \tag{2.2.41}$$

Which can be rewritten to:

$$\frac{v^2}{8}\,(g\,W_3^\mu - q_Y B^\mu)\,(g\,W_{3\mu} - q_Y B_\mu) \tag{2.2.42}$$

At this point one can note that the $W_3^\mu B_\mu$ terms are basically interactions where a W_3 can turn into a B and vice versa without cost. What is actually happening is that these two fields mix, and the mass matrix is non-diagonal. The fields that physically propagate are those with a diagonal mass matrix, so we need to change the basis so this is the case. These fields will then have a mass equal to the eigenvalues of this mass matrix. Rewriting equation 2.2.42 we get:

$$\frac{v^2}{8}\,(W_3^\mu, B^\mu)\begin{pmatrix} g^2 & -q_Y g \\ -q_Y g & q_Y^2 \end{pmatrix}\begin{pmatrix} W_{3\mu} \\ B_\mu \end{pmatrix} \tag{2.2.43}$$

Determining the eigenvalues, λ, of the resulting matrix one finds:

$$\lambda_1 = 0,\ \lambda_2 = g^2 + q_Y^2 \tag{2.2.44}$$

We thus end up with one physical boson that is massless, and one that has a mass of $\frac{1}{2}v\sqrt{g^2+q_Y^2}$. These bosons are known as the photon, A^μ, and the Z boson respectively. All interaction terms between the photon and Z boson fields also disappear, meaning the Z boson has no electric charge unlike the two remaining W bosons. The Higgs mechanism thus succeeds in its goal of granting a mass to the W and Z bosons. For completeness we should also note that via expression 2.2.38 it also mixes the W_1 and W_2 fields into the massive $W^\pm$ where $W^\pm \equiv \frac{1}{\sqrt{2}}\,(W_1 \mp iW_2)$.

Giving masses to the fermions proceeds by adding an interaction with the ϕ field to the mass terms:

$$m_e\left(\bar{\psi}_{eR}\psi_{eL} + \bar{\psi}_{eL}\psi_{eR}\right) \rightarrow \mathcal{Y}_e m_e\left(\bar{\psi}_{eR}\phi\psi_{eL} + \bar{\psi}_{eL}\phi^\dagger\psi_{eR}\right) \tag{2.2.45}$$

The presence of the ϕ field can absorb any $SU(2)_L$ action on each of the terms because it itself is also an $SU(2)$ doublet. After the spontaneous symmetry breaking occurs the ϕ field can again be expanded around the vacuum expectation value giving rise to terms containing the two fermion chiralities and the combination $v+h$. The terms without h then become the mass terms, those with the h fields become couplings between the Higgs fields and the fermions in question. The factor $\mathcal{Y}_e$ is known as the Yukawa coupling strength between the fermion (in this case electron) and the ϕ field. It can be chosen in such a way that the resulting mass one will obtain for the electron corresponds to experimental measurements. It is therefore important to stress that the Higgs mechanism does not actually predict what the masses of the fermions will be, these are taken from the Yukawa couplings, which are essentially free parameters tuned to experimental data. It does, however, allows for the introduction of mass terms into an otherwise massless $SU(2)_L$ symmetric model.

2.3 Motivation for measuring the Higgs potential

SEVERAL years ago a new scalar boson was discovered at the LHC, bearing all the hallmarks of the long-sought-after Higgs boson [1, 2]. It remains to be seen however, whether this new particle is responsible for the spontaneous breaking of the electroweak symmetry at lower energies as the Higgs mechanism [17, 18] foresees. The simplest Higgs potential portrayed by expression 2.2.24 might not be the one Nature choose to realize in reality. And it is important to understand the actual shape of this potential since it might also have consequences for cosmology in areas such as baryogenesis [19], inflation [20], and even the stability of the Universe as we know it [13].

The Higgs mechanism gives a face to an existential threat that hangs as a dark shadow over the Universe itself. This is because the Higgs mechanism, if it is shown to be completely correct, provides an example of spontaneous symmetry breaking, and thus indicates that Nature made a decision to include unstable potentials in her Laws. This opens up the possible question whether what we think is the groundstate of the vacuum is the true groundstate, or whether we live in a so-called false vacuum, where the possibility remains open that one day this false vacuum might decay to a lower energetic state. This decay could then change the numerical values or even definitions of particles[†] in such a way as to wipe out life as we know it, and probably all other kinds of matter as well. Such a vacuum decay initiated at a single point in space will free up a lot of energy, which can allow the Higgs field in the immediate surrounding to decay as well by using the generated energy to tunnel to the lower vacuum state. This will set off a chain reaction with a bubble of the alternative vacuum advancing with the speed of light. Since this "death" bubble travels at the speed of light its arrival cannot be seen from afar, especially because its ignition required a stochastic quantum effect[‡].

In addition to this dark prospect there are other reasons for wanting to explore the exact mechanism of the electroweak symmetry breaking further. For example it is unknown what the order is of this phase transition, which might affect questions like baryogenesis [22, 23]. A modified Higgs potential may also be responsible for cosmic inflation [24].

A different Higgs potential can be parametrized by adding higher order terms to the Higgs potential, keeping in mind to maintain the instability so that spontaneous symmetry breaking will actually happen. Various examples of this can be found in literature [25, 26, 27]. At the current particle accelerators and detectors we can only probe the Higgs potential in a narrow area around the groundstate, and we are thus blind to shape differences further away. In a sense we can only probe the mass term at present. After expanding the ϕ field around the ground state the potential of the h field is:

$$V(h) = \frac{1}{2}m_h h^2 + \frac{1}{6}\lambda_{3h}h^3 + \frac{1}{24}\lambda_{4h}h^4 + ... \tag{2.3.1}$$

[†]Recall that the photon and Z boson did not even exist as distinct physical particles before the electroweak symmetry breaking.

[‡]One might propose to set a limit on the rate of occurrence of these death bubbles by using the fact that we are still alive. But this approach inevitably runs afoul of the anthropic principle. In addition one could get more philosophical on this and regard it as one big quantum immortality experiment, where we are basically Schrödinger's cat ourselves, and we can only observe worlds where we yet survive. For more information on quantum immortality see for example [21].

Process	LO [fb]	NLO [fb]	NNLO [fb]
$gg \to hh$ (gluon-gluon fusion)	$16.5^{+4.6}_{-3.5}$ [29]	$39.56^{+1.74}_{-2.4}$ [30]	$40.2^{+3.2}_{-3.5}$ [31]
$qq \to qqhh$ (vector-boson fusion)	$1.81^{+0.16}_{-0.14}$ [32]	$2.01^{+0.03}_{-0.02}$ [33]	
$qq \to W^{\pm}hh$ (associated production)	$0.43^{+0.005}_{-0.006}$ [32]		$0.57^{+0.0006}_{-0.002}$ [33]
$qq \to Zhh$ (associated production)	$0.27^{+0.004}_{-0.004}$ [32]		$0.42^{+0.02}_{-0.02}$ [33]

Table 2.4.1: *Important production cross-sections for Standard Model double Higgs production at $\sqrt{s} = 14\,TeV$. The listed errors include scale uncertainties where the scale was varied between $\sqrt{\hat{s}}/2$ and $2\sqrt{\hat{s}}$ only.*

where the lowest order forms the mass term with a bare mass of m_h, given by $m_h = \sqrt{\lambda v^2}$. Any deviation of the simplest potential used in the Standard Model can be parametrized by deviations in the higher order couplings[†]. Measuring these becomes progressively more difficult as the terms become smaller[‡]. The lowest hanging fruit should be analyzed first, which is λ_{3h}, also known as the Higgs trilinear coupling. The Standard Model predicts its value to be:

$$\lambda_{3h} = 3\frac{m_h^2}{v} \tag{2.3.2}$$

Any deviation from this value indicates new physics. Measuring λ_{3h} requires measuring the cross-sections of processes where a single Higgs boson splits into two others. We will review such processes further in section 2.4. It should be noted as further motivation, that the Higgs trilinear coupling is quite sensitive to the existence of yet-undiscovered massive particles. This is because such particles should couple strongly to the Higgs boson and might feature in a 1-loop correction to the trilinear coupling. We will investigate potential measurements of λ_{3h} in the next section, which was published as [28].

2.4 Theory predictions for double Higgs production

THERE are various processes that contribute to the double Higgs signature. Several of the more important ones are listed in table 2.4.1. The leading production channel is gluon-gluon fusion, having a cross-section exceeding the rest by at least an order of magnitude at approximately 40 fb. At leading (1-loop level) order gluon-gluon fusion production of double Higgs events are described by two Feynman diagrams, shown in figure 2.4.1. Only the "triangle" diagram is sensitive to the trilinear self-coupling and unfortunately the "box" diagram decreases the overall cross-section due to destructive

[†]There are also models where the Higgs potential is modified in a non-analytical way (i.e. it cannot be written as a power series). See for example [27]. These models might still show deviations when attempting to measure λ_{3h} though.

[‡]And also the complexity of the Feynman graphs describing the processes to be measured increases. This typically also lowers the cross-section, because it increases the number of (off-shell) particle propagators in those graphs.

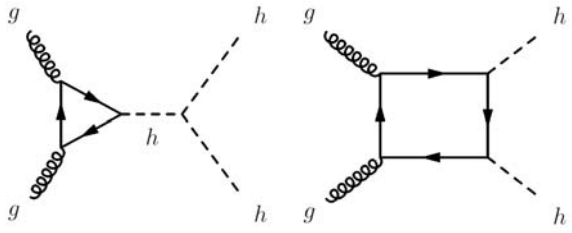

Figure 2.4.1: *Leading order Feynman diagrams for double Higgs production via gluon-gluon fusion.*

interference. The partonic cross-section is given by [29, 34]:

$$\hat{\sigma}^{(LO)}_{gg\to hh} = \int d\hat{t} \frac{\alpha^2 \alpha_S^2}{2^{15}\pi M_W^4} \left| C_\triangle F_\triangle + C_\square F_\square \right|^2 \tag{2.4.1}$$

where $C_\triangle F_\triangle$ and $C_\square F_\square$ represent the contributions from the triangle and box diagrams separately. The scale, α_S is set to the invariant mass of the two incoming partons, $\sqrt{\hat{s}}$. Here $\hat{t}$ represents a Mandelstam variable:

$$\hat{t} = -\frac{1}{2}\left[\hat{s} - 2m_h^2 - \hat{s}\sqrt{1 - \frac{4m_h^2}{\hat{s}}}\cos\theta\right] \tag{2.4.2}$$

where θ is the angular separation between the two final state Higgs bosons in the center of mass frame. The diagram specific factor:

$$\left| C_\triangle F_\triangle + C_\square F_\square \right|^2 \tag{2.4.3}$$

contains the form factors $F_\triangle$ and $F_\square$ coming from the loop calculation. The coefficients $C_\triangle$ and $C_\square$ describe the resonance behavior of the Higgs boson propagators. The loop calculation can be performed either exactly or by setting the top quark mass to infinity, which leads to an effective field theory (EFT). At the time the work that this part of the thesis describes was performed EFT based event generators were readily available whilst ones using exact calculations did not exist. However, exact calculations were available for very specific channels and observables. As such we compared the performance of these to a general EFT generator and see how well the latter agrees with the exact calculations.

In the exact case the expression for $F_\triangle$ is given by [34, 35]:

$$F_\triangle = \tau_q \left[1 + (1 - \tau_q) f(\tau_q)\right] \tag{2.4.4}$$

where τ_q is defined as $\frac{4m_q^2}{\hat{s}}$ and m_q is the mass of the quark in the loop. Only the top quark is considered as the contributions from the other quarks are negligible due to the small coupling between them and the Higgs boson. The function $f(\tau_q)$ is given by:

$$f(\tau_q) \quad = \quad \begin{cases} \arcsin^2\left(\frac{1}{\sqrt{\tau_q}}\right) & \tau_q \geq 1 \\ -\frac{1}{4}\left[\log\frac{1+\sqrt{1-\tau_q}}{1-\sqrt{1-\tau_q}} - i\pi\right] & \tau_q < 1 \end{cases} \tag{2.4.5}$$

The expression for $F_\Box$ is given by:

$$S = \frac{\hat{s}}{m_q^2}$$

$$\begin{aligned} F_\Box &= \frac{m_q^4}{\hat{s}^2}\left\{\frac{4\hat{s}}{m_q^2} + 8\hat{s}C_{ab} - 2\hat{s}m_q^2\left(\frac{\hat{s}}{m_q^2} + 2\frac{m_h^2}{m_q^2} - 8\right)(D_{abc} + D_{bac} + D_{acb})\right\} \\ &\quad + [\left(\hat{t} - m_h^2\right)C_{ac} + \left(\hat{u} - m_h^2\right)C_{bc} + \left(\hat{u} - m_h^2\right)C_{ad} + \left(\hat{t} - m_h^2\right)C_{bd} \\ &\quad - \left(\hat{t}\hat{u} - m_h^4\right)D_{acb}] \cdot \frac{m_q^4}{\hat{s}^2}(\rho_c + \rho_d - 8) \end{aligned} \tag{2.4.6}$$

where $\hat{u}$ is a Mandelstam variable comparable to $\hat{t}$ and:

$$C_{ij} = \int \frac{1}{(q^2 - m_q^2)\left[(q+p_i)^2 - m_q^2\right]\left[(q+p_i+p_j)^2 - m_q^2\right]}\frac{d^4q}{i\pi^2} \tag{2.4.7}$$

$$\begin{aligned} D_{ijk} &= \int \frac{1}{(q^2 - m_q^2)\left[(q+p_i)^2 - m_q^2\right]} \\ &\quad \frac{1}{\left[(q+p_i+p_j)^2 - m_q^2\right]\left[(q+p_i+p_j+p_k)^2 - m_q^2\right]}\frac{d^4q}{i\pi^2} \end{aligned} \tag{2.4.8}$$

and where the numbering corresponds to $g_a g_b \to h_c h_d$. The propagator coefficients are given by:

$$C_\triangle = \frac{\lambda_{3h}v}{\hat{s} - m_h^2} \tag{2.4.9}$$

$$C_\Box = 1 \tag{2.4.10}$$

In the case the mass of the quark in the loop is taken to be infinite the form factors reduce to [35]:

$$F_\triangle = \frac{2}{3} \tag{2.4.11}$$

$$F_\Box = -\frac{2}{3} \tag{2.4.12}$$

Entering all these factors into expression 2.4.3 one gets:

$$|C_\triangle F_\triangle + C_\Box F_\Box|^2 = \left|\frac{\lambda_{3h}v}{\hat{s} - m_h^2}\cdot\frac{2}{3} - \frac{2}{3}\right|^2 = \frac{4}{9}\left(\frac{\lambda_{3h}^2 v^2}{\left(\hat{s} - m_h^2\right)^2} - \frac{2\lambda_{3h}v}{\hat{s} - m_h^2} + 1\right) \tag{2.4.13}$$

The behavior of this factor as a function of $\sqrt{\hat{s}}$ is plotted in figure 2.4.2 along with isolated contributions from only the triangle diagram and of only the box diagram. From this figure one observes that the contribution of the Higgs trilinear coupling, which is present only in the triangle diagram, drops off rapidly when $\sqrt{\hat{s}} > 400$ GeV, whilst the box contribution remains strong over a far bigger range. This implies sensitivity to the trilinear coupling is to be found close to the on-shell production threshold of $2m_h$. Unfortunately the total cross-section drops steeply around this value due to a strong destructive interference between the box and triangle diagrams. It can also be seen in the picture that the effective field theory approach overestimates the triangle

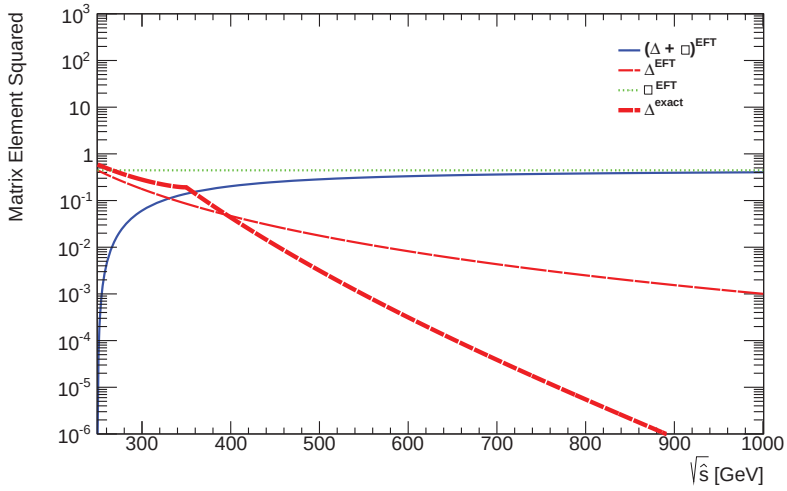

Figure 2.4.2: *Behavior of the form factor part of the matrix element squared as a function of* $\sqrt{\hat{s}}$. $(\triangle + \square)^{EFT} \equiv \left|C_{\triangle}F_{\triangle}^{EFT} + C_{\square}F_{\square}^{EFT}\right|^2$, $\triangle^{EFT} \equiv \left|C_{\triangle}F_{\triangle}^{EFT}\right|^2$, $\square^{EFT} \equiv \left|C_{\square}F_{\square}^{EFT}\right|^2$, *and* $\triangle^{exact} \equiv |C_{\triangle}F_{\triangle}|^2$.

contribution massively at larger energies and does not reproduce the kink around $\sqrt{\hat{s}} \approx 2m_q$. The agreement improves with increasing m_q and, as expected, becomes perfect when $m_q \to \infty$. The kink is missed due to the use of the flat $F_{\triangle} = \frac{2}{3}$ instead of the full expression (2.4.4). The reason that the triangle contribution drops faster than the box contribution is that the triangle diagram contains a third Higgs boson which is forced far off-shell already at the threshold production energy of $2m_h$. This Higgs propagator carries an additional energy dependence in the denominator that is absent in the box diagram.

2.4.1 Numerical predictions of double Higgs production

Partonic cross-section

Examining the evolution of the triangle contribution as a function of $\sqrt{\hat{s}}$ we see that any sensitivity to the trilinear coupling must be found in the region where $\sqrt{\hat{s}} < 2m_h$. This is below the on-shell production threshold and therefore the final state Higgs bosons are forced off-shell. It also means that expression 2.4.3 is no longer valid. We can, however, analyze the behavior in this regime by using numerical calculations performed with an effective field theory model implemented in MADGRAPH5 [36]. To allow for the final state Higgs boson to be off-shell we force one to decay to $\gamma\gamma$ and the other one to $b\bar{b}$. The partonic cross-section for $gg \to hh \to b\bar{b}\gamma\gamma$ as a function of $\sqrt{\hat{s}}$ is shown in figure 2.4.3. In this figure the following kinematic cuts† were applied:

- ► $|\eta| < 2.5$ for both the photons and the b-quarks

†The coordinate system used by ATLAS is fully explained in section 3.2. One might skip ahead and read that section before returning here if some of the concepts like p_T and η are unfamiliar.

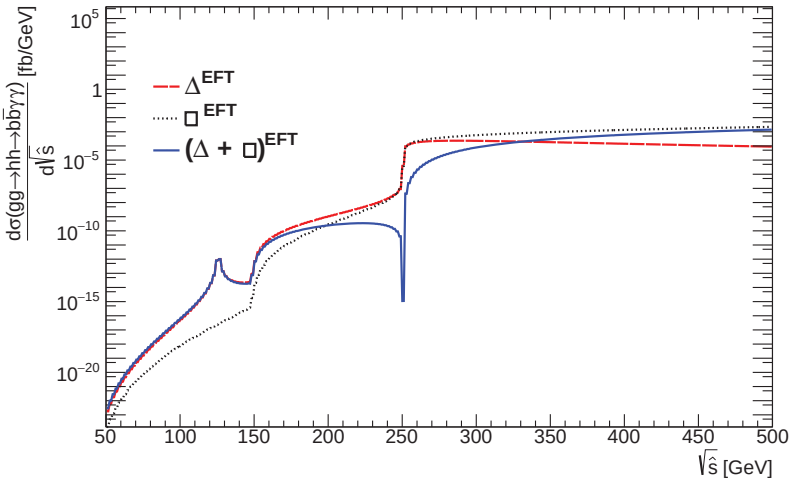

Figure 2.4.3: *The fiducial partonic cross-section of $gg \to hh \to b\bar{b}\gamma\gamma$ as a function of $\sqrt{\hat{s}}$. For $\triangle^{EFT}$ only the triangle diagram was activated, likewise for $\square^{EFT}$ only the box diagram was used. Both, an their interference, were used for $(\triangle + \square)^{EFT}$.*

- $p_T > 10$ GeV for both the photons and the b-quarks
- $\Delta R(b, \bar{b})$, $\Delta R(b, \gamma)$, $\Delta R(\bar{b}, \gamma)$, and $\Delta R(\gamma, \gamma) > 0.4$.

Where $\Delta R(x, y)$ represents the angular separation between particles x and y and is defined as:

$$\Delta R(x, y) \equiv \sqrt{\Delta\phi_{x,y}^2 + \Delta\eta_{x,y}^2} \tag{2.4.14}$$

and where η denotes the pseudorapidity of a particle and ϕ its azimuthal angle. The figure also shows that the partonic cross-section spans several orders of magnitude with the low end being the region where all Higgs propagators are forced off-shell, $\sqrt{\hat{s}} < 125$ GeV. In the region defined by $m_h < \sqrt{\hat{s}} < 2m_h$ at least one Higgs propagator is allowed to be on-shell and this boosts the cross-section massively. It should be noted, and it can be seen in the figure in the region $125 < \sqrt{\hat{s}} < 145$ GeV, that because of the transverse momentum cuts on the final state particles the phase-space is severely constrained until there is enough energy to satisfy the cuts and also have one propagator on-shell. This causes the sloping behavior and the kink just after the single Higgs mass peak. It can be seen that the triangle diagram is the dominant contribution around $\sqrt{\hat{s}} \approx m_h$, with the box diagram only obtaining comparable size around $\sqrt{\hat{s}} \approx 2m_h$, where the strong destructive interference between the two results in a sharp drop in the overall cross-section. Beyond the double Higgs mass peak the cross-section is mostly carried by the box diagram and the sensitivity to the Higgs trilinear coupling drops radically. Unfortunately, focusing on the behavior of the total cross-section, the part of the spectrum that is sensitive to the coupling is orders of magnitude smaller than the rest, meaning the total cross-section is influenced very little by the triangle diagram.

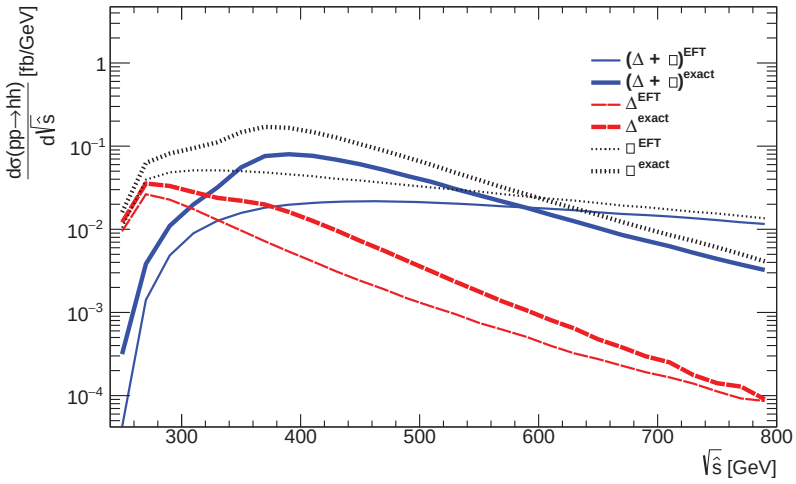

Figure 2.4.4: *The hadronic cross-section for $pp \to hh$ for both effective field theory and exact calculations.*

Hadronic cross-section

To see how the situation changes when moving to the hadronic cross-section, we add the CTEQ6L [37] parton density functions and examine double Higgs boson production from proton-proton collisions, $pp \to hh$ and analyze both effective field theory and exact predictions. The effective field theory calculations were made using MADGRAPH5 as before. At the time of this study the EFT model in MADGRAPH5 did not contain the $gghh$ coupling in its EFT model, and we implemented ourselves into the framework using the Feynman rules given in [29]. By setting the appropriate coupling constants to zero we switched off the triangle or box contributions as required. The exact calculations were made using the public code of [38]. In both cases the factorization and renormalization scales were set to $\sqrt{\hat{s}}$. In figure 2.4.4 the differential cross-section with respect to $\sqrt{\hat{s}}$ is shown. It is known that the gluon-density increases with decreasing x_{Bj} but this is not enough to elevate the cross-section at low $\sqrt{\hat{s}}$ much, mainly due to the destructive interference. Because the kinematics of the box diagram are not well modeled in the effective field theory it underestimates its contribution lower $\sqrt{\hat{s}}$ but overestimates it beyond $\sqrt{\hat{s}} > 590$ GeV, whilst for the triangle contribution the difference is mainly found in the normalization. As was seen before the EFT calculation does not reproduce the kink at the double top quark mass.

In figure 2.4.5 we analyze the kinematics of the process a bit more. It can be seen that the angle between the two Higgs bosons is in agreement between EFT and exact calculations for the triangle contributions up to the normalization. However, the box kinematics are mismodelled in EFT with respect to the exact calculation. Since the box contribution is dominant in the total cross-section this signals a mismodelling of the total process, at the very least at higher $\sqrt{\hat{s}}$; the exact calculations have a stronger preference than EFT for smaller angles. Studying the right figure in figure 2.4.5 one observes that the exact calculations also predict a broader Higgs boson pseudorapidity distribution, which means that the difference in x_{Bj} between the two colliding protons

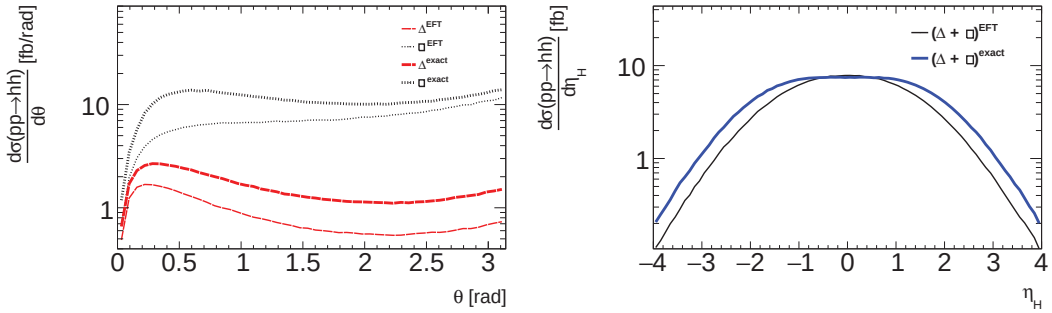

Figure 2.4.5: *The angle between the two produced Higgs bosons (in the lab frame) for the triangle and box contributions separately for both EFT and exact calculations. Shown on the right is the pseudorapidity distribution of every Higgs boson.*

is larger more often in the exact case as opposed to the EFT case. We conclude that EFT does not provide good predictions if a precise description of the kinematics is important.

Final state	Branching ratio	8 TeV	14 TeV	100 TeV
$b\bar{b}b\bar{b}$	32.5%	23(60)	$1.6 \cdot 10^4$ $(3.3 \cdot 10^4)$	$8.5 \cdot 10^5$ $(1.3 \cdot 10^6)$
$b\bar{b}WW$	23.9%	17(44)	$1.2 \cdot 10^4$ $(2.4 \cdot 10^4)$	$6.3 \cdot 10^5$ $(9.7 \cdot 10^5)$
$bb\bar{Z}Z$	3.0%	2.2(5.5)	$1.5 \cdot 10^3$ $(3.0 \cdot 10^3)$	$8.0 \cdot 10^4$ $(1.2 \cdot 10^5)$
$b\bar{b}\gamma\gamma$	0.26%	0.19(0.48)	128 (264)	6800 (10500)
$\gamma\gamma\gamma\gamma$	0.001%	0(0)	0.25 (0.53)	14 (21)

Table 2.4.2: *Expected number of events for double Higgs boson production from gluon fusion at various energy levels. The cross-sections and luminosities used can be found in the main text. The calculations are provided at leading order [38] and at next-to-leading order [29] (within parentheses). The branching ratios were taken from [32].*

$\sqrt{\hat{s}}$ [GeV]	Leading order [fb^{-1}]	Next-to-leading order [fb^{-1}]
8	3.58	9.22
14	16.23	33.86
100	877	1350

Table 2.4.3: *Cross-sections for double Higgs boson production from gluon fusion at leading order [38] and next-to-leading order [29].*

2.4.2 Irreducible backgrounds to double Higgs boson production

In this section we will only consider exact calculation as we have seen in the preceding section that effective field theory is woefully inadequate to describe the relevant kinematics. We will move on to investigate various different decay channels for the Higgs bosons and the backgrounds that will plague any analysis of them. In table 2.4.2 we show hypothetical event yields for several interesting decay channels. The cross-sections used are found in table 2.4.3. For 8 TeV we used a total integrated luminosity of $20\,\text{fb}^{-1}$ which is comparable to the total dataset collected by ATLAS and CMS for 8 TeV in 2012. For 14 and 100 TeV we use $3000\,\text{fb}^{-1}$ which is the expected end-of-lifetime luminosity for the upgraded LHC.

In terms of statistics the $b\bar{b}b\bar{b}$ channel definitely wins out, but it is hampered by large regular and combinatorial backgrounds. Also the requirement of four b-tags drops the acceptance considerably. Likewise $b\bar{b}WW$ and $b\bar{b}ZZ$ have large backgrounds. By contrast the quadruple photon channel, $\gamma\gamma\gamma\gamma$, has a very clean signature because the invariant masses of photon pairs can be determined quite accurately by the detector[†]. Unfortu-

[†]This is because the photon is directly detected by ATLAS, whereas b-jets and vector bosons are only seen

nately this channel performs really bad statistics-wise, having next to no events even at $3000\,\text{fb}^{-1}$ with 14 TeV, making an LHC probe of this channel essentially hopeless. The intermediate channel, $b\bar{b}\gamma\gamma$, is much more interesting; the inclusion of the two photons allows for a strong invariant mass cut, and removes the combinatorial background from the b-jets. With 264 expected events at the end of the HL-LHC it has the potential to be explored if a reasonable cut acceptance can be obtained, and it should definitely be worthwhile at 100 TeV with $3000\,\text{fb}^{-1}$ with a future hadron collider.

It is prudent to mention that, since the time the work we here describe was performed, several new experimental searches for double Higgs production (but no λ_{3h} measurements) have been published. ATLAS has published 8 TeV searches in $b\bar{b}\gamma\gamma$, $b\bar{b}b\bar{b}$, $WW\gamma\gamma$, and $b\bar{b}\tau\bar{\tau}$, as well as a $b\bar{b}b\bar{b}$ 13 TeV analysis [39, 40, 41, 42]. There was also an 8 TeV CMS search in $b\bar{b}b\bar{b}$ [43]. All of these either found no double Higgs events, or were consistent with Standard Model predictions within uncertainties.

For the remainder of this study we will select $pp \to hh \to b\bar{b}\gamma\gamma$ as our channel of choice and study the irreducible background. Specifically this means we will not be incorporating detector inefficiencies[†] or final and initial state radiation. And while there have been several studies of this channel (see for example [33], [44], and [45]) we will continue our practice of examining the triangle and box contributions separately in order to distill a measure of the sensitivity to the trilinear Higgs coupling. We follow loosely [44] in applying the following cuts:

- $p_T(b) > 45$ GeV, $p_T(\gamma) > 20$ GeV
- $|\eta(b)| < 2.5$, $|\eta(\gamma)| < 2.5$
- $\Delta R(b, b) > 0.4$, $\Delta R(\gamma, \gamma) > 0.4$
- $|m_{b\bar{b}} - m_h| < 20$ GeV, $|m_{\gamma\gamma} - m_h| < 2.3$ GeV

where $m_{b\bar{b}}$ denotes the invariant mass of the two selection b-jets, and $m_{\gamma\gamma}$ the invariant mass of the photon pair. These cuts are intended to simulate typical cuts an actual analysis might make. For example, the cut on the pseudorapidity restricts us to events where the selected particles are within the reach of the inner detector of ATLAS. The invariant mass cuts are designed to cut away a lot of the irreducible background, with the photon pair cut being more tight compared to that on the b-jets, because the photon pair invariant mass is easier to measure.

To assess the sensitivity to the trilinear coupling we will examine the $pp \to b\bar{b}\gamma\gamma$ process without double Higgs boson production (i.e. the irreducible background), the $pp \to hh \to b\bar{b}\gamma\gamma$ signal, and the signal with only the triangle diagram, and hence the trilinear coupling, contributing. In figure 2.4.6 we show the ΔR separation between the two photons and the minimum ΔR between a photon and a b-jet. The full process was generated using EFT in MADGRAPH5 [36]. Because the double Higgs boson signal was removed from it (by setting the relevant coupling constants to zero) we do not need to use exact calculations for it. It should be stressed that it still contains diagrams with a single Higgs boson present.

indirectly through their decay products.

[†]This includes for example particles being reconstructed wrongly, or mis-identified. It also includes detector resolution effects, which combined with cuts can lead to particles be included/excluded erroneously.

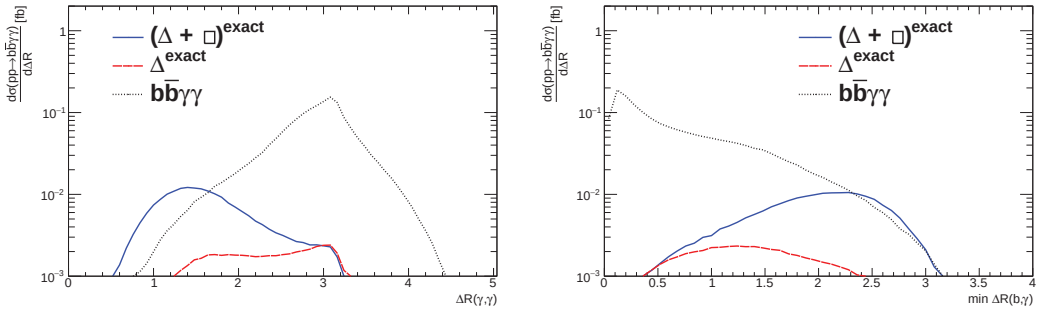

Figure 2.4.6: *Differential cross-section distribution of the separation in* $\sqrt{\Delta\phi^2 + \Delta\eta^2}$ *for the photon pair, and the minimum separation between a* b*-jet and a photon, for the* $pp \to b\bar{b}\gamma\gamma$ *process without double Higgs bosons (*$b\bar{b}\gamma\gamma$*), for the signal (*$(\triangle + \square)^{exact}$*), and for just the triangle contribution (*$\triangle^{exact}$*).*

It can be seen that the double Higgs production process favors smaller $\Delta R(\gamma, \gamma)$ because the decay of the heavy Higgs boson provides a significant boost to the photon pair. And this is less pronounced in the irreducible background where the photon pair might not form the decay product of a single heavy particle. In fact, it dominates the background for $\Delta R(\gamma, \gamma) < 1.6$. However, the triangle contribution prefers the region of $1.6 < \Delta R(\gamma, \gamma) < \pi$ because unlike the box diagram it has three Higgs boson propagators it needs to satisfy. It is not possible to satisfy all three at the same time, and when one of the outgoing Higgs bosons goes off-shell its decay products have close to zero total momentum resulting in a back-to-back decay.

Inspecting the second plot in figure 2.4.6 we see that the background maxes out at very small values of $min\,\Delta R(b, \gamma)$ because most of the photons are radiated off b-jets. The distribution is much more spread out for the signal since the b-jets and photons are guaranteed to come from different Higgs boson parents. The triangle contribution is more centralized again due to the triple propagator effect. Overall we can recognize that whilst the signal might be distinguishable from the irreducible background, the sensitivity to the trilinear coupling is often hidden underneath the bulk of the background.

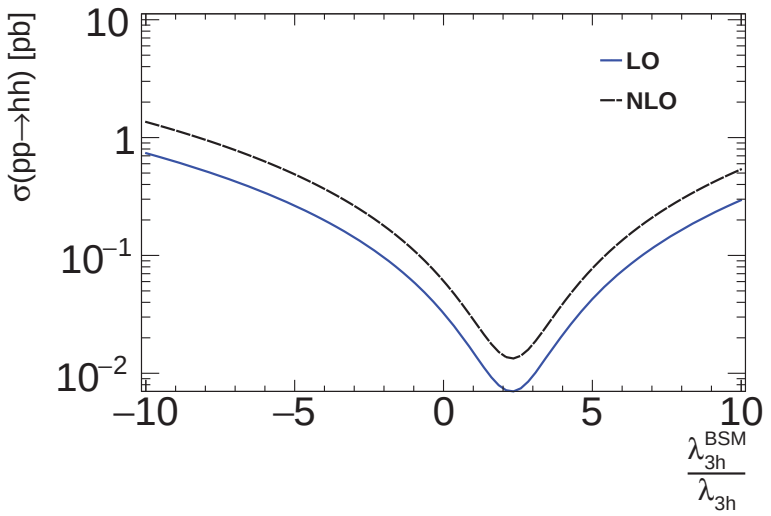

Figure 2.4.7: *Leading and next-to-leading cross-sections [29, 35] of* $pp \to hh$ *as a function of* ξ *at* $\sqrt{\hat{s}} = 14$ *TeV.*

2.4.3 Sensitivity to Non-Standard Model values of λ_{3h}

We saw in the preceding section that it is going to be very hard to kinematically optimize for sensitivity to the trilinear coupling since cuts designed to select for double Higgs boson production are typically not sensitive to this coupling, and the actual sensitive areas are below the background. In this section we are going to study the effects of λ_{3h} having non-Standard Model values. In what follows we shall use λ_{3h} to denote the Standard Model value and introduce λ_{3h}^{BSM} as a free parameter. For brevity we also introduce:

$$\xi \equiv \frac{\lambda_{3h}^{BSM}}{\lambda_{3h}} \tag{2.4.15}$$

where $\xi = 1$ implies λ_{3h}^{BSM} has the Standard Model value. The triangle contribution scales with ξ^2 because the coupling features once in the diagram and is then squared to get to the cross-section. The interference term between the triangle and the box scales only with ξ because the box diagram does not depend on ξ. As such the interference term flips sign if ξ does, meaning that the process will obtain a constructive interference and the cross-section will rise accordingly. At $\xi = 0$ there is no interference, and indeed no triangle diagram, which also results in a cross-section bigger than the Standard Model prediction. This can be seen in figure 2.4.7. It can also be noted from that figure that the cross-section is larger in the next-to-leading calculation with respect to the leading one. This is visualized further in figure 2.4.8. Around the Standard Model value the ratio between next-to-leading order and leading order cross-sections is about 1.92, but this decreases to 1.46 at $\sqrt{\hat{s}} = 100$ TeV. At 100 TeV the production cross-section is two orders of magnitude larger than at 14 TeV, but the sensitivity to ξ is diminished as can be seen in figure 2.4.9.

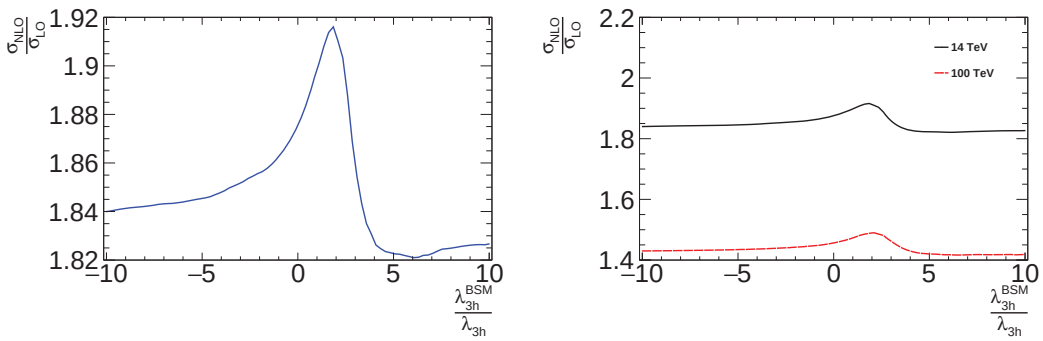

Figure 2.4.8: *Relative variation of the cross-section between leading and next-to-leading order calculations as a function of ξ at $\sqrt{\hat{s}} = 14$ TeV. On the right the variation is shown for two values of $\sqrt{\hat{s}}$. Cross-sections were obtained using* HPAIR *[46].*

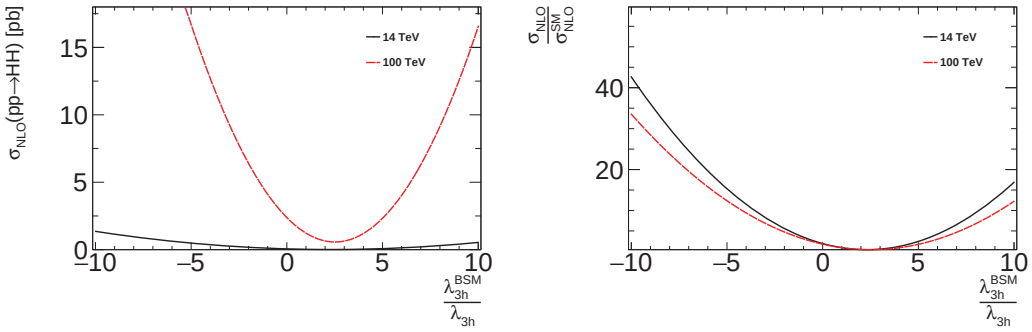

Figure 2.4.9: *The next-to-leading order cross-section using* HPAIR *[46] as a function of ξ for $\sqrt{\hat{s}} = 14$ TeV and for $\sqrt{\hat{s}} = 14$ TeV both as an absolute value and as a ratio with respect to the Standard Model prediction.*

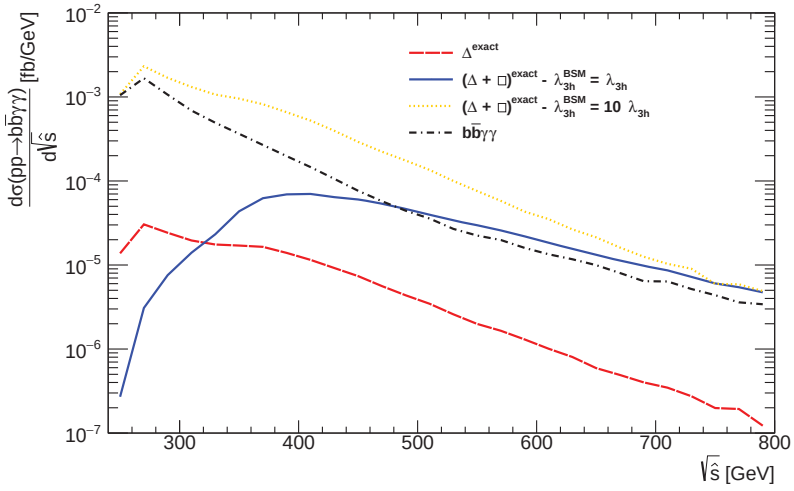

Figure 2.4.10: *Differential cross-sections for $\xi = 1$ and $\xi = 10$ as well as the triangle contribution and irreducible background as a function of $\sqrt{\hat{s}}$.*

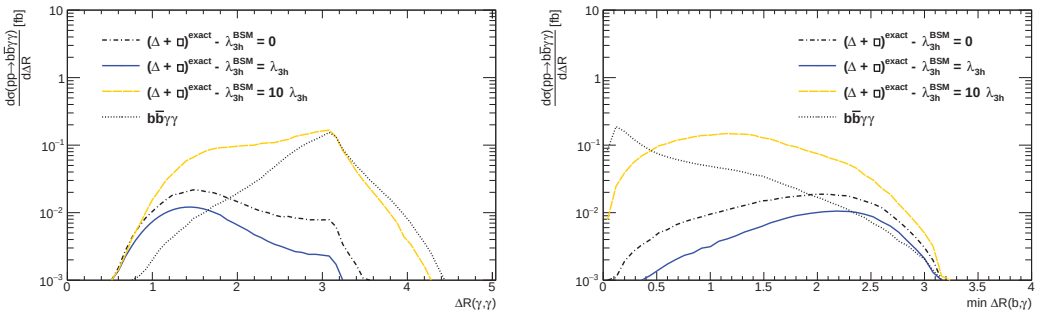

Figure 2.4.11: *Differential cross-sections as a function of ΔR separation between the photon pair, and the minimum ΔR separation between a photon and a b-jet, for different values of ξ and including the irreducible background.*

In figure 2.4.10 we return to the kinematic cuts of the previous section and examine the behavior of the differential cross-section when $\xi = 10$. This large increase from the Standard Model prediction is most pronounced at lower energy levels; at $\sqrt{\hat{s}} > 750$ GeV all sensitivity is lost because the triangle diagram becomes completely swallowed by the box contribution. Another unfortunate feature is that the largest increase from having $\xi = 10$ as opposed to $\xi = 1$ happens around $\sqrt{\hat{s}} = 270$ GeV, which is also where the largest irreducible background can be found. In figure 2.4.11 we again show the differential cross-sections as a function of $\Delta R(\gamma, \gamma)$ and $min\,\Delta R(b, \gamma)$, but this time for different values of ξ. This allows us to gauge where the sensitivity to ξ resides. Glancing at the plots we see that the sensitivity is located at the lower end of $\Delta R(\gamma, \gamma)$ and at the lower end of $min\,\Delta R(b, \gamma)$. However, at the higher end of $min\,\Delta R(b, \gamma)$ there is a potential to disentangle the signal from the irreducible background that is located mostly at the low end. For example, for $\xi = 0$ the signal surpasses the background when $min\,\Delta R(b, \gamma) > 2$.

Cut	B	$S_{\xi=1}$	$S_{\xi=10}$	$\frac{S_{\xi=1}}{\sqrt{B}}$	$\frac{S_{\xi=10}}{\sqrt{B}}$
No cuts	-	128	2280	-	-
C_K	166950	48	726	0.12	1.78
C_I	395	48	726	2.43	36.53
A_1	26	35	189	6.80	36.85
A_2	41	39	256	6.10	40.16
A_3	240	48	679	3.08	43.84

Table 2.4.4: *The number of events passing the cuts discussed in the main text at* $\sqrt{\hat{s}} = 14$ *TeV and a total integrated luminosity of* $3000\,fb^{-1}$ *for the signal,* S *for* $\xi = 1$ *and* $\xi = 10$*, and for the irreducible background,* B*.*

We thus propose the following cut variants:

A_1 : $\Delta R(\gamma,\gamma) < 2$, $\Delta R(b,\gamma) > 1$

A_2 : $\Delta R(\gamma,\gamma) < 2.2$, $\Delta R(b,\gamma) > 0.4$

A_3 : No $\Delta R(\gamma,\gamma)$ cut, $\Delta R(b,\gamma) > 0.4$

where cut A_1 is the one proposed in [44] to improve sensitivity to double Higgs boson production. However, as we have seen, this does not necessarily imply sensitivity to the trilinear coupling. Cut A_3 was chosen to optimize sensitivity to $\xi = 10$, and cut A_2 is an intermediate version. In table 2.4.4 we show how these cuts affect the signal-to-noise ratios for the measurement of double Higgs boson production. Cuts C_K (kinematic) and C_I (invariant mass) are those mentioned in the previous section:

C_K $p_T(b) > 45$ GeV, $p_T(\gamma) > 20$ GeV, $|\eta(b)| < 2.5$, $|\eta(\gamma)| < 2.5$, $\Delta R(b,b) > 0.4$, $\Delta R(\gamma,\gamma) > 0.4$

C_I $|m_{b\bar{b}} - m_h| < 20$ GeV, $|m_{\gamma\gamma} - m_h| < 2.3$ GeV

The background events could not be generated without the invariant mass cuts (C_K) due to phase-space singularities. From the table we observe that cut A_3 has a much better signal-to-noise ratio for the enhanced ξ, while cut A_1 performs better at the Standard Model value. We see that the cuts need to be optimized for the particular value of ξ one is interested in.

2.4.4 Conclusion

We investigated at the phenomenology and feasibility of measuring the Higgs trilinear self-coupling, λ_{3h}, especially for the $b\bar{b}\gamma\gamma$ production channel from gluon fusion. As can

be seen from table 2.4.3 on page 35 we only have 0.5 expected events in the 2012 ATLAS dataset, but it should be mentioned that an ATLAS analysis examining double Higgs boson production in this channel found a 2σ excess [40]. In any case, due to the low statistics, a measurement of λ_{3h} is going to depend crucially on knowledge of the precise kinematics in order to be able to cut away as much as possible from the background whilst retaining the signal. In this spirit we show that effective field theory is woefully inadequate, even though it might get the total cross-sections correct. At energy levels around double the top quark mass large deviations are seen in the kinematics.

We also analyzed the triangle and box diagrams and how their relative contributions depend on kinematics. We saw that the sensitivity to λ_{3h} is located around the single Higgs mass peak, but that the contribution of this region to the overall cross-section is beyond tiny. However, it is advised for any analysis hoping to explore λ_{3h} to restrict themselves to lower values of $\sqrt{\hat{s}}$. Finally we analyzed a system of cuts proposed in [44] to study the double Higgs boson signal in $b\bar{b}\gamma\gamma$, and found that an improvement could be made if one is interested in measuring λ_{3h} and not only detecting double Higgs boson production.

In addition we have shown that $pp \to b\bar{b}\gamma\gamma$ is a promising channel, and it is therefore useful to study it as a background to a future measurement. With this in mind we will move to the last section of this chapter and examine it.

2.5 Theory predictions for the $pp \to b\bar{b}\gamma\gamma$ channel

In this section we shall investigate what the Standard Model has to say on the topic of $pp \to b\bar{b}\gamma\gamma$. We will use a Monte Carlo event generator to produce relevant distributions and we will later on come back to these when we compare the analysis results with the theory predictions at the end of chapter 6.

Leading order (LO) theory predictions were obtained using a combination of MADGRAPH5 [36] and MC@NLO [47] and using the CT10 parton distribution functions [48]. All Standard Model processes were included, which means that processes involving the Higgs boson were also included, but these contribute only marginally due to the small cross-section of such processes. The simulation was restricted to parton level, no particles were decayed or hadronized, and was based on 250.000 events. Next-to-leading order (NLO) simulated events were also generated with the same tools but using the NN23NLO parton distribution functions [49], and, owing to the longer time needed to generate NLO events, only 50.000 were generated. Both LO and NLO events are used later in a comparison with the analysis results.

The fiducial volume for the analysis is defined by the following cuts:

Photons At least two photons with $p_T > 22$ GeV and $|\eta| < 2.37$.

Jets At least two jets with $p_T > 25$ GeV and $|\eta| < 2.5$. This applies equally to b-jets and other jet types.

Diphoton invariant mass The diphoton invariant mass for the two leading photons needs to satisfy $m_{\gamma\gamma} > 60$ GeV.

Separation We require $\Delta R(\gamma,\gamma) > 0.4$ for the two leading photons, $\Delta R(j,j) > 0.4$ for the two leading jets, and $\Delta R(\gamma,j) > 0.6$ for any combination of the two leading jets with the two leading photons.

This volume corresponds with the fiducial cuts applied to the data, which are applied for various reasons that are explained in sections 5.1 and 5.2. Generally speaking this volume was chosen to be close to various experimental cuts that were needed to reduce detector effects, the need to simplify the unfolding process, and in the case of the invariant mass the pragmatic reason that no simulations were available lacking this cut. In most cases the cuts follow another analysis [50] that studied diphoton events with a variable number of jets. To assess the effects of this choice of cuts we also generated 50.000 NLO events using a much looser variety of cuts where the transverse momentum requirements were lowered to 11 GeV and the $|\eta|$ range was increased to 2.5 for photons[†]. In addition the photon invariant mass cut was removed. All the angular separation requirements were lowered to only 0.05. This is listed as "small cuts"[‡] in the plots.

It is also interesting to contemplate the difference between the $pp \to b\bar{b}\gamma\gamma$ process and the related $pp \to jj\gamma\gamma$ process, where j stands for a gluon or any quark except for a top or bottom one. As we will see in chapter 6 the unfolding will need to be performed on $pp \to jj\gamma\gamma$ due to lack of suitable simulation samples so one needs to assess whether there are any large kinematical differences between these two processes. Although of course the ultimate test is seeing if the unfolding itself is sensitive to such differences, but that will be investigated at in section 6.4.2. For now we show the cross-sections thus obtained in table 2.5.1. It is immediately clear from the numbers in that table that there are significant k-factors at play here with NLO cross-sections larger than the LO ones; this reflects the theoretic difficulty this channel has, and can be taken as additional motivation to spend time to measure it experimentally. Normalized distributions comparing $pp \to b\bar{b}\gamma\gamma$ with $pp \to jj\gamma\gamma$ and the small cut versions are shown in figures 2.5.1 to 2.5.4. From these figures it can be seen that the shapes of most distributions are similar between $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$, a given that is important later on in the main analysis.

In a single proton-proton interaction it is in rare cases possible to have more than one parton-parton interaction. Such an event is known as a multiple parton interaction (MPI), the most common of which is known as a double parton interaction (DPI). A potential background to our $pp \to b\bar{b}\gamma\gamma$ process is therefore formed by a DPI event where we have one $pp \to \gamma\gamma$ interaction and one $pp \to b\bar{b}$ one. We can estimate the cross-section for this process by considering the (single parton-interaction) cross-sections of these two processes separately and combining them using the following formula [51]:

$$\sigma_{DPI} = \frac{\sigma_{pp\to\gamma\gamma}\sigma_{pp\to b\bar{b}}}{\sigma_{eff}} \tag{2.5.1}$$

Here σ_{eff} is the so-called effective cross-section that deals with the matter distribution inside a proton. Its experimental value[⊙] lies around 15 mb [52]. Finding fiducial cross-sections around 10.24 pb and $6.00 \cdot 10^5$ pb for $pp \to \gamma\gamma$ and $pp \to b\bar{b}$ respectively using

[†]The reason for the larger range for non-b-jets is that the ATLAS inner detector is not crucial for the reconstruction of such jets. See chapter 3 for more information.

[‡]Some cuts are always needed in event generators to prevent any singularities in the cross-section from dominating.

[⊙]It is devilishly difficult to calculate from first principles.

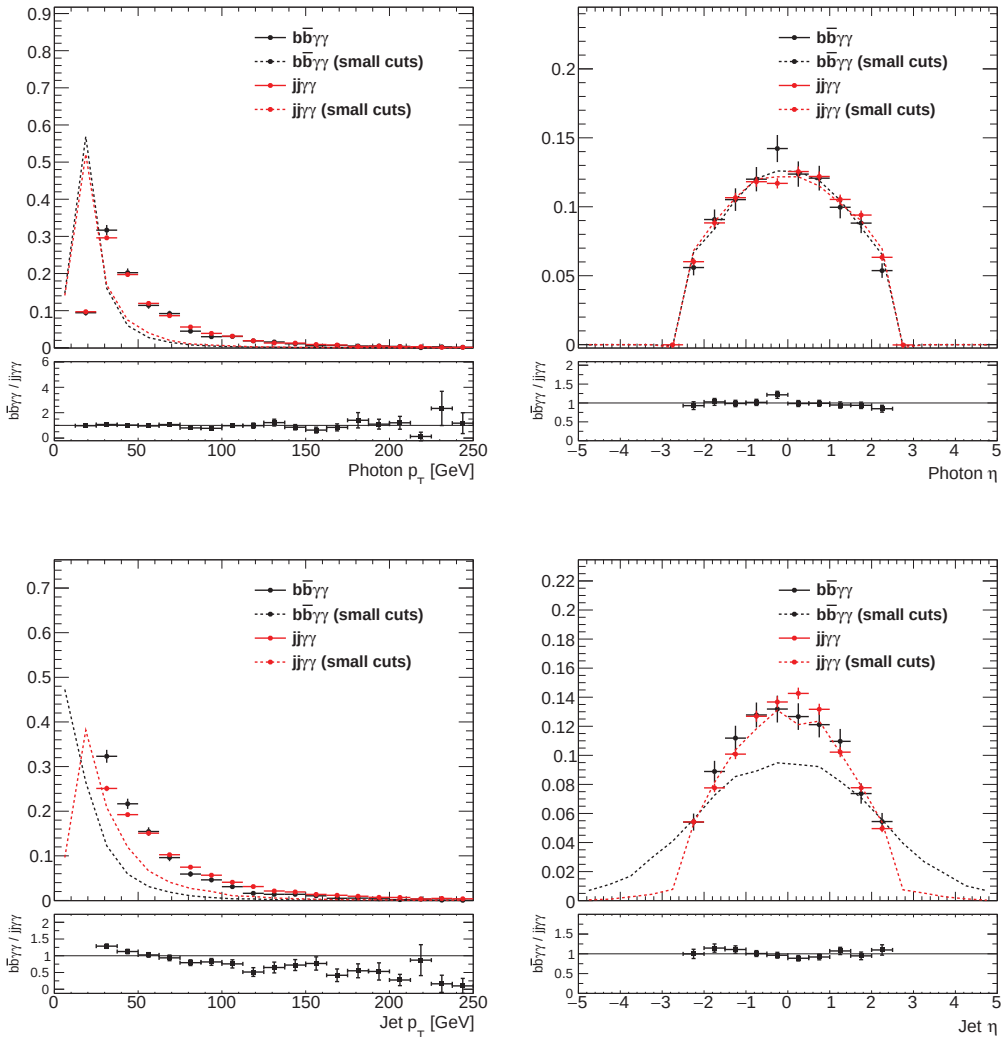

Figure 2.5.1: *Normalized production cross-sections as a function of the transverse momentum and pseudorapidity of the jets and photons for the* $pp \to b\bar{b}\gamma\gamma$ *and* $pp \to jj\gamma\gamma$ *processes. In the case of* $b\bar{b}\gamma\gamma$ *both the* b *and* $\bar{b}$ *are classified as a jet. For readability the small cuts versions have their error bars suppressed. Below each graph a ratio plot shows the difference between* $pp \to b\bar{b}\gamma\gamma$ *and* $pp \to jj\gamma\gamma$.

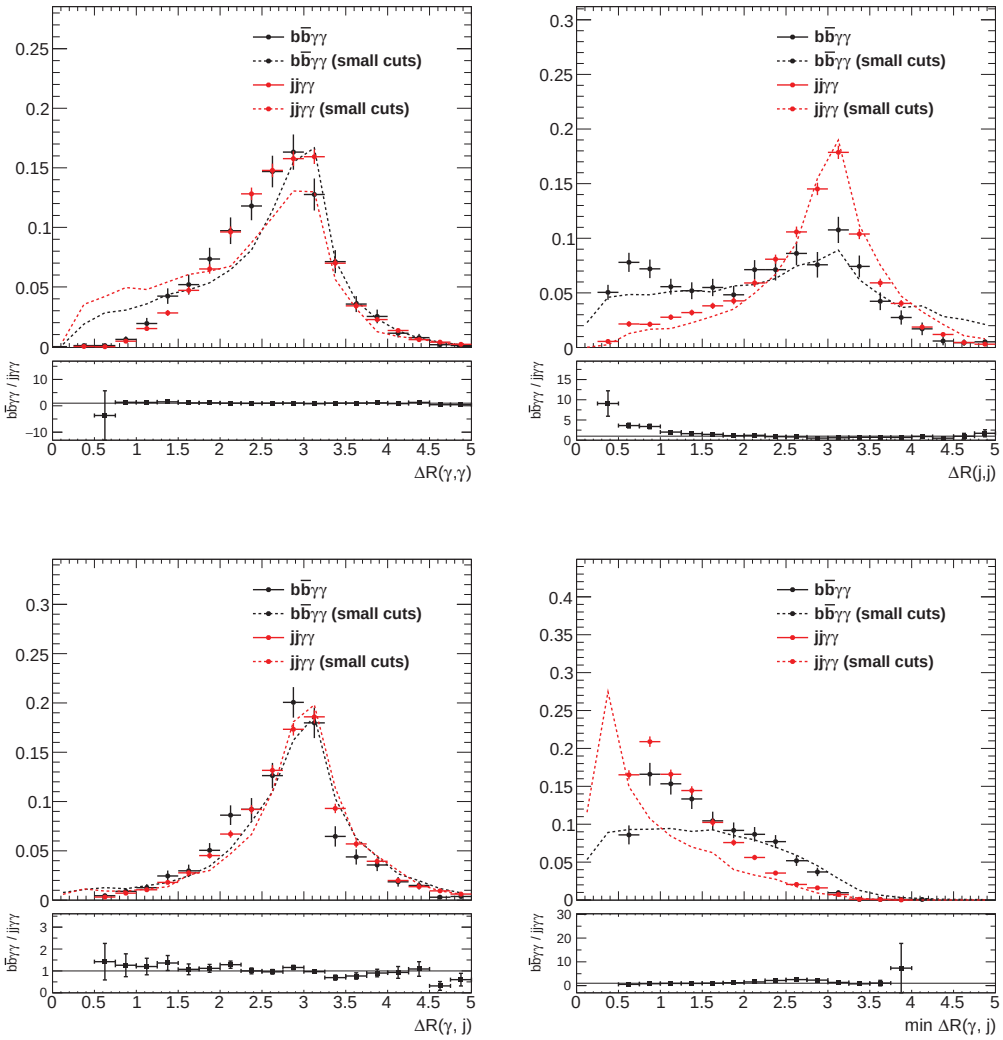

Figure 2.5.2: *Normalized production cross-sections in terms of the angular distances between the photons and jets. The top two plots show the differences for between the photons and jets. The bottom left plot shows the angular difference between the leading photon and the leading jet. The bottom right plot shows the angular difference between the photon and jet pair that are the closest together in that event. For readability the small cuts versions have their error bars suppressed. Below each graph a ratio plot shows the difference between $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$.*

Process	Order	σ[pb]
$pp \to b\bar{b}\gamma\gamma$	LO	$0.0138 \pm 3.04 \cdot 10^{-5}$
$pp \to b\bar{b}\gamma\gamma$	NLO	0.0347 ± 0.00124
$pp \to b\bar{b}\gamma\gamma$ (small cuts)	NLO	4.96 ± 0.0413
$pp \to jj\gamma\gamma$	LO	2.53 ± 0.00576
$pp \to jj\gamma\gamma$	NLO	3.43 ± 0.0518
$pp \to jj\gamma\gamma$ (small cuts)	NLO	150 ± 1.80

Table 2.5.1: *The partonic cross-sections as obtained from simulations described in the main text. The quoted uncertainty is purely statistical in nature and is based on 250.000 simulated events for LO and 50.000 for NLO. Note the large k-factor between LO and NLO cross-sections.*

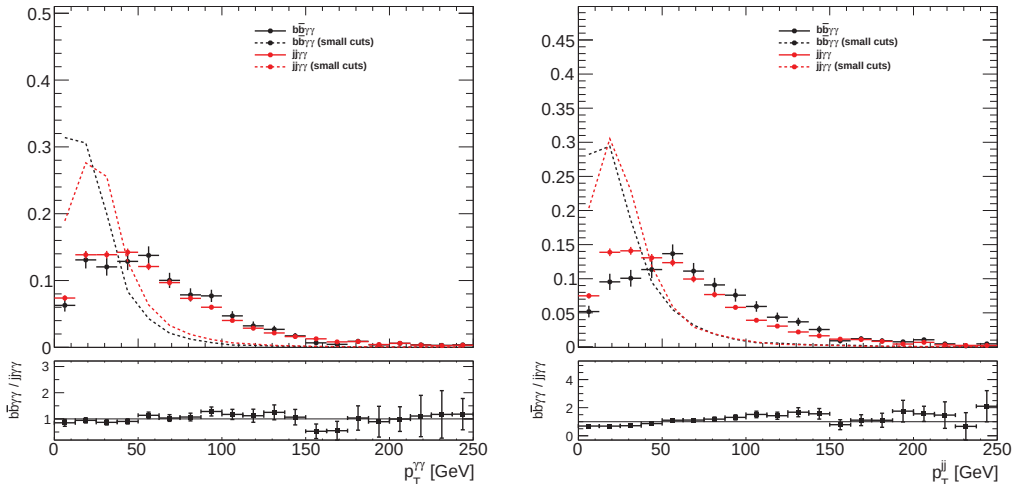

Figure 2.5.3: *Normalized production cross-sections as a function of the transverse momentum of the diphoton and dijet pair for the $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$ processes. In the case of $b\bar{b}\gamma\gamma$ both the b and $\bar{b}$ are classified as a jet. For readability the small cuts versions have their error bars suppressed. Below each graph a ratio plot shows the difference between $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$.*

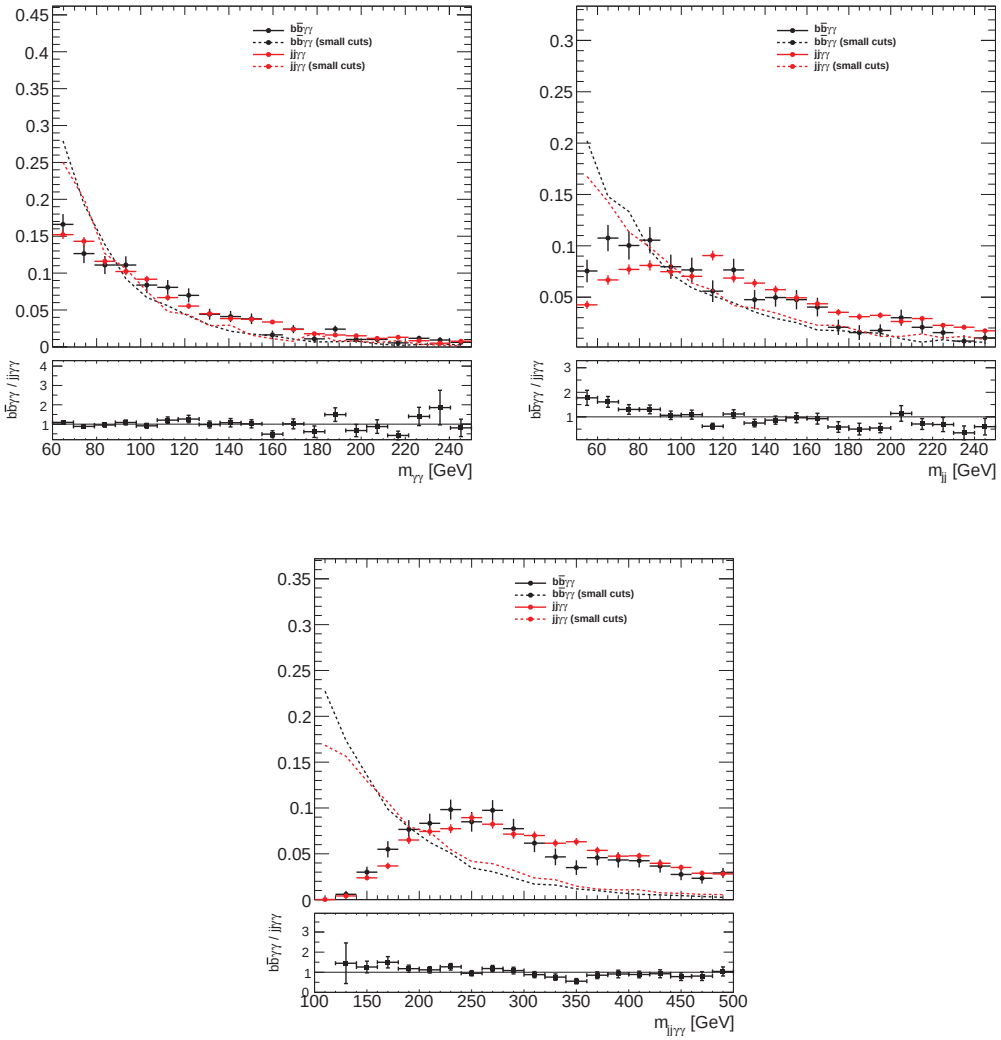

Figure 2.5.4: *Normalized partonic cross-sections as a function of $m_{\gamma\gamma}$, m_{jj}, and $m_{jj\gamma\gamma}$. For readability the small cuts versions have their error bars suppressed. Below each graph a ratio plot shows the difference between $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$.*

LO Madgraph5 simulations and the NN23NLO parton distribution functions we can estimate the DPI cross-section to be around $4.07 \cdot 10^{-4}$ pb. This is much smaller than the cross-section we find for the nominal $pp \to b\bar{b}\gamma\gamma$ process and we therefore will neglect DPI in the rest of this thesis.

This is the end of our discussion on the theory and motivations. We will now proceed with two chapters dedicated to describing the ATLAS detectors and the LHC; the tools with which the data used in this thesis was collected. Whilst the first of these two chapters focuses on the hardware, the second one will focus on the software, and describe the reconstruction software, jet clustering algorithms and ways to find out of a jet finds its origin in a b-hadron or not. The latter is important to us since we need to find at least two such jets in our events. After that we move on to the two chapters describing the actual analysis.

CHAPTER 3

The Large Hadron Collider and the ATLAS experiment

"No amount of experimentation can ever prove me right; a single experiment can prove me wrong."

- Albert Einstein

THE most powerful particle accelerator and physics experiment, and also the largest machine humanity has ever built, is located on the Franco-Swiss border near Geneva. It is housed at CERN, the European Organization for Nuclear Research[†], which was established in 1954. CERN represents a collaborative enterprise spanning the generations and scientists from virtually every nation on Earth partake in its many research endeavors. Many are CERN's accomplishments, from the discovery of the W and Z bosons in 1983 to the creation of the first anti-hydrogen atoms in 1995.

In this thesis we shall be concerned with Large Hadron Collider (LHC) and one of the four principal experiments located there, the ATLAS experiment. The other three experiments are CMS, which is similar to ATLAS in being a general-purpose detector, LHCb, mainly concerned with processes involving b-quarks and is sensitive to various CP-violating processes, and ALICE, which investigates quark-gluon plasmas, which are thought to have dominated the Universe in its very early stages.

In this chapter we will discuss the CERN accelerator complex that is being used to accelerate protons, and the ATLAS detector itself, which has been designed to investigate high-energy collisions between accelerated protons. Data from the ATLAS detector is used throughout this thesis.

[†]The abbreviation "*CERN*" derives from the French "*Conseil Européen pour la Recherche Nucléaire*".

3.1 The CERN accelerator complex and the LHC

CERN hosts several accelerators supplying particles to a variety of experiments, as can be seen in figure 3.1.1. Protons enter the LHC via a supply chain which starts with hydrogen atoms that are stripped from their electrons. They are then accelerated to an energy of 50 GeV in the LINAC 2, a linear accelerator. After this they are injected into the first and smallest circular accelerator called the Proton Synchrotron Booster, which brings their energy to up 1.4 GeV. Then they are moved to the Proton Synchrotron (PS), which accelerates the protons to an energy of 26 GeV, after which the Super Proton Synchrotron (SPS) increases this to 450 GeV. The protons then arrive in the LHC where they are accelerated from an insertion energy of 450 GeV to a final energy of 6.5 TeV. The entire chain of accelerators can also be used to accelerate lead ions.

The LHC is 27 km in circumference and lies underground at a depth ranging from 50 to 175 meters depending on the landscape. It occupies a tunnel that was previously used by an electron and positron accelerator called the Large Electron-Positron collider (LEP) which was decommissioned in 2001 to make room for the LHC. Inside the LHC protons are circulated both ways around the ring in two separate vacuum pipes. Thousands of superconducting magnets restrict a beam to a pipe and keep it focused. The cooling of the magnets happens with helium at an operating temperature of 1.9 K and makes the LHC the largest cryogenic machine using helium in the world. The protons are accelerated using radio-frequency cavities (RF-cavities) which contain oscillating electric fields. The idea of these fields is that protons that are going too slow are accelerated, and protons that are going too fast are decelerated. The combined effect is a restorative force that keeps the protons going at the ideal energy (to which the magnetic fields are dynamically tuned).

In a head-on proton-proton collision[†] the total center of mass energy is 13 TeV. However, the data used for this thesis was obtained in 2012 when the proton beams had an energy of only 4 TeV each, implying a total center of mass energy of 8 TeV.

The beams are organized in bunches, collections of protons traveling the ring together, and each bunch contains around a hundred billion individual protons. The protons are kept together inside a bunch via the RF-cavities described above. It should also be noted that bunches are organized in bunch trains, with every train being spaced out more than the individual bunches are. Typically protons are injected into the LHC as a train of bunches rather than as individual bunches. The separation between each bunch in arrival time at a collision point is 25 ns, but it was 50 ns in 2012 when the data used in this thesis was recorded. At the four collision points at the LHC, located at the sites of each of the four major experiments, the beams are squeezed together and made to collide. Since protons are very tiny particles only very few of the billions of protons in each bunch actually collide, most of them continue on and get the opportunity to collide later on. In 2012 the average number of inelastic proton-proton collisions per bunch crossing was about 30 as can be seen on the left in figure 3.1.2. Because of this the LHC can run for many hours before it needs to refill its supply of protons[‡].

[†]Note that the collision will not be exactly head-on. The beams are tilted slightly to make them collide, since the two beams do not share the same vacuum pipe.

[‡]And even then the LHC will not be empty. It is just that by depleting the bunches a bit the average number of collisions per bunch crossing decreases and at some point it is more efficient to refill rather than to continue.

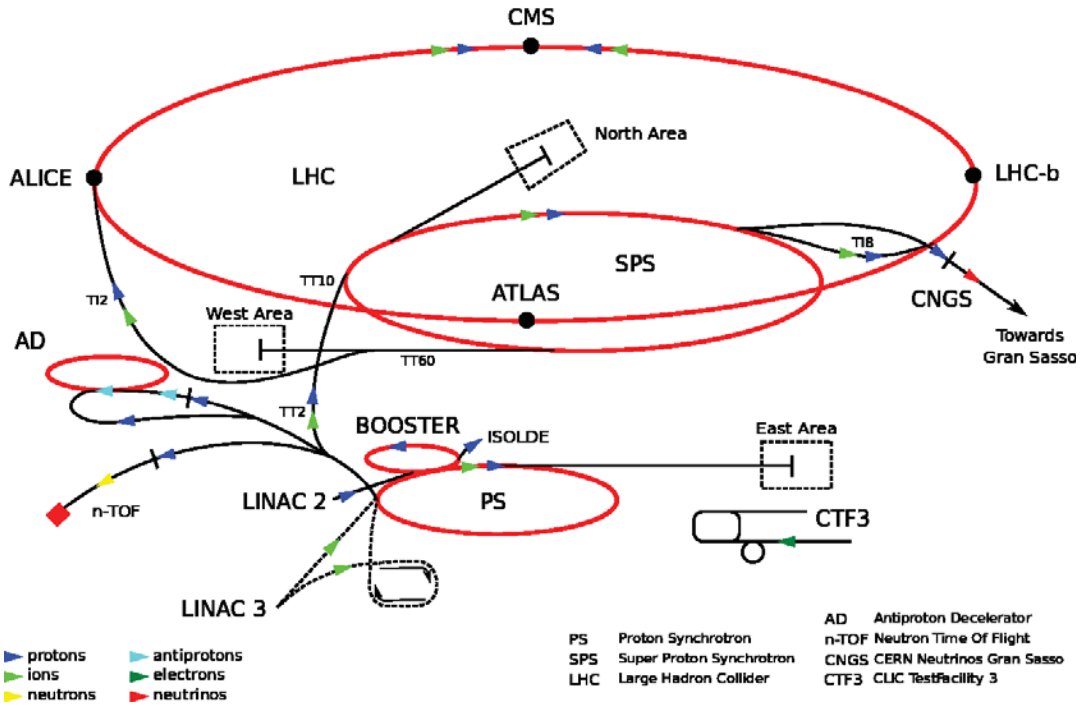

Figure 3.1.1: *The accelerator complex at CERN. The LHC is visible at the top and houses four major experiments: ALICE, ATLAS, CMS, and LHCb. The color of the triangles on the lines denote the types of particle being produced / accelerated at the various facilities.*

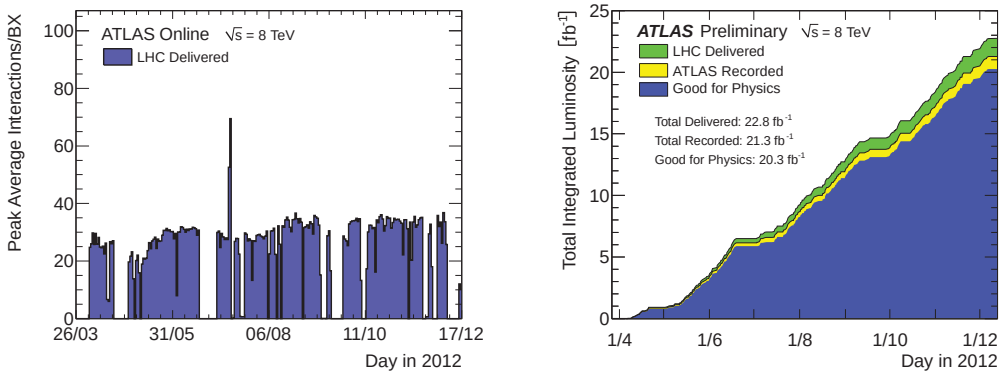

Figure 3.1.2: *On the left, the average number of proton-proton interactions per bunch crossing during 2012 as seen by ATLAS. On the right, the integrated luminosity delivered by the LHC, the amount recorded by ATLAS, and the amount suitable for physics analysis. Both plots were taken from [53].*

In 2012 the total amount of data collected by the ATLAS detector amounts to a total integrated luminosity of 22.7 fb^{-1}, of which 20.3 fb^{-1} is usable for physics analyses [54, 55]. This is visualized in the right plot of figure 3.1.2. The main reason for data not being able to be used is because parts of the detector were switched off or malfunctioning during the data taking. Another reason could be that the magnetic field was switched off.

3.2 The ATLAS detector

THE ATLAS [56] detector derives its moniker from the very original and clever acronym of "A Toroidal LHC ApparatuS"[†]. It is the largest detector at CERN measuring 46 meters in length and 25 meters in height. It weighs 7 million kilogrammes, and lies at a depth of a 100 meters below the surface, on the other side of the street from the CERN Meyrin site, next to the Globe of Science and Innovation[‡]. Like a proverbial onion the ATLAS detector consists of multiple specialized layers surrounding a collision point. Each layer is geared towards the detection of a certain class of particles or determining certain characteristics of particles traversing it. Because of the size of ATLAS and indeed of each layer, each layer can be regarded as a subdetector that is mostly self-contained. By combining all these subdetectors it is possible to detect a wide range of particles allowing for a good reconstruction of a collision event. A schematic overview of ATLAS can be found in figures 3.2.1 and 3.2.2.

The innermost part of the detector is aptly known as the inner detector whose task it is to measure the direction, momentum, and charge of particles. As we shall see in its specific section later on it is actually made up from three smaller sections which are the pixel detector, the semi-conductor tracker, and the transition-radiation tracker. The inner detector surrounds the beam pipe in which the protons travel and collide. Outside the inner detector are two calorimeters that specialize in measuring the energy of particles by estimating how much energy they lose traversing the calorimeter material. It consists of two separate parts: the liquid argon calorimeter which is meant mostly for photons and electrons and the tile hadronic calorimeter which, as the name implies, is optimized to measure the energy of hadronic particles. Most particles do not survive beyond the calorimeters, only neutrinos (which cannot be measured by ATLAS directly) and muons can escape. The muons are taken care of by the outermost layer of ATLAS, the muon spectrometer, which measures the momenta of said muons. The detector is also embedded in a powerful magnetic field that curves the trajectory of charged particles, which aids in the determination of the momentum of these particles. The magnet system consists of two parts which produce respectively a solenoidal and a toroidal magnetic field.

The ATLAS detector can expect to witness around one billion proton-proton collision events every second, which means every subdetector has to be able to deal with the influx of large amounts of particles, and also to be resistant to damage caused by radiation. Storing all the information from every such event is impossible; the needed data

[†]The wonderful world of experimental physics is full of exciting acronyms like "OverWhelmingly Large telescope" (OWL) and "A Large Ion Collider Experiment" (ALICE).

[‡]The Globe, incidentally, gives a nice idea of the scale of the ATLAS detector. With its height of 40 meters and width of 27 meters it is comparable in size to the ATLAS detector.

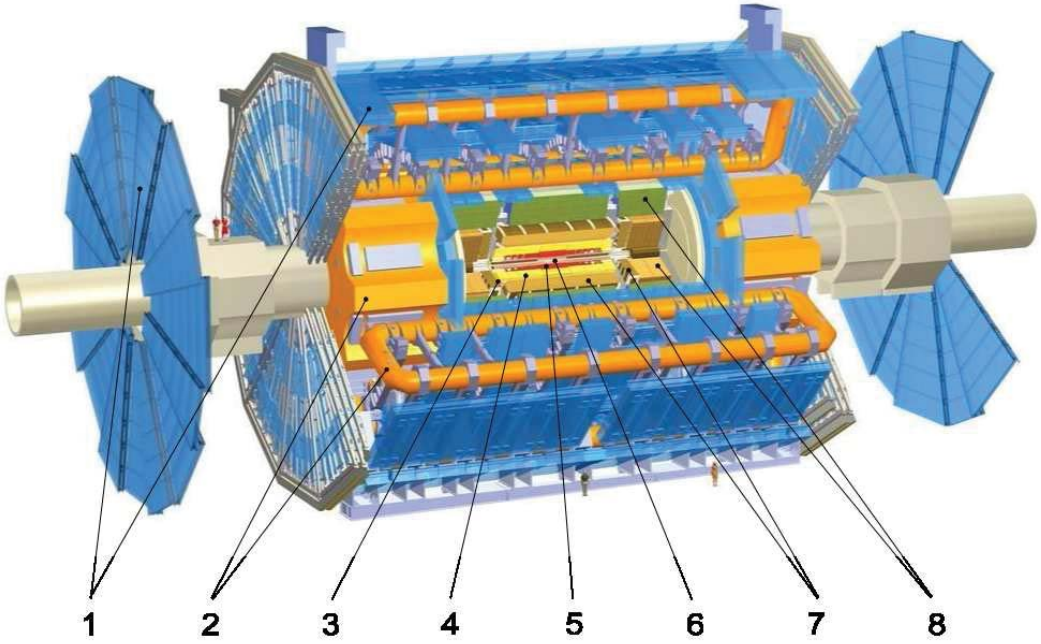

Figure 3.2.1: *A computer drawing of the ATLAS detector. 1) The muon spectrometer, 2) the toroidal magnet system, 3) the solenoidal magnet system, 4) the transition-radiation tracker, 5) the semi-conductor tracker, 6) the pixel detector, 7) the liquid argon calorimeter, and 8) the tile hadronic calorimeter.*

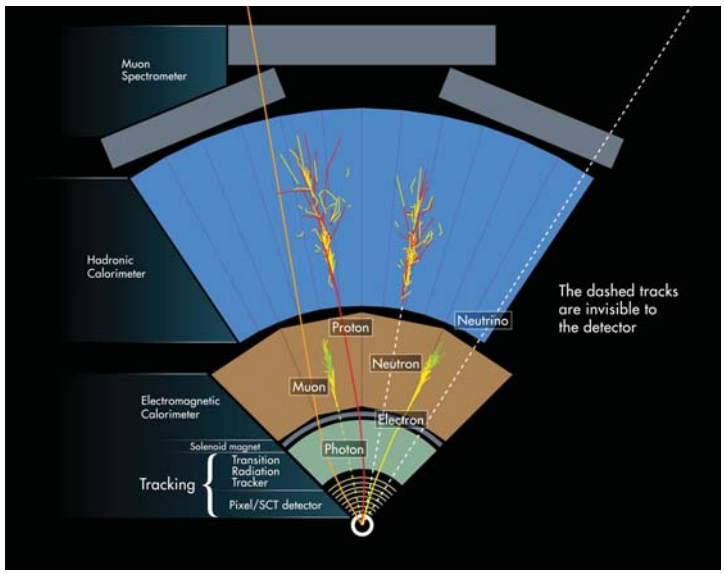

Figure 3.2.2: *A computer drawing showing various detector components and where different particles are detected.[57]*

rate would somewhere around 60 terabytes every second[†]. Evidently we need make a choice and not store all events, but only a subset, preferably a subset containing events with interesting characteristics. This choice is made by a trigger system that consists of multiple levels. These levels are needed because the decision making process takes time. And when you have a billion collisions happening each second you do not have much time to decide on each and every one of them if they are interesting enough to store or not. The first level, known as the Level-1 Trigger needs to make the quickest decisions and does this in less than two microseconds after each event. There is no time to integrate information from all subdetectors and as such only partial information is available to the trigger algorithm. Therefore, the Level-1 Trigger can only make rough cuts, but still manages to reduce the event rate to around a hundred thousand per second. The Level-2 Trigger has a bit more time due to the reduced data rate of events surviving Level-1. It still can only use partial data from an event, but due to the bigger time allotment is able to make more informed decisions. It cuts the event rate down to a few thousand per second. Finally the Level-3 Trigger, which runs on computer cluster outside the actual detector, has access to the information coming from all the subdetectors and can therefore reconstruct events completely. Only about two hundred events survive Level-3 and only those events are stored. A lot of care has to be taken when designing the trigger algorithms; mistakes can cause irreversible data loss, an event that was not stored is lost forever.

We shall discuss the various parts of the detector in more detail in the several dedication sections of this chapter, but it is useful first to discuss the coordinate system employed by ATLAS so it will be more convenient to relay information about the various sizes of the detector components. The origin of this system lies in the center of the detector, in the middle of the beam pipe in which the collisions occur. The z-direction is defined as running alongside the beam pipe and is tangent to the LHC ring. It follows the tradition in mathematics in that it is positive when pointing in a counter clockwise fashion, taking a bird's eye view of the LHC. The positive x-direction points to the center of the LHC, the positive y-direction points upwards away from the Earth's center. The x-y plane is called the transverse plane, being transverse to the beamline. Because the momentum along this transverse plane is the easiest to measure, it is often convenient to adopt a cylindrical coordinate system where the azimuthal angle in this transverse plane is called the azimuthal angle, ϕ, where $\phi = 0$ corresponds to the direction of the positive x-axis. The direction away from the transverse plane is called the polar angle, θ, which is the angle as measured from the positive z-axis.

In analyses we often use the transverse momentum, which is simply the momentum as measured in the transverse plane and which is denoted by p_T. Another thing that is often used instead of the polar angle is the pseudorapidity, as shown in figure 3.2.3, and which is defined as:

$$\eta \equiv -\ln\left(\tan\frac{\theta}{2}\right) \tag{3.2.1}$$

Pseudorapidity is related to a quantity called rapidity which is used in the theory of

[†]To give an indication of how much data 60 TB represents, this is 2,400 single layer Blue-ray discs or more than one million hours of music. As another indication of the scale: a 2010 study[58] showed that the total information found in all the 26 million books in the world's largest library, the US Library of Congress, amounts to only 10 TB. ATLAS deals with about six such libraries worth of information *every* second.

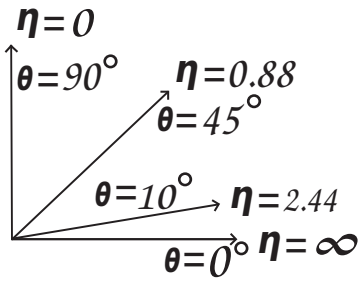

Figure 3.2.3: *Relation between the pseudorapidity and the polar angle. A pseudorapidity of infinity represents the direction along the beam pipe.*

special relativity, and which is defined as:

$$w = \frac{1}{2}\ln\frac{E + |\vec{p}|\,c}{E - |\vec{p}|\,c} \tag{3.2.2}$$

where E is a particle's energy, $\vec{p}$ its momentum, and c the speed of light in vacuo. Its use is mainly derived from the fact that, when applied in a single direction, it is additive. So whereas adding two velocities in special relativity requires one to use the relativistic velocity addition formula, adding their rapidities is straightforward:

$$w_{1+2} = w_1 + w_2 \tag{3.2.3}$$

The rapidity of light and other massless objects is infinity which can be easily seen when one applies the energy-momentum relation that holds for massless particles:

$$E = |\vec{p}|\,c \tag{3.2.4}$$

Which causes the denominator in the natural logarithm to go to zero. In particle physics the rapidity one uses is usually the one along the z-direction and is denoted by a different letter, namely y:

$$y = \frac{1}{2}\ln\frac{E + p_z c}{E - p_z c} \tag{3.2.5}$$

Apart from the preference for the z-direction, along the beam pipe, note also that y no longer has the property that massless particles always have an infinite rapidity. Instead this is only true if such a massless particle travels along the beam pipe and can actually be negative infinity in the case where it flies in the negative z-direction. For massless particles y depends only on the polar angle, and no longer on the actual energy of the particle. This can be seen by substituting in the before-mentioned energy-momentum relation for massless particles:

$$y = \frac{1}{2}\ln\frac{|\vec{p}|\,c + p_z c}{|\vec{p}|\,c - p_z c} \tag{3.2.6}$$

After which one easily obtains expression 3.2.1. To move to η for massive particles one pretends a particle is massless and use the same formula. Finally there is a variable, ΔR, typically called the angular distance or angular separation between two particles and which was already defined in expression 2.4.14 on page 32. It is helpful to realize that one can obtain a cylindrical coordinate system using ϕ and η. In such a system ΔR forms the shortest distance on this cylinder between two points.

3.2.1 Inner detector

The inner detector is responsible for tracking and identifying charged particles very close to the interaction point. It is also very important in finding the vertices these particles were created in, which is especially relevant for b-tagging since b-hadrons live long enough to form a secondary vertex a bit removed from the primary proton-proton interaction point when decaying. In fact this is one of the distinguishing features of b-quarks with respect to the heavier top quark, and the lighter up, bottom, and strange quarks.[†] The inner detector is embedded in a magnetic field of 2 T that points along the beam-axis so that particles traveling outwards into the detector are bent by a Lorentz force into a spiral whose curvature allows one to infer the particle's momentum[‡]. The inner detector has a cylindrical shape and is around 7 meters in length and has a radius of 1.15 meters, but part of that is taken up by the cryostat. It consists of three separate systems: the pixel detector, the semi-conductor tracker (SCT), and the transition radiation tracker (TRT), which are arranged as the shells of an onion around each other, as can be seen in the overviews found in figures 3.2.4 and 3.2.5. Each system has a barrel region for the center, and an end-cap region to detect particles with a high pseudorapidity. The inner detector is capable of tracking particles up to a pseudorapidity of 2.5 as can also be seen in figure 3.2.5.

Pixel detector

The pixel detector is the innermost part of the inner detector and consists of three layers and two end-caps that each also consist of three layers, see figure 3.2.5. The first layer is located at a distance of 50.5 mm from the center of the beampipe, and the final one 122.5 mm. The end-caps start 495 mm away from the detector's center. The detector consists of 1744 silicon sensors in total of which 288 are located in the end-caps. All of these are identical and have a size of 19x63 mm^2 and each contain 47232 pixels, most of which have a size of 50x400 mm^2. As can be seen in figure 3.2.4 the sensors are angled slightly and overlap to make sure that no particles can slip through between them. The accuracy of the pixel detectors is 14x115 μm^2 for both the barrel and the end-caps.

Whilst not relevant for the analyses discussed in this thesis, because they use 2012 data, it is worth mentioning that the pixel detector was upgraded in between Run 1 and Run 2 by adding an additional layer of pixel detectors. This layer is known as the insertable B-layer (IBL) [59] and will greatly help in making b-tagging more efficient.

Semi-conductor tracker (SCT)

The semi-conductor tracker is the next layer in the inner detector. It consists of 4088 silicon sensors spread out over four separate layers in the barrel region and 18 end-cap discs. The sensors are arranged in such a way that each charged particle is expected to leave four hits as long as its pseudorapidity is below 2.5. Just as with the pixel detector the sensors are angled to provide hermetic coverage. The accuracy in the $R - \phi$ (lateral) plane is about 17 μm, in the direction along the z-axis the resolution is 580 μm.

[†]Hadrons formed from charm quarks mimic this longevity that b-quarks enjoy, making them the most tenacious background to finding and identifying b-jets.

[‡]Particles with high momentum curve less than particles with a low momentum; for a given magnetic field strength the momentum is proportional to the radius of curvature.

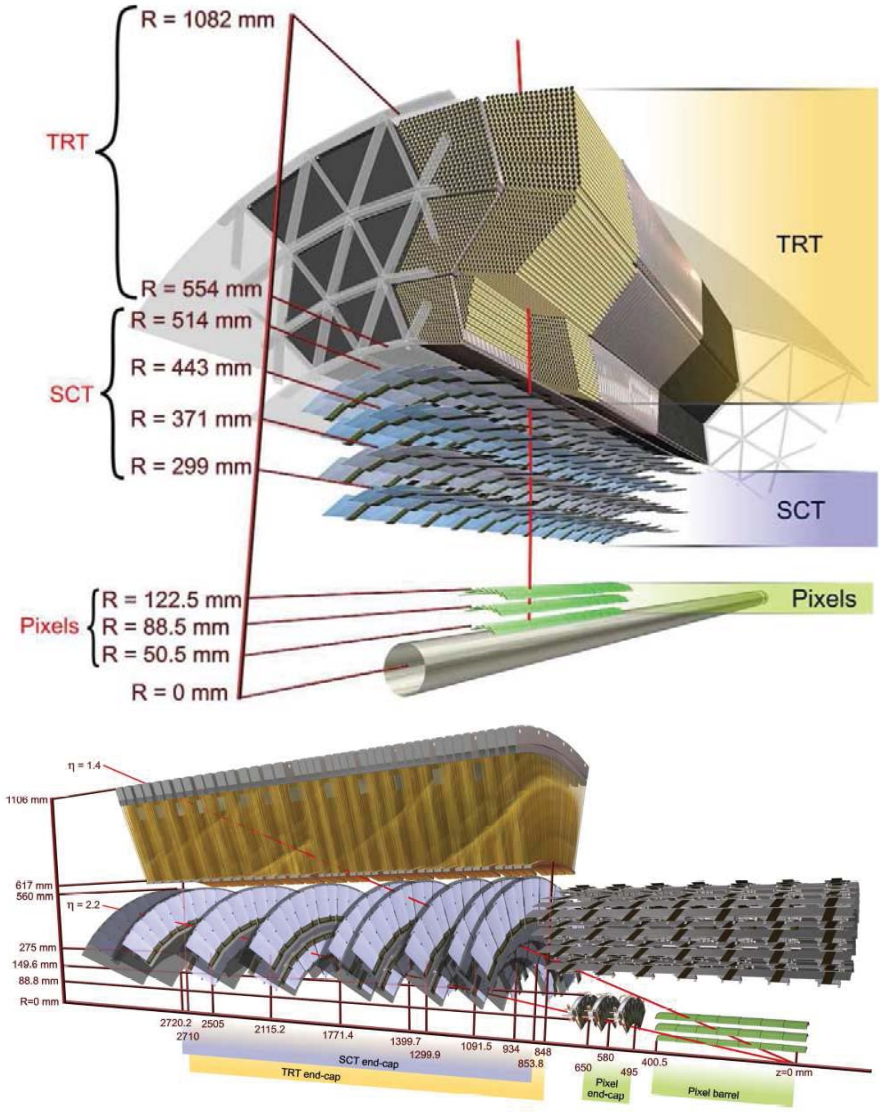

Figure 3.2.4: *Overview of the ATLAS inner detector. Shown are the pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT).*

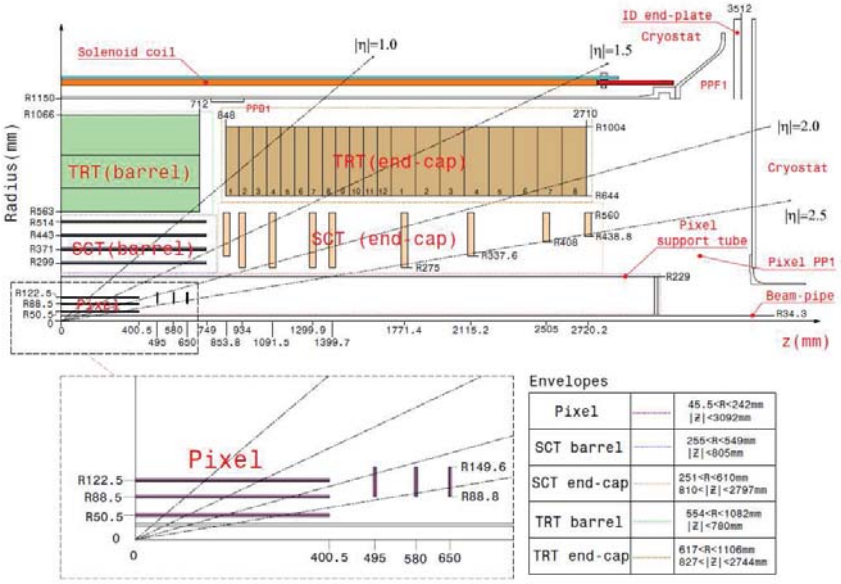

Figure 3.2.5: *Schematic overview of the ATLAS inner detector along with pseudorapidity (η) angles.*

Transition radiation tracker (TRT)

The outer shell of the inner detector is formed by the transition radiation tracker. It is built of polyimide drift tubes† that are 144 cm long for the barrel and 37 cm for the end-caps. The tubes have a diameter of about 4 mm, and contain a tungsten wire with a diameter of 31 μm coated with a layer of gold about 0.6 μm thick. The tubes are filled with a gas that is a mixture of 70% xenon, 27% CO_2 and 3% oxygen. The physical principle upon which these tubes work is that charged particles that enter the tube emit transition radiation whose intensity depends on the particle's Lorentz factor, γ. Because the Lorentz factor of a particle with a given energy depends on its rest mass, the intensity of the transition radiation can be a powerful discriminant to separate particles based on their mass. This principle is employed to separate electrons from pions and kaons. The transition radiation photons ionize atoms in the gas. Because the outside of the tube is kept at a large negative voltage of -1530 V, there is a powerful electric field in the tube pointing away from the central wire which means the electrons originating from the ionization get accelerated towards the tungsten wire. The acceleration is large enough for the electrons to start ionizing gas particles themselves leading to a cascade of electrons impinging upon the central wire where they are counted. In total there are 50000 of these tubes in the barrel region and 250000 in both end-caps.

3.2.2 Calorimeters

The calorimeters are responsible for measuring the energy of particles by absorbing them completely in detector material. Most particles are stopped by this; only muons and neutrinos escape. There are two separate calorimeters inside the ATLAS detector,

†Sometimes referred to as straw tubes, or even just as straws.

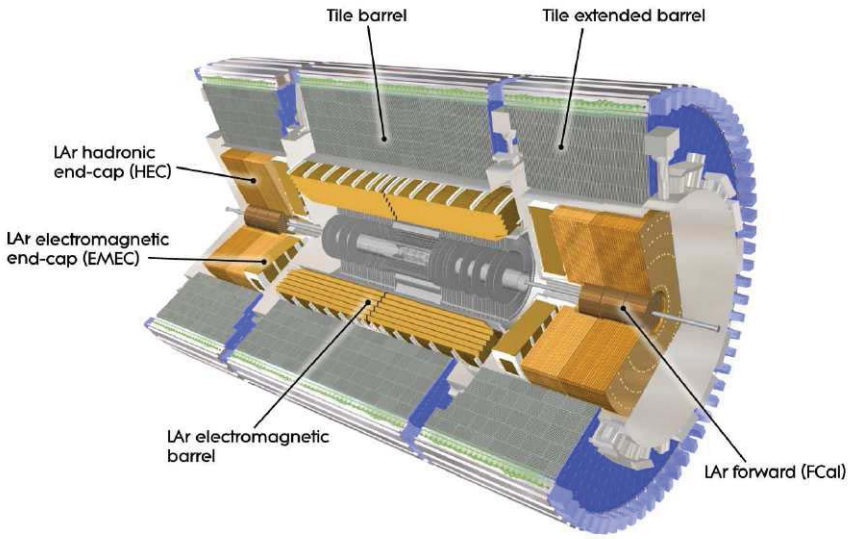

Figure 3.2.6: *Zoomed-in computer drawing of the ATLAS calorimeter systems.*

the electromagnetic calorimeter which is optimized to measure the energy of photons and electrons, and the hadronic calorimeter which measures the energy of nuclear particles such as gluons and quarks. These latter two hadronize and form showers of hadrons very soon after the proton-proton collision that spawned them. There is also a forward calorimeter but it is usually counted along with the hadronic calorimeter. As can be seen in figures 3.2.1 and 3.2.6 the electromagnetic liquid argon calorimeter forms the inner layer with the hadronic tile calorimeter forming the outer part of the combined system. The calorimeters work by offering absorption material for particles to lose their energy in interwoven with particle detectors to count how many were lost in each layer to determine the energy of the incident particle. They are also meant to completely stop the particle, not only to be able to determine fully its energy, but also to stop it from punching-through into the muon spectrometer which is located outside the calorimeter system.

Liquid argon calorimeter

Like all other subdetectors the electromagnetic calorimeter consists of a barrel region and two end-caps. The barrel region covers the area with pseudorapidity below 1.475 and extends radially from 1.4 m to 2 m. The two end-caps each consist of two wheels covering $1.375 < |\eta| < 2.5$ and $2.5 < |\eta| < 3.2$ respectively and both 63 cm thick and extend radially from 0.33 m to approximately 2.1 m. The absorption material used is lead, and the particles detectors are liquid argon (LAr) sampling detectors where a charged particle ionizes the gas and due to an applied potential difference the free cascading electrons produce a signal in the electronics. The lead and LAr detectors are interleaved, with an additional LAr detector, 11 mm thick, at the front to correct for particles losing energy in the inner detector.

The barrel detector has a thickness which varies from at least 22 to 33 radiation lengths

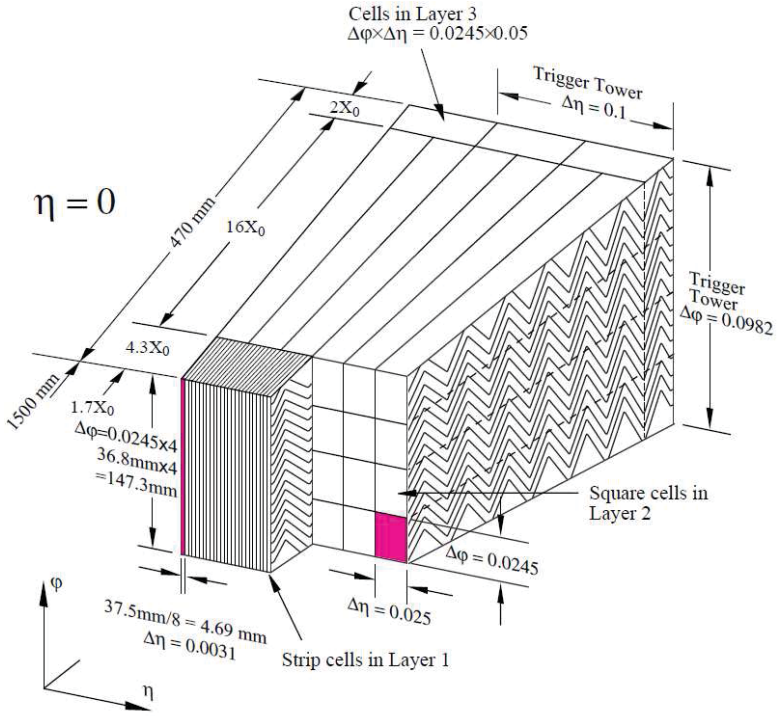

Figure 3.2.7: *Schematic overview of a part of the barrel region of the electromagnetic calorimeter. Clearly visible is the accordion geometry. Also indicated is the* $\phi \times \eta$ *granularity.*

(X_0), for the end-caps it is at least 24 to 38 radiation lengths[†]. The calorimeter has an accordion geometry that provides hermetic coverage as can be seen in figure 3.2.7. The barrel is made from two half-barrels, each containing 1024 layers of absorption material. Most of the detector in this region has a resolution in $\phi \times \eta$ of 0.0245×0.025 which increases to 0.0982×0.1 further away from the interaction point. The leading LAr detector has a much better granularity; 0.0245×0.0031.

As stated before each end-cap is made up of two separate wheels, each having interleaved lead absorption layers and liquid argon detection layers. There are 256 such layers in the inner wheel, and 768 layers in the outer wheel. As with the barrel the resolution in $\phi \times \eta$ varies, 0.025×0.1 closest to the interaction point, increasing to 0.025×0.025, and then decreasing again to 0.025×0.05. It is important to note that the η-region defined by $1.37 < |\eta| < 1.56$ falls in the overlap region of the barrel and end-caps and this region has very limited resolution and photons and electrons found there are typically ignored.

[†]A radiation length is a common way to define the thickness of a calorimeter. In very dense materials (such as lead) particles lose energy mainly through brehmsstrahlung or (in the case of photons) electron-positron pair production. In this case a single radiation length corresponds to the average distance over which a particle loses $\frac{1}{e}$ of its energy through radiation and it is also $\frac{7}{9}$ of the mean free path length for a photon.[60]

Tile calorimeter

Like the electromagnetic calorimeter the hadronic calorimeter consists of a barrel and two end-caps. The absorption material in the barrel region is steel with scintillator tiles equipped with photomultiplier tubes as particle samplers. In the end-cap the absorption material is copper, and the detection works via liquid argon samples. The barrel consists of one primary detector covering $|\eta| < 1.0$ and two extended barrels (see figure 3.2.6) which cover $0.8 < |\eta| < 1.7$ and all of which extend radially from 2.28 m to 4.25 m and have a $\phi \times \eta$ resolution of approximately 0.1×0.1 except for the outer layer which has a resolution of 0.1×0.2. The thickness can also be expressed in terms of hadron interaction lengths[†], in which case the thickness is 9.7 at $\eta = 0$. The end-caps again consist of two wheels each and covers $1.5 < |\eta| < 3.2$ which gives it a partial overlap with the barrels. The resolution in the end-caps is 0.1×0.1 for $1.5 < |\eta| < 2.5$ and 0.2×0.2 for $2.5 < |\eta| < 3.2$.

Forward calorimeter

The forward calorimeter consists of three parts and is situated at $3.1 < |\eta| < 4.9$. Its resolution[‡] varies from 3.0×2.6 cm^2 to 5.4×4.7 cm^2. All parts use liquid argon samplers for detection. The first part uses copper as an absorption material, the other two use tungsten. The first part is mainly meant for electromagnetic calorimetry whilst the final two are optimized for hadronic calorimetry. The entire forward calorimeter is about ten hadron interaction lengths deep.

3.2.3 Muon spectrometer

The outer part and largest part of the ATLAS detector is formed by the muon spectrometer which extends radially from 5 m to 10 m, see figure 3.2.8. As its moniker suggests it is meant to detect muons. It covers the region of $|\eta| < 2.7$ and consists of a set particle trackers and also a set of detectors whose express purpose it is to provide a way to trigger on muons. The magnetic field is azimuthal due to the toroidal magnets (see the next section) and bends the muons allowing for momentum measurements after tracking and measuring the curvature of the paths and has a field strength that varies between 0.5 T and 1 T.

The trackers are monitored drift tubes (MDTs) and the cathode strip chambers (CSCs). The MDTs are present over the whole $|\eta| < 2.7$ area except for the first layer where the CSCs covers $2.0 < |\eta| < 2.7$. The actual shapes of the detectors are irregular due to the need to correct for acceptance losses caused by the presence of the magnet system, and also the dimensions of them increase the further they are away from the interaction point, so that every muon "sees" the same amount of detector material regardless of its pseudorapidity. The MDTs individually are drift tubes with a diameter of around 30 mm filled with a gas that is a mixture of argon (93%) and CO_2 (7%) with a 50 μm thick central anode wire made of tungsten and rhenium. The MDTs have a spatial resolution of about 80 μm.

[†]This is the mean distance over which a hadron loses $\frac{1}{e}$ of its energy.

[‡]Being so far forward makes $\phi \times \eta$ a messy coordinate system for expressing the resolution, which is why we use centimeters here.

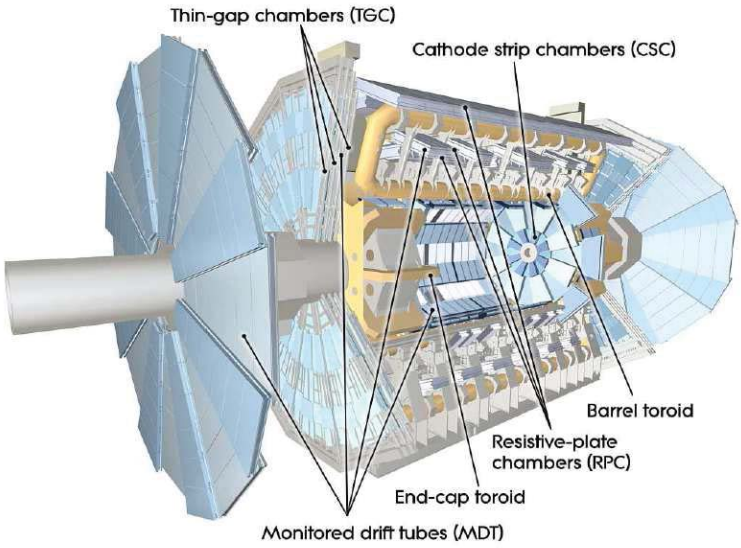

Figure 3.2.8: *Computer drawing of the muon spectrometer as a part of the ATLAS detector.*

In the region of $|\eta| > 2.0$ one expects more than 150 hits per cm^2 which is the reason for using CSCs instead of MDTs in this region. The CSCs have a tolerance for up to a thousand hits per cm^2 and consists of two end-cap disks. They are multiwire proportional chambers with a resolution of 60 μm which increases to 5 mm in the azimuthal direction.

The other detectors are the resistive plate chambers (RPCs) found in the barrel region, and thin gap chambers (TGCs) in the end-caps. The RPCs are gaseous plate detectors where two parallel plates are kept at a high potential difference and distance of 2 mm. Particles entering the detector leave ionization avalanches. The TGCs are multiwire proportional chambers. The main selling-point of the RPCs and the TGCs is the great time resolution making them suitable to use as triggers but they are also used to measure the ϕ coordinate of tracks.

3.2.4 Magnet systems

The magnetic field inside ATLAS is a hybrid toroidal-solenoidal field and is generated by four superconducting magnet systems, a barrel solenoid, a barrel toroid, and two toroidal end-caps. The barrel solenoid is located around the inner detector and generates a 2 T field that points along the beamline. The 0.5 T to 1 T azimuthal magnetic field in the muon spectrometer is created by the barrel and end-cap toroids which give the ATLAS detector its distinctive shape and name. The whole system is 22 m in diameter and 26 m in length. A picture of the magnets can be found in figures 3.2.1 and 3.2.9.

3.2.5 Trigger systems

During the 2012 data taking the proton bunches were separated in time by 25 ns, meaning that about 40 million bunches cross each other at the ATLAS collision point each

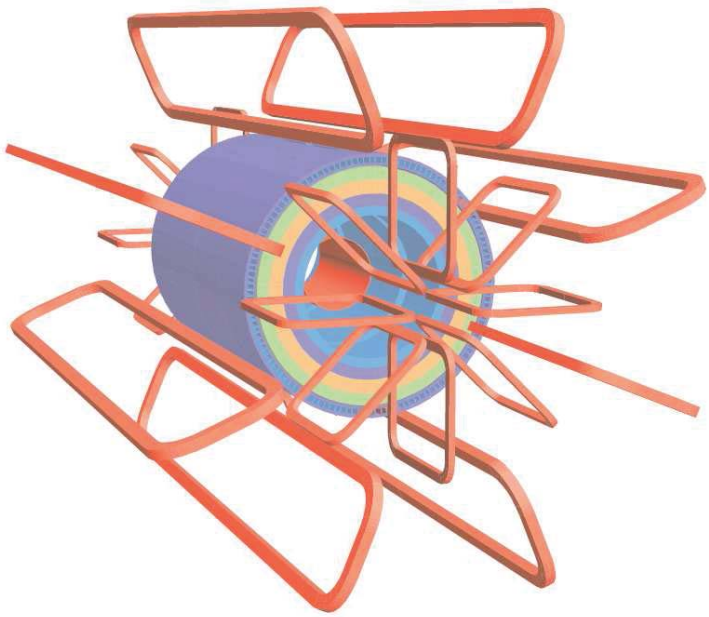

Figure 3.2.9: *Computer drawing of the ATLAS magnet system. Shown are the large barrel and end-cap toroidal magnets as well as the inner solenoidal magnet.*

second. As can be seen in figure 3.1.2 the number of actual collisions in such a crossing in 2012 could be over 35, which translates to about 1.4 billion collisions occuring inside the ATLAS detector each second. With a typical event requiring a few megabytes to store this leads to data streams in excess of thousands of terabytes per second, far beyond what is reasonable to process and store. To tackle this the ATLAS trigger system is devised to select only those few events worth keeping and ignoring the rest, bringing the event rate down to a comfortable 200 events per second.

The trigger system operates in three levels, which are the level 1 trigger, the level 2 trigger, and the event filter. Each subsequent trigger deals with a lower data rate than the previous one and therefore has a longer time period available to make a decision. This means that the higher level triggers can be more precise and elaborate than the lower ones. It is very important to ensure no physics-relevant events are rejected because they will not be stored and would lost forever in that case.

The level 1 trigger operates at hardware level and uses information gained from the calorimeters and from the trigger systems in the muon spectrometer, and is thus mostly sensitive to objects that show up in these detectors, which includes electrons, muons, photons, hadronic showers or jets, and hadronic τ decays. Only 2.5 μs are available to consider each event at level 1 and to save time the calorimeter is read-out with a reduced $\phi \times \eta$ resolution. The exact criteria that the trigger selects on varies, but it mostly involves objects of high transverse momentum. Some triggers let through too many events, and so events are then randomly rejected to reduce the data flow. This is known as prescaling the trigger. The level 1 trigger also defines so-called regions of interest (RoI), which are areas defined in $\phi \times \eta$ space where interesting things were seen (usually a highly energetic object), and are used as partial input to the level 2 trigger. After the level 1 trigger the data rate should be reduced to only 75000 events per second.

The level 2 trigger uses the regions of interest defined by the level 1 trigger as input, and has access to the full detector response inside this area (which is about 2% of the total event data), which means that both the inner detector and calorimeter information is now available at full resolution. The data rate is reduced by this trigger layer from 75000 events per second to 3500 events per second and it has about 40 ms per event to decide whether it is worth keeping.

The final trigger layer is known as the event filter. It is run offline on a computer cluster and has access to all of the information stored about the event. At this point the data rate is low enough to be no longer a concern. However, it is still not possible to store all events because that will take up to much storage space. The event filter uses analysis algorithms and takes about four seconds per event to complete its task. The data rate is reduced to only 200 events per second which are then stored on disk. Events are also sorted into data streams based on the prevalent objects in the event. For example events with highly energetic photons or electrons end up in the EGAMMA stream. The aim here is to pre-sort the events so offline physics analyses (like the ones described in this thesis) have an easier time accessing the type of events they are interested in. Events can be sorted into more than one stream in some cases.

CHAPTER 4

Simulation and reconstruction

"Mathematics catalogues everything that is not self-contradictory; within that vast inventory, physics is an island of structures rich enough to contain their own beholders."

- Greg Egan

HAVING described the experimental setup and the Standard Model in the previous two chapters we now turn to describing the computer algorithms used to translate theory predictions into simulated events. Each such event contains a list of particles and their kinematical properties that describe a group of proton-proton collisions. These can then be directly compared with measurements made by the detector. Events are simulated via the use of Monte Carlo methods and resultant particles traverse a simulated ATLAS detector. They are then put through the same reconstruction software as real data events are, ensuring that data events and simulated events can be directly compared.

In this chapter we discuss how these simulations are performed. The simulations have been thoroughly tested and calibrated on real data, and their descriptions are for the most part very accurate. This not only strengthens our confidence in these simulations, but also in the Standard Model as a whole. Usually a systematic error is given to any experimental result for uncertainties involved with the simulation. This is typically done by performing the simulation using different methods and quoting any differences as a systematic. First we shall discuss the actual event generation and afterwards we will discuss the way ATLAS reconstructs events from detector output, as well as some algorithms meant to identify jets originating from the decay of b-hadrons in the data, since these will have a particular relevance to the analyses described in this thesis.

4.1 Event generation

GENERATING an event typically involves five separate stages, the first of which is to use the knowledge of the Standard Model to generate a single proton-proton collision at parton level[†]. This comes close to doing a Feynman diagram calculation by hand. In the second stage charged or colored particles are allowed to emit various radiation (typically photons, light quarks, and gluons). Technically it is possible to include this in the calculation of the relevant Feynman diagram describing the proton-proton collision but in reality this is very complex and an exact calculation will take too much effort and time[‡]. Instead several approximations are made here. The foremost one is to apply factorization theorems that allow one to separate the Feynman diagram description from the so-called parton showering. After this the colored particles are combined into mesons and baryons in a process called hadronization. This process cannot be calculated analytically and again some approximations are used. The next step is to add additional proton-proton collisions to serve as a background since in the actual ATLAS detector every bunch crossing can yield multiple separate proton-proton interactions. The final step involves simulating the propagation of particles through the detector taking into account interactions with material.

In order for a collection of simulated events (called a simulation sample), to be useful it should cover the entire available kinematical phase-space. Some parts of phase-space have, however, a low cross-section and will thus end up being poorly described if one generates events in numbers proportional to their cross-sections. This is typically fixed by generating all different types of events a sufficient amount of times, but then reweighing individual events so that rare events receive a lower weight than the more common ones, but still span the entire phase-space.

In a simulated event the analyzer has access to much more information than when handling an event from a data sample. For instance, one can access the underlying event and their precise object types and kinematical values. This type of information is often called truth information or truth data. This is useful because it allows one to assess how well the detector is able to measure certain values or reconstruct certain particles. To do this one simply compares the truth information to the actual reconstructed information. We will use truth information several times in this thesis. For instance to determine how often a signal photon is correctly identified.

Different event generators differ in the details of their way of handling various stages of the event generation. Because of this, systematic errors are often defined where the results of an analysis are recalculated using a different event generator. This procedure allows one to assess the dependence of the results on these details. In addition the various free parameters are tuned to match experimental data. In the rest of this section we will describe the various stages of event generation in more detail.

[†]This means we can talk about particles such as quarks as distinct particles. In reality such particles quickly hadronize or decay and will thus never be "seen" directly in a detector.

[‡]When millions of events are needed for many different channels computation time becomes a relevant concern.

4.1.1 Generating a collision

The first part of generating an event is generating a hard collision event between two high energy partons. In most cases these are simple processes, but more advanced ones, including hypothetical particles, might be used to study the kinematics of rare events. An example is a quark and an anti-quark fusing together to produce a muon and anti-muon pair, or a process where a hypothetical heavier cousin of the Higgs boson decays to two photons.

Because the energy scale is very high (in our case around 8 TeV), the collision event can be described perturbatively with the help of Feynman diagrams, see chapter 2. The incoming particles are typically quarks or gluons, and the incoming and outgoing particles are given random kinematical values and quantum numbers (although they are of course conserved in the process). Unfortunately the internal structure of the proton cannot be easily computed from first-principles[†], and the probability to "collide" a particular parton[‡] with a particular energy inside a proton is taken from experimentally obtained parton distribution functions.

Leading order[⊙] parton distribution functions (PDFs) describe the probability for a given parton to be found in a proton with a given longitudinal momentum fraction, usually called x, at an energy scale Q^2. They are obtained by fitting the functions to experimental data, and there are many PDFs in circulation, see [61] for an overview. A transverse momentum is typically added, the specifics of which depend on the actual generator used.

4.1.2 Parton showering

Due to the acceleration particles experience as a result of the hard-scattering they can emit other particles in a process comparable to Brehmsstrahlung. In some cases the particles emitted this way can emit further particles leading to a cascade.

The amplitude of this process can in principle be described by using Feynman diagrams with more particles emitted from colored or charged particles that either enter or leave the hard scattering as a higher-order process. In practice this is very laborious when the number of such emitted particles increases. Another problem is that which each splitting the energy per particle decreases and so the energy scale of each subsequent splitting becomes lower and lower. At some point QCD becomes non-perturbative, and even before that a tree-level approximation might not be adequate anymore further complicating things.

In order to solve this most generators approximate this process by replacing the external particles by splitting where the particle emits a certain other particle and the

[†]The main reason is that the internal structure of a proton is ignorant of the actual energy of the proton the lab frame in accordance with special relativity. Therefore, the relevant energy there is quite low. And this makes perturbative methods unsuitable, since the internal structure is heavily dependent upon the strong nuclear force which becomes non-perturbative at low energies.

[‡]The name given to any particle that is "part" of a proton or more generally any hadron. The most common ones are gluons and valence quarks, the quarks that give a hadron its quantum numbers. But it can also contain other quarks due to quantum corrections. This can be seen as quarks popping in and out of existence inside the proton.

[⊙]Next-to-leading order parton distribution functions are bit more complicated since their output cannot be directly interpreted as a probability. We shall not go into this intricacy here.

four-vectors of all the particles are later changed to conserve energy and momentum. This proceeds in a stochastic manner where a splitting has a probability of happening at a certain energy scale, comparable with how one can calculate the probability of a radioactive atom decaying in a set time interval. Eventually at some low energy scale on the order of a few GeV the procedure is terminated. All such split particles are then subsequently also processed, leading to a very large tree of particles.

It should be mentioned that typically a distinction is made between the processes where particles are radiated off the incoming particles to the hard-scattering, and those where the particles are radiated off the outgoing particles. The former is known as Initial State Radiation (ISR) and the latter as Final State Radiation (FSR) and the two can be separated from each other and from the hard process via factorization theorems. Extra care has to be taken when dealing with ISR because the ingoing particles have to retain the correct energy and momenta after being evolved backwards from the hard process scale to the scale of ~2-3 GeV where PDFs enter.

4.1.3 Hadronization

After showering all the particles that carry a color need to be collected into color-neutral composite particles such as baryons and mesons, because as is mentioned in chapter 2 colored particles cannot be directly observed due to the color confinement property of the strong nuclear force. This process is known as hadronization and cannot be calculated analytically[†], so once again various different techniques are used to approximate this step. In this thesis the most commonly used event generators are PYTHIA [62, 63] and SHERPA [64], which differ in the way they approach hadronization. PYTHIA uses the Lund string model [65], whereas SHERPA favors the cluster model [66]. Both operate in the so-called large color limit, where the number of colors is taken to be infinite as opposed to three. In these models quarks carry one color, whereas gluons carry two. This approach has the advantage that each parton is only interacting with at most two other partons, simplifying things enormously.

Lund string model

In the string model, the large color limit is exploited by creating structures known as color strings between each colored particle and its color partner. These strings can be thought of as a cloud of virtual gluons. This strings will resemble tiny tubes stretched between the two particles due to the gluon's self-interactions. As the particles move apart the potential energy trapped in the string will increase until the moment where the energy is large enough to create a new light $q\bar{q}$ pair with each of them linking to a different original particle. At this point the string will snap, forming two new smaller strings. The process can be repeated, snapping and forming new strings all the time until some low energy level is reached. See figure 4.1.1 for a diagrammatic representation.

[†]The reason again is that the energy scale here is very small and so the theory of strong interactions becomes non-perturbative.

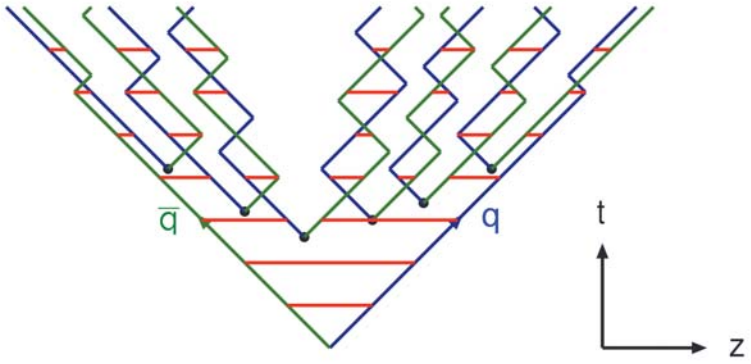

Figure 4.1.1: *Time evolution in the Lund string model. The original $q\bar{q}$ pair is linked by a string that snaps when the two quarks move far enough apart. Subsequent strings snap subsequentially. Figure taken from [67].*

Cluster model

The cluster model has its basis in a property of QCD that was discovered back in the late seventies called preconfinement [68]. In short, this property entails that partons created in a hard-scattering event already cluster in color singlet groups in a way that does not depend on the details of the hard-scattering process, at energy scales much less that that of the initial event. These clusters can then be taken as seeds to form hadrons, which is what the cluster model does.

4.1.4 Particle decay

Many of the hadrons and other particles produced in the hard-scattering are unstable and decay rapidly without ever interacting with the detector itself. This also includes hadrons in excited states that decay to the groundstate. Unstable particles with a lifetime long enough to reach parts of the detector are dealt with in the detector simulation. Typically these can depend on the spin or parity of the decaying particles, and the branching ratios[†] are tuned to the best available experimental and theoretical data.

4.1.5 Pile-up, the underlying event and MPI

When two protons collide and a hard-scattering occurs this is typically a process involving two partons, one selected from each individual proton. But the rest of the proton also collides with the other one. This can lead to all kinds of low energetic activity that is collectively known as the underlying event. The underlying event can be experimentally studied using so-called minimum-bias events; proton-proton collisions without any hard-scattering and the event generators can be tuned to the results thus obtained.

The cross-section for several interaction processes can in the limit of very small momentum transfer grow larger than the proton-proton collision cross-section itself. This can

[†]A branching ratio describes which fraction of the parent particles decay to a certain configuration of child particles. All branching ratios should sum up to unity.

be conveniently interpreted as the theory predicting more than one such interaction inside a single proton-proton collision. This is the basis of multiple parton interactions (MPI) where different pairs of partons inside the same two protons undergo a hard interaction. At generator level this is typically handled by adding these new hard interactions on top of the original one.

Another origin of additional particles comes from additional proton-proton collisions happening in the same bunch-crossing. These can lead to separate hard-scatterings (although most are soft) at different locations inside the beampipe. This effect is known as pile-up because in a sense the detector has not yet reset itself from the previous collision so the different collisions "pile-up" and one gets a picture with several separate collision events inside. This is easily modeled in event generators by producing additional collision events and adding them. It should be noted that this can be computationally costly though. All particles produced by these different effects will be fully parton showered, hadronized, decayed and fed along with the main collision products into the detector simulation.

4.1.6 Detector simulation

The effects the ATLAS detector has on particles traversing it are modeled using GEANT4 [69]. Using this software package a fully digital model of ATLAS has been constructed [70] which allows for the simulation of particle decays, energy losses in detector material and the bending of particle tracks by a magnetic field. This simulation also includes the detector response; when particles enter parts of the detector where actual detection occurs the output of these interactions are recorded in the same way the actual real ATLAS detector does when taking data. This means that the output of the detector simulation has the same format as the output from the real detector. As such simulated and real recorded data can both be put through the same event reconstruction machinery.

4.2 Event reconstruction

THE data that comes from the ATLAS detector or its simulation is a collection of electronic responses from the various sub-detectors. It needs to be converted into a physical picture of what was measured to be useful. The responses need to be converted into so-called physics objects, which are the objects one performs the eventual analysis on, like electrons, photons, and jets. Of course this reconstruction is not perfect; sometimes objects fail to be reconstructed, or they are wrongly reconstructed into a different species. Or feature distorted kinematics. All these things are part of the reconstruction efficiency which is an essential part of the systematic uncertainty of any final analysis.

In this section we shall discuss how the various physics objects are reconstructed using the ATLAS reconstruction software. Typically only stable particles that end up in the detector are considered to be physics objects. So for example, a Higgs boson, which decays long before reaching any detector part, will not be reconstructed in this way. But its decay products, two photons say, will. It is the task of the analysis to combine these reconstructed photons into a reconstructed Higgs boson later. Another important example are b-hadrons which are of particular importance in the analyses described in

this thesis. These hadrons typically decay forming large cascades of particles known as jets. These jets are physics objects, but the b-hadrons themselves are not. Instead, as we will discuss later in section 4.2.2, an algorithm is used to give a probability that a particular jet originated from a b-hadron, and not some other object.

The objects essential to the analyses in this thesis are electrons, photons, and jets. We will discuss the particularities of each in what follows. We will also discuss the way jets are defined, clustered, and b-tagged.

4.2.1 Electrons and photons

The reconstruction of photons and electrons are similar because both are identified mainly through information gained from the electromagnetic calorimeter, and in the case of electrons, tracks left in the inner detector. In addition photons can convert to an electron-positron pair when interacting with detector material. In this case you still want to reconstruct the photon as a photon and not as an electron. Such a photon is known as a converted photon and likewise a photon that remains unconverted is called an unconverted photon. About 40% of all photons converts into an electron-positron pair before reaching the electromagnetic calorimeter [56].

The reconstruction starts by finding clusters of energy in the calorimeter of size 0.025×0.025 in $\phi \times \eta$ coordinates with a minimal energy deposit of 2.5 GeV. Electrons are distinguished from unconverted photons by matching the cluster to tracks in the inner detector. The tracks are inspected for possible secondary vertices to distinguish between electrons and converted photons. The location of the cluster and the energy measured therein are subject to various calibrations. For example the energy measurement can be calibrated by studying $Z \to e^+e^-$ or $\mathrm{J}/\psi \to e^+e^-$ events.

Identification of these objects proceeds mainly through variables depending on the so-called shower shapes in the electromagnetic calorimeter and in some cases in the hadronic calorimeter, though in the case of electrons the track information is also used. This is also meant to reject as much as possible hadronic showers being wrongly reconstructed as either electrons or photons. The physics objects are issued with a quality flag[†] which gives some indication to the likelihood that that object corresponds with a genuine particle.

In the analyses used described in this thesis we use three different electron flags: LOOSE++, MEDIUM++, and TIGHT++, and two different ones for photons: LOOSE and TIGHT. The tighter variants typically incorporate the looser ones. The exact settings are optimized by the EGAMMA performance group inside ATLAS, but the cuts are based on the number of hits in various parts of the inner detector, the width and evolution of the shower inside the calorimeter, any possible leakage into the hadronic calorimeter, and track matching parameters [71].

4.2.2 Jets

Jets are very complicated objects, a large collimated cascade of particles typically finding their origin in a single colored particle created in a hard interaction. Since they are,

[†]This is a boolean that describes whether the physics object passes certain criteria or not.

by construction, a multitude, one does not seek to identify each individual particle, but rather treating the whole shower as one composite object with a direction and energy. This approach is also vindicated by considering that we are only actually interested in the original colored particle, and not its shower products, which are largely random. The parameters and algorithms used here define what we actually mean by a jet of that particular type. And care should be taken; for example taking a larger radius for the jet clustering algorithms can lead to one finding a single big jet which was actually created by two different partons which were close together, whilst taking a too small a radius might cause on to miss energy.

Reconstructing jets

Inside ATLAS there are two ways of reconstructing jets. The most commonly used ones are the so-called calorimeter jets (sometimes called calo-jets), which are seeded from topological clusters (groups of calorimeter cells). A topological cluster starts with a calorimeter cell for which the signal-to-noise ratio exceeds four, where the noise is the typical ambient energy flux. From this starting cell additional cells are added iteratively if they have a signal-to-noise ratio of at least two. The topological clusters thus reconstructed are then fed into a jet clustering algorithm which creates the actual jet physics objects.

The second type of jets are so-called track-jets. They are seeded not from calorimeter clusters, but instead by tracks that are left in the inner detector by charged particles. Track-jets are less often used than calorimeter jets, but do have some benefits, such as a higher resistance to pile-up. They are, however, typically harder to use; matching to calorimeter clusters is required to obtain a good energy estimation. They are also less useful at very low p_T (lack of solid tracks) or very high p_T (more difficult to disentangle tracks). In this thesis we describe a b-tagging calibration on track-jets in appendix A, but all other jets mentioned in this thesis are calorimeter jets.

Jet clustering algorithms

Jet clustering algorithms combine objects such as the topological clusters and tracks into actual objects we call jets. There are some requirements that such a clustering algorithm should satisfy to be usable in a physics analysis. The main requirement is that the jet preserves as much as possible of the original parton's momentum and energy. In addition it should be robust against pile-up effects, and not depend too much on the exact way a parton was hadronized since this is mostly a stochastic effect and as said before only the original parton is of any interest usually. Another crucial aspect is that the algorithm is infrared-safe. This means that adding tiny low-energetic clusters should not have a large effect on the resulting jet. The algorithm should also be collinear-safe which means that splitting clusters into smaller ones should not change their combined energy and direction.

There are various algorithms. In the past cone-type algorithms were favored for their computational ease, but since they are not typically infrared- or collinear-safe they have been mostly discontinued as increased available computing power made their advantage obsolete. Nowadays most people use sequential clustering algorithms such as

the Cambridge/Aachen (CA), the k_t algorithm, and the anti-k_t algorithm. They are all infrared- and collinear-safe, and are also favored by theorists. The only downside is that the jet area, which is the region the jet occupies in $\phi \times \eta$ space, is not trivial as is the case for a cone-type algorithm.

All analyses in this thesis use the anti-k_t algorithm [72], the ATLAS standard, which is a variant of the older k_t algorithm [73]. For the three main sequential clustering algorithms CA, k_t, and anti-k_t one uses the following distance measure[†]:

$$d_{ij} \equiv \min\left(p_{T,i}^{2p}, p_{T,j}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2} \tag{4.2.1}$$

where $p_{T,i}$ is the transverse momentum of cluster i, R is a free parameter called the radius parameter, p is free parameter governing the relative weight of the momentum factor, and the geometric distance ΔR_{ij} is defined:

$$\Delta R_{ij} \equiv \sqrt{(\phi_i - \phi_j)^2 + (y_i - y_j)^2} \tag{4.2.2}$$

The radius R is important as any jet created from these algorithm is separated from other jets by at least R, making it suitable as an indication of jet radius. Taking a large value makes the measurement of a jet's kinematic values more precise as more particles are included, but it also makes it more susceptible to wrongly including particles which maybe should be clustered into a different jet, or are coming from pile-up effects.

In addition to this one defines the beam distance as:

$$d_{iB} \equiv p_{T,i}^{2p} \tag{4.2.3}$$

The parameter p distinguishes between the various algorithms. In the case that $p = 0$ the expressions simplify to the CA case where only the geometric distance plays a role. The regular k_t algorithm is recovered when $p = 1$, and the anti-k_t algorithm is obtained for $p = -1$. The algorithm proceeds in the following way. First it finds d_{ij} and d_{iB} for all clusters and then finds the smallest one. If this happens to be a d_{ij} it sums up the four-momenta of clusters i and j to create a new cluster k. Clusters i and j are then removed, and essentially combined into the new cluster k. If this happens to be a d_{iB} then cluster i is removed , and a jet is defined with its four-momentum equal to that of the removed cluster i. The algorithm then starts anew and continues until there are no more clusters left.

Because of the weight $\min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2}\right)$ in the anti-k_t case clusters with a small amount of energy will prefer to cluster first with clusters that already have a lot of energy as two clusters of low energy will have a much greater distance, d_{ij}, than a hard and soft cluster at the same geometric distance. This makes the algorithm handle high pile-up better than the k_t algorithm, where soft clusters tend to cluster amongst themselves first.

The algorithm works the same way for track-jets, except that, as mentioned before, one uses tracks instead of calorimeter clusters as input to the algorithm.

[†]The transverse momentum in this expression, p_T, is typically written as k_t which shines some light onto the name k_t algorithm.

b-tagging algorithms

The process of tagging b-jets consists investigating a collection of jets and deciding whether each jet has a good probability to originate from a b-hadron. The jet in question is then said to be b-tagged. There are various methods to do this which are maintained and calibrated by the ATLAS b-tagging performance group. We will take some time here since b-tagging is a very important part of most of this thesis. A later chapter features a calibration of a b-tagging algorithm known as MV1 [74, 75] and we shall also be using MV1 a lot in the main analysis of $b\bar{b}\gamma\gamma$.

The MV1 algorithm is a multivariate algorithm that returns a number between zero and one on whose value one can apply a cut. Jets originating from light quarks or gluons, called light jets, tend to gravitate towards zero, jets originating from b-hadrons, named b-jets, tend to have high values towards 1, and those coming from c-hadrons, called charm jets, are located somewhere in between. The algorithm cannot distinguish between b-jets finding their origin in b-quarks and those coming from $\bar{b}$-quarks. Both of these are called b-jets, which means that technically our analysis of $pp \to b\bar{b}\gamma\gamma$ also contains bb and $\bar{b}\bar{b}$ combinations. Such processes are not allowed in the Standard Model though, except in the rare situation where their partners escape in the beam pipe. MV1 combines the likelihood output of three other algorithms into one single output variable using an artificial neural network. These algorithms are SV1, IP3D, and JetFitter which we shall describe here.

IP3D IP3D exploits the fact that b-hadrons are relatively long-lived compared to other (unstable) ones, on the order of a few picoseconds. Combined with the time dilation caused by their near-light speed they can travel up to several millimeters in the detector before decaying. Typically what one sees is a vertex for the hard-scattering event, the so-called primary vertex, and then within a few millimeters a secondary vertex coming from the decay of the b-hadron. The presence of such a secondary vertex is a good sign a b-hadron was present since as mentioned before their relatively long lifetime is rather unique.

More specifically IP3D is an impact-parameter-based tagger, using the impact parameters of the tracks associated to a jet with respect to the primary vertex. Since b-hadrons come from secondary vertices the tracks for b-jets tend to have bigger impact parameters. There are two such impact parameters; a longitudinal one, typically called z_0, which is the difference between the z coordinates of the primary vertex and the location of the closest approach of the track in question, and a transverse one, typically called d_0 which is the distance to this closest approach and the primary vertex in the $r \times \phi$ plane. The d_0 parameter is also signed to increase sensitivity. The sign is set to be positive when the track crosses the jet axis in front of the primary vertex, and negative when it crosses behind it. The intersection is calculated in the $r \times \phi$ plane, and the jet axis is the jet direction as taken from the jet clustering algorithm. If however a secondary vertex is found inside the jet then the jet axis is taken as the line stretching from the primary vertex to this secondary vertex.

For tracks originating from the primary vertex the sign of d_0 is essentially random and driven by the limited resolution of the inner detector. For b-jets there is a preference for positive d_0 since the b-hadron decay should happen along the flight path of the jet.

IP3D takes these impact parameters into account and considers a light-jet versus a b-jet hypothesis using a log-likelihood approach for each track individually and then combines them into one weight for the whole jet.

SV1 Like IP3D SV1 is a lifetime-based tagger, but instead of relying on the impact parameters, SV1 is vertex-based. A vertex finding algorithm tries to reconstruct vertices from tracks that are displaced from the primary vertex, rejecting tracks originating from photon conversions, or decays of K_s or Λ mesons. SV1 uses a log-likelihood approach similar to IP3D taking into account three vertex properties in order to identify secondary vertices that indicate b-hadron decays. The first of these is the ratio of the sum of the energies of the tracks associated to the vertex and the total energy of all tracks associated with the jet under consideration. For b-jets most of the tracks should come from this vertex and therefore this ratio tends towards unity. The second is the invariant mass of all the combined tracks associated to the vertex, which tends to be bigger for b-jets. And the third is the number of vertices reconstructed with only two tracks associated to them. The ΔR separation between the jet axis and a line drawn from the primary vertex to the secondary vertex is also used.

JetFitter JetFitter uses a neural network with a large number of input variables and outputs the likelihood that a jet corresponds to a light, charm, or b-jet. The algorithm proceeds by first finding a line connecting the primary vertex, the secondary vertex, and a third vertex that finds its origin in the decay of the c-hadron, which itself originated from the decay of the b-hadron. In this way it attempts to fully reconstruct the flight trajectory of the jet. It then takes eight input variables, two of which are the transverse momentum and the absolute pseudorapidity of the jet under investigation. It also takes two variables that are the same as in SV1: the vertex invariant mass, and the energy ratio. In addition it takes four novel ones: the number of vertices with at least two associated tracks, the combined number of all tracks at all such vertices, the number of single track vertices along the reconstructed flight trajectory, and the averaged vertex decay length.

As mentioned before, the ATLAS default tagger during Run 1, MV1, works via an neural network and combines the output of JetFitter, SV1, and IP3D, into a single number that acts as a discriminant and which falls between zero and one. In our analyses a cut is simply applied to this discriminant and any jet that has an MV1 discriminant above this cut is considered to be a b-jet. Unfortunately while b-jets tend to have a high discriminant exceptions do occur.

b-tagging is a complex process and often one has to make the choice of preferring quality or quantity. Making the cut very high will reject more light and charm jets, but it will also reject a lot of proper b-jets, thus leading to a smaller sample of jets to work with. Reducing the cut increases the sample size but contaminates it more. Typically every individual analysis needs to optimize this parameter for their use-case. However, b-tagging needs to be calibrated; simulated events need to be reweighted in order to match the b-tagging performance in data. This calibration depends on the exact cut value used and so the the b-tagging supports only a handful of these values.

These values are typically referred to as working points and they are labeled by the nominal efficiency at that cut value. In the analysis described in this thesis the 70%

working point is used, which means that on average 70% of the b-jets in a sample will end up exceeding the cut, which is in this case is close to 0.8. This choice was made a balance between having a somewhat pure sample of b-jets whilst still keeping the statistics high, which is a real concern in that analysis. It should be stressed that this 70% is the nominal efficiency, the efficiency still varies with a jet's transverse momentum. We shall investigate this for our analysis in section 6.3.

CHAPTER 5

Defining the $jj\gamma\gamma$ dataset

"The atoms or elementary particles themselves are not real; they form a world of potentialities or possibilities rather than one of things or facts."

- Werner Heisenberg

THE goal of this chapter is to define a sample that is as pure as possible in $jj\gamma\gamma$ events. It will serve as a basis for our $pp \to b\bar{b}\gamma\gamma$ cross-section measurement. Hence we will discuss procedures to extract photons and jets from the data. The dataset used is the full $\sqrt{s} = 8$ TeV proton-proton dataset that was collected by ATLAS in 2012.

The fiducial volume of the analysis is chosen to match the experimental limitations imposed on us by the detector. We require p_T cuts on the jets and photons to reduce contributions from pile-up and to obtain physics objects of sufficient quality. In addition we require a diphoton invariant mass cut because the simulation samples we use contain such a cut.

As mentioned in section 2.5 on page 42 the fiducial volume of our $pp \to b\bar{b}\gamma\gamma$ analysis is:

Photons At least two photons with $p_T > 22$ GeV and $|\eta| < 2.37$.

Jets At least two jets with $p_T > 25$ GeV and $|\eta| < 2.5$. This applies equally to b-jets and other jet types.

Diphoton invariant mass The diphoton invariant mass for the two leading photons needs to satisfy $m_{\gamma\gamma} > 60$ GeV.

Separation We require $\Delta R(\gamma, \gamma) > 0.4$ for the two leading photons, $\Delta R(j, j) > 0.4$ for the two leading jets, and $\Delta R(\gamma, j) > 0.6$ for any combination of the two leading jets with the two leading photons.

An important variable in this chapter is the photon isolation efficiency. The photon isolation is a variable that describes how much energy is measured around a photon hit in the calorimeter. By applying a cut and requiring that this energy is minimal one can reduce backgrounds originating from jets being wrongly reconstructed as photons. The photon isolation efficiency codifies how often a proper photon ends up passing such a cut.

In the first two sections of this chapter we will discuss the cuts and object definitions we shall apply in our analysis to approximate this volume. After this we will need to investigate the two main backgrounds for selecting diphoton events: electrons and jets faking photons. To start we will use a data-driven method for extracting the photon isolation efficiency. This is done in section 5.3. We will use the extracted efficiency in section 5.4 to estimate the background coming from electrons being wrongly reconstructed as photons, and in section 5.5 to do the same but for jets faking photons. In section 5.6 we will provide a description of general systematic uncertainties and corrections applied to the analysis. Even though we moved the description of these uncertainties to the end of this chapter for readability, we do apply them everywhere. They will also apply to the main $pp \to b\bar{b}\gamma\gamma$ analysis detailed in the next chapter. We close by giving an estimate of the $jj\gamma\gamma$ content of our dataset, which is used when we investigate $pp \to b\bar{b}\gamma\gamma$.

In this chapter we will not analyze differential cross-sections and instead consider all events passing the cuts inclusively. In the next chapter we will revisit the background estimation methods again and apply them bin-by-bin.

5.1 Object definitions

In this section we will define the physics objects that we will use in our analysis. They are based on the objects reconstructed by the ATLAS reconstruction software as explained in chapter 4 and are purified by applying further cuts on their characteristics.

Good photon A good photon is defined as a photon with the following characteristics:

- p_T >20 GeV.
- $|\eta| < 2.37$ where the region $(1.37,\ 1.56)$ is also excluded.
- Passes standard cleaning cuts.
- Passes the LOOSE photon identification [76] criteria. Specifically the LOOSE4 criterion is used following [50]. This is important because this definition avoids any correlation between this identification criterion and later isolation criteria.

The region $(1.37,\ 1.56)$ in $|\eta|$ corresponds to the gap between the barrel and endcap electromagnetic calorimeters. The standard cleaning cuts are small cuts recommended by the ATLAS EGAMMA performance group to increase the quality of the object. Likewise the LOOSE identification criterion is recommended, and it is based on how the shower of the photon develops in the electromagnetic calorimeter. More information can be found in [76]. The photons have their energy and calorimeter-based

isolation variables calibrated following the recommendations of the ATLAS EGAMMA performance group. In simulated samples it was found that some variables having to do with the photon identification are mismatched with the data. This was corrected in this analysis, also per recommendations by the aforementioned performance group via the use of some data/simulation scale factors. The p_T cut is to remove soft photons which are less likely to be prompt[†].

Good electron Electrons are not directly of interest to the analysis. They are however used to determine the effects of misconstructing photons as electrons and of electrons faking photons. Therefore, their selection closely mimics that of the good photons so the two can be compared with the least amount of problems:

- p_T >20 GeV.
- $|\eta| < 2.37$ where the region $(1.37, 1.56)$ is also excluded.
- Passes standard cleaning cuts.
- Passes the MEDIUM++ or TIGHT++ electron identification criteria[‡].
- Cannot be within ΔR of 0.05 of a reconstructed photon.

Electrons thus selected are reconstructed as photons are so that their kinematics and properties can be directly compared. This also allows us to apply the TIGHT photon identification criterion[⊙] [76] both to photons and to electrons. A cut requiring the leading photons to pass this criterion is included in the analysis as detailed in section 5.2, and is also used in the estimation of a major background in section 5.5. Electrons are also calibrated with respect to their energy and calorimeter isolation, in the same fashion as photons.

Good jet Jets were reconstructed using the ANTIKT cluster algorithm with a radius parameter of 0.4 and were seeded using calorimeter clusters. The jets were calibrated using the specifications set by the JETETMISS ATLAS performance group. The calibration used is not the latest one, but rather the penultimate one. This is because the simulation samples used in this analysis lacked the required information to perform the latest calibration properly. The following cuts were applied:

- p_T >25 GeV.
- p_T >50 GeV if $|\eta| > 2.4$.
- $|\eta| < 4.4$.
- Standard cleaning cuts.

[†]Prompt means that the photon was created directly in the hard-interaction, as opposed to, for example, being radiated after the fact by charged particle.

[‡]There is a slight overlap between both cases which is why we use the OR of the two.

[⊙]The TIGHT photon flag is used to reject hadronic jet background by imposing requirements on nine discriminating variables computed from the energy leaking into the hadronic calorimeter and the lateral and longitudinal shower development in the electromagnetic calorimeter.

- Cannot be with ΔR of 0.4 of a selected photon (i.e. the leading or sub-leading one).
- If p_T <50 GeV and $|\eta| < 2.4$ the jet vertex fraction[†] (JVF) with respect to the selected primary vertex must be 0.5 or more. The selected primary vertex is the one that the diphoton system points to. This is further addressed in section 6.1.2.

5.2 Event selection

HAVING defined our physics objects we move on to selecting events. The goal is to select two photons and two jets. We will get there in three separate stages. First we have the preselection. We remove events that had detector problems and also reject events without any hard proton-proton interaction. In addition we make our first step towards selecting for $jj\gamma\gamma$ by requiring at least two reconstructed photons.

After this we move on to the jet preselection. This stage is to finalize the photon cuts and prepare for the cuts selecting for jets which will come afterwards. We improve the quality of the selected diphoton sample here by placing additional constrains on the photons such as isolation and angular separation. To have clean $jj\gamma\gamma$ events we also reject any event in which there are jets close to the two leading photons.

Finally in the pretag stage we require two jets. Also in the pretag stage we require that the two leading photons "point" towards the most energetic primary vertex. This is because b-tagging is only calibrated for this vertex. We will discuss this in more detail in section 6.1.2. We are not yet applying any b-tagging algorithms in this chapter, but we will restrict our $jj\gamma\gamma$ sample to the most energetic vertex nonetheless because the sample would be useless to us later without it. After listing these different cutflow stages we view various background channels and see how they are affected by the cuts.

In this chapter and the next the two photons with the highest transverse momentum (leading photons) and the two leading jets are used as the signal particles.

5.2.1 Cutflow for $jj\gamma\gamma$

Preselection

To start with a so-called photon-preselection sample was made, which requires events to have at least two reconstructed photons, a hard-scattering vertex, and the absence of any detector defects. The following cuts are made to this sample:

Trigger Events were required to pass the EF_2G20VH_MEDIUM trigger. This trigger selects two photons with a minimum p_T of 20 GeV at the event level.

GRL The standard *All Good* run list was used to remove events that were recorded when there were problems with the detector (for example the magnetic field was disabled).

[†]The jet vertex fraction quantifies which fraction of the tracks associated to the jet can be matched to a given primary vertex. A low fraction increases the probability that the jet is not coming from that primary vertex and can be considered to be background noise. A high fraction increases the confidence one has in that the jet was produced in that particular primary vertex. This cut is restricted to the inner detector because you need tracks for it to work, and it is also restricted to jets with low transverse momentum since they form more than 99% of pile-up jets.

Cleaning Standard cleaning cuts were applied. This includes for example removing events that involve a damaged part of ATLAS' liquid argon detector.

Primary vertex Events were required to have at least one reconstructed primary vertex with at least three associated tracks. This is mainly to reject events without any hard interactions, or where a cosmic ray interacted with the detector material.

Photons At least two reconstructed photons were required.

Jet preselection

At this stage further quality cuts are performed. Specifically the photons are now required to be of the "good" quality defined before in section 5.1.

Good photons At least two good photons were required. The leading two are selected as our signal photons.

Angular separation The two signal photons should have a ΔR of at least 0.4.

Transverse momentum The signal photons were required to have p_T >22 GeV to move away from the trigger threshold of E_T > 20 GeV.

Invariant mass The invariant mass of the diphoton system was required to be at least 60 GeV. This is because a cut of 55 GeV is present in the SHERPA signal simulation sample used.

Identification The photons were required to pass the TIGHT identification. This is aimed at removing other particles faking photon signatures.

Isolation The photons were required to be isolated. Specifically it was required that TOPOETCONE40 < 4 GeV. For brevity we shall call this variable E_T^{iso}, and it roughly describes how much energy is measured in the calorimeter in a cone of ΔR of 0.4 around the photon (excluding the center, which is to be occupied by said photon). Lower values thus mean that not much happened around the photon and thus the photon is "isolated" from other activity. This rejects jets faking photons and also makes the energy measurement of the photon more reliable. More on this is discussed in section 5.3.

Jet separation No good jet can be within ΔR of 0.6 of the selected photons. Note that jets within ΔR of 0.4 are not considered at for this selection since such jets are not good jets according to the definition of a good jet. Such jets will overlap the photon, typically making the photon fail the isolation requirement. This cut improves the quality of the photon isolation.

Pre-tag selection

This is the stage where we select the jets. We also require the two leading photons to point to the most energetic primary vertex. This is required for the b-tagging that is to happen afterwards in the next chapter. The vertex pointing method is discussed in more detail in section 6.1.2.

Good primary vertex The vertex to which the diphoton system points is required to be the most energetic one.

Double jet Require at least two good jets. The leading two form our potential signal and must satisfy $|\eta| < 2.5$.

5.2.2 Cutflow numbers

We compare the behavior of the data with two signal diphoton samples, SHERPA and PYTHIA, as well as a few commonly used ATLAS Standard Model background samples, including Z and W bosons, and $t\bar{t}$ processes. We also include a simulation sample containing events featuring a single Higgs boson.

The first of the two signal samples are referred to as the SHERPA signal sample and was generated using SHERPA [64] combined with the CT10 [48] parton distribution functions. The second one is called the PYTHIA signal sample, and is based on PYTHIA8 [63] with the AU2 tune and the CTEQ6L1 [37] PDFs. Both these samples include all Standard Model processes but were filtered for diphoton events. Two samples were used to simulate $t\bar{t}$ events. One pure $t\bar{t}$ simulation generated by MC@NLO [47] interfaced with JIMMY [77], and using the CT10 PDFs. It was supplemented by a group of samples containing $t\bar{t}$ events with associated photon or diphoton production. These were generated by MADGRAPH5 [36] and showered by PYTHIA8 using the AUET2B [78] tune and the CTEQ6L PDFs. Single Higgs boson vector boson fusion processes were added using various POWHEG+PYTHIA8 [79] samples using the CT10 PDFs and the AU2 tune. Added to this were non-VBF processes generated by PYTHIA8 using the AU2 [80] tune and the CTEQ6L PDFs. Various processes involving W bosons were generated using ALPGEN+JIMMy [81] and MADGRAPH5+PYTHIA8, both using the AUET2 [82] tune and CTEQ6L PDFs. Finally Z bosonic processes were generated using SHERPA and the CT10 parton distribution functions. All simulation samples were processed using the official ATLAS detector simulation [70] based on GEANT4 [69]. In addition pile-up events were added to generated events using PYTHIA8, and the events were reweighted so that the average number of interactions per bunch crossing matched the data.

The amount of data and simulated events passing our cuts are listed in table 5.2.1. The simulation samples were weighted to the total integrated luminosity of the 2012 8 TeV ATLAS dataset, but we did not apply any scale factors meant to make the simulation samples match the number of events found in data. As such the normalization is totally off (notice for example the Z sample overshooting the data in the preselection). We notice, however, that the backgrounds coming from W, Z, single h, and $t\bar{t}$ are all reduced massively by the cuts. Even with the much larger (incorrect) number of events in the Z sample it ends up considerably below the data in the jet preselection. This means that the only real background is the one that consists of wrongly reconstructed objects typically called fakes. Because of this we decided not to spend time on tracking down the exact normalizations and will ignore these background in the rest of this chapter. A total of 60162 data events passes through all our cuts.

Note the difference in total cutflow acceptance between the PYTHIA and SHERPA signal samples. We will study the performance of Pythia versus Sherpa later on during the unfolding stage of the analysis, see section 6.4 on page 151.

5.3 Photon isolation

USING the isolation of a photon is a critical step to reject backgrounds from jets. Because the simulation is not perfect we will use a method to extract the isola-

	data	SHERPA	PYTHIA	W	Z	$t\bar{t}$	single h
preselection	19465200	5879580	5573290	333799	39141200	30539	781
2 good photons	10527700	5711390	5413490	33613	284774	2076	746
$\Delta R(\gamma,\gamma) > 0.4$	10205500	5658510	5407690	31478	281368	1995	746
$p_T^\gamma > 22$ GeV	7724440	4600170	4413410	25690	259566	1725	737
$m_{\gamma\gamma} > 60$ GeV	5693880	3109730	3261780	18347	243086	1169	737
TIGHT identification	1678590	2485210	2594570	7560	155268	356	642
photon isolation	814639	2131010	2144470	3732	121856	147	566
$\Delta R(j,\gamma) > 0.6$	799927	2079120	2110800	3600	120642	137	561
good primary vertex	545096	1425810	1364270	2859	94878	135	413
at least 2 jets	60162	193334	91122	514	4851	118	71

Table 5.2.1: *Event numbers after the different cuts of the cutflow for data and various simulation samples.*

tion dynamically from the data. Ultimately we shall use this to extract the isolation efficiency, ϵ_I.

The isolation definition used in this analysis, TOPOETCONE40, is computed from energy clusters in the electromagnetic calorimeter in a cone of ΔR of 0.4 around the photon. It should be noted that the energy deposited very close to the photon (0.125×0.175 in $\eta \times \phi$) is ignored and the energy in the isolation cone is corrected for additional proton-proton interactions. It is also corrected for energy from the photon leaking outside the 0.125×0.175 area allowing TOPOETCONE40 to be negative [83] in some cases. In this analysis we consider photons to be isolated if TOPOETCONE40 < 4 GeV. For brevity we shall refer to TOPOETCONE40 as E_T^{iso}.

There is an alternative isolation definition that is used shortly. It is based not on isolation in the calorimeter but on isolation in the inner detector and focuses on whether the photon candidate is isolated from charged tracks. In this case one counts the p_T of tracks within a cone of ΔR of 0.2 around the photon where a track needs to have at least $p_T > 1$ GeV. We shall call this variable p_T^{iso}.

5.3.1 Extracting the photon isolation distribution from data

A sample of photons can be divided into two groups using the TIGHT identification criterion. One group passes this criterion and is denoted by T, the other group fails it and is contained in $\bar{T}$. $T + \bar{T}$ thus contains all photons and the particles faking them.

The final cross-section measurement is performed with a diphoton sample where both photons are required to be TIGHT. However, even with this identification requirement

applied there are still jets and electrons present that fake photons. We want to know the distribution of E_T^{iso} for the photons and thus need to remove the influence of these fakes, because we need to find the isolation efficiency for actual photons.

The first thing to note is that the behavior of the isolation is quite similar for electrons and photons, because both are stable and do not create large hadronic showers. We therefore only need to worry about the jet contribution at this stage and as such we will add the photon and electron contributions together. A second thing to note is that the distribution of E_T^{iso} is the same whether the photon candidates pass the TIGHT criteria or not because the identification is not correlated to the isolation. The last thing is caused by our choice of identification flag used in the definition of a good photon, namely LOOSE4, see [50].

We can summarize these facts in a few expressions. If we call the E_T^{iso} distribution, ξ, we can state:

$$\xi^T = \xi^T_{\gamma+e} + \xi^T_j \tag{5.3.1}$$

where ξ^T is the distribution for the TIGHT sample, and $\xi^T_{\gamma+e}$ and ξ^T_j are the photon+electron and fake jet components respectively. Similarly one can state for $\bar{T}$, the NON-TIGHT sample:

$$\xi^{\bar{T}} = \xi^{\bar{T}}_{\gamma+e} + \xi^{\bar{T}}_j \tag{5.3.2}$$

Then we can describe the similarity in the isolation distributions between the identification as:

$$\xi^T_{\gamma+e} = \alpha\xi^{\bar{T}}_{\gamma+e} \tag{5.3.3}$$

$$\xi^T_j = \beta\xi^{\bar{T}}_j \tag{5.3.4}$$

where the normalization differences are captured in two constants we call α and β. From these expressions we can solve our desired $\xi^T_{\gamma+e}$:

$$\xi^T_{\gamma+e} = \frac{\xi^T - \beta\xi^{\bar{T}}}{1 - \frac{\beta}{\alpha}} \tag{5.3.5}$$

We can measure ξ^T and $\xi^{\bar{T}}$ directly, and we can ignore α because it only affects the overall normalization which will not effect the isolation efficiency because that depends only on the ratio of photons passing or failing the isolation cut and is therefore insensitive to an overall factor. The fake normalization factor, β, can be estimated by examining a sample of TIGHT and NON-TIGHT photons where the jet fakes dominate, in which case we have:

$$\beta = \frac{\xi^T}{\xi^{\bar{T}}} \tag{5.3.6}$$

Here we consider two possible ways of defining such a sample. One way is use the tail of E_T^{iso} where fakes dominate. There is a struggle here between choosing a high value of E_T^{iso} to ensure there is no contamination from genuine photons and choosing a somewhat lower value to ensure one has enough statistics. Another option is to use the tail of the track-based isolation, p_T^{iso}.

A comparison made in a $\gamma\gamma + jets$ analysis [50] finds that the p_T^{iso} based method reproduces the photon isolation efficiency better in simulations and we will therefore adopt

the p_T method. Specifically we take β to be the ratio of the photon counts in the TIGHT and NON-TIGHT samples in the region where $10 < p_T^{iso} < 25$ GeV. However, as a systematic to the final cross-section we also derive β using the tail of the E_T^{iso} distribution, taking the region where $10 < E_T^{iso} < 25$ GeV.

After calculating β one easily retrieves $\xi_{\gamma+e}^T$ (up to a factor depending on α). Of course p_T^{iso} and E_T^{iso} are correlated, but one might worry if we are being completely consistent here. To illustrate this let us take an extreme example. Let us assume that all photons end up in $10 < p_T^{iso} < 25$ GeV but the E_T^{iso} has a normal distribution. In this case β will end up being the ratio of the TIGHT and NON-TIGHT photons and $\xi_{\gamma+e}^T$ will come out as zero because we will have labeled all photons as fakes. By construction this can never happen with the E_T^{iso} method. Now, the real world is of course far from that extreme, but there could be a small effect with photons that have a very different p_T^{iso} compared to E_T^{iso}. Fortunately we can report that we studied this, and the effects were found to be completely negligible, and are, in any case, captured in the systematic we defined above.

5.3.2 Extracting the photon isolation efficiency, ϵ_I

The photon isolation efficiency, denoted by ϵ_I, is the efficiency for a signal photon to pass the isolation criterion. This efficiency can be obtained by measuring how much of the isolation distribution ends up in the region of $E_T^{iso} < 4$ GeV. The procedure is done separately for the leading (ϵ_{I1}) and sub-leading (ϵ_{I2}) photon and for every bin for every observable we are interested in. In figures 5.3.1 the E_T^{iso} distributions are shown for the leading photon, and in addition the extracted signal distributions, $\xi_{\gamma+e}^T$ is also shown for both ways to extract β. The distributions for the sub-leading photon are very similar and are therefore omitted. In the plots α was chosen such that the signal distribution goes to zero in the region of $10 < E_T^{iso} < 25$ GeV.

As can be seen in these figures the signal distribution follows the TIGHT distribution initially but then rapidly drops off as more fakes enter the game at higher values of E_T^{iso}. It can also be seen that distribution does not change significantly between the jet preselection and the pretag stage. In figure 5.3.2 a comparison can be found between using p_T^{iso} and E_T^{iso} for finding β. It can be seen that the difference is located mainly in the tail, and will affect the final determination of the isolation efficiency by varying the ratio of isolated photons versus non-isolated ones. In table 5.3.1 the extracted isolation efficiencies are listed. We find 0.82 ± 0.00058(stat) $\pm\, 0.039$(syst) for the leading photon and 0.79 ± 0.0021(stat) $\pm\, 0.1$(syst) for the sub-leading photon in the pretag selection.

The following systematic uncertainties were considered when extracting the photon isolation efficiency:

Background boundary This uncertainty assesses the effects of our choice of the domain that we assume to be dominated by background events. Nominally this is $10 < p_T^{iso} < 25$ GeV, but the lower bound was varied from 8 to 12 GeV and the isolation efficiency recomputed.

Using E_T^{iso} As mentioned before there is the option of deriving β using the region defined by $10 < E_T^{iso} < 25$ GeV instead of using $10 < p_T^{iso} < 25$ GeV. To capture the

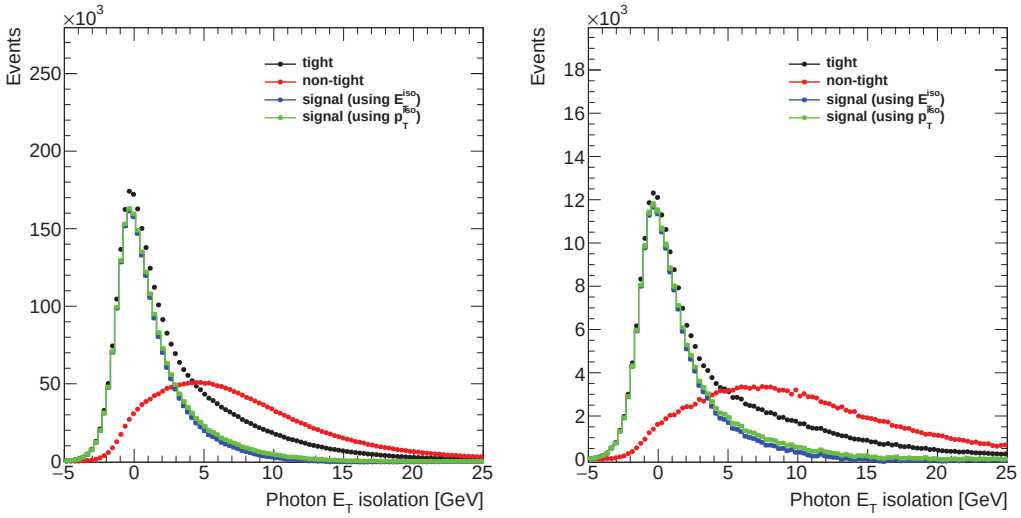

Figure 5.3.1: *Leading photon isolation distribution for photons passing the* TIGHT *cut or not passing it, along with the extracted signal distributions. Left is after the jet preselection, right after the pretag stage.*

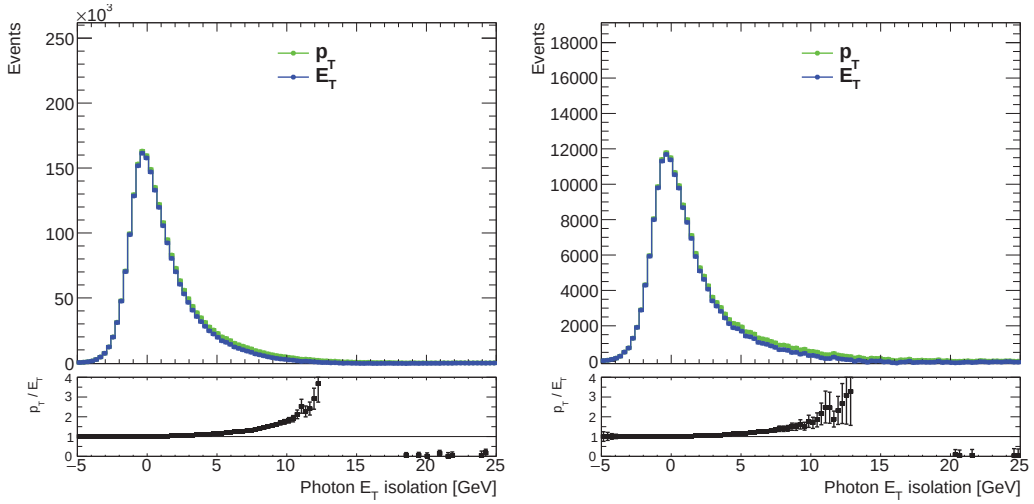

Figure 5.3.2: *The isolation distributions of the leading photon. Shown are photons passing or not passing the* TIGHT *requirement and the distribution for* TIGHT *photons corrected for fakes using the two methods described in the main text. On the left the jet preselection, on the right the pretag sample.*

ϵ_{I1}	jet preselection	pretag
nominal value	0.86 ± 0.00058	0.82 ± 0.018
background boundary up	-0.013	-0.016
background boundary down	$+0.018$	$+0.02$
using E_T^{iso} for background	± 0.035	± 0.06
total systematic uncertainty	± 0.039	± 0.063
ϵ_{I2}		
nominal value	0.77 ± 0.00078	0.79 ± 0.0021
background boundary up	-0.018	-0.016
background boundary down	$+0.026$	$+0.023$
using E_T^{iso} for background	± 0.1	± 0.097
total systematic uncertainty	± 0.1	± 0.1

Table 5.3.1: *Extracted isolation efficiencies from data events for the leading and sub-leading photons along with the statistical and systematic uncertainties.*

sensitivity on this choice we recomputed ϵ_I using the E_T^{iso} domain, took the difference with the nominal value and symmetrized it as a systematic. As can be seen from table 5.3.1 this systematic is dominant.

These systematics are propagated down everywhere ϵ_I is used.

5.4 Electron background

THE first of the two backgrounds we shall consider to the diphoton signal is created by electrons being misidentified as photons, for example because the track in the inner detector was not reconstructed properly or the electron emits a hard brehmsstrahlung photon. In addition genuine signal events may be lost the other way around due to photons being wrongly reconstructed as electrons. To gauge the effects of this background it is required to measure the probability of an electron being reconstructed as a photon and vice versa and also to do this in bins of any observables we might be interested in.

We start with a method to estimate the electron background in section 5.4.1. To make this work we will need to extract certain 'fake rates' which we will do in section 5.4.2. In section 5.4.3 we will then combine all this into a quantity that we will name the "electron fraction". This quantity will make it easy to gauge the electron background in a bin-by-bin way which we will use in chapter 6. It will also prove convenient when we combine the electron background with the jet background described in section 5.5.

5.4.1 Quantifying the electron background

Generally one can divide potential diphoton events into one of four categories: $\gamma\gamma$, γe, $e\gamma$, and ee. Where in $\gamma\gamma$ we select events with two good photons without regard to any electrons that might be present. In γe and $e\gamma$ we select events with exactly one good photon and a good electron, where in the γe case the photon has the higher transverse momentum, and in the $e\gamma$ case it is the electron. Events without any good photons and that have at least two good electrons end up in the ee category. No event can end up in more than one category. If we call $N_{\gamma\gamma}$ the number of events in the $\gamma\gamma$ category and that pass the selection cuts we can state:

$$N_{\gamma\gamma} = \epsilon_\gamma^2 n_{\gamma\gamma} + \epsilon_\gamma \rho_{e\to\gamma} \left(n_{\gamma e} + n_{e\gamma}\right) + \rho_{e\to\gamma}^2 n_{ee} \tag{5.4.1}$$

where n_{xx} is the number of true events of type xx that pass the selection cuts, ϵ_γ is the probability a photon is correctly reconstructed and classified as a photon, and $\rho_{e\to\gamma}$ is the probability an electron is reconstructed wrongly as a photon. Note that this means we are assuming $\rho_{e\to\gamma}$ (and the soon to be introduced $\rho_{\gamma\to e}$) is independent of the kinematical properties of the electron/photon and specifically whether the electron leads the photon or not. We believe that this approach is acceptable as the electron background will turn out to be small and its effects much smaller than the statistical uncertainty at the end of the analysis, and also the method for extracting $\rho_{e\to\gamma}$ used in this analysis does not allow for kinematic dependencies, but is rather simple and intuitive. We thus make the pragmatic choice to go with this method. We can measure $N_{\gamma\gamma}$ directly but to determine the cross-section we need to know $n_{\gamma\gamma}$, but for this we need to know the values of the other quantities in this equation. We can however measure $N_{\gamma e}$, $N_{e\gamma}$, and N_{ee} as well, each of which can be expressed using mostly the same quantities:

$$\begin{aligned}
N_{\gamma e} &= \epsilon_\gamma \rho_{\gamma\to e} n_{\gamma\gamma} + \epsilon_\gamma \epsilon_e n_{\gamma e} + \rho_{\gamma\to e}\rho_{e\to\gamma} n_{e\gamma} + \epsilon_e \rho_{e\to\gamma} n_{ee} && (5.4.2)\\
N_{e\gamma} &= \epsilon_\gamma \rho_{\gamma\to e} n_{\gamma\gamma} + \rho_{\gamma\to e}\rho_{e\to\gamma} n_{\gamma e} + \epsilon_\gamma \epsilon_e n_{e\gamma} + \epsilon_e \rho_{e\to\gamma} n_{ee} && (5.4.3)\\
N_{ee} &= \rho_{\gamma\to e}^2 n_{\gamma\gamma} + \epsilon_e \rho_{\gamma\to e} \left(n_{\gamma e} + n_{e\gamma}\right) + \epsilon_e^2 n_{ee} && (5.4.4)
\end{aligned}$$

where $\rho_{\gamma\to e}$ is the probability that a photon is wrongly reconstructed as an electron, and ϵ_e the probability to correctly classify and reconstruct an electron. Note that $n_{\gamma e}$ and $n_{e\gamma}$ can be combined in $N_{\gamma\gamma}$ and N_{ee}, this leads to the simplification of combining $N_{e\gamma}$ and $N_{\gamma e}$ into $N_{\gamma e+e\gamma}$:

$$\begin{aligned}
N_{\gamma e+e\gamma} &\equiv N_{\gamma e} + N_{e\gamma} && (5.4.5)\\
&= 2\epsilon_\gamma \rho_{\gamma\to e} n_{\gamma\gamma} + (\epsilon_e \epsilon_\gamma + \rho_{\gamma\to e}\rho_{e\to\gamma}) n_{\gamma e+e\gamma} + 2\epsilon_e \rho_{e\to\gamma} n_{ee} && (5.4.6)
\end{aligned}$$

where $n_{\gamma e+e\gamma}$ is defined as $n_{\gamma e} + n_{e\gamma}$. One can conveniently summarize the resulting three equation into one matrix equation:

$$\begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e+e\gamma} \\ N_{ee} \end{pmatrix} = \begin{pmatrix} \epsilon_\gamma^2 & \epsilon_\gamma \rho_{e\to\gamma} & \rho_{e\to\gamma}^2 \\ 2\epsilon_\gamma \rho_{\gamma\to e} & \epsilon_\gamma \epsilon_e + \rho_{\gamma\to e}\rho_{e\to\gamma} & 2\epsilon_e \rho_{e\to\gamma} \\ \rho_{\gamma\to e}^2 & \epsilon_e \rho_{\gamma\to e} & \epsilon_e^2 \end{pmatrix} \begin{pmatrix} n_{\gamma\gamma} \\ n_{\gamma e+e\gamma} \\ n_{ee} \end{pmatrix} \tag{5.4.7}$$

Whilst we are interested in knowing $n_{\gamma\gamma}$ we cannot use it directly. This is because $n_{\gamma\gamma}$ is sensitive to detector effects which we shall deal with during the unfolding process later on. We can introduce new quantities W_{xy} that are defined as:

$$W_{xy} \equiv \epsilon_x \epsilon_y n_{xy} \tag{5.4.8}$$

we thus absorb some of the unknowns into a new definition. It is the task of the unfolding process to revert $W_{\gamma\gamma}$ back to $n_{\gamma\gamma}$. Using this definition we can rewrite the matrix equation:

$$\begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e+e\gamma} \\ N_{ee} \end{pmatrix} = \begin{pmatrix} 1 & \frac{\rho_{e\to\gamma}}{\epsilon_e} & \left(\frac{\rho_{e\to\gamma}}{\epsilon_e}\right)^2 \\ 2\frac{\rho_{\gamma\to e}}{\epsilon_\gamma} & 1+\frac{\rho_{\gamma\to e}\rho_{e\to\gamma}}{\epsilon_e\epsilon_\gamma} & 2\frac{\rho_{e\to\gamma}}{\epsilon_e} \\ \left(\frac{\rho_{\gamma\to e}}{\epsilon_\gamma}\right)^2 & \frac{\rho_{\gamma\to e}}{\epsilon_\gamma} & 1 \end{pmatrix} \begin{pmatrix} W_{\gamma\gamma} \\ W_{\gamma e+e\gamma} \\ W_{ee} \end{pmatrix} \tag{5.4.9}$$

for brevity we can then introduce two new quantities, which we shall call fake rates:

$$f_{\gamma\to e} \equiv \frac{\rho_{\gamma\to e}}{\epsilon_\gamma} \tag{5.4.10}$$

$$f_{e\to\gamma} \equiv \frac{\rho_{e\to\gamma}}{\epsilon_e} \tag{5.4.11}$$

This then brings the matrix equation to this:

$$\begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e+e\gamma} \\ N_{ee} \end{pmatrix} = \begin{pmatrix} 1 & f_{e\to\gamma} & f_{e\to\gamma}^2 \\ 2f_{\gamma\to e} & 1+f_{\gamma\to e}f_{e\to\gamma} & 2f_{e\to\gamma} \\ f_{\gamma\to e}^2 & f_{\gamma\to e} & 1 \end{pmatrix} \begin{pmatrix} W_{\gamma\gamma} \\ W_{\gamma e+e\gamma} \\ W_{ee} \end{pmatrix} \tag{5.4.12}$$

Inverting this equation in order to extract $W_{\gamma\gamma}$ leads to:

$$\begin{pmatrix} W_{\gamma\gamma} \\ W_{\gamma e+e\gamma} \\ W_{ee} \end{pmatrix} = \frac{1}{\left(f_{\gamma\to e}f_{e\to\gamma}-1\right)^2} \begin{pmatrix} 1 & -f_{e\to\gamma} & f_{e\to\gamma}^2 \\ -2f_{\gamma\to e} & 1+f_{\gamma\to e}f_{e\to\gamma} & -2f_{e\to\gamma} \\ f_{\gamma\to e}^2 & -f_{\gamma\to e} & 1 \end{pmatrix} \cdot \begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e+e\gamma} \\ N_{ee} \end{pmatrix} \tag{5.4.13}$$

Thus, the coveted $W_{\gamma\gamma}$ can be calculated using the following expression:

$$W_{\gamma\gamma} = \frac{N_{\gamma\gamma} - f_{e\to\gamma}N_{\gamma e+e\gamma} + f_{e\to\gamma}^2 N_{ee}}{\left(f_{\gamma\to e}f_{e\to\gamma}-1\right)^2} \tag{5.4.14}$$

So we need to calculate the number of $\gamma\gamma$, $\gamma e + e\gamma$, and ee events in a certain observable bin and calculate $W_{\gamma\gamma}$ accordingly. We do, however, first need to extract the fake-rates.

An electron is reconstructed in a different way than a photon is. The principal difference is that electrons have charge and can therefore be traced in the inner detector as a track that is curved, whereas for photons there will only be a deposit in the calorimeter.

A photon can however interact with detector material and convert into an electron-positron pair. Such a photon is known as a converted photon. Likewise photons that do not convert are called unconverted photons. Electrons in this study were reconstructed as electrons but then reconstructed again as photons so they can be directly compared to photons and the cuts applied directly[†].

[†]Not doing this would make the analysis more sensitive to specific differences in reconstruction techniques beyond the simple requirement that there should be a reconstructed electron in the first place.

There is some ambiguity whether to reconstruct the electron as a converted or unconverted photon. For this study we have chosen to reconstruct them as converted photons. An option is to reconstruct them as both converted and unconverted photons and then take an average between the two options afterwards. However, this complicates things; if we consider ee events we have four possibilities, and when we consider eee events, which we shall do a bit later on, this number raises to eight. Because the electron background will turn out to be non-dominant anyway we will go with the converted option as this is the most likely fake candidate. We will explore using electrons reconstructed as unconverted photons as a systematic variation later on.

5.4.2 Extracting the fake rates from Z decays

While the case can certainly be made that $f_{e\to\gamma}$ and $f_{\gamma\to e}$ depend on the topology of an event and on its kinematical properties, in our analysis we shall assume the fake-rates to be constant together with $\rho_{e\to\gamma}$ and $\rho_{\gamma\to e}$ introduced before. Note that this, by virtue of expressions 5.4.10 and 5.4.11, also means we take the reconstruction efficiencies, ϵ_e and ϵ_γ to be constant. We do this for the pragmatic reason that there is no easy way to extract these dynamically as a function of the kinematical traits of an event. By deriving these fake-rates here, instead of retrieving them from another analysis we are at least making sure they are consistent with the various cuts and corrections made to the electrons and photons in this analysis.

Extracting $f_{e\to\gamma}$

Assuming $f_{e\to\gamma}$ is constant it can be conveniently extracted from the number of diphoton events in a control region where the invariant mass of the photon pair is close to the Z mass. This is because the Z boson does not decay directly to two photons, but readily decays to an electron-positron pair, so the overwhelming majority of such events are electrons faking photons coming from $Z \to e^+e^-$. In this case we can assume $W_{\gamma\gamma} = W_{\gamma e+e\gamma} = 0$ and the matrix equation 5.4.12 changes to:

$$N_{\gamma\gamma} = f^2_{e\to\gamma} W_{ee} \tag{5.4.15}$$
$$N_{\gamma e+e\gamma} = 2 f_{e\to\gamma} W_{ee} \tag{5.4.16}$$
$$N_{ee} = W_{ee} \tag{5.4.17}$$

One can simply extract $f_{e\to\gamma}$ by dividing the bottom two equations[†]:

$$f_{e\to\gamma} = \frac{N_{\gamma e+e\gamma}}{2N_{ee}} \tag{5.4.18}$$

To extract $N_{\gamma e+e\gamma}$ and N_{ee} the m_{ee} and $m_{\gamma e+e\gamma}$ spectra are taken and the continuous background is subtracted using a Landau distribution fitted to two sidebands around the Z peak in the regions $(66, 76)$ and $(121, 131)$ GeV. The Z peak is then fitted using a

[†]One might notice that $f_{e\to\gamma}$ can be extracted in a different way, by dividing the $N_{\gamma\gamma}$ equation by any of the two others. Theoretically this is true, but because the fake rate is not very big there will not be so many $\gamma\gamma$ events near the Z peak compared to ee events. Therefore, in the interest of obtaining the smallest statistical error it is best to go with N_{ee} and $N_{\gamma e+e\gamma}$.

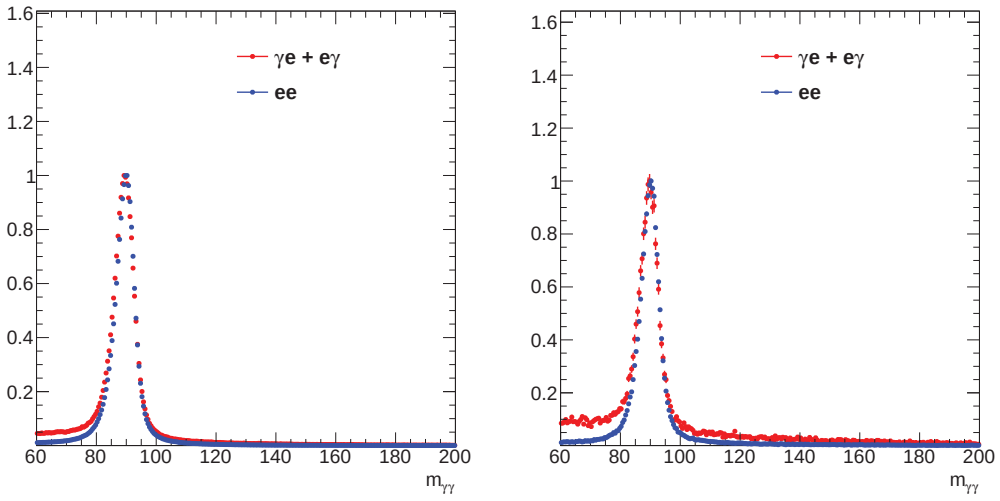

Figure 5.4.1: *Diphoton invariant mass spectra for ee and $\gamma e + e\gamma$ events. For visibility the distributions have been scaled so that the peaks have a height of one. On the left the distribution after the jet preselection, on the right after the pretag stage.*

Gaussian in an interval of ± 20 GeV around the Z mass ($m_Z = 91.1876$ GeV [8]) where the Gaussian's mean is set to the Z mass. The events, after the background subtraction found within the region of $(m_Z - 1.5\sigma, m_Z + 1.5\sigma)$, where σ is the fitted width of the Gaussian, are taken to be associated to the peak and are used as N_{ee} and $N_{\gamma e + e\gamma}$. Since the peak is fitted for ee and $\gamma e + e\gamma$ events separately any shift in the peak's location (due to reconstruction differences between electrons and photons) is made unimportant.

The invariant mass spectrum for ee and $\gamma e + e\gamma$ can be found in figure 5.4.1. The decompositions into the Z peak and background can be found in figures 5.4.2 and 5.4.3. We find an $e \to \gamma$ fake rate of $0.063 \pm 0.00034\,(\text{stat})$ in the pretag selection and $0.070 \pm 6.9 \cdot 10^{-5}\,(\text{stat})$ in the jet preselection.

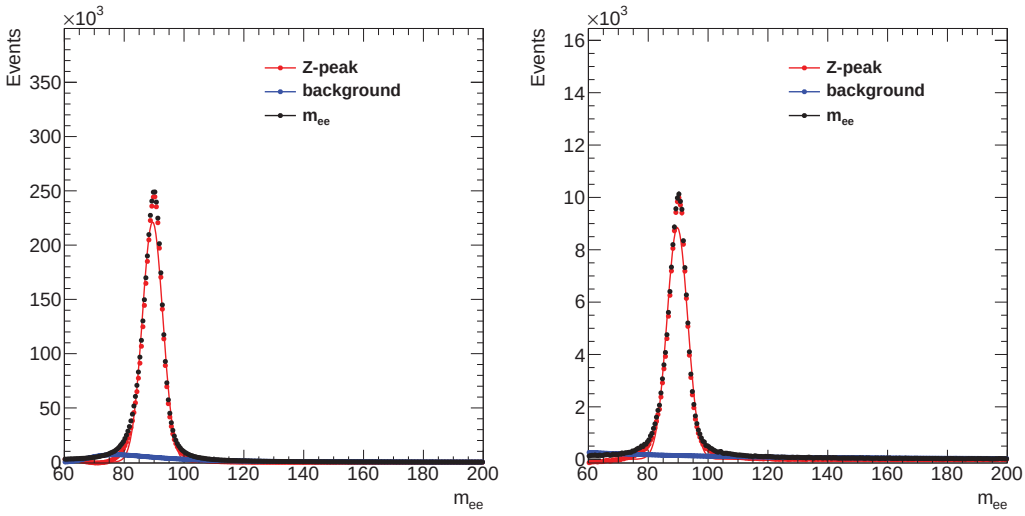

Figure 5.4.2: *Invariant mass spectrum for ee events with the decomposition into the Z peak and background. The fit made to the Z peak is also shown as a solid red line. The left plot was made after the jet preselection, the right one after the pretag stage.*

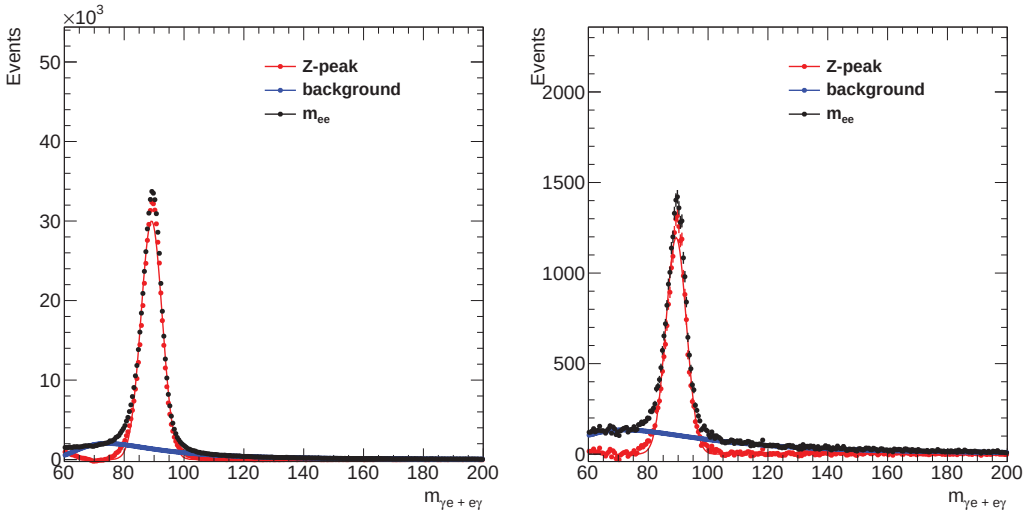

Figure 5.4.3: *Invariant mass spectrum for $\gamma e + e\gamma$ events with the decomposition into the Z peak and background. The fit made to the Z peak is also shown as a solid red line. The left plot was made after the jet preselection, the right one after the pretag stage.*

	Jet preselection [%]	Pretag [%]
Z mass interval	± 5.92	± 4.31
Sideband interval	± 4.74	± 9.09
Using unconverted electrons	< 0.1	< 0.1
Total	± 7.58	± 10.06

Table 5.4.1: *Systematics as a percentage of the nominal value on the extraction of $f_{e\to\gamma}$ from the measured data. The total systematic error listed in this table was calculated by adding each type in quadrature.*

The following two systematic errors on the fake rate were considered and they were added in quadrature:

Sideband-intervals The effects of our specific choice of sidebands was evaluated by shifting the sideband regions up and down by 5 GeV. So for example, for the up variation the sidebands were $(71, 81)$ and $(126, 136)$ GeV. We found that the variation up and down often has the same sign, which was traced to how the sidebands will sometimes dip into negative values after the background subtraction. Since we do not want to be sensitive to this artificial effect we will take the larger absolute variation between up and down, and symmetrize it around the nominal value.

Z mass interval The peak interval was chosen to be $(m_Z - 1.5\sigma, m_Z + 1.5\sigma)$, but we should assess the effects of our choice of 1.5σ. As a systematic we include a shift down to 1.0σ and up to 2.0σ which will alter the extracted N_{ee} and $N_{\gamma e+e\gamma}$. Like the sideband-intervals the largest variation was symmetrized around the nominal value.

Unconverted photons With this systematic we assess the effect of having our electrons reconstructed as unconverted photons as opposed to converted photons as explained before. The difference with the nominal result is symmetrized. The variation caused by this was found to be extremely tiny.

The size of these systematics on $f_{e\to\gamma}$ can be seen in table 5.4.1.

Extracting $f_{\gamma\to e}$

The photon to electron fake rate, $f_{e\to\gamma}$ can be extracted by examining the $m_{ee\gamma}$ and m_{eee} spectra around the Z peak. Because the peak is dominated by $Z \to e^+e^-\gamma$ decays we can state:

$$N_{ee\gamma} = W_{ee\gamma} \tag{5.4.19}$$
$$N_{eee} = f_{\gamma\to e} W_{ee\gamma} \tag{5.4.20}$$

	Systematic uncertainty [%]
Z mass interval	± 3.70
Sideband interval	± 93.18
Using unconverted electrons	< 0.1
Total	± 93.26

Table 5.4.2: *Systematics as a percentage of the nominal value on the extraction of $f_{\gamma\to e}$ from the measured data. The total systematic error listed in this table was calculated by adding the larger of the up/down values for each type in quadrature.*

where N_{xxx} is the number of events near the Z peak of type xxx and $W_{ee\gamma}$ is analogous to the W quantities introduced before. The fake-rate can then be calculated as:

$$f_{\gamma\to e} = \frac{N_{eee}}{N_{ee\gamma}} \tag{5.4.21}$$

Because we now select three objects the cutflow is a little different. The differences are that we now require the p_T cut on all three objects, we also require all three to be separated by an ΔR of at least 0.4. Events entering in the eee category were not allowed to have any spurious photons and likewise events in the $ee\gamma$ category were not allowed a third electron. Finally the invariant mass cut was changed to $45 < m_{ee} < 85$ GeV, for the two leading electrons, in order to remove contributions from $Z \to e^+e^-$. The Z peak is fitted in the same way as was done for the extraction of $f_{e\to\gamma}$ and similar systematics are considered.

Even so the number of triple electron events passing the cuts beyond the jet preselection are tiny, and it was impossible to get a stable fit for the other cutflow points. As such we will use the jet preselection and there we find a value of 0.012 ± 0.0022 (stat) for $f_{\gamma\to e}$. The systematics are summarized in table 5.4.2. The invariant mass spectrum can be found in the top of figure 5.4.4 and the fits to the Z peak in the bottom. The extracted value has large systematic uncertainties because of the small number of events, but agrees with the value found in [50], and is quite small; so fortunately its effect on the analysis is negligible.

Summary

In this section we extracted $f_{e\to\gamma}$ and $f_{\gamma\to e}$ from leptonic Z decays. We found that both are quite tiny: 0.063 ± 0.00034 (stat) ± 0.0063 (syst) and 0.012 ± 0.022 (stat) ± 0.011 (syst) respectively. For $f_{e\to\gamma}$ we were able to determine this for the pretag selection. For $f_{\gamma\to e}$ we were unable to find a value beyond the jet preselection due to a shortage of statistics.

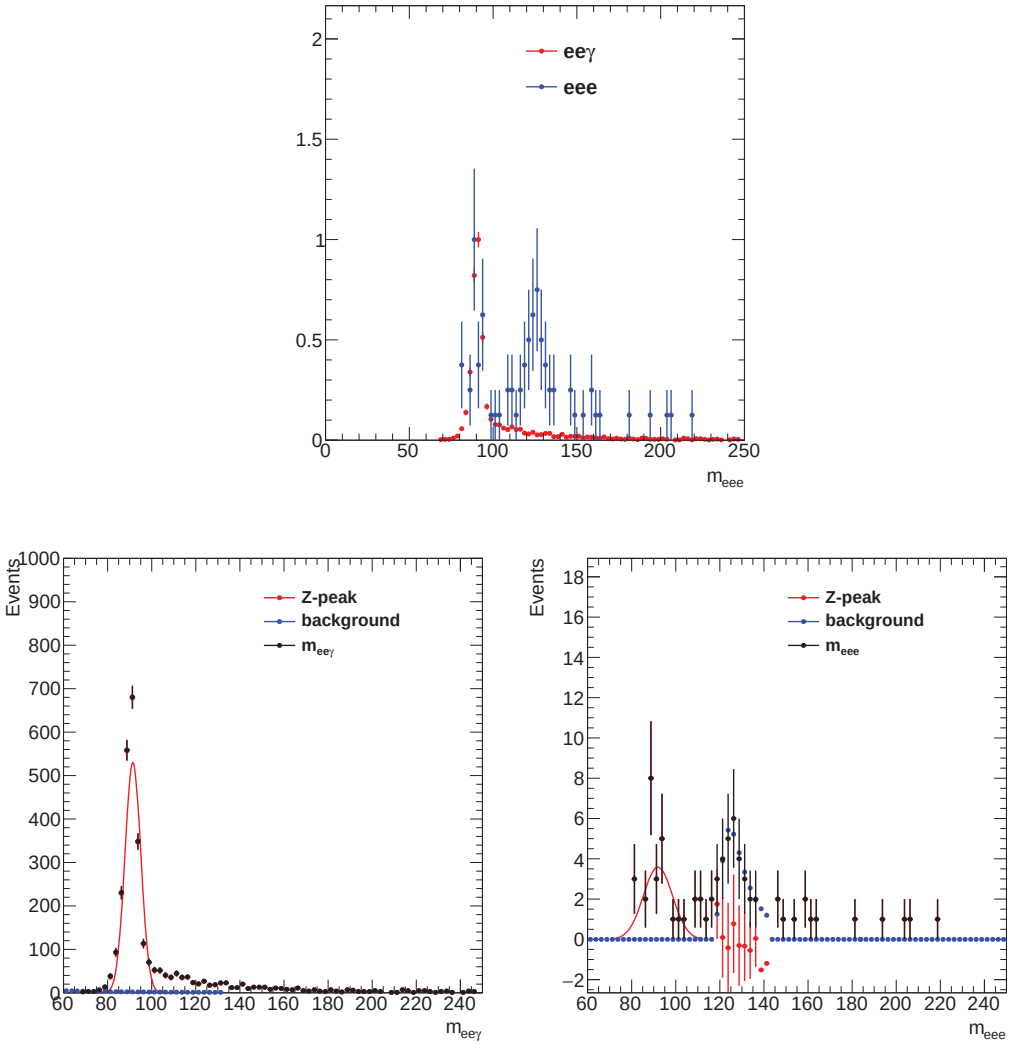

Figure 5.4.4: *In the top the tri-electron invariant mass spectra for eee and $ee\gamma$ events in the jet preselection. For visibility the distributions have been scaled so that the peaks have a height of one. In the bottom the $ee\gamma$ (left) and eee (right) invariant mass spectra after the jet preselection with the decomposition into the Z peak and background. The fit made to the Z peak is also shown as a solid red line.*

5.4.3 Electron fraction

Now that we have determined the fake rates we are now in the position to calculate $W_{\gamma\gamma}$ using:

$$W_{\gamma\gamma} = \frac{N_{\gamma\gamma} - f_{e\to\gamma}N_{\gamma e+e\gamma} + f_{e\to\gamma}^2 N_{ee}}{\left(f_{\gamma\to e}f_{e\to\gamma} - 1\right)^2} \tag{5.4.22}$$

which we derived before in section 5.4.1. We assumed that the fake rates do not depend on our observables, but we do need to find $W_{\gamma\gamma}$ in bins of these observables in order to obtain differential cross-sections later. This means we need to find $N_{\gamma\gamma}$, $N_{\gamma e+e\gamma}$, and N_{ee} also for each bin in any observable we might care to investigate. It is useful to quantify the effects of the electron background in a bin-by-bin quantity that we shall call the electron fraction, Γ_e, which we shall relate to $W_{\gamma\gamma}$ in the following way:

$$W_{\gamma\gamma} = N_{\gamma\gamma}\left(1 - \Gamma_e\right) \tag{5.4.23}$$

We can derive the functional form of Γ_e by equating expressions 5.4.22 and 5.4.23:

$$N_{\gamma\gamma}\left(1 - \Gamma_e\right) = \frac{N_{\gamma\gamma} - f_{e\to\gamma}N_{\gamma e+e\gamma} + f_{e\to\gamma}^2 N_{ee}}{\left(f_{\gamma\to e}f_{e\to\gamma} - 1\right)^2} \tag{5.4.24}$$

$$1 - \Gamma_e = \frac{1 - f_{e\to\gamma}\frac{N_{\gamma e+e\gamma}}{N_{\gamma\gamma}} + f_{e\to\gamma}^2\frac{N_{ee}}{N_{\gamma\gamma}}}{\left(f_{\gamma\to e}f_{e\to\gamma} - 1\right)^2} \tag{5.4.25}$$

$$\Gamma_e = \frac{f_{e\to\gamma}\frac{N_{\gamma e+e\gamma}}{N_{\gamma\gamma}} - f_{e\to\gamma}^2\frac{N_{ee}}{N_{\gamma\gamma}} - 1}{\left(f_{\gamma\to e}f_{e\to\gamma} - 1\right)^2} + 1 \tag{5.4.26}$$

As mentioned in section 5.4.2 the value found for $f_{\gamma\to e}$ is tiny and it was not possible to extract it for each cutflow point. We can estimate what the effect of this tiny $f_{\gamma\to e}$ is. Note that it only appears in the denominator of Γ_e and is multiplied by $f_{e\to\gamma}$ which is also tiny. The Taylor expansion of Γ_e around the point $f_{\gamma\to e}f_{e\to\gamma} = 0$ is:

$$\Gamma_e \approx f_{e\to\gamma}\frac{N_{\gamma e+e\gamma}}{N_{\gamma\gamma}} - f_{e\to\gamma}^2\frac{N_{ee}}{N_{\gamma\gamma}} + 2\left[f_{e\to\gamma}\frac{N_{\gamma e+e\gamma}}{N_{\gamma\gamma}} - f_{e\to\gamma}^2\frac{N_{ee}}{N_{\gamma\gamma}} - 1\right]\cdot f_{\gamma\to e}f_{e\to\gamma} + \mathcal{O}\left(f_{\gamma\to e}^2 f_{e\to\gamma}^2\right) \tag{5.4.27}$$

With the values we find for the fake rates the combination $f_{\gamma\to e}f_{e\to}$ has a value on the order of $\mathcal{O}\left(10^{-4}\right)$ which means the second term is already much smaller than the first one. Therefore we will neglect $f_{\gamma\to e}$ from now on. We note that we will see later that another uncertainty on the final analysis result will have a far larger impact than this simplification. Γ_e now simplifies to the first term of the Taylor expansion:

$$\Gamma_e = f_{e\to\gamma}\frac{N_{\gamma e+e\gamma}}{N_{\gamma\gamma}} - f_{e\to\gamma}^2\frac{N_{ee}}{N_{\gamma\gamma}} \tag{5.4.28}$$

The electron fraction can be found in table 5.4.3 along with the systematic errors that have been propagated forward from the determination of $f_{e\to\gamma}$. We find a value of $0.014\pm 5.2\cdot 10^{-5}$ (stat) $\pm\, 1.7\cdot 10^{-5}$ (syst) in the pretag selection with the largest uncertainty coming from the shifting sidebands and Z mass interval used in the extraction of $f_{e\to\gamma}$. Thus, only a very tiny fraction of our events come from electron fakes. Note that the statistical uncertainties on the event yields have been propagated into the statistical uncertainty on Γ_e. The statistical uncertainties on the fake rates have been added as a systematic on Γ_e.

Γ_e	Jet preselection	Pretag
Nominal value	$0.013 \pm 1.8 \cdot 10^{-5}$	$0.014 \pm 5.2 \cdot 10^{-5}$
$f_{e\to\gamma}$ Z mass interval	± 0.00044	± 0.0005
$f_{e\to\gamma}$ sideband interval	± 0.00036	± 0.001
$f_{e\to\gamma}$ fake rate statistics	$\pm 7.6 \cdot 10^{-6}$	$\pm 6.2 \cdot 10^{-5}$
Using unconverted electrons	~0	~0
Total systematic error	± 0.00061	± 0.0012

Table 5.4.3: *Extracted electron fraction for all events in the jet preselection and the pretag selection. The total systematic error listed in this table was calculated by adding the larger of the up/down values for each type in quadrature.*

5.5 Jet background

APART from electrons being wrongly reconstructed as photons, there is a remaining sizable background from jets being wrongly reconstructed as photons. Therefore, it is prudent to estimate the contribution of such fake photons. In section 5.5.1 we will detail the method chosen to do so. In this method we will neglect electrons faking photons. In other words with this simplication we assume only jets fake photons. In section 5.5.2 we will fix this by combining the jet background with the electron background described before in section 5.4.

5.5.1 Estimating the jet background from a fit

A measured diphoton event may be sorted into one of four classes: $\gamma\gamma$ for genuine diphoton events, γj and $j\gamma$ for events where either the sub-leading or the leading photon is actually a jet faking a photon, and finally jj for double jet events where both are wrongly reconstructed as photons. The aim is to extract the number of events in the $\gamma\gamma$ category. The events we measure can be subdivided into sixteen categories, depending on whether the leading photons pass or fail either the identification or isolation requirements. Each photon candidate falls into one of four classes, A, B, C, or D, see figure 5.5.1. The two are then combined into one of sixteen categories. So for example an event with the leading photon candidate passing both requirements, but with the sub-leading photon candidate failing the isolation criterion will end up in category AB.

Events passing the jet preselection will end up in category AA, because it requires the leading photons to pass both the isolation and the identification requirements. Most of the signal is in this category, but some of the signal can leak into the other categories since the efficiency of the identification and isolation cuts is not 100%. At the same time there are $j\gamma$, γj, and jj events inside category AA. If we call the number of events of type yy before requiring isolation or TIGHT identification, W_{yy} and if we call the number

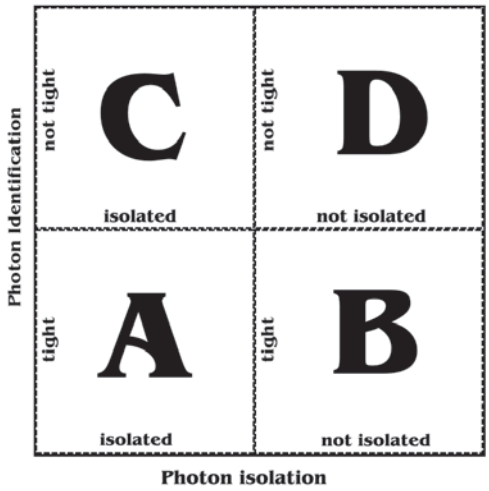

Figure 5.5.1: *The four different classifications into which each photon candidate can be sorted.*

of events measured in category AA, N_{AA} we can write:

$$\begin{aligned} N_{AA} &= \epsilon_{I1}\epsilon_{T1}\epsilon_{I2}\epsilon_{T2}W_{\gamma\gamma} \\ &+ \epsilon_{I1}\epsilon_{T1}f_{I2}f_{T2}\phi_2^2 W_{\gamma j} \\ &+ f_{I1}f_{T1}\epsilon_{I2}\epsilon_{T2}\phi_1^2 W_{j\gamma} \\ &+ g_{I1}g_{T1}g_{I2}g_{T2}\zeta^2\phi_1^2\phi_2^2 W_{jj} \end{aligned} \tag{5.5.1}$$

where ϵ_{I1} is the isolation efficiency of the leading photon candidate, and ϵ_{I2} the same but for the sub-leading candidate. The TIGHT identification efficiency is denoted by ϵ_{T1} and ϵ_{T2} for the leading and sub-leading candidates respectively. The f's are similar quantities but involve the efficiency of a jet faking a photon passing a particular cut. The g's are the same as the f's but are specific for jj events. The reason is that the efficiencies depend strongly on whether they involve a dijet event or not. The ϕ_1 and ϕ_2 encode a potential correlation between the identification and isolation efficiencies of the leading or sub-leading jet fakes. Finally ζ encodes a possible correlation between the isolation of the photon candidates in jj events[†]. Using the same quantities we can express all sixteen N_{xx} categories, of which we only show a few here[‡]:

$$\begin{aligned} N_{AB} &= \epsilon_{I1}\epsilon_{T1}(1-\epsilon_{I2})\epsilon_{T2}W_{\gamma\gamma} \\ &+ \epsilon_{I1}\epsilon_{T1}(1-\phi_2 f_{I2})f_{T2}W_{\gamma j} \\ &+ f_{I1}f_{T1}(1-\epsilon_{I2})\epsilon_{T2}\phi_1^2 W_{j\gamma} \\ &+ g_{I1}g_{T1}(1-\zeta\phi_2 g_{I2})g_{T2}\phi_1^2 W_{jj} \end{aligned} \tag{5.5.2}$$

[†]For example if one of the jets is isolated the probability of the other jet also being isolated increases by a factor of ζ. We assume this factor does not distinguish between leading and sub-leading jets.

[‡]The rest may be reconstructed in an analogous manner.

$$\begin{aligned} N_{AC} &= \epsilon_{I1}\epsilon_{T1}\epsilon_{I2}(1-\epsilon_{T2})W_{\gamma\gamma} \\ &+ \epsilon_{I1}\epsilon_{T1}f_{I2}(1-\phi_2 f_{T2})W_{\gamma j} \\ &+ f_{I1}f_{T1}\epsilon_{I2}(1-\epsilon_{T2})\phi_1^2 W_{j\gamma} \\ &+ g_{I1}g_{T1}g_{I2}(1-\phi_2 g_{T2})\zeta^2\phi_1^2 W_{jj} \end{aligned} \tag{5.5.3}$$

$$\begin{aligned} N_{AD} &= \epsilon_{I1}\epsilon_{T1}(1-\epsilon_{I2})(1-\epsilon_{T2})W_{\gamma\gamma} \\ &+ \epsilon_{I1}\epsilon_{T1}(1-f_{I2})(1-f_{T2})W_{\gamma j} \\ &+ f_{I1}f_{T1}(1-\epsilon_{I2})(1-\epsilon_{T2})\phi_1^2 W_{j\gamma} \\ &+ g_{I1}g_{T1}(1-\zeta g_{I2})(1-g_{T2})\phi_1^2 W_{jj} \end{aligned} \tag{5.5.4}$$

$$\begin{aligned} N_{DC} &= (1-\epsilon_{I1})(1-\epsilon_{T1})\epsilon_{I2}(1-\epsilon_{T2})W_{\gamma\gamma} \\ &+ (1-\epsilon_{I1})(1-\epsilon_{T1})f_{I2}(1-\phi_2 f_{T2})W_{\gamma j} \\ &+ (1-f_{I1})(1-f_{T1})\epsilon_{I2}(1-\epsilon_{T2})W_{j\gamma} \\ &+ (1-\zeta g_{I1})(1-g_{T1})g_{I2}(1-\phi_2 g_{T2})W_{jj} \end{aligned} \tag{5.5.5}$$

The sixteen equations have nineteen unknowns. We will fix seven. The isolation efficiencies, ϵ_{I1} and ϵ_{I2}, are taken from the E_T^{iso} distribution that is derived from data using the method described in section 5.3. The identification efficiencies, ϵ_{T1} and ϵ_{T2}, are taken from Monte Carlo simulations as detailed below, and ϕ_1, ϕ_2, and ζ are taken to be equal to unity for now, but will later be independently varied by $\pm 10\%$ to gauge how much possible correlations affect the outcome, and enter the analysis as a systematic uncertainty. This fitting procedure can also be done in a bin-by-bin fashion which we shall do in the next chapter.

The remaining twelve unknowns are then extracted via a fit method and the number of genuine $\gamma\gamma$ events is then calculated as:

$$N_{\gamma\gamma} = \epsilon_{I1}\epsilon_{T1}\epsilon_{I2}\epsilon_{T2}W_{\gamma\gamma} \tag{5.5.6}$$

Extracting the identification efficiencies is done by examining how often the leading or sub-leading photon passes the TIGHT criterion in the signal simulation samples. No requirement is made on the other photon or on the isolation of the photon in question. The extracted ϵ_{T1} and ϵ_{T2} were 0.93 ± 0.00056(stat)± 0.025(syst) and 0.88 ± 0.00072(stat)± 0.0027(syst) respectively in the pretag selection. See also table 5.5.1. The SHERPA sample was used for the nominal value, the difference to the value to the one obtained by using the PYTHIA sample instead was symmetrized and taken as a systematic to gauge the dependence on the simulation method.

The number of events counted in each event class events can be found in table 5.5.2. A signal purity can now be defined:

$$\mathcal{P}_{\gamma\gamma} \equiv \frac{N_{\gamma\gamma}}{N_{\gamma\gamma} + N_{j\gamma} + N_{\gamma j} + N_{jj}} \tag{5.5.7}$$

$$\begin{aligned} &= \epsilon_{I1}\epsilon_{T1}\epsilon_{I2}\epsilon_{T2}W_{\gamma\gamma}[\epsilon_{I1}\epsilon_{T1}\epsilon_{I2}\epsilon_{T2}W_{\gamma\gamma} + \phi_2^2 f_{I1}f_{T1}\epsilon_{I2}\epsilon_{T2}W_{j\gamma} \\ &+\phi_1^2\epsilon_{I1}\epsilon_{T1}f_{I2}f_{T2}W_{\gamma j} + \phi_1^2\phi_2^2 g_{I1}g_{T1}g_{I2}g_{T2}\zeta^2 W_{jj}]^{-1} \end{aligned} \tag{5.5.8}$$

where the fit errors of the variables are propagated forward into the $\gamma\gamma$ purity, $\mathcal{P}_{\gamma\gamma}$. This then describes which fraction of our events can be considered real diphoton events. Its

ϵ_{T1}	jet preselection	pretag
nominal value	0.93 ± 0.00018	0.93 ± 0.00056
using PYTHIA	± 0.0033	± 0.025
ϵ_{T2}		
nominal value	0.88 ± 0.00022	0.88 ± 0.00072
using PYTHIA	± 0.0081	± 0.0027

Table 5.5.1: *Extracted identification efficiencies from simulation for the leading and subleading photons along with the statistical and systematic uncertainties.*

event class	jet preselection	pretag
AA	799927	60162
AB	401761	29230
AC	363024	20501
AD	680684	53062
BA	210190	28122
BB	216891	18770
BC	124256	11989
BD	422745	36674
CA	196940	11800
CB	136768	7250
CC	125464	5645
CD	247738	13659
DA	309796	47485
DB	363217	33700
DC	193109	20743
DD	718721	65273
total	5511231	464065

Table 5.5.2: *Event counts per event class.*

value and the associated uncertainties can be found in table 5.5.3. We observe that by far the largest uncertainty is caused by the two different photon isolation spectra extraction methods. However, its significance drops from 11% to 5.3% after the pretag cuts. The contributions of the various background types can be seen in table 5.5.4. The values of the other variables present in the fit are presented in table 5.5.5.

As expected we find 60162 events in the AA class in the pretag selection, which corresponds to all data events in that selection. Similar event counts are found in the other fifteen classes, the largest being DD with 65273 events. From this we find that the $\gamma\gamma$ purity is 0.90 ± 0.0012(stat) $\pm\ 0.055$(syst), which indicates that about one-tenth of our events are jet fakes.

$\gamma\gamma$ purity	jet preselection	pretag
Nominal value	0.88 ± 0.00053	0.90 ± 0.0012
(jet) Using E_T^{iso} for ϵ_I [%]	± 11.0	± 5.3
(jet) Using PYTHIA for ϵ_T [%]	± 0.18	± 0.73
(jet) ϵ_{I1} background boundary up [%]	1.0	0.4
(jet) ϵ_{I1} background boundary down [%]	-1.3	-0.5
(jet) ϵ_{I2} background boundary up [%]	1.6	0.71
(jet) ϵ_{I2} background boundary down [%]	-2.1	-0.96
(jet) ϕ_1 up [%]	-1.3	-1.2
(jet) ϕ_1 down [%]	0.94	0.97
(jet) ϕ_2 up [%]	-3.3	-2.0
(jet) ϕ_2 down [%]	2.3	1.6
(jet) ζ up [%]	0.18	0.18
(jet) ζ down [%]	-0.24	-0.21
(general) Pile-up reweighting up [%]	-0.13	-0.12
(general) JVF cut up [%]	< 0.1	-0.9
(general) JVF cut down [%]	< 0.1	-0.59
Total systematic error [%]	± 12.0	± 6.1

Table 5.5.3: *Signal purity obtained from the fit. The total systematic uncertainty was calculated by taking maximum of the up / down values of each systematic and adding them in quadrature. The word in parentheses is to remind the reader of the general source of the uncertainty. Only systematics which have an effect above 0.1% are mentioned.*

	jet preselection	pretag
$N_{\gamma\gamma}$	699510.0 ± 1539.4	54133.4 ± 354.5
$N_{j\gamma}$	28189.9 ± 259.5	2204.2 ± 42.9
$N_{\gamma j}$	62574.1 ± 346.5	3408.6 ± 58.6
N_{jj}	9085.3 ± 34.2	363.5 ± 5.6
total	799359.3	60109.7

Table 5.5.4: *Composition of the sample in event classes as determined by the fit. The quoted error is the fit uncertainty only. Note that the numbers in the top row correspond closely to the numbers found for class AA in table 5.5.2, as expected.*

	jet preselection	pretag
f_{I1}	0.339 ± 0.00101 (stat) ± 0.00225 (syst)	0.143 ± 0.00208 (stat) ± 0.0284 (syst)
f_{I2}	0.306 ± 0.00155 (stat) ± 0.0941 (syst)	0.206 ± 0.00216 (stat) ± 0.0141 (syst)
f_{T1}	0.191 ± 0.00146 (stat) ± 0.086 (syst)	0.244 ± 0.0027 (stat) ± 0.0598 (syst)
f_{T2}	0.203 ± 0.000959 (stat) ± 0.112 (syst)	0.219 ± 0.00256 (stat) ± 0.0868 (syst)
g_{I1}	0.226 ± 0.000439 (stat) ± 0.0131 (syst)	0.14 ± 0.00134 (stat) ± 0.0176 (syst)
g_{I2}	0.19 ± 0.000414 (stat) ± 0.0118 (syst)	0.196 ± 0.00134 (stat) ± 0.00176 (syst)
g_{T1}	0.316 ± 0.000452 (stat) ± 0.0206(syst)	0.288 ± 0.00163 (stat) ± 0.0306(syst)
g_{T2}	0.291 ± 0.000457 (stat) ± 0.0296 (syst)	0.265 ± 0.00169 (stat) ± 0.0451 (syst)

Table 5.5.5: *Fit values along with statistical and systematic uncertainties. These are the efficiencies for a jet faking a photon to pass the photon isolation or identification requirements.*

5.5.2 Combining with the electron background

Combining the jet background with the electron background requires some careful thought. The problem arises because during the determination of the electron fraction one assumes there are no jet fakes whilst during the determination of the jet fakes one assumes there are no electron fakes. Neither assumption of course bears out in reality, strictly speaking.

However, the fact that during the electron subtraction the photon sample contains jets faking photons is not a problem. One can easily reinterpret the extracted fake rates to describe how often electrons can fake an object that looks like a photon or like a jet that looks like a photon. One then takes care of these jets that look like photons during the jet background subtraction. More problematic are the jets faking electrons because these inflate the fake rates and are not taken care of later. One can appreciate the problem by assuming all jets fake electrons (and electrons only). In that situation the fake rates, $f_{e\to\gamma}$ and $f_{\gamma\to e}$ will be largely unaffected since these are derived around the Z peak and when the continuum background is subtracted off. It does however affect the event counts of type ee and $\gamma e + e\gamma$. Let us examine the expression for the electron fraction (expression 5.4.28 on page 96):

$$\Gamma_e = f_{e\to\gamma}\frac{N_{\gamma e+e\gamma}}{N_{\gamma\gamma}} - f_{e\to\gamma}^2\frac{N_{ee}}{N_{\gamma\gamma}} \tag{5.5.9}$$

Many jets faking electrons will increase $N_{\gamma e+e\gamma}$ and N_{ee}, which in turn will affect Γ_e. Now of course the situation we are dealing with outside this hypothetical scenario is not as bad as this, but there is still some effect on Γ_e from jets faking electrons.

In practice, however, this does not matter too much. One can calculate the electron fraction whilst there are still jet fakes in the sample. We interpret the result slightly differently. Instead of obtaining photon to electron or electron to photon fake rates one obtains fake rates between a sample of photons and jets faking photons, and a sample of electrons and jets faking electrons. The electron fraction is therefore actually the fraction of events where you have electrons faking photons or jets faking electrons.

One way, perhaps, to be more correct is to subtract one background, then the other one, and then go back to the first one, refining it, etc. and repeat it for a few cycles until stable values are reached. In this analysis we opt not to choose this route. This is because only marginal gains would be made that are of no real consequence anyway since the electron background is tiny, especially when compared to other uncertainties that we shall discuss later.

To proceed we calculate Γ_e for each of the sixteen event types and then replace the event counts accordingly:

$$N_{XX} \to N_{XX}\left(1 - \Gamma_{e,XX}\right) \tag{5.5.10}$$

where $\Gamma_{e,XX}$ is the electron fraction for a particular event class (e.g. AA, CD, etc.). After which we use the fitting method described in the previous section normally where all systematic uncertainties on the Γ_e's are propagated forward into the eventual fit output. The same identification and isolation cuts are made on the electrons as are made on the photons.

This procedure has some effect on the results obtained before, but the numbers do not vary much, see table 5.5.6.

	jet preselection	pretag
$N_{\gamma\gamma}$	670488.0 ± 1487.6	52443.8 ± 345.1
$N_{j\gamma}$	28628.3 ± 254.1	2201.8 ± 42.5
$N_{\gamma j}$	63284.0 ± 340.0	3385.1 ± 57.8
N_{jj}	9025.5 ± 34.0	360.9 ± 5.5
total	771425.8	58391.5

Table 5.5.6: *Composition of the sample as determined by the fit. The quoted error is the fit uncertainty only.*

5.6 Systematic uncertainties and corrections

In this section we will discuss the systematic uncertainties not directly connected to any specific scheme used in the analysis but intrinsic to the ATLAS detector and reconstruction software. I will also present various corrections made to the simulated events in order to make them correspond better to the data. For all these corrections the recommendations of the relevant ATLAS performance group were followed. This list applies equally to the $jj\gamma\gamma$ selections made in this chapter as well as the $b\bar{b}\gamma\gamma$ selections of the next.

Pile-up reweighting The average number of proton-proton interactions per bunch crossing varies between simulation and real data. Therefore, the simulation events are reweighted in order to match the average measured in data. This reweighting is not perfect and the effects of this reweighting are assessed by varying this weight $\pm 10\,\%$ as a systematic.

Electron reconstruction efficiency Electrons are not directly used by the analysis, but they can sometimes be incorrectly reconstructed as photons. This is a background we will tackle in section 5.4. In the method we employ there we will also require electrons, and therefore we need to correct these objects and inspect how the uncertainties specific to electrons affect the greater analysis. The electron reconstruction efficiency governs how well the detector correctly reconstructs electrons. Since this is efficiency is not the same in data and simulations it is corrected by reweighting events using a scale factor obtained from dedicated studies within ATLAS. The uncertainty contained in this scale factor are propagated to the analysis by varying the scale factor up and down by this uncertainty and checking its effect on the analysis.

Electron identification efficiency The electrons used in the analysis have are required to pass certain identification criteria that involve for example shower shapes in the electromagnetic calorimeter and tracking information. These criteria are set by ATLAS in order to remove electrons with poor reconstruction. Again the efficiency of this identification is different in data and simulations and therefore

needs to be corrected with a scale factor. The scale factor is varied up and down within its uncertainty and the effects propagated into the analysis.

Photon identification efficiency Likewise there is an efficiency associated with the identification flag attached to photons which is corrected in simulated events via a scale factor. This scale factor is varied up and down and the analysis rerun to gauge the effects.

Electron energy correction The energy of an electron as measured by the electromagnetic calorimeter needs to be calibrated in data and simulation events. These calibrations are obtained by baseline measurements on well understood processes, for example $Z \to e^+e^-$, and depend also on the geometric model used in the reconstruction software. The various uncertainties contained in these procedures were combined into a single uncertainty that is then varied up and down.

Photon energy correction This correction operates in a way analogous to the electron energy correction.

Jet smearing The detector has a finite energy resolution that typically does not corresponds to that in simulated events. Therefore, simulated events often have the four-momenta of the jets in each event randomly smeared a bit to lower the resolution and to make it more in line with the actual data. However, in the simulated samples used in this analysis the simulated resolution agrees with the data within uncertainty bands and so was left unsmeared. However, as a systematic the smearing was applied with a 1 sigma variation up and down.

JVF scale factor The jet vertex fraction varies between data and simulation events. Therefore, the events in simulated samples are reweighted so the total weight of events passing and failing a given JVF cut are the same in data and simulations. The reweighting is varied with its uncertainty up and down as a systematic.

JVF cut In our analysis we use a JVF cut of 0.5 on the jets. The effects of choosing this particular value[†] for the cut is assessed by varying the cut value up and down. The exact value with which this happens depends on the kinematics of the jet and uses the recommended values set by an ATLAS performance group.

In addition to these there are also a lot of systematic uncertainties that have to do with jets and their energy scale (sometimes abbreviated to JES - jet energy scale), mainly because jets are complicated composite objects. Another reason is that jets are measured in the calorimeters are the electromagnetic scale and need to be corrected to the hadronic scale and these corrections have sizable systematics attached to them. All these systematics affect the four-momenta of jets to correct for various uncertainties. These systematics are listed below. More information on jet calibrations within ATLAS can be found in [84].

The JES is calibrated and its uncertainties assessed by a large range of mutually supporting analyses based on real data. This setup necessarily spawns a large number of uncertainties, all of which should be propagated into this analysis as a systematic. Most of the uncertainties can fortunately be combined into a small number of nuisance parameters.

[†]A good pragmatic reason for choosing this cut is that the b-tagging is only really calibrated (i.e. scale factors are available) for this particular JVF cut.

Jet nuisance parameters This is a reduced list of parameters that were present in the in-situ analyses which were used to calculate the jet uncertainties. The total list should contain about fifty such nuisance parameters, but there have been reduced by the ATLAS JETETMISS performance group to six. We will not list the exact sources of these parameters here.

The so-called η intercalibration analysis has uncertainties that must be processed separately. This particular analysis works by balancing jets in a dijet event. Because back-to-back jets should have mirrored kinematics they can be used to fine-tune calibrations. The systematics we need to consider are related to the modeling of jets in event generators and the available statistics and will therefore be called "Jet η-cal modeling", and "Jet η-cal statistics".

Jet single particle high p_T This uncertainty has to do with the propagation of single hadron uncertainties in at high p_T.

Jet relative non-closure This uncertainty stems from using simulation samples that are not configured in the same way as the ones that were used to derive the jet uncertainties[†].

Jet NPV Related to corrections applied for dealing with pile-up jets. Such jets can add energy to jets coming from proper hard interactions and thus affect the energy scale. This one is needed to assess for uncertainties in the jet kinematics due to the number of primary vertices in the event.

Jet μ Also related to pile-up. This time needed for uncertainties in terms of the average number of proton-proton interactions per bunch crossing.

Jet close-by The kinematics of jets can be disturbed by other jets that lie close to the original jet. Uncertainties arising from this are estimated by this systematic.

5.7 Results

IN this chapter we have defined a sample of $jj\gamma\gamma$ events upon which to base the main $b\bar{b}\gamma\gamma$ analysis. We have described the two main backgrounds to the diphoton selection: electrons and jets wrongly reconstructed as photons. We have also discussed our methods for estimating these backgrounds and correcting our event counts. The electron background was estimated by examining the photon fake rates around the Z peak, an area where electrons dominate. The jet background was estimated by subdividing events into sixteen classes by using photon identification and isolation criteria and then using a fit. We have also combined these estimates into one whole.

These methods have been applied here in an inclusive way but can also be used in a bin-by-bin approach, which is what we shall do in the next chapter. From the 60162

[†]Typically large number of simulation samples inside ATLAS are created together sharing a lot of settings, making it easier to combine them. From time to time a new such project starts. The jet calibrations were performed on a group called MC12a, so using samples from a different group requires this systematic as a way to assess the effects of using a sample from a different group than the one used to derive the JES uncertainties.

events passing our cutflow to the pretag stage we estimate that the true number of $jj\gamma\gamma$ events within our fiducial cross-section is:

$$52443.8 \pm 345.1 \,\text{(stat)}$$

We quote only the statistical uncertainty here. The systematic uncertainty is revisted in the next chapter.

CHAPTER 6

Measurement of the $b\bar{b}\gamma\gamma$ cross-section with the ATLAS detector

"The boldness of asking deep questions may require unforeseen flexibility if we are to accept the answers."

- Brain Greene

WITHIN this chapter we will detail a cross-section measurement of the production of two b-jets and two photons from a single proton-proton collision at the LHC with the ATLAS detector using the 2012 8 TeV ATLAS dataset. In addition we will provide differential cross-sections of this channel in terms of various kinematic variables.

These variables are mentioned below and were chosen because they are of interest in a future search for double Higgs production or to measure λ_{3h}. In these instances the photon pair is originating from one Higgs boson and the b-jet pair from the other one, making the invariant mass combinations of these objects powerful discriminators to select for Higgs boson mediators. In addition we will examine the transverse momentum of the pair combinations, those of the leading object in each combination, and various angle separations. These are all popular variables investigated by di-Higgs analyses.

$m_{\gamma\gamma}$ The invariant mass of the two leading photons.

m_{jj} The invariant mass of the two leading jets.

$m_{jj\gamma\gamma}$ The invariant mass of the two leading photons and the two leading jets taken together.

$p_T^{\gamma\gamma}$ The transverse momenta of the two leading photons. It is equal to the p_T of the combined Lorentz vectors.

p_T^{γ} The transverse momentum of the leading photon.

p_T^{jj} The transverse momenta of the two leading jets. It is equal to the p_T of the combined Lorentz vectors.

p_T^{j} The transverse momentum of the leading jet.

$\Delta R(j,j)$ The angular separation between the two leading jets.

$\Delta R(\gamma,\gamma)$ The angular separation between the two leading photons.

$\Delta R(j,\gamma)$ The angular separation between the leading jet and the leading photon.

The starting point of the analysis is the $jj\gamma\gamma$ data sample we described in the previous chapter, and the used simulation samples are the same as those used there. Here we will build upon it by adding b-tags to the jets. Another novelty is formed by updating the background estimates to work in a bin-by-bin fashion so they can be used on the bins of our observables defined above.

Because we want these observables to describe both $jj\gamma\gamma$ and $b\bar{b}\gamma\gamma$ we write them in terms of jets and not specifically b-jets. Cuts made in the full selection will ensure the leading jets are b-tagged in the case of $b\bar{b}\gamma\gamma$.

We start in section 6.1 by extending the cutflow previously discussed in section 5.2. After this we shall discuss in section 6.2 how the electron and jet backgrounds to photons behave after the additional cuts, and how the various efficiencies defined there vary as functions of our observables. A description of how we deal with jets that are incorrectly tagged as b-jets is presented in section 6.3. In section 6.4 we discuss the unfolding; the method of deconvoluting detector effects from our observables to arrive at a fiducial cross-section useful for theorists. After that section we will summarize the results. The conclusion and outlook are given in chapter 7.

6.1 Event selection

In this first section we shall extend the cutflow to include b-tags, thus selecting $b\bar{b}\gamma\gamma$ events. In addition we will discuss the little snag first mentioned in section 5.2, namely that we need to restrict ourselves to events where the diphoton system 'points' to the most energetic primary vertex. This constraint will have a significant effect on how many events pass our cutflow. Finally we discuss several new systematic uncertainties dealing with flavor tagging and which we shall now have to include after b-tagging.

6.1.1 Extended cutflow and object definitions

We begin by defining a good b-jet:

Good b-jet

	data	SHERPA	PYTHIA	W	Z	$t\bar{t}$	single h
$jj\gamma\gamma$ sample	60162	193334	91122	514	4851	118	71
at least 1 b-jet	4837	21515	6444	35	568	80	7
at least 2 b-jets	277	863	263	0	5	19	1

Table 6.1.1: *Event numbers after the different cuts of the cutflow for data and various simulation samples. The number given for the $jj\gamma\gamma$ sample does not include the background estimates from the previous chapter; it is the bare amount of events passing all the cuts defined there.*

- ► Is a good jet
- ► $|\eta| < 2.5$
- ► MV1 > 0.7892 (see section 4.2.2 on page 71)

The second requirement the jet to the inner detector which is vital for b-tagging[†]. The MV1 cut corresponds to the 70% working point for calorimeter jets of the type used in the analysis. In addition to this we can now define two new cutflow points involving the two leading good jets:

Single tag selection

Require a b-tag on at least one of the two leading good jets.

Full selection

Require that both leading good jets are b-tagged.

We update table 5.2.1 on page 83 to incorporate these new cuts. The new table 6.1.1 shows that only very few events, 277, survive the full selection. As we saw in the previous chapter the different background processes are negligible. However, the small amount of data events passing the full selection will have a severe impact on the analysis as some parts of the background cannot be estimated from the full selection directly. This is explained in more detail when we discuss the backgrounds.

6.1.2 Effects of the diphoton pointing requirement on b-tagging

In order to increase the likelihood that the jets and the photons come from a single primary vertex a technique is used to determine the most likely primary vertex from which the photons are coming, and then a JVF cut (see page 80) is applied on the jets with respect to this vertex. This technique is more fully explained in [85]. In this technique the trajectories of the two leading photons are determined. If the photon is

[†]In fact any jet outside this range will automatically fail to be b-tagged. The cut is listed here mainly for clarity.

a converted one, one can simply trace a straight line from the calorimeter cluster to the conversion vertex to find the photon's origin. If the photon is unconverted instead the centers of the energy deposits in the first and second layer of the calorimeter are used. This information is then combined with tracking information and a likelihood fit determines the most likely primary vertex for the two photons to originate from.

In our analysis we run into a specific limitation of this method. The b-tagging algorithms are only calibrated on the most energetic primary vertex. Consequently this is the only primary vertex we can use, and so we must discard all events where the diphoton system does not point to the most energetic vertex. A quick study using the shows only in around 68% of the cases a data event satisfies this condition. We thus lose about a third of our events on this requirement.

In the following we will investigate if the selection of the most energetic primary vertex can be loosened. The diphoton vertex pointing method has a finite spatial resolution. If the diphoton vertexing method selects a primary vertex that is not the most energetic one but its distance to the most energetic primary vertex is smaller than the spatial resolution of the vertexing method, it makes sense to accept the event anyway. To study this we need to examine the spatial resolution.

The SHERPA signal sample was used to study how often the diphoton vertexing method obtains the correct result (i.e. the actual diphoton vertex obtained from the truth information). In figure 6.1.1 we show the distribution of the distance between the correct vertex and the one selected by the vertexing method. As can be seen there is a very high peak at a distance of zero, meaning the diphoton vertex method does obtain the correct primary vertex most of the time, but there is also a considerable tail. By integration one can calculate what the 2σ distance is, i.e. within which distance the method obtains the correct primary vertex more than 95% of the time. For events within the fiducial volume (green in the figure) this distance was found to be approximately 30 mm.

To assess the effects of this resolution on the final analysis we added a systematic to the main analysis where we relaxed the good primary vertex cut to one where the vertex selected by the diphoton vertexing method only has to be within 30 mm of the most energetic primary vertex. This systematic has been named 'lenient good PV cut' and was symmetrized around the nominal result. The effects were small.

6.1.3 Flavor tagging systematic uncertainties

Since we now apply b-tagging we need to add a few corrections to the simulation samples, and these come with their own associated systematic uncertainties. These are:

Jet flavor composition The energy measured for jets compared to their truth energy differs depending on the flavor of the jet. For example gluon and light jets differ in their fragmentation patterns. Thus, deriving the JES corrections is done for a sample with certain flavor fractions. If one then uses those corrections in a sample with a different flavor content there is going to be some error. This error is captured the flavor composition, which deals with a difference in the calibration sample and the actual sample used in the analysis in terms of flavor content. This systematic is not applied to b-jets.

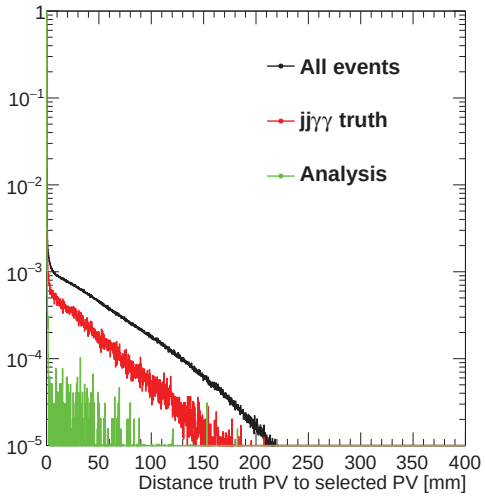

Figure 6.1.1: *The distance between the primary vertex indicated by the method versus the actual primary vertex that the diphotons originate from. Shown are all events in the sample, all events passing the $jj\gamma\gamma$ truth selection used in the unfolding, and all events passing the full selection (Analysis).*

Jet flavor response The other systematic related to the flavor content of the sample. This one takes into account that different event generators handle gluon jets slightly differently. Just as for the flavor composition this systematic is not applied to b-jets.

Jet b-JES This systematic takes the place of the flavor composition and flavor response systematics defined for non-b-jets, in the case of b-jets.

Flavor tagging Since the tagging of jets is a very complicated process depending on a large number of variables it is very unlikely that the simulation samples describe this in a way that perfectly harmonizes with the data. Specifically the efficiencies to tag the various different jet flavors differs between data and simulation. To correct for this events are reweighted by so-called scale factors (see appendix A for more information). The uncertainties on these scale factors needs to be propagated into the analysis as systematic uncertainties. The ATLAS flavor tagging performance group provides a way to encode all these different systematic uncertainties into a smaller group of uncertainties using an eigenvector approach. Instead of mentioning each of these in turn we combine them into one big uncertainty by varying all of them and quoting the largest effect on the analysis.

6.2 Backgrounds to photons

We will use the methods defined in previous chapters to gauge the size of the jet and electron backgrounds which can fake photons. The fake rate $f_{e\to\gamma}$ will still be taken to be constant, but all other quantities are rederived in a bin-by-bin fashion. This

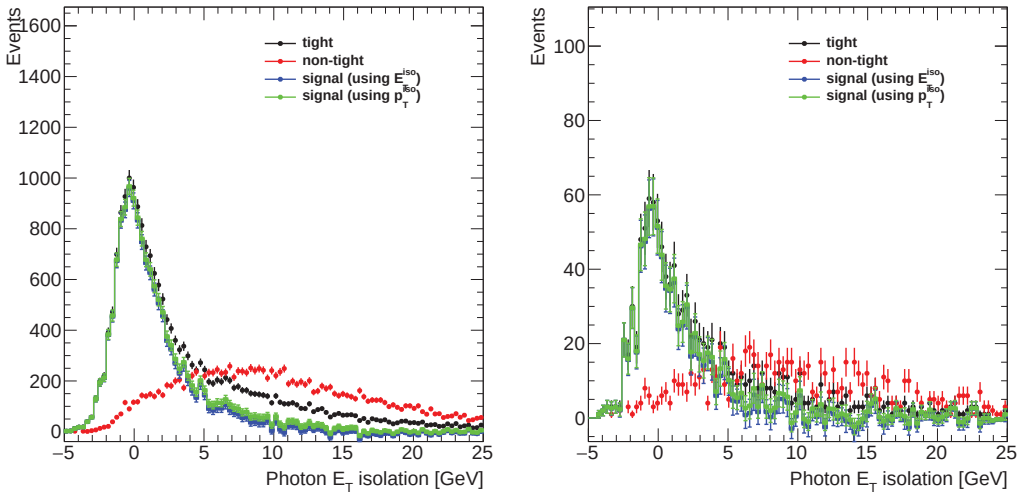

Figure 6.2.1: *Leading photon isolation distribution for photons passing the* TIGHT *cut or not passing it, along with the extracted signal distributions. On the left the single tagged selection, on the right the full selection.*

transition is relatively easy; we take only the events found inside a given observable bin into consideration when calculating these quantities.

In this chapter we will quickly revisit all background subtraction methods introduced in chapter 5 and examine how they fare with the extended cutflow, and how they vary as a function of our observables. In the first section we will discuss the data-driven extraction of the photon isolation efficiency. After this we will discuss the electron background and the jet backgrounds. Finally in section 6.2.4 we will discuss their combination.

6.2.1 Photon isolation

The inclusive photon isolation efficiencies extracted at the pretag stage do not vary much when we move to more advanced stages of our cutflow. We observe in figure 6.2.1 that the data-driven isolation distributions do not change from the ones for the distributions we had for earlier stages in the cutflow. The only exception is that the statistics became worse, especially towards the full selection. We find an inclusive value of 0.81 for both ϵ_{I1} and ϵ_{I2}. The systematic uncertainties can be seen in table 6.2.1.

The low number of events passing the full selection means that it is impossible to obtain the photon isolation efficiency in a bin-by-bin fashion. Because the inclusive photon isolation efficiencies are pretty stable across the cutflow it is reasonable to use the isolation efficiencies found at the single tagged stage when we do the bin-by-bin extraction. Since we are discussing a property of photons the existence of jets that happen to be b-tagged should only have a marginal effect. We also find that the distributions after the full selection are fully compatible within uncertainties with the distribution after requiring a single b-tag, and there are no wild fluctuations. In addition the distributions of ϵ_{I2} were very similar to those of ϵ_{I1}.

ϵ_{I1}	single tag	full selection
nominal value	0.81 ± 0.006	0.81 ± 0.025
background boundary up	-0.015	-0.014
background boundary down	$+0.024$	$+0.034$
using E_T^{iso} for background	± 0.067	± 0.073
total systematic uncertainty	± 0.071	± 0.08
ϵ_{I2}		
nominal value	0.79 ± 0.0071	0.81 ± 0.031
background boundary up	-0.017	-0.013
background boundary down	$+0.023$	$+0.018$
using E_T^{iso} for background	± 0.10	± 0.12
total systematic uncertainty	± 0.11	± 0.12

Table 6.2.1: *Extracted isolation efficiencies from data events for the leading and sub-leading photons along with the statistical and systematic uncertainties.*

One thing to note is an increase in the isolation efficiency around the Z peak in $m_{\gamma\gamma}$ which is visible in the jet preselection, but which disappears into the statistical uncertainty after requiring at least one b-tag. This can be seen in figure 6.2.2. This increase is caused by an increase of electrons faking photons from $Z \to e^+e^-$ which makes them more relevant than jets faking photons in that bin. Since electrons tend to be more often isolated than jets this increases the extracted photon isolation efficiency. For the other observables we show only the distribution at the pretag selection in figures 6.2.3 and 6.2.4 to show the overall clear shape, but we stress that it is the distributions after a single b-tag that are going to be used in the analysis and these are consistent with those found in the pretag selection.

In section 5.3.2 we also discussed using p_T^{iso} versus E_T^{iso}. We found that adding the b-tags did not alter the situation we saw there.

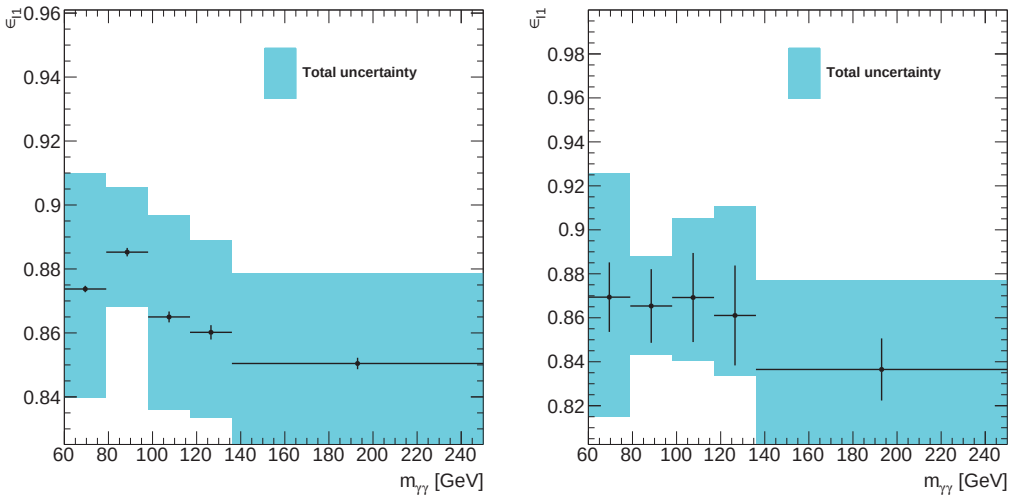

Figure 6.2.2: *Leading photon isolation efficiency as a function of $m_{\gamma\gamma}$ after the jet preselection (left) and after requiring a single b-tag (right).*

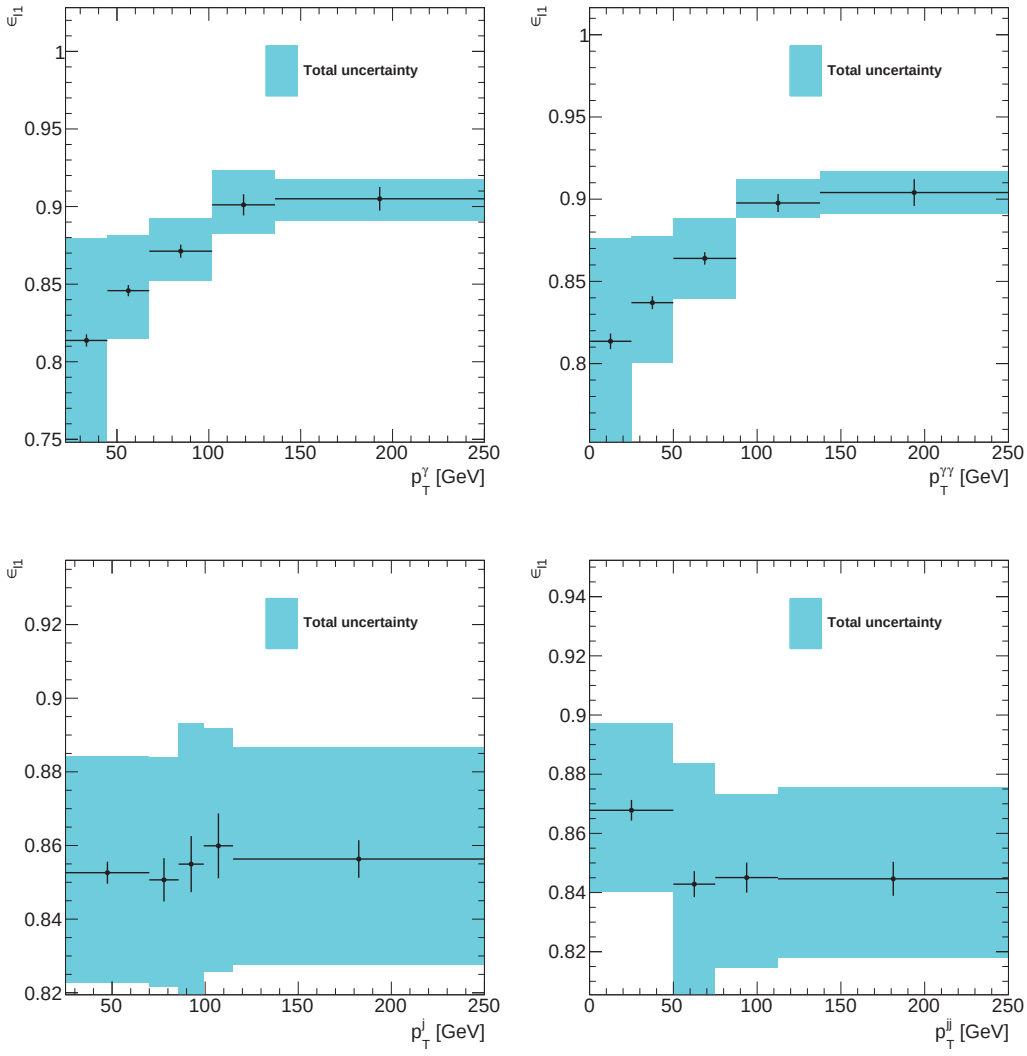

Figure 6.2.3: *Leading photon isolation efficiency as a function of p_T^{γ} (top left), $p_T^{\gamma\gamma}$ (top right), p_T^{j} (bottom left), and p_T^{jj} (bottom right) after the pretag selection.*

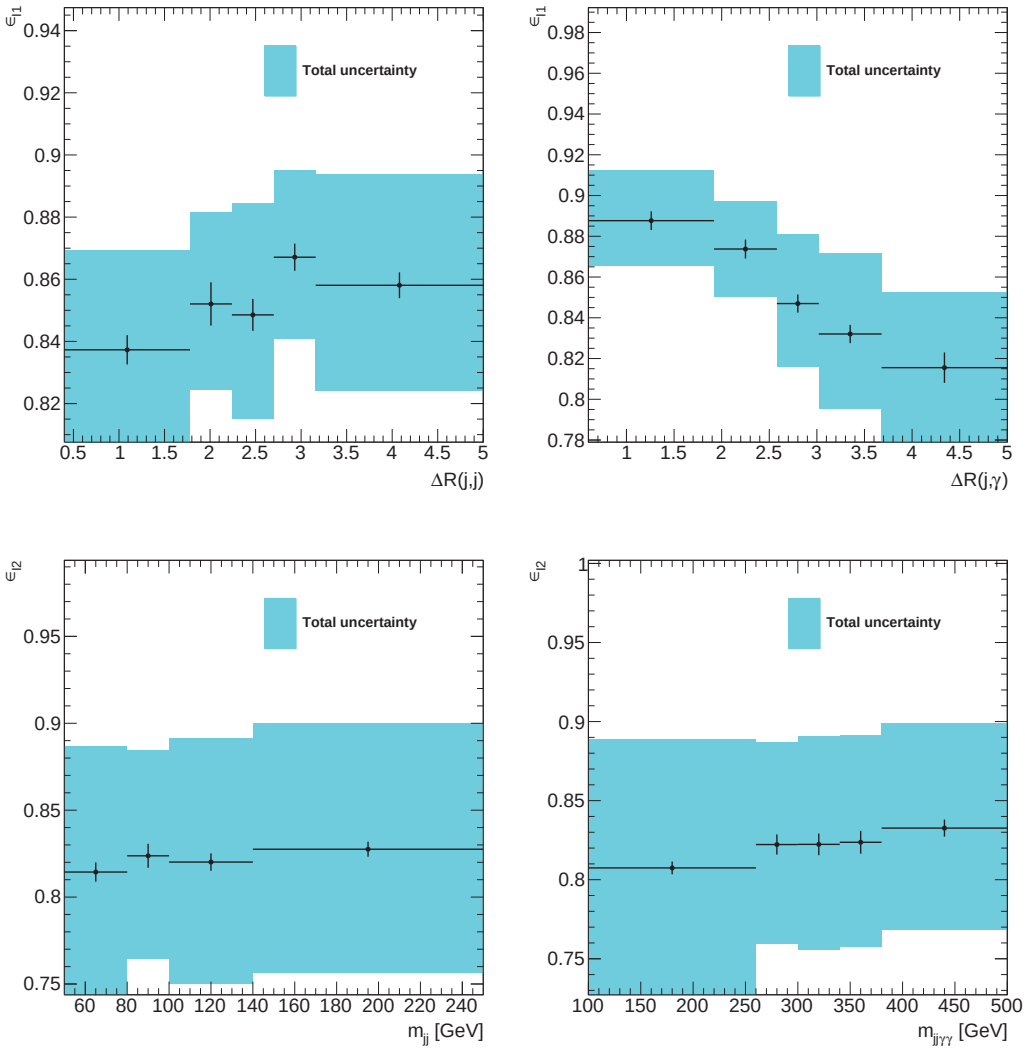

Figure 6.2.4: *Leading photon isolation efficiency as a function of $\Delta R(j,j)$ (top left), $\Delta R(j,\gamma)$ (top right), m_{jj} (bottom left), and $m_{jj\gamma\gamma}$ (bottom right) after the pretag selection.*

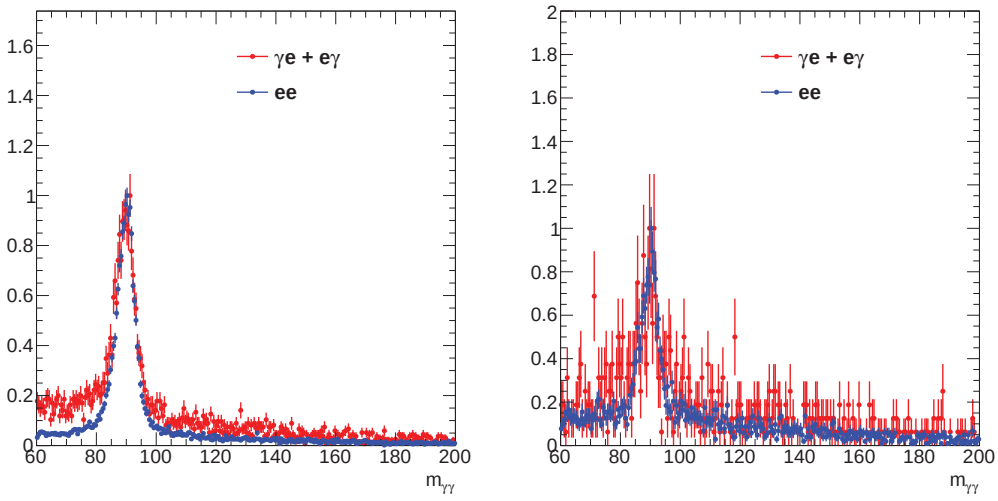

Figure 6.2.5: *Diphoton invariant mass spectra for ee and $\gamma e + e\gamma$ events. For visibility the distributions have been scaled so that the peaks have a height of one. On the left the distributions with the single tag selection, on the right with the full selection.*

6.2.2 Electron background

For the electron background we first need to extract $f_{e\to\gamma}$. To do so we employ the same method as discussed in section 5.4.2 except we now add the additional cuts. The diphoton invariant mass spectra we obtained are found in figure 6.2.5. As can be seen in that picture we suffer from degrading statistics when applying stronger and stronger cuts. In fact for the full selection the number of events is so low no reliable fake rate could be extracted at all. We will therefore extract the fake rates without the full selection and use instead the single b-tagged sample analogous to the way the photon isolation efficiencies were extracted also at this stage. We show the decomposition into the Z peak and background for the single tag selection in figure 6.2.6. The $e \to \gamma$ fake rate obtained from comparing the number of ee and $\gamma e + e\gamma$ events under the Z peak can be was found to 0.069 ± 0.0012(stat). For completeness we should note that we still consider $f_{\gamma\to e}$ to be negligible.

The next thing to investigate is the electron fraction, Γ_e. Now that we have the fake rates we can use expression 5.4.28 on page 96 and apply it bin-by-bin. The Γ_e we thus find can be seen in table 6.2.2. One can clearly see in that table that there is an effect from requiring b-tags in that the electron fraction increases with each additional tag requirement. This is because b-tags also select for t-jets which may decay semileptonically. For the full selection we find that about 5.3% of the events in the full selection are tainted by electron fakes.

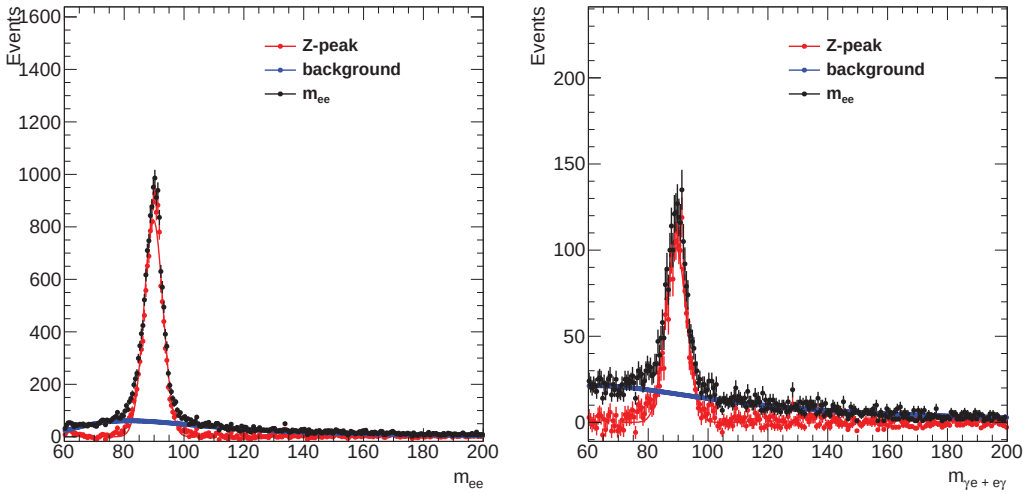

Figure 6.2.6: *Invariant mass spectrum for ee (left) and $\gamma e + e\gamma$ (right) events with the decomposition into the Z peak and background in the single tag selection. The fit made to the Z peak is also shown as a solid red line.*

Γ_e	Single tag	Full selection
Nominal value	0.028 ± 0.003	0.053 ± 0.0021
$f_{e\to\gamma}$ Z mass interval	± 0.00039	± 0.00068
$f_{e\to\gamma}$ sideband interval	± 0.0055	± 0.0098
$f_{e\to\gamma}$ fake rate statistics	± 0.00044	± 0.00078
Using unconverted electrons	~0	~0
Total systematic error	± 0.0055	± 0.0098

Table 6.2.2: *Extracted electron fraction for all events passing the cuts up including the requirement of one or two b-tags. The total systematic error listed in this table was calculated by adding the larger of the up/down values for each type in quadrature.*

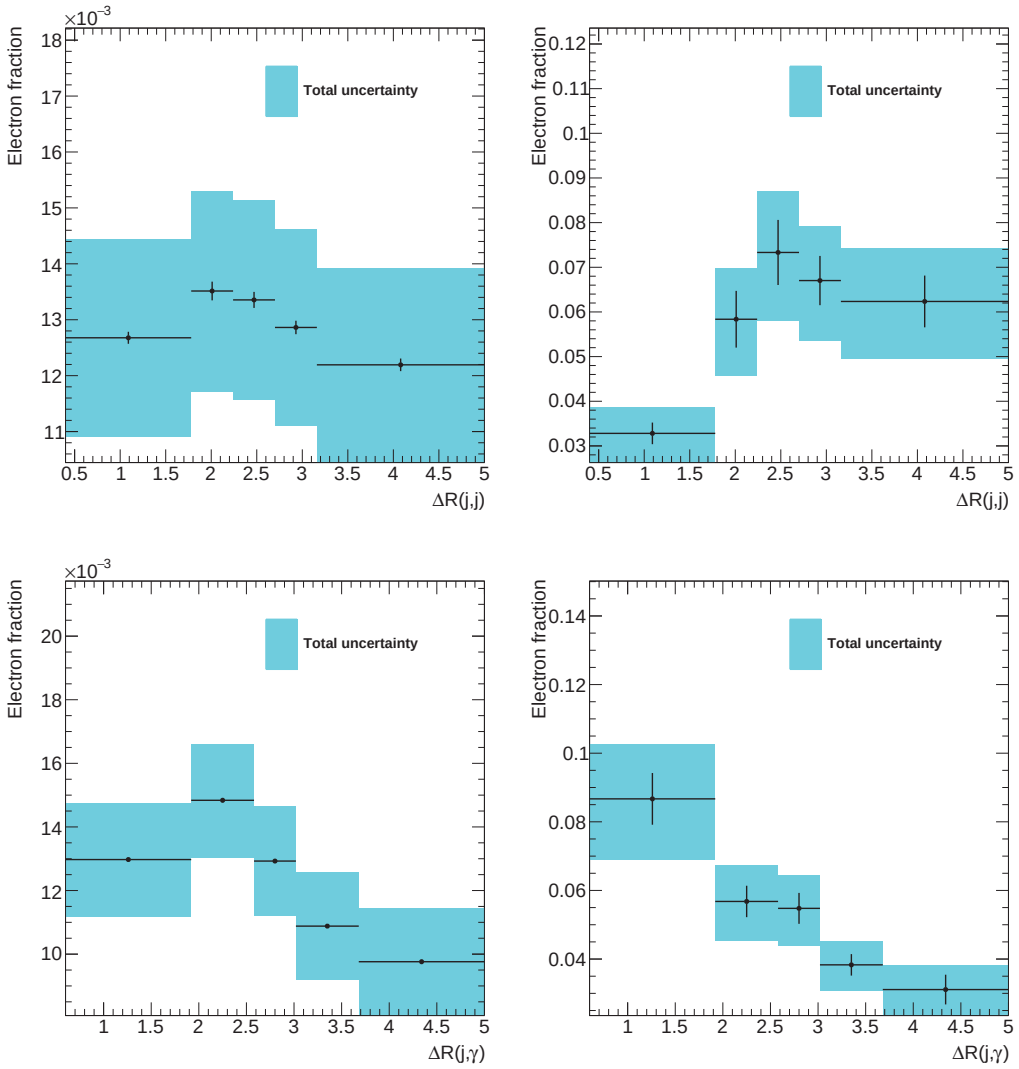

Figure 6.2.7: *Electron fraction after the jet preselection (left) and after the full selection (right) including systematic errors versus $\Delta R(j,j)$ (top) and $\Delta R(j,\gamma)$ (bottom).*

We also find that the distributions of Γ_e as a function of our observables roughly follows the jet preselection after every additional cut. However, some differences in shape do occur, most significantly in the angular separations $\Delta R(j,j)$ and $\Delta R(j,\gamma)$, and in the invariant masses. The catalyst of these changes seems to be again the b-tagging requirements as the shapes deform from the pretag stage to the full selection with each additional b-tag. Thus, the single b-tagged distributions are in between the pretag and final selection ones. The differences can be seen in figures 6.2.7 to 6.2.9 where we show a comparison of the Γ_e distribution in the jet preselection and in the full selection as a function of the angular separations and invariant masses. For the transverse momenta we show only the full selection in figure 6.2.10.

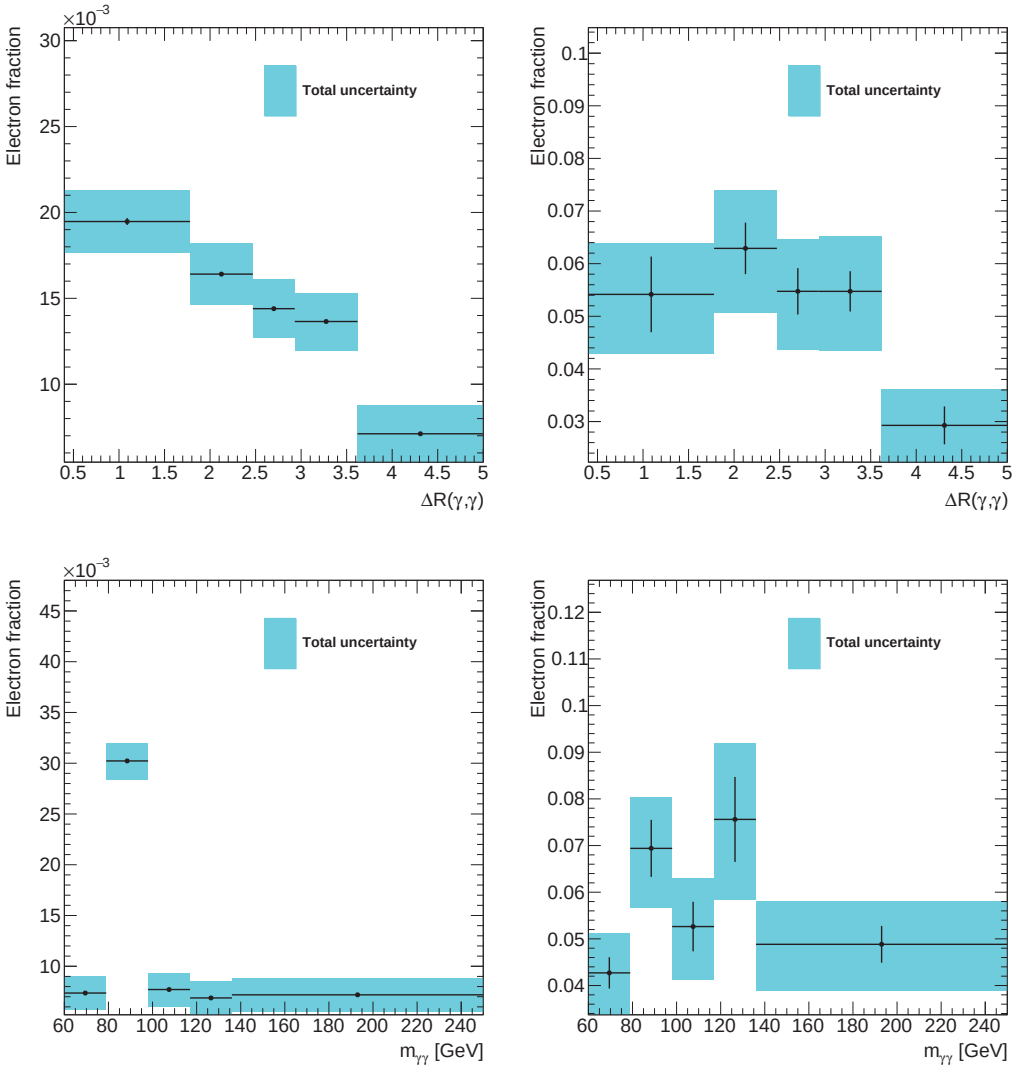

Figure 6.2.8: *Electron fraction after the jet preselection (left) and after the full selection (right) including systematic errors versus $\Delta R(\gamma,\gamma)$ (top) and $m_{\gamma\gamma}$ (bottom).*

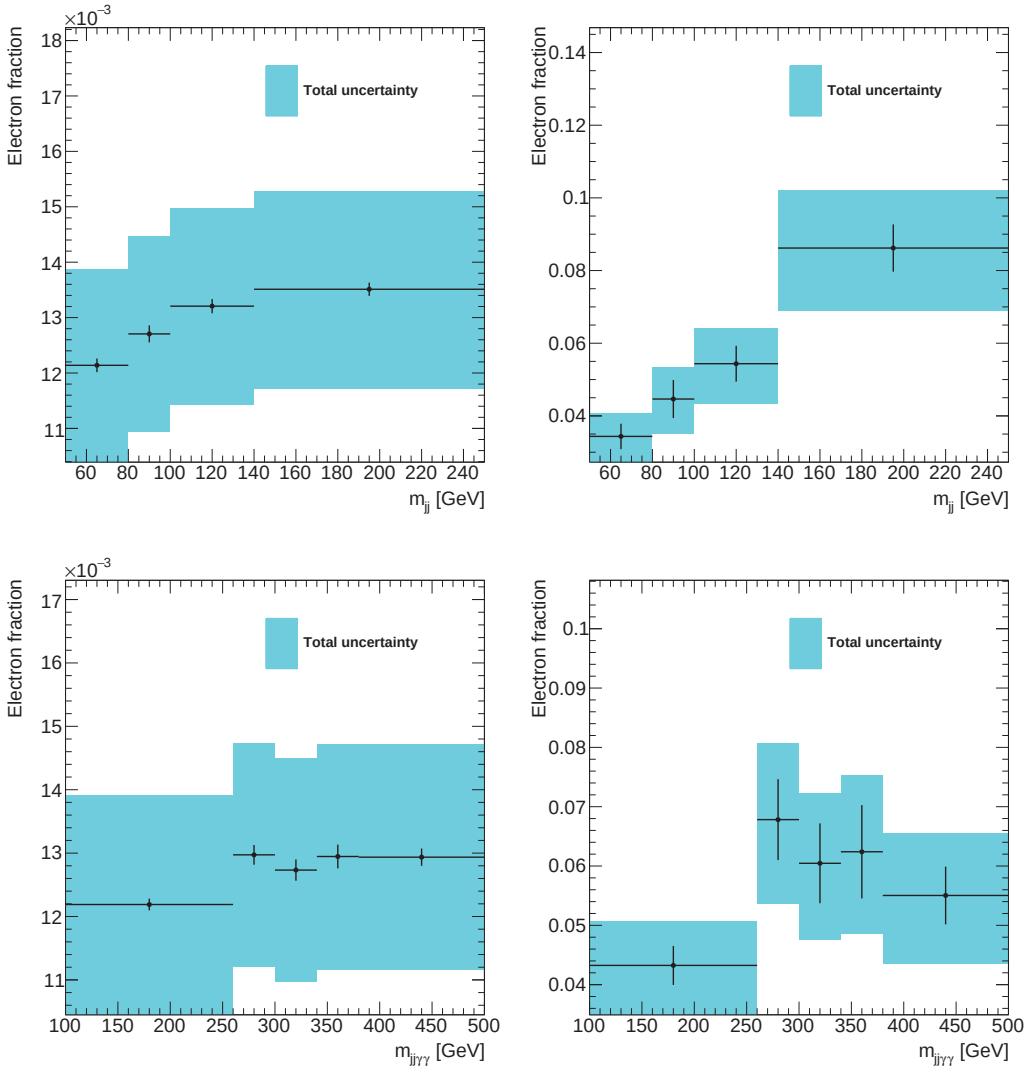

Figure 6.2.9: *Electron fraction after the jet preselection (left) and after the full selection (right) including systematic errors versus m_{jj} (top) and $m_{jj\gamma\gamma}$ (bottom).*

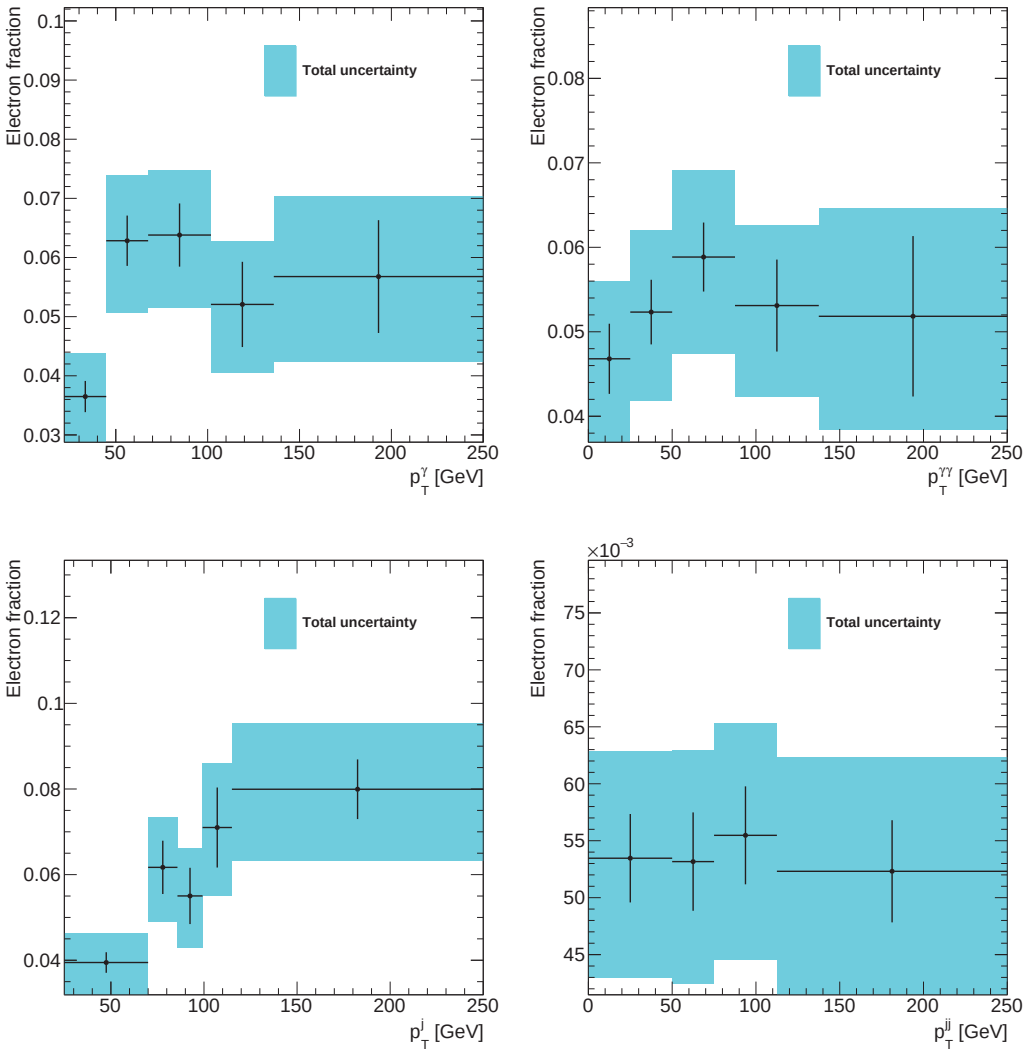

Figure 6.2.10: *Electron fraction after the full selection including systematic errors versus p_T^{γ} (top left), $p_T^{\gamma\gamma}$ (top right), p_T^{j} (bottom left), and p_T^{jj} (bottom right).*

ϵ_{T1}	single tag	full selection
nominal value	0.94 ± 0.0017	0.95 ± 0.0083
using PYTHIA	± 0.019	± 0.045
ϵ_{T2}		
nominal value	0.89 ± 0.0022	0.91 ± 0.011
using PYTHIA	± 0.0057	± 0.021

Table 6.2.3: *Extracted identification efficiencies from simulation for the leading and subleading photons along with the statistical and systematic uncertainties.*

6.2.3 Jet background

By defining the sixteen classes from section 5.5 on page 97 for every observable bin separately and performing the fit in each bin anew we can easily apply the jet background estimation bin-by-bin. Recall that we extracted one of the fit variables, the photon identification efficiencies (ϵ_T), from simulations. We repeat now the extraction of ϵ_T with the additional b-tags applied. In the inclusive case we find $0.95 \pm 0.0083\,(\text{stat}) \pm 0.045\,(\text{syst})$ for ϵ_{T1} and $0.91 \pm 0.011\,(\text{stat}) \pm 0.021\,(\text{syst})$ for ϵ_{T2}. The systematics are found in table 6.2.3. As with the electron background discussed before we have low statistics in the full selection, this makes it impossible to extract ϵ_T for every bin separately after all cuts are applied. To solve this we extract ϵ_T from the single tagged selection and we present the distributions in figures 6.2.11 to 6.2.13. The shapes do not vary significantly between cuts, and there is also not much shape difference between ϵ_{T1} and ϵ_{T2} with respect to our observables and therefore we only show ϵ_{T1}.

The values we find for the $\gamma\gamma$ purity, $\mathcal{P}_{\gamma\gamma}$, are listed in table 6.2.4, and its dependency on our observables in figures 6.2.14 to 6.2.16. The inclusive result in the full selection implies that only 10% of the events contain jets faking one of the two leading photons. The largest uncertainty is still coming from the photon isolation efficiency method.

We find that there are slight differences between the jet preselection and the final selection. The effect is similar to that seen in the electron background; the shape is the same for the jet preselection and in the pretag selection, but then starts to deform with each additional required b-tag on a jet. Overall most distributions are consistent with being flat within the uncertainties, but the $\gamma\gamma$ purity is increased when the orientation of the two leading jets or the two leading photons is back-to-back.

The contributions of the various background types can be seen in table 6.2.5, and the relative contributions of each type are constant over our observables, even though we extract these on a bin-by-bin basis. The values of the other variables present in the fit are found in table 6.2.6 for all events inclusively. The fit variables do vary as a function of our observables with respect to different stages in the cutflow, but only in overall size, the shapes are more-or-less constant and the distributions after the full selection are (within uncertainties), for most observables, compatible with those found in the jet preselection. As an example we show the distributions with respect to the leading photon transverse momentum in figure 6.2.17.

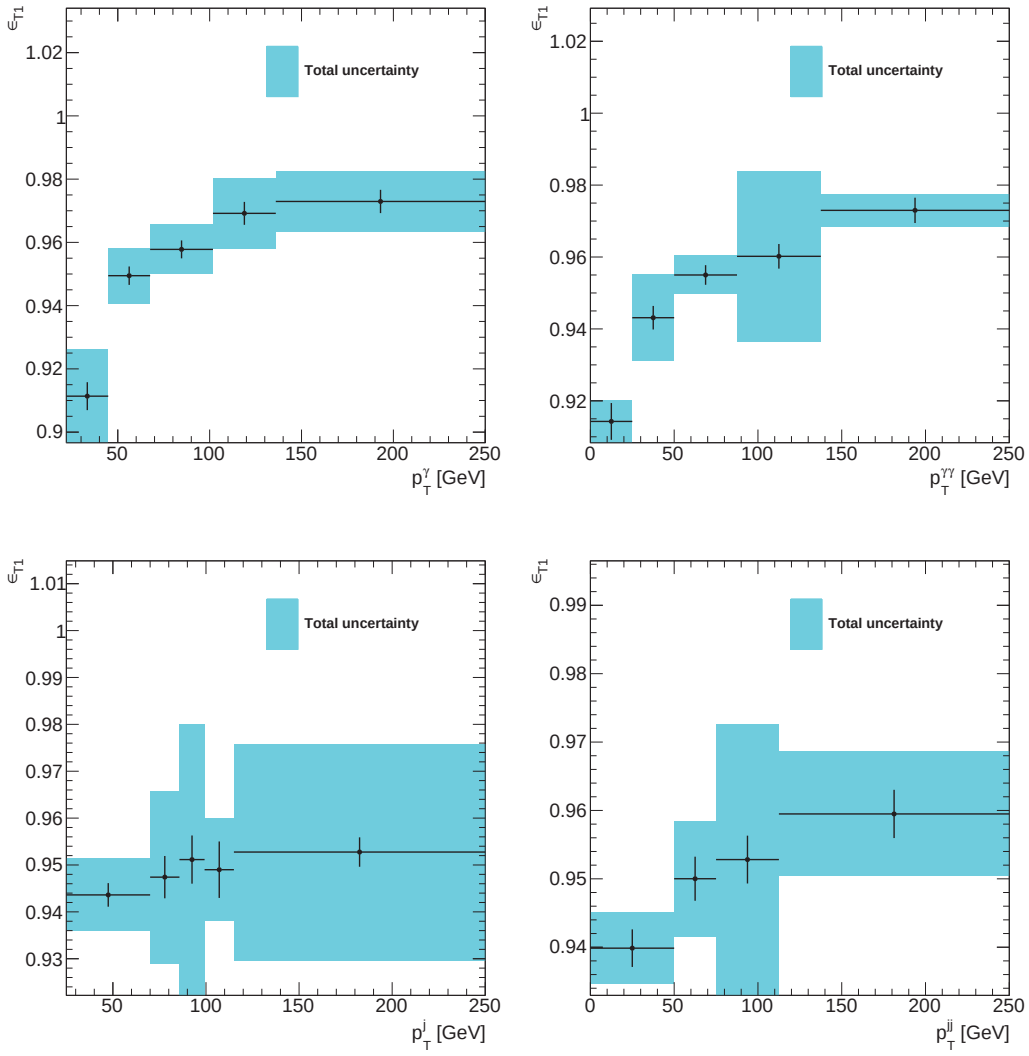

Figure 6.2.11: *Leading photon identification efficiency as a function of transverse momentum with systematics at the single b-tag stage. Shown are p_T^{γ} (top left), $p_T^{\gamma\gamma}$ (top right), p_T^{j} (bottom left), and p_T^{jj} (bottom right).*

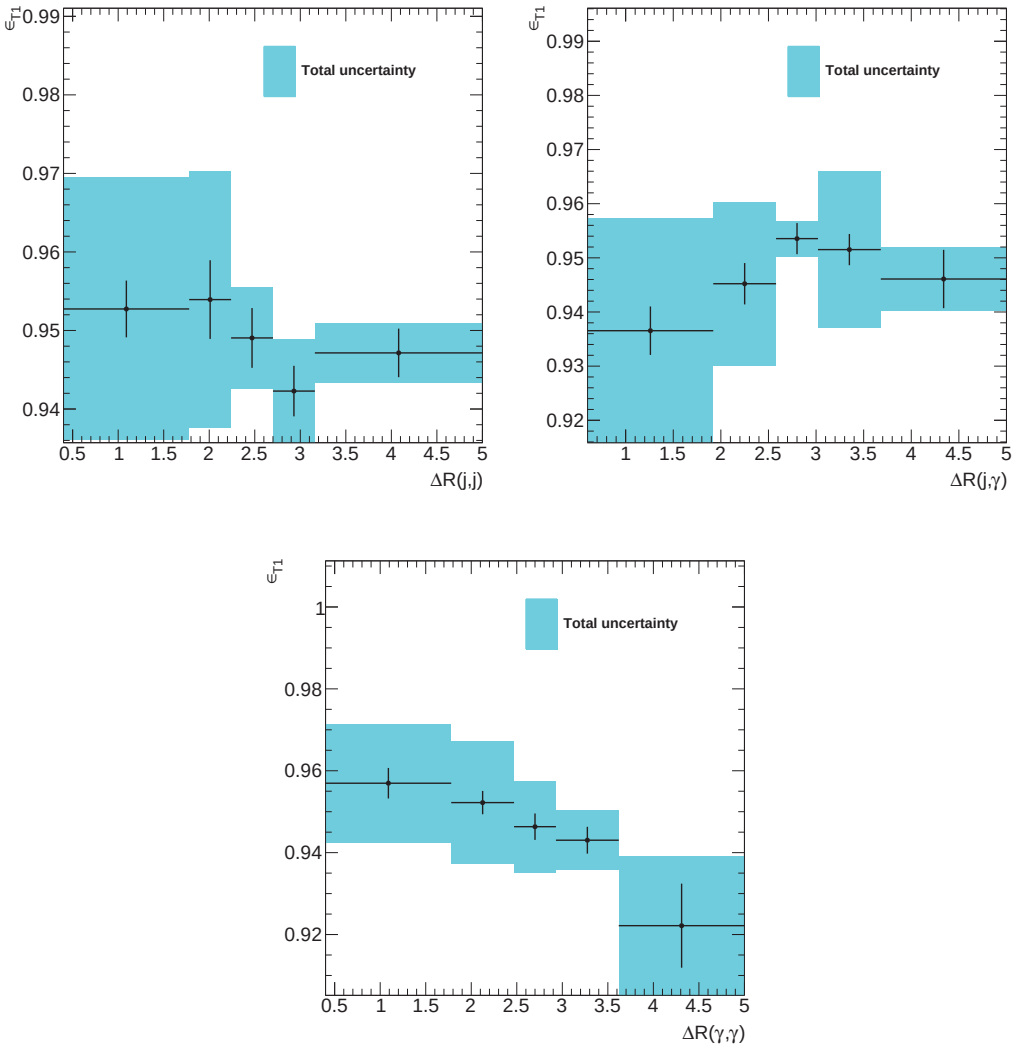

Figure 6.2.12: *Leading photon identification efficiency as a function of angular separation with systematics at the single b-tag stage. Shown are $\Delta R(j,j)$ (top left), $\Delta R(j,\gamma)$ (top right), and $\Delta R(\gamma,\gamma)$ (bottom).*

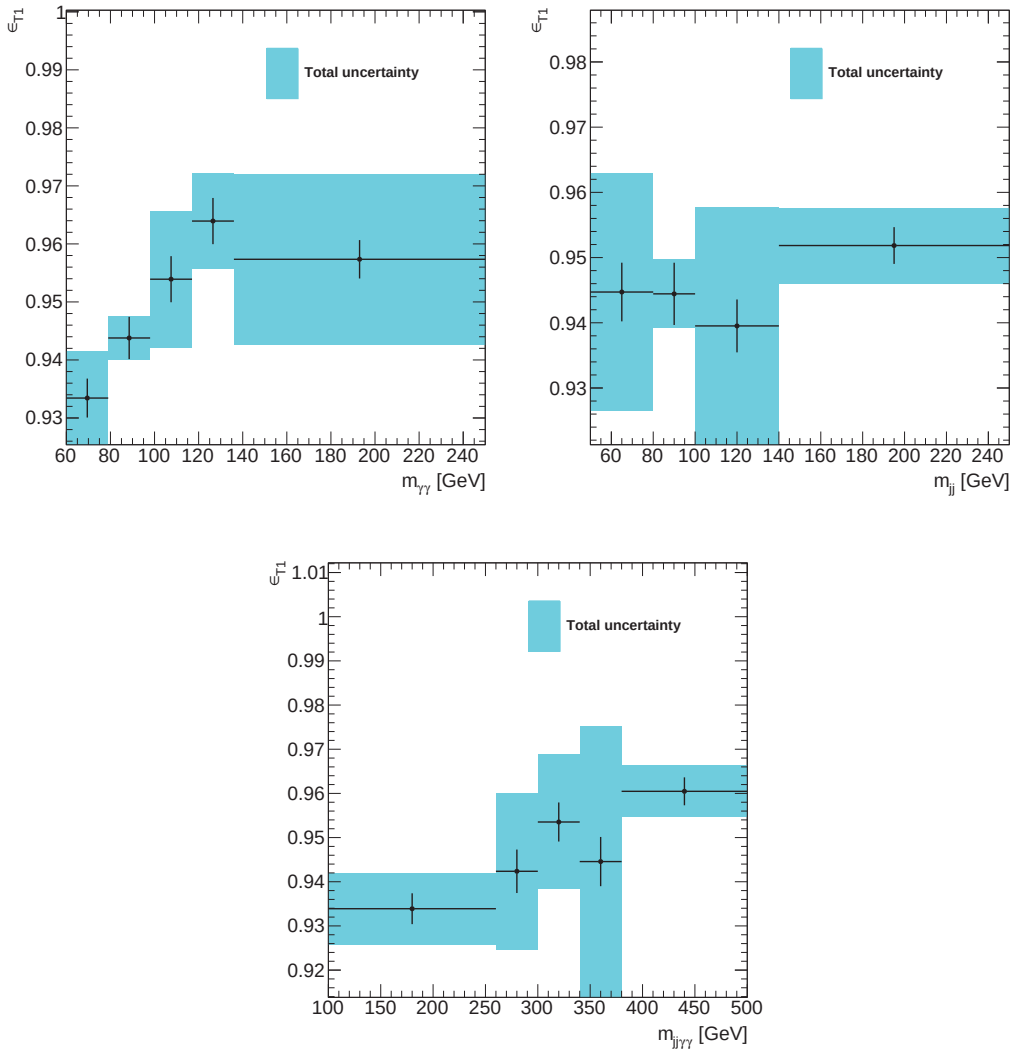

Figure 6.2.13: *Leading photon identification efficiency as a function of invariant mass with systematics at the single b-tag stage. Shown are $m_{\gamma\gamma}$ (top left), m_{jj} (top right), and $m_{jj\gamma\gamma}$ (bottom).*

$\gamma\gamma$ purity	single tag	full selection
Nominal value	0.90 ± 0.0045	0.90 ± 0.017
(jet) Using E_T^{iso} for ϵ_I [%]	± 5.8	± 5.1
(jet) Using PYTHIA for ϵ_T [%]	± 0.48	± 0.50
(jet) ϵ_{I1}background boundary up [%]	0.36	0.40
(jet) ϵ_{I1}background boundary down [%]	-0.55	-0.61
(jet) ϵ_{I2}background boundary up [%]	0.76	0.63
(jet) ϵ_{I2}background boundary down [%]	-1.0	-0.82
(jet) ϵ_{I2} statistical error up [%]	-0.31	-0.26
(jet) ϵ_{I2} statistical error down [%]	0.31	0.26
(jet) ϕ_1 up [%]	-1.2	-1.4
(jet) ϕ_1 down [%]	0.93	1.1
(jet) ϕ_2 up [%]	-2.4	-2.0
(jet) ϕ_2 down [%]	1.8	1.5
(general) Flavor tagging scale factors [%]	± 0.20	± 0.21
(general) Pile-up reweighting up [%]	-0.24	-0.25
(general) JVF cut up [%]	-0.43	< 0.2
(general) JVF cut down [%]	< 0.20	0.56
Total systematic error [%]	± 6.2	± 5.6

Table 6.2.4: *Signal purity obtained from the fit with uncertainties. The total systematic uncertainty was calculated by taking maximum of the up/down values of each systematic and adding them in quadrature. The word in parentheses is to remind the reader of the general source of the uncertainty. Only systematics with an effect of 0.2% or more are listed.*

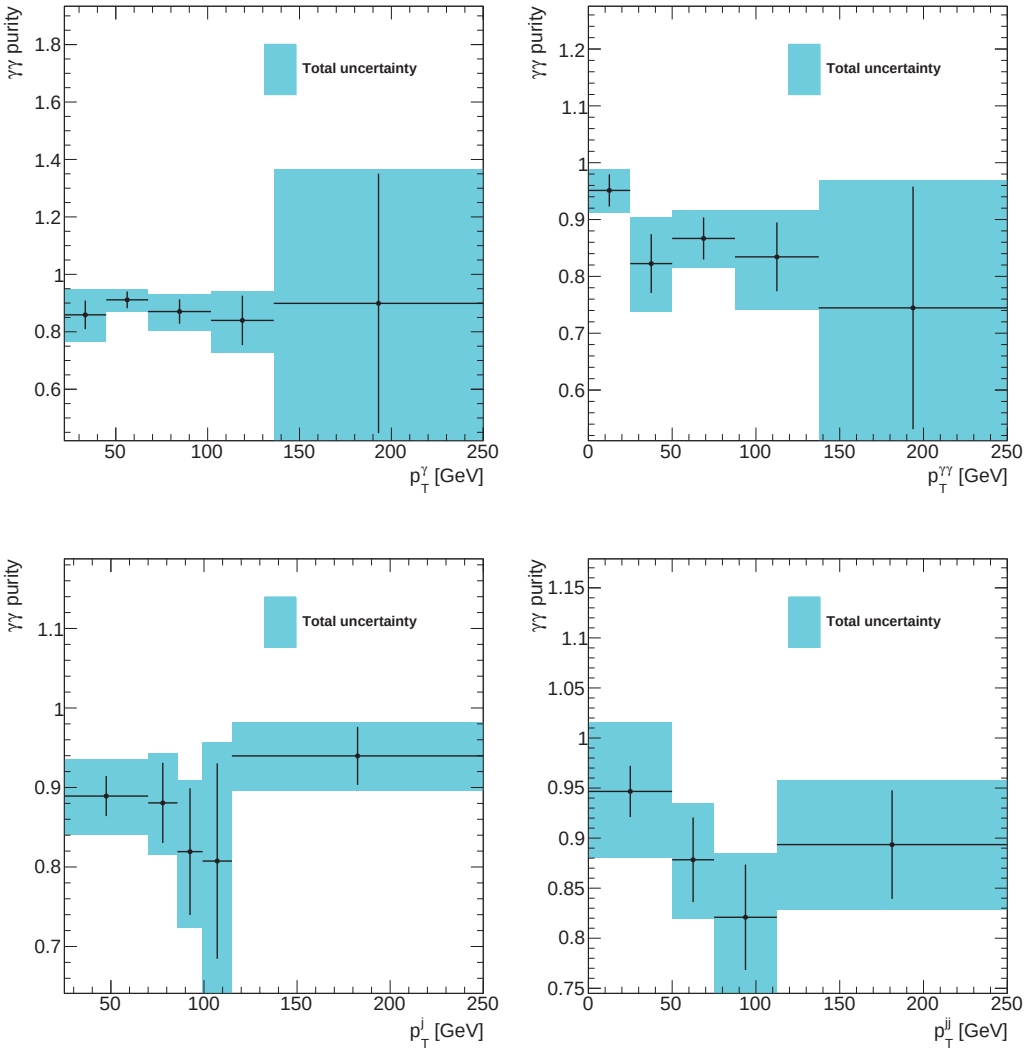

Figure 6.2.14: *Signal purity with systematics for the full selection as a function of the transverse momenta p_T^{γ} (top left), $p_T^{\gamma\gamma}$ (top right), p_T^{j} (bottom left), and p_T^{jj} (bottom right).*

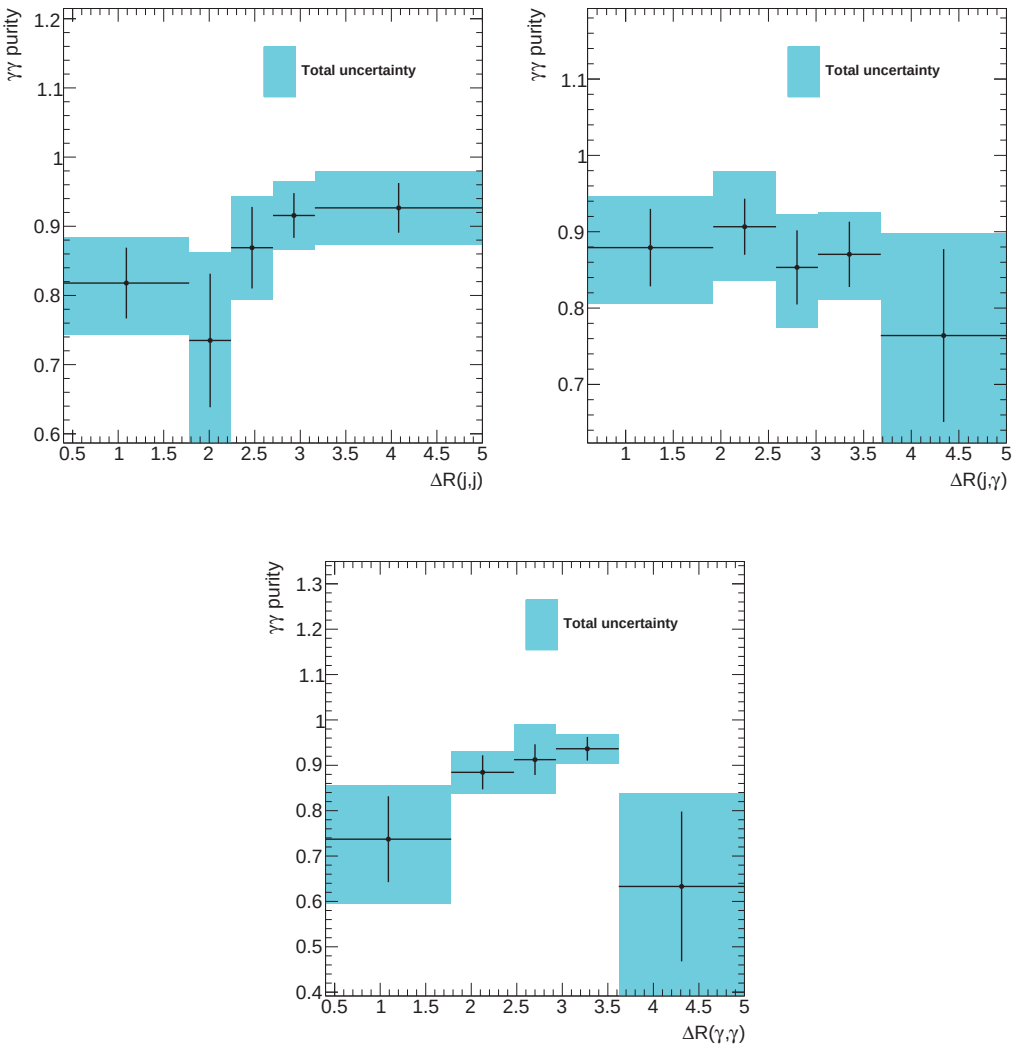

Figure 6.2.15: *Signal purity with systematics for the full selection as a function of the angular separations $\Delta R(j,j)$ (top left), $\Delta R(j,\gamma)$ (top right), and $\Delta R(\gamma,\gamma)$ (bottom).*

	single tag	full selection
$N_{\gamma\gamma}$	4332.3 ± 99.3	249.1 ± 20.2
$N_{j\gamma}$	169.6 ± 11.8	11.3 ± 3.0
$N_{\gamma j}$	309.1 ± 17.6	15.3 ± 3.7
N_{jj}	21.8 ± 1.3	0.9 ± 0.3
total	4832.8	276.7

Table 6.2.5: *Composition of the sample as determined by the fit. The quoted error is the fit uncertainty only.*

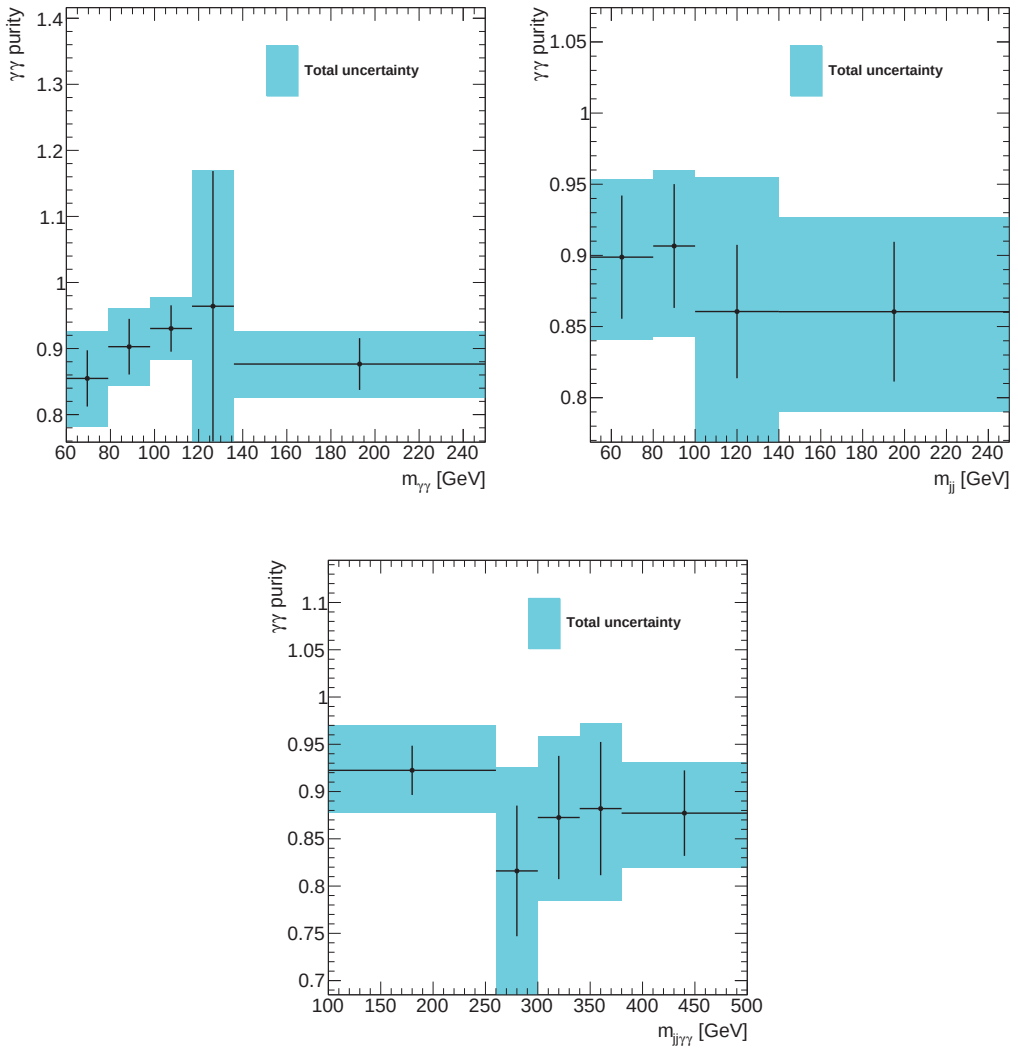

Figure 6.2.16: *Signal purity with systematics for the full selection as a function of the invariant masses $m_{\gamma\gamma}$ (top left), m_{jj} (top right), and $m_{jj\gamma\gamma}$ (bottom).*

	single tag	full selection
f_{I1}	$0.133 \pm 0.00715\,(\text{stat}) \pm 0.0256\,(\text{syst})$	$0.143 \pm 0.0292\,(\text{stat}) \pm 0.0242\,(\text{syst})$
f_{I2}	$0.207 \pm 0.00752\,(\text{stat}) \pm 0.0128\,(\text{syst})$	$0.180 \pm 0.03033\,(\text{stat}) \pm 0.0119\,(\text{syst})$
f_{T1}	$0.253 \pm 0.00945\,(\text{stat}) \pm 0.0673\,(\text{syst})$	$0.265 \pm 0.0381\,(\text{stat}) \pm 0.0636\,(\text{syst})$
f_{T2}	$0.242 \pm 0.00926\,(\text{stat}) \pm 0.0892\,(\text{syst})$	$0.236 \pm 0.0362\,(\text{stat}) \pm 0.0884\,(\text{syst})$
g_{I1}	$0.126 \pm 0.00478\,(\text{stat}) \pm 0.0143\,(\text{syst})$	$0.116 \pm 0.0205\,(\text{stat}) \pm 0.0172\,(\text{syst})$
g_{I2}	$0.185 \pm 0.00547\,(\text{stat}) \pm 0.00942\,(\text{syst})$	$0.208 \pm 0.0239(\text{stat}) \pm 0.00987\,(\text{syst})$
g_{T1}	$0.279 \pm 0.00606\,(\text{stat}) \pm 0.0393(\text{syst})$	$0.263 \pm 0.0267\,(\text{stat}) \pm 0.0467\,(\text{syst})$
g_{T2}	$0.265 \pm 0.00631\,(\text{stat}) \pm 0.0526\,(\text{syst})$	$0.229 \pm 0.0287\,(\text{stat}) \pm 0.064\,(\text{syst})$

Table 6.2.6: *Fit values along with statistical and systematic uncertainties.*

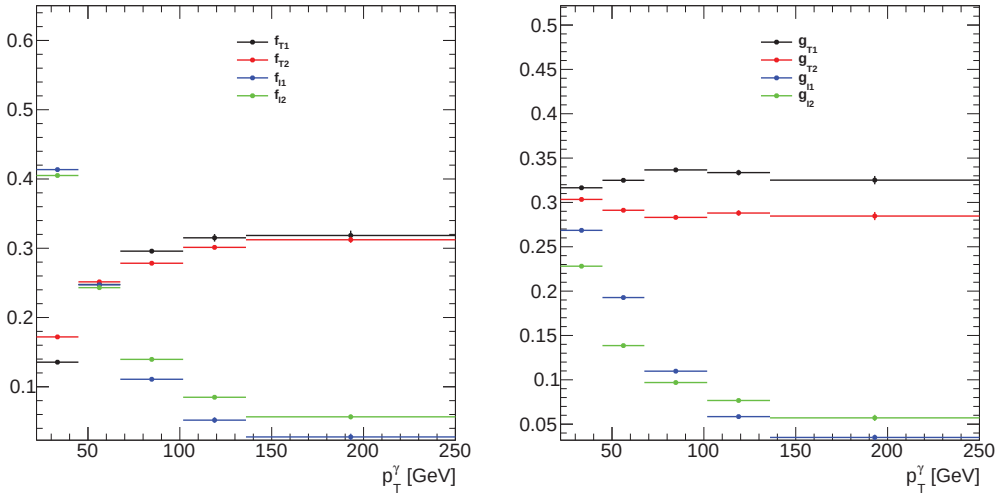

Figure 6.2.17: *Efficiencies for jets faking photons passing the isolation and identification criteria (left) and the same for jet-jet events (right) as a function of leading photon p_T for the jet preselection.*

6.2.4 Combining the two backgrounds to photons

The combination of the two backgrounds proceeds exactly as in section 5.5.2 on page 104 but is now performed in a bin-by-bin manner. The $\gamma\gamma$ purity can be found in table 6.2.7, and the event class counts in table 6.2.8. The other fit variables and the resultant observable distributions were found to differ only slightly from the case where only the jet background subtraction was included. Since the $\gamma\gamma$ purity is only very mildly perturbed by the inclusion of the electron background we will not show its dependence on our observables again. It's inclusive value drops from 0.90 to 0.89. The largest systematic uncertainty is still the extraction method variation for the photon isolation.

To conclude this section we remark that from the 277 events passing the full selection in data we estimate the $jj\gamma\gamma$ contribution to be:

$$215.2 \pm 18.4$$

In the next section we will discover how many of these are $b\bar{b}\gamma\gamma$ events.

$\gamma\gamma$ purity	single tag	full selection
Nominal value	0.89 ± 0.0046	0.89 ± 0.019
(electron) $f_{e\to\gamma}$ sideband interval up [%]	< 0.2	0.22
(jet) Using E_T^{iso} for ϵ_I [%]	± 5.9	± 5.0
(jet) Using PYTHIA for ϵ_T [%]	± 0.49	± 0.59
(jet) ϵ_{I1}background boundary up [%]	0.37	0.39
(jet) ϵ_{I1}background boundary down [%]	-0.56	-0.58
(jet) ϵ_{I2}background boundary up [%]	0.77	0.63
(jet) ϵ_{I2}background boundary down [%]	-1.0	-0.83
(jet) ϵ_{I2} statistical error up [%]	-0.32	< 0.2
(jet) ϵ_{I2} statistical error down [%]	0.32	< 0.2
(jet) ϕ_1 up [%]	-1.2	-1.5
(jet) ϕ_1 down [%]	0.98	1.2
(jet) ϕ_2 up [%]	-2.5	-2.2
(jet) ϕ_2 down [%]	1.9	1.7
(jet) ζ down [%]	-0.20	< 0.2
(general) Flavor tagging scale factors [%]	-0.20	-0.21
(general) Pile-up reweighting up [%]	-0.24	-0.25
(general) JVF cut up [%]	-0.43	0.35
(general) JVF cut down [%]	< 0.2	0.80
Total systematic error [%]	± 6.6	± 5.9

Table 6.2.7: *Signal purity obtained from the fit with uncertainties. The total systematic uncertainty was calculated by taking maximum of the up/down values of each systematic and adding them in quadrature. The word in parentheses is to remind the reader of the general source of the uncertainty. Only systematics with an effect of 0.2% or larger are listed.*

	single tag	full selection
$N_{\gamma\gamma}$	4096.8 ± 94.8	215.2 ± 18.4
$N_{j\gamma}$	168.6 ± 11.6	11.0 ± 3.0
$N_{\gamma j}$	301.6 ± 17.1	14.5 ± 3.6
N_{jj}	21.3 ± 1.3	1.0 ± 0.5
total	4588.2	241.7

Table 6.2.8: *Composition of the sample as determined by the fit. The quoted error is the fit uncertainty only.*

6.3 Non-b-jet background

THE b-tagging algorithm used for this study MV1, described in section 4.2.2 on page 71, has a nominal efficiency of 70% for the working point we chose. This means we only tag about 49% of genuine bb events. In addition a sizable amount of events are selected where one of the two jets is actually not a b-jet at all.

Our task will therefore be to estimate the amount of non-b-jets in our selected sample, and to correct the number of bb events we might have missed due to the low efficiency of simultaneously tagging two jets. Note that the odds are stacked slightly in our disfavor; we need to tag both jets for an event to enter our signal sample, but only one of them has to be wrongly tagged for it to be considered a background.

We considered briefly using another working point, in essence changing the b-tagging efficiency of the MV1 algorithm. However, we found that choosing a working point with a better efficiency let through too many non-b-jets (a problem compounded by the fact that only a single b-jet needs to be a fake for the event to be non-signal) and choosing a lower efficiency drastically reduced the number of signal events at the end of the cutflow. Therefore, we settled on the 70% working point.

6.3.1 Estimating the mistagging background

It is possible to retrieve the tagging efficiency for MV1 on a per-jet basis for simulated events. This is important because the reported 70% is an average, and might not describe our signal region all that well. The efficiency has also been derived on a certain class of events (for example $t\bar{t}$ events), which might differ from the event topology in our analysis. This tagging efficiency can be obtained for actual b-jets but also for non-b-jets. Non-b-jets are organized into two distinct classes, c-jets, which originate from c-hadrons, and light jets, which originate from gluons and the light quarks: up, down, and strange. Light jets are denoted by the letter l.

For the b and c classes the average MV1 tagging efficiency of either of the two leading jets as a function of jet-related observables can be found in figure 6.3.1. These were obtained via the SHERPA signal sample for every observable bin separately and with the pretag selection cuts applied. We left out the light contribution in these plots since their efficiency is so tiny that it is not visible on a linear scale[†]. The distributions with respect to the other observables was largely flat, and in table 6.3.1 the inclusive numbers are provided, also including the l case.

All these numbers have been determined in the pretag selection on a per-jet basis using a tool developed by the ATLAS b-tagging group which gives the probability of a certain jet to be tagged in a certain way depending on the kinematics of the jet, mainly the transverse momentum[‡]. The numbers thus obtained were then averaged per observable bin. It is important to note that this tool only considers the characteristics of the one jet it is asked about. It knows nothing about other things happening in the event. As such the exercise of obtaining the efficiency versus various kinematical variables of different objects is tantamount to checking for a correlation between that variable

[†] On a logarithmic scale it is impossible to see the small variations in the more important b and c tagging efficiencies.

[‡] The numbers the tool produces are of course based on calibrations performed by the b-tagging group. See chapter A for an example of such a calibration in which the author partook.

Flavor	Tagging efficiency
b	0.67 ± 0.024
c	0.20 ± 0.0033
l	$0.0069 \pm 2.4 \cdot 10^{-5}$

Table 6.3.1: *Tagging efficiencies for the different jet flavors from simulation samples in the pretag stage. The listed errors are statistical in nature.*

and the p_T of the jet in question because only for variables with a heavy correlation does the efficiency vary significantly. Such strong correlation was not found amongst our non-jet observables, but the average efficiency is smaller than the nominal 70% by approximately 3%.

Our method to remove the non-b-jet background is similar to the one employed in an earlier analysis that measured the inclusive bb cross-section at 7 TeV[86]. This involves determining the flavor fractions in the pretag selection, and then use tagging efficiencies to extrapolate to the double b-tag stage. The crux of the strategy is to take a distribution that is suitably sensitive to the presence of b-jets, which in our case is the MV1 discriminant. Then we extract shape templates for the signal bb events and for the background using simulation samples. We can then perform a template fit on the distribution of the MV1 discriminant in data to find the quantity of bb events. Technically there should be one signal template, bb, and up to eight background templates, but as discussed in [86] some of these backgrounds have very similar shapes with regard to b-tagging variables, and so we follow [86] in defining four distinct classes:

bb This is our signal and the two leading jets are supposed to be genuine b-jets.

bX In this class there is a single genuine b-jet but the other one is either a light jet or a c-jet.

cX Here there are no b-jets but there is at least one c-jet.

ll In this class we have two light jets.

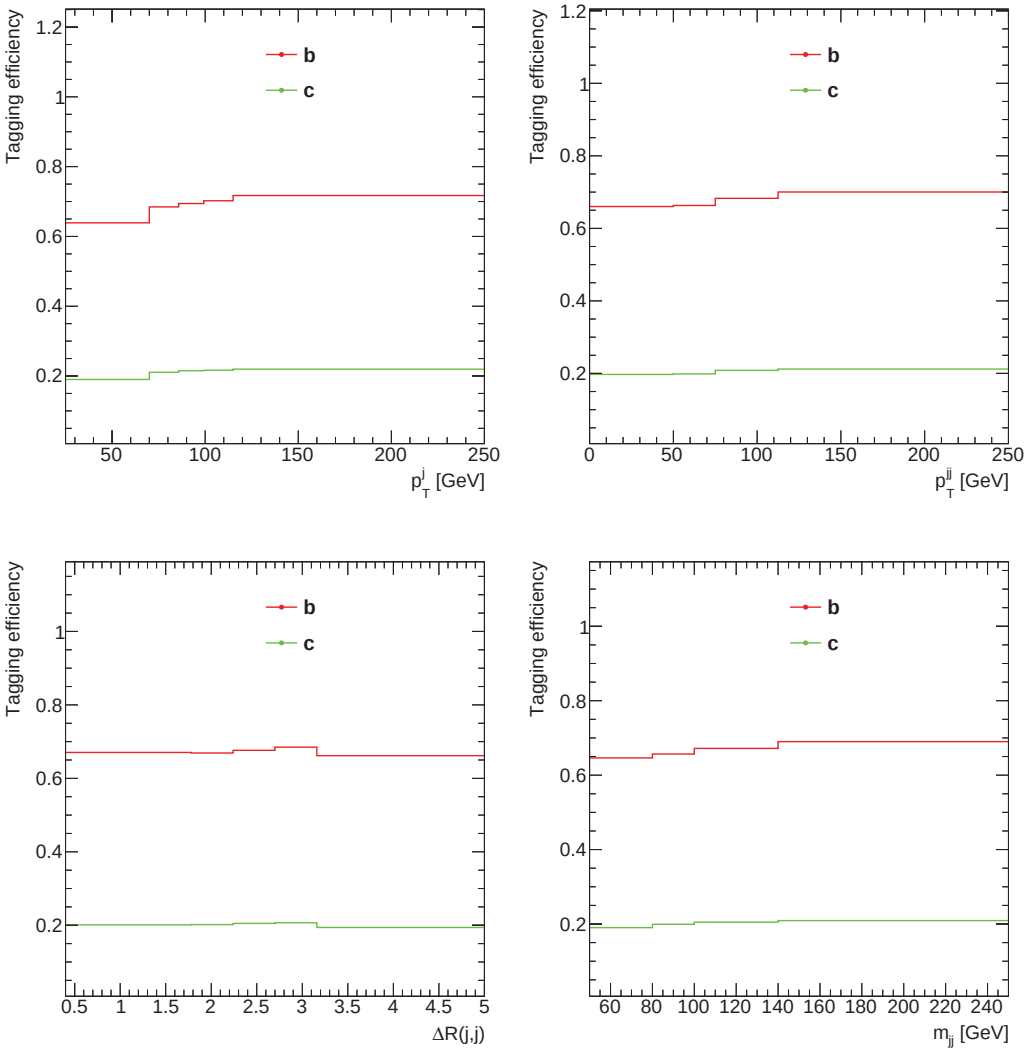

Figure 6.3.1: *Flavor c-tagging and b-tagging efficiencies as a function of the transverse momentum of the leading jet (top left) and of the leading dijet system (top right), and of $\Delta R(j,j)$ (bottom left) and m_{jj} (bottom right) in simulation samples in the pretag selection.*

Class	Tagging efficiency
bb	0.45
bX	0.024
cX	0.01
ll	0.005

Table 6.3.2: *Tagging efficiencies for the different flavor classes from simulation samples in the pretag stage. The listed errors are statistical in nature.*

Each of these classes has a separate efficiency for being actually tagged. These efficiencies were obtained by running over simulation samples and requiring that events pass all cuts up and including the pretag stage. Then the event was sorted into one of the four classes and its tagging efficiency taken to be the multiplication of the tagging efficiencies of the two leading jets. This procedure can also be done in bins of each observable. The results are found in table 6.3.2 where we show that the tagging efficiency for bb is 45%. For bb and bX we also show them as a function of the jet-related observables in figure 6.3.2. Apart from a dependence on the jet p_T, where the b-tagging efficiency drops at low values, the distributions are flat. The cX and ll classes were left out of these plots to increase readability.

To create the template the MV1 discriminant values of both jets are combined to uniquely determine one of sixteen bins. This procedure starts by placing each jet into an MV1 bin. The boundaries of these bins were chosen to comply with the b-tagging scale factors issued by the ATLAS b-tagging group. These scale factors correct for differences between b-tagging performance in data and simulation and are only measured in these specific bins:

$$[0,\ 0.3900,\ 0.8119,\ 0.9867,\ 1.0]$$

Thus, bin 1 runs from 0 to 0.39, bin 2 from 0.39 to 0.8119 and so forth. Jets with an MV1 discriminant above 0.8119 are considered b-tagged as this is the 70% working point. This corresponds to bins 3 and 4. The two bins are then combined into a single bin in a 2D MV1 spectrum via the following formula:

$$bin_{2D} = 4\,(bin_{lead} - 1) + bin_{sub} \tag{6.3.1}$$

Where bin_{2D} is the resultant bin, bin_{lead} is the MV1 bin for the leading jet and bin_{sub} the MV1 bin for the sub-leading bin. Bins 11, 12, 15, and 16 are those containing events in which we have two b-tagged jets. The inclusive MV1 discriminant shapes for every group in the pretag selection can be found on the left in figure 6.3.3. The templates are rederived for every bin for every observable separately and are taken from from simulations as before.

Using these templates we can perform a fit to find the flavor fractions in data in the pretag selection. The results are found in table 6.3.3. We find that the percentage of bb events in the pretag selection is 1.8%, with about 1096 expected events in total. In figure 6.3.3, on the right side, the fit is visualized for the inclusive case. The fit is stable and works well, and this also extends to fits made in bins of our observables. The flavor composition in terms of the jet-related observables can be found in figure 6.3.4. We find that the relative composition does not vary much as a function of these observables, but only very few events fall into the bb category.

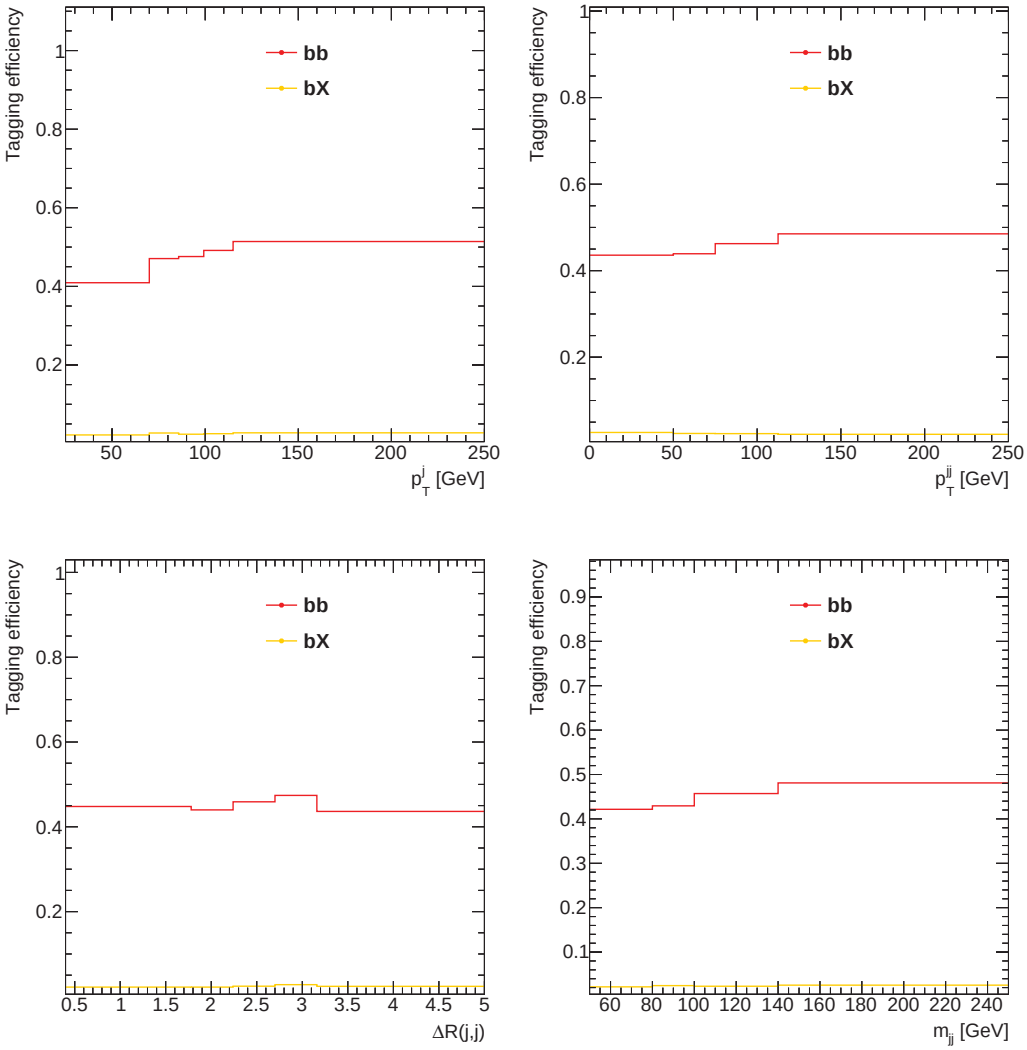

Figure 6.3.2: *Class flavor tagging efficiencies as a function of the transverse momentum of the leading jet (top left), leading dijet system (top right), $\Delta R(j, j)$ (bottom left), and the leading dijet invariant mass (bottom right), in simulation samples in the pretag selection.*

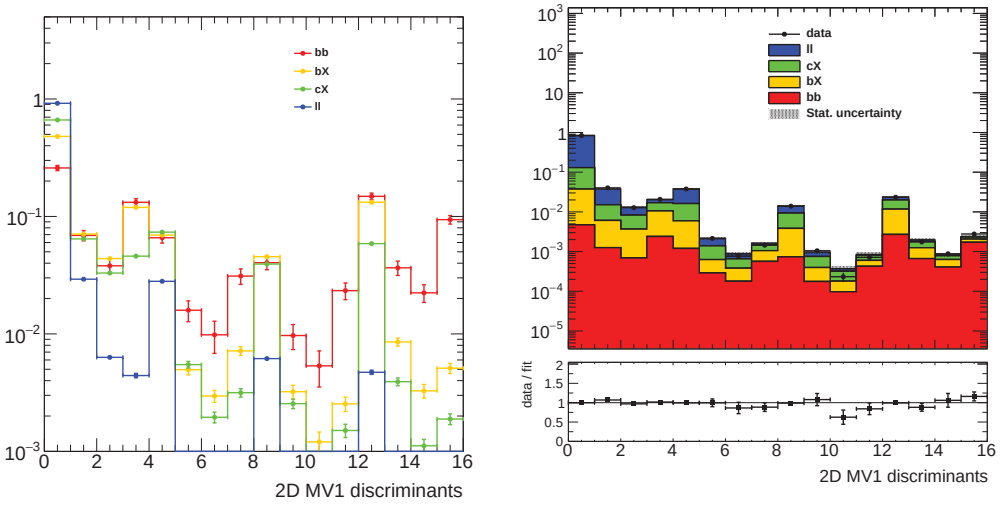

Figure 6.3.3: *On the left: The distribution of events in the 16 MV1 bins. The binning itself is explained in the main text. The distributions have been normalized to unity. On the right: Fit to the 2D MV1 discriminants in data in the pretag selection.*

Class	Flavor fraction	Events
bb	0.018 ± 0.0019	1095.8 ± 114.2
bX	0.068 ± 0.0091	4104.3 ± 547.5
cX	0.14 ± 0.016	8395.7 ± 984.4
ll	0.77 ± 0.0093	46539.6 ± 592.9
Total		60135.4 ± 1278.0
Data		60162

Table 6.3.3: *Fitted flavor fractions and event numbers in the pretag selection compared to the number of events found in data at the same stage in the cutflow. The listed error is the fitting error which takes into account only the statistical uncertainty on the measured shape in data.*

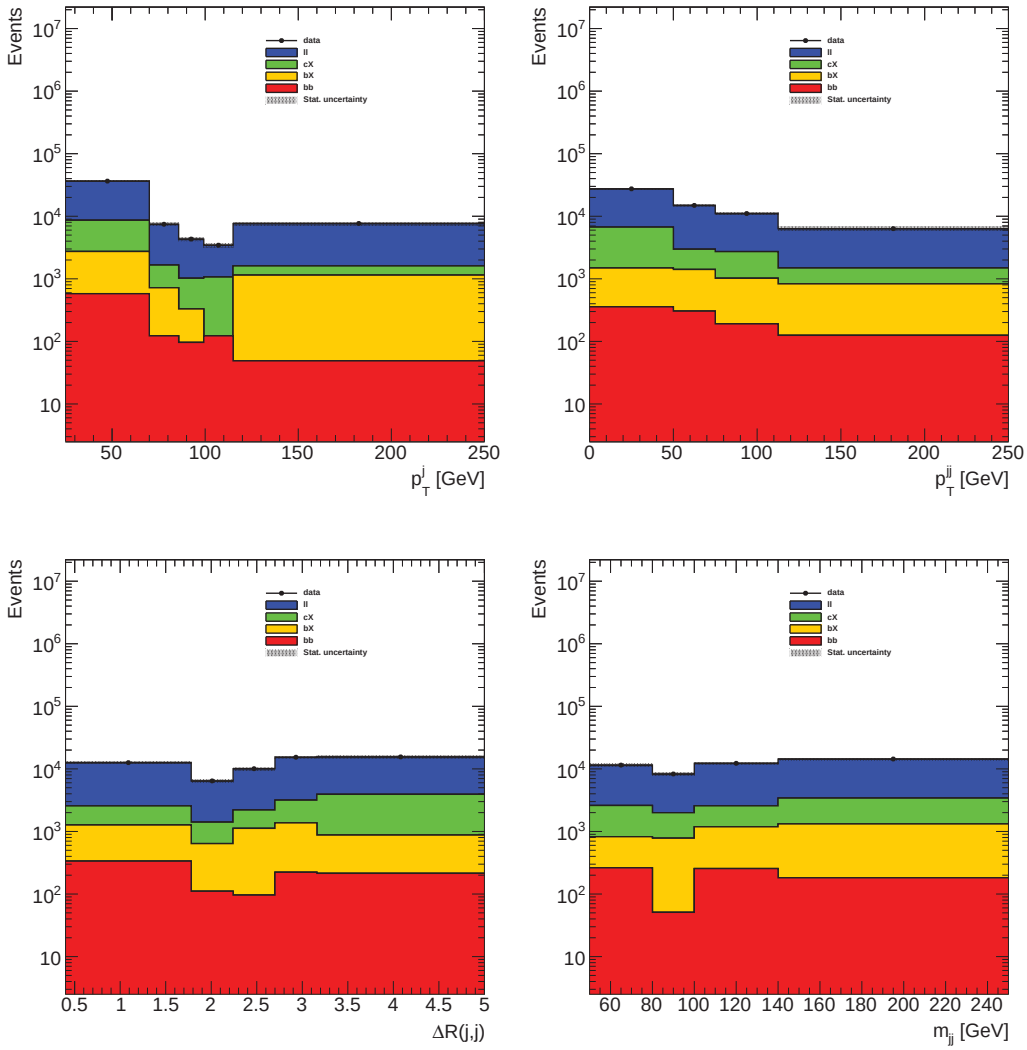

Figure 6.3.4: *Fitted flavor composition as a function of the jet-related observables. Shown are the leading jet transverse momentum (top left), transverse momentum of the leading jet pair (top right), $\Delta R(j,j)$ (bottom left), and the dijet invariant mass (bottom right).*

We can now move on to the actual signal region where we require the two leading jets to be b-tagged. It is possible to estimate the bb fraction in the full selection sample by combining the knowledge of the flavor composition at the pretag stage with the flavor tagging efficiencies we obtained before from simulations.

Let us define f_{xx}^{pretag}:

$$f_{xx}^{pretag} \equiv \frac{N_{xx}^{pretag}}{N^{pretag}} \tag{6.3.2}$$

This quantity is the pretag flavor fraction and quantifies which fraction of events in the pretag selection is of class xx, where xx stands for bb, bX, cX, or ll. This quantity is known to us from the template fits made before. N^{pretag} is the number of events in the pretag stage as measured in data and N_{xx}^{pretag} the number of xx events in the pretag selection.

The number of events of class xx estimated to be present in the full selection, where we require two b-tags, is then:

$$N_{xx}^{2b} \equiv f_{xx}^{pretag} \epsilon_{xx} N^{pretag} \tag{6.3.3}$$

Where ϵ_{xx} is the tagging efficiency for events of class xx, which we obtained from the aforementioned simulations. We can also defined the full selection flavor fraction as follows:

$$f_{xx}^{2b} \equiv \frac{N_{xx}^{2b}}{N_{bb}^{2b} + N_{bX}^{2b} + N_{cX}^{2b} + N_{ll}^{2b}} \tag{6.3.4}$$

We can conveniently rewrite this using the N_{xx}^{2b} defined before[†]:

$$f_{xx}^{2b} = \frac{f_{xx}^{pretag} \epsilon_{xx} N^{pretag}}{N^{pretag} \sum_{i \in classes} f_i^{pretag} \epsilon_i} \tag{6.3.5}$$

$$= \frac{f_{xx}^{pretag} \epsilon_{xx}}{\sum_{i \in classes} f_i^{pretag} \epsilon_i} \tag{6.3.6}$$

Where the sum sums over the four flavor classes: bb, bX, cX, and ll. If we multiply this quantity with the number of events in the full selection, N^{2b} by the xx flavor fraction, f_{xx}^{2b} we obtain the number of xx events in the full selection:

$$N_{xx}^{2b} = f_{xx}^{2b} N^{2b} \tag{6.3.7}$$

$$= \frac{f_{xx}^{pretag} \epsilon_{xx}}{\sum_{i \in classes} f_i^{pretag} \epsilon_i} N^{2b} \tag{6.3.8}$$

However, we also need to correct for any events of type xx which we missed by failing to tag them properly. The probability to tag such an event correctly is given by ϵ_{xx}. If we divide by ϵ_{xx} we thus correct for these missing events. This leads to the following expression for the number of events of type xx:

$$N_{xx}^{true} \equiv \frac{f_{xx}^{pretag}}{\sum_{i \in classes} f_i^{pretag} \epsilon_i} N^{2b} \tag{6.3.9}$$

[†]At this point an astute reader might ask why we did not replace $N_{bb}^{2b} + N_{bX}^{2b} + N_{cX}^{2b} + N_{ll}^{2b}$ by the number of measured double b-tagged events, N^{2b}. The reason is that this will make f_{xx}^{2b} dependent on N^{pretag} since it will not cancel any more. And this will then make combining the non-b-tagged background with the other backgrounds very complicated since N^{pretag} and N^{2b} both will then need corrections whereas now only N^{2b} does.

Class	C_{xx}	f_{xx}^{2b}
bb	1.2 ± 0.13	0.542 ± 0.043
bX	4.5 ± 0.83	0.108 ± 0.021
cX	9.2 ± 1.59	0.096 ± 0.017
ll	50.8 ± 6.54	0.254 ± 0.033

Table 6.3.4: *Extracted flavor correction factors, C_{xx}, and flavor fractions, f_{xx}^{2b} obtained from the pretag stage to be applied to the doubly b-tagged sample. The listed error is the fitting error which takes into account only the statistical uncertainty on the measured shape in data.*

This means that to arrive at the true number of events of type xx we need to multiply the number of events in the full selection by a so-called flavor correction factor which we shall name C_{xx}:

$$C_{xx} = \frac{f_{xx}^{pretag}}{\sum_{i \in classes} f_i^{pretag} \epsilon_i} \tag{6.3.10}$$

We thus extract the necessary C_{xx} from the pretag stage and simply apply it after requiring two b-tags. The values found for f_{xx}^{2b} and C_{xx} for all events in the pretag stage can be found in table 6.3.4. We list both because f_{xx}^{2b} gives a feel of the flavor fraction inside our doubly b-tagged sample, whilst C_{xx} gives a measure of the true number of events one should measure with a perfect detector. From the table one can read that the full selection contains about one half genuine bb events, which combined with the 45% tagging efficiency means a flavor correction factor of about 1.2. The distributions of the C_{xx} with respect to our observables were found to be largely flat and are omitted. The following systematic uncertainties were considered:

Template uncertainties To assess the effects of the statistical uncertainties in the class templates the fit was repeated but the values in each bin in the MV1 templates were varied in a specific way. For the up variation every even bin was varied up according to the statistical uncertainty, and every odd bin was varied down. For the down variation the procedure was repeated with the roles of the odd and even bins being swapped.

Tagging efficiency The uncertainty on the extracted tagging efficiencies are obtained together with the tagging efficiencies themselves. As mentioned before these are based on various tagging calibrations performed by the ATLAS b-tagging group. We vary the efficiencies within their uncertainties.

In the next section we will combine this background with the other two and discuss the size of these systematics.

6.3.2 Combining with the jet and electron backgrounds

The backgrounds concerning the diphoton system, namely jets and electrons faking photons, can be conveniently separated from the background involving the dijet system, which are non-b-jets that end up being b-tagged. This is possible because both pertain to a different part of the event. As such the combination of the backgrounds is rather straightforwards. In section 5.5.2 on page 104 we calculated the fraction of the fake events in each of the isolation and identification categories, and performed the fit to remove the jet background with these modified numbers. The determination of the flavor correction factors for the non-b-jet background is blind to the photons however and we can therefore multiply the bin-by-bin yields coming out of the jet fit with these correction factors[†].

The resultant distributions for our observables after the combined total background subtraction can be found in figures 6.3.5 to 6.3.7. The total number of events found to pass all criteria and corrected for the various backgrounds can be found in table 6.3.5 along with a list of the systematic uncertainties and their sizes. The largest systematic (~10%) is formed by the b-tagging MV1 template uncertainties.

In the inclusive case we expect the following number of genuine $b\bar{b}\gamma\gamma$ events in the full selection if we had a perfect detector:

$$257 \pm 47.9\,(\text{stat}) \pm 29.30\,(\text{syst})$$

[†]This is equivalent to first applying the flavor correction factors and then inputting the yields into the jet fit. It shall only be different if we calculate the correction factors for each of the sixteen categories, used in the jet fit, separately. However, this might introduce correlations which the sixteen equations are not equipped to handle, and our final $b\bar{b}\gamma\gamma$ yield is too small to study this in further detail.

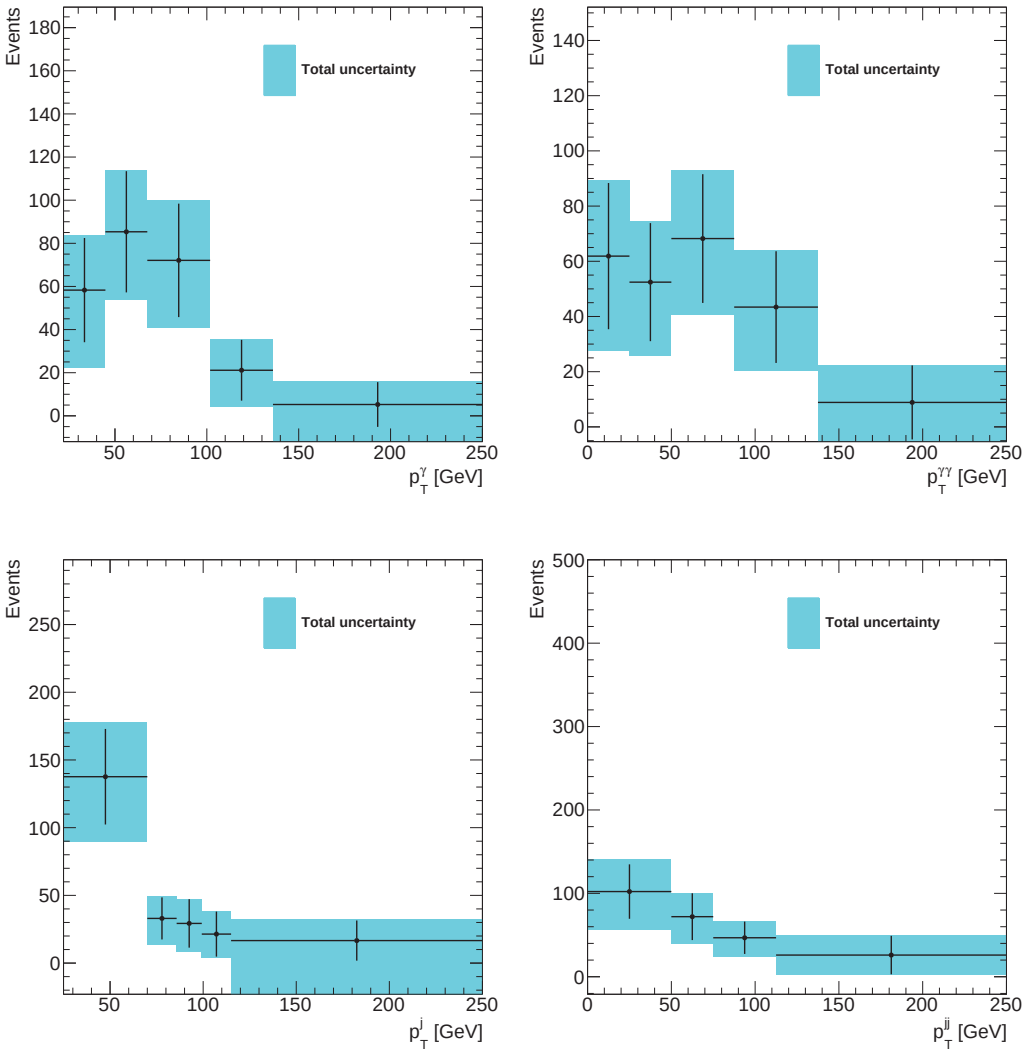

Figure 6.3.5: *Distribution of the transverse momenta of the leading photon (top left), leading diphoton system (top right), leading jet (bottom left), and leading dijet system (bottom right), after correction for all backgrounds and after requiring two b-tags on the two leading jets including systematic and statistical uncertainties.*

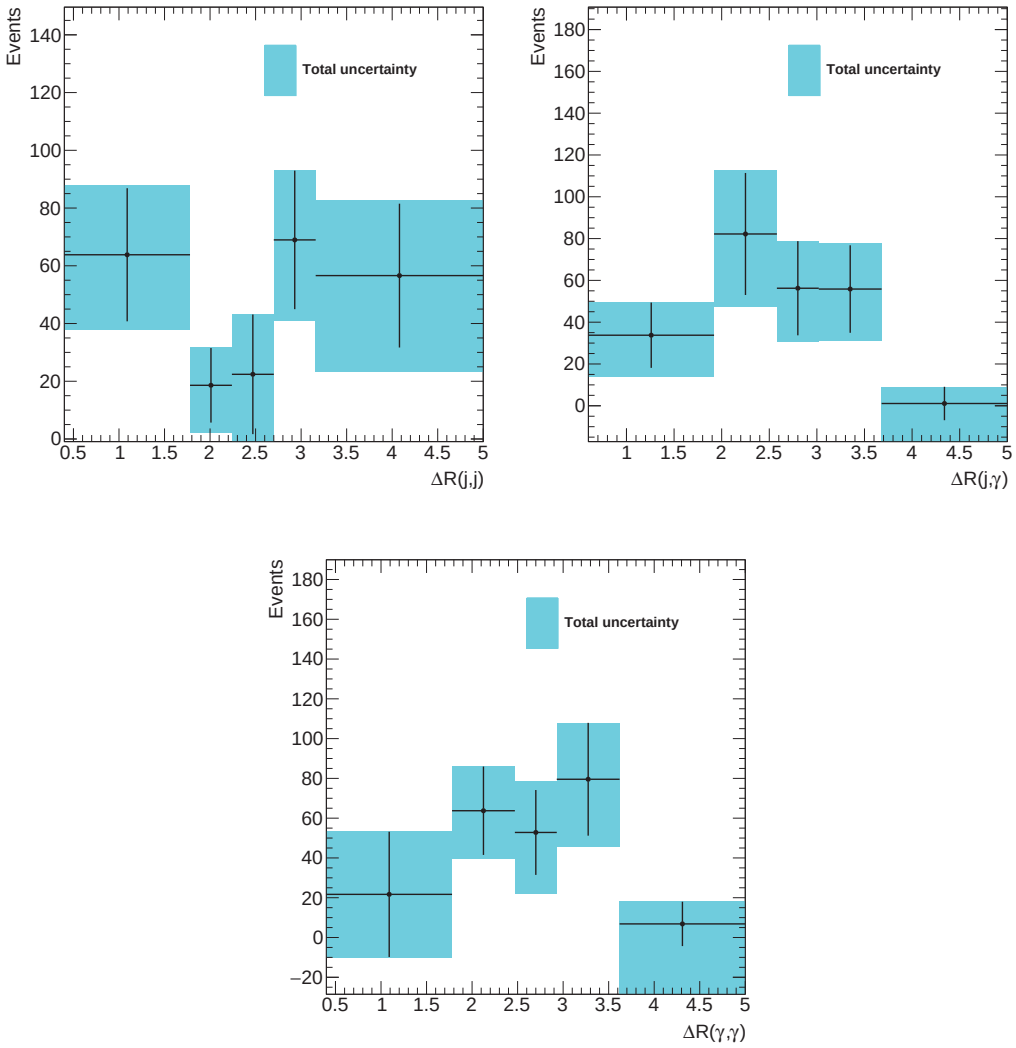

Figure 6.3.6: *Distribution of the angular separations $\Delta R(j,j)$ (top left), $\Delta R(j,\gamma)$ (top right), and $\Delta R(\gamma,\gamma)$ (bottom), after correction for all backgrounds and after requiring two b-tags on the two leading jets including systematic and statistical uncertainties.*

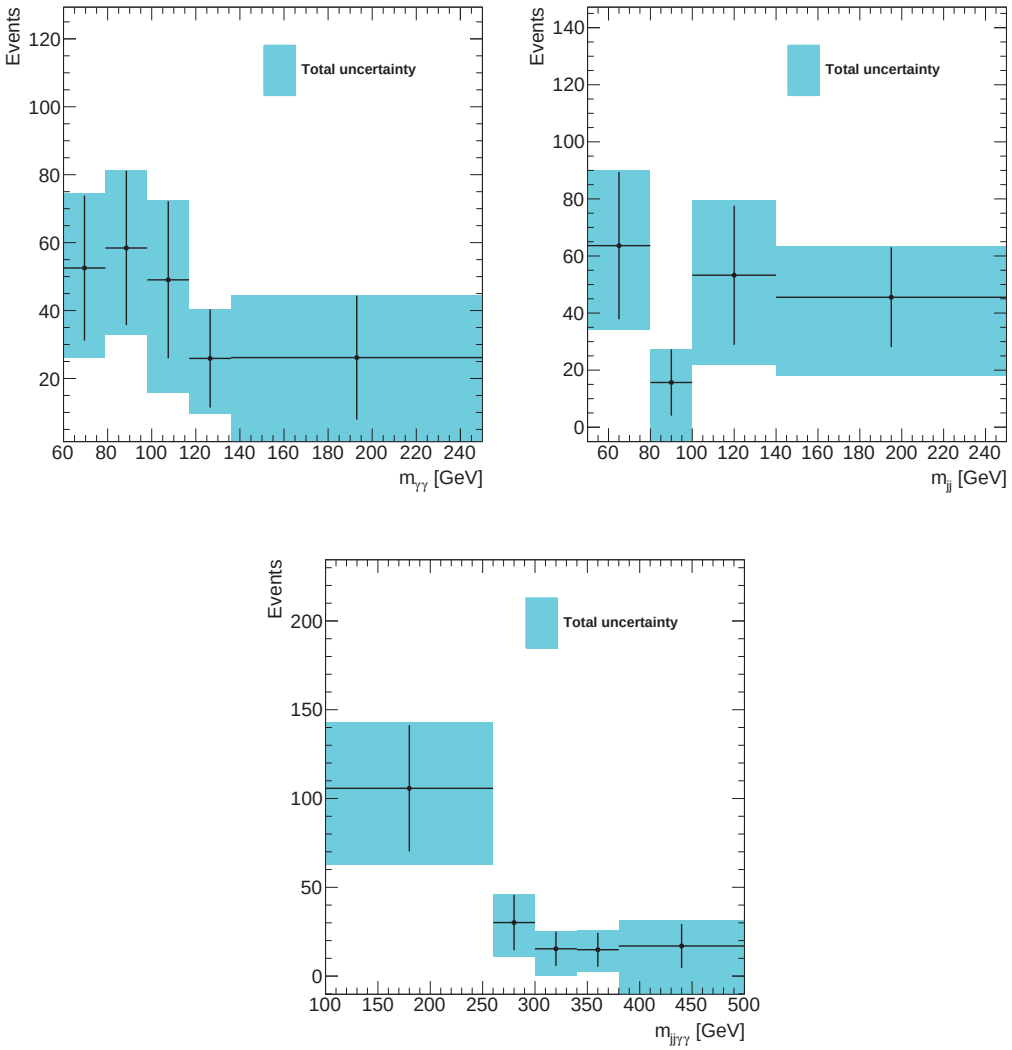

Figure 6.3.7: *Distribution of the invariant masses of the leading diphoton system (top left), leading dijet system (top right), and the combined system (bottom), after correction for all backgrounds and after requiring two b-tags on the two leading jets including systematic and statistical uncertainties.*

Events passing the full selection	
Nominal value	257 ± 47.9
(non-b-tag) Template uncertainties up [%]	-9.47
(non-b-tag) Template uncertainties down [%]	4.87
(electron) $f_{e\to\gamma}$ sideband shift up [%]	-0.99
(electron) $f_{e\to\gamma}$ sideband shift down [%]	1.87
(jet) Using E_T^{iso} for ϵ_I[%]	± 4.97
(jet) Using PYTHIA for ϵ_T[%]	± 0.59
(jet) ϵ_{I1}background boundary down [%]	-0.58
(jet) ϵ_{I2}background boundary up [%]	0.62
(jet) ϵ_{I2}background boundary down [%]	-0.82
(jet) ϕ_1 up [%]	-1.54
(jet) ϕ_1 down [%]	1.23
(jet) ϕ_2 up [%]	-2.16
(jet) ϕ_2 down [%]	1.68
(general) Lenient good PV cut [%]	0.43
(general) JVF cut up [%]	1.20
(general) JVF cut down [%]	1.65
Total systematic error [%]	± 11.4

Table 6.3.5: *Inclusive number of events passing the full selection and corrected for the backgrounds. The total systematic uncertainty was calculated by taking maximum of the up / down values of each systematic and adding them in quadrature. The word in parentheses is to remind the reader of the general source of the uncertainty. Only systematics with an effect of 0.4% or more are listed.*

6.4 Unfolding

In order to be able to compare our results to those obtained by other physics experiments, or to theory calculations, we need to correct for artifacts imposed on the results by the detector. This includes the limited acceptance of the detector where events are wrongly reconstructed or fall outside the fiducial cross-section. It also includes bin-migration where measured observables like the energy of a particle are mismeasured and thus an event might end up in a different histogram bin.

There are different ways of dealing with this[†], but the basic strategy is virtually the same: one generates a large number of simulated events and runs the full detector simulation and reconstruction code on them. By comparing the output from this to the original simulated sample one can extract the detector effects, and then use this information to correct the actual measured data. The method we shall be employing in this analysis is known as SVD unfolding [88] where SVD stands for Singular Value Decomposition, which is a concept in linear algebra. We shall be explaining this method in the next section before moving on and applying it to the analysis in section 6.4.2.

6.4.1 SVD unfolding

In this section we will present singular value decomposition before moving on to describe the unfolding procedure. It is defined for both real and complex matrices, but since we will only be dealing with real matrices for the unfolding we shall restrict ourselves to the slightly easier real case. In the same spirit we shall also restrict ourselves to square matrices. The actual unfolding method is described in detail in [88] but we shall outline the general procedure here.

SVD decomposition

Mathematically SVD decomposition is the factorization of a, in our case, real square matrix, M, into the following form:

$$M = U\,\Sigma\,V^T \tag{6.4.1}$$

where X^T indicates the transpose of a matrix X and U, Σ, and V are matrices with specific properties. U and V are unitary matrices[‡] that can be regarded as rotations. Σ is a diagonal matrix with non-negative values on the diagonal which are known as the singular values of the matrix M. The columns of U form the set of orthonormal eigenvectors of $M\,M^T$, and the columns of V for the set of orthonormal eigenvectors of $M^T M$. Determining these eigenvectors is the principal way of finding U and V and thus Σ. Examples of actual algorithms can be found in the literature [89, 90], but the important thing is that this factorization is always possible.

[†] See [87] for example for an overview.

[‡] This means that in the case of a real-valued matrix the transpose of the matrix is its inverse, i.e. for a real unitary matrix U the following holds: $UU^T = U^T U = I$ where I is the identity matrix.

There is a bit of freedom in how singular values are to be ordered because there is freedom in how to order the orthonormal eigenvectors. The usual convention is to order them in a descending order. So that:

$$\Sigma_{ii} \geq \Sigma_{jj} \text{ for } i < j \tag{6.4.2}$$

This is the convention we adopt as well.

Unfolding using SVD decompositions

The problem that the unfolding aims to solve can be elegantly stated as a matrix equation:

$$\hat{A}\vec{x} = \vec{b} \tag{6.4.3}$$

where $\vec{b}$ represents a measured observable. This means that if a histogram codifies a measured distribution the contents of bin i are described by the i'th element of $\vec{b}$, and so forth. Since our signal has a low cross-section the number of bins is small in order to keep down the statistical uncertainties. The actual truth distribution is captured in $\vec{x}$ which, due to detector effects and cut acceptance, can differ significantly from the measured $\vec{b}$. The detector and acceptance effects are then stored in the matrix $\hat{A}$ which is a probability matrix; an element $A_{ij} \in \hat{A}$ denotes the probability that an event that sits in bin j of the truth distribution (i.e. element j of $\vec{x}$) ends up in bin i of the measured distribution. To do any unfolding one first determines $\hat{A}$ by generating a sample of simulated events, store them in $\vec{x}$, and then keeping track of in which bins (if any) these events end up after being reconstructed and handled by the analysis cuts. Then for the actual measurement one simply determines the inverse of $\hat{A}$ and solves for $\vec{x}$:

$$\vec{x} = \hat{A}^{-1}\vec{b} \tag{6.4.4}$$

Whilst this is technically correct it will not work well in practice. The main reason is that the matrix inverse is very sensitive to small fluctuations in its elements. As an example [88] imagine the following form of $\hat{A}$:

$$\hat{A}_{example} = \frac{1}{2}\begin{pmatrix} 1+q & 1-q \\ 1-q & 1+q \end{pmatrix} \tag{6.4.5}$$

where q is a parameter that describes the quality of the detector. If $q = 1$ we get the identity matrix and so $\vec{b}$ and $\vec{x}$ are equal; the detector is perfect. If $q = 0$ the detector cannot distinguish between the two bins of our measured distribution anymore and so is a very poor detector. The inverse of this matrix is:

$$\hat{A}^{-1}_{example} = \frac{1}{2q}\begin{pmatrix} q+1 & q-1 \\ q-1 & q+1 \end{pmatrix} \tag{6.4.6}$$

Imagine we measured the following number of events per bin:

$$\vec{b}_{example} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{6.4.7}$$

This would mean that:

$$\vec{x}_{example} = \hat{A}^{-1}_{example}\vec{b}_{example} = \frac{b_1 + b_2}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{b_1 - b_2}{2q}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{6.4.8}$$

The first term is sensitive to the total number of measured events, whilst the second one is sensitive to the difference in events measured in the two bins. The trouble arises from this second term, and more specifically from the division by q. If $q = 0$ the solution is not even well defined, but even if q is very small the solution is problematic. Imagine for example that in our measurement the difference between b_1 and b_2 is not statistically significant, then $b_1 - b_2$ is essentially a random number, which divided by a small q could easily cause the second term to dominate in the expression. This in turn would mean our $\vec{x}$ becomes a random number as well. Clearly the solution is to set the second term to zero, and report only the first one whilst quoting an error for the second term. This is a regularization of the statistical fluctuations, which we shall examine next as an alternative to simply taking the inverse of the probability matrix.

The method put forward in [88] starts by modifying expression 6.4.3 in several ways. The first is to introduce a new vector, $\vec{w}$, that contains a series of weights which contrast the unfolded distribution $\vec{x}$ to the one initially used to construct $\hat{A}$, $\vec{x}_{ini}$:

$$w_i \equiv \frac{x_i}{x_{ini,i}} \tag{6.4.9}$$

The reason behind this change is that the bin-to-bin variations in $\vec{w}$ are likely to be smaller than those in $\vec{x}$ since $\vec{x}$ and $\vec{x}_{ini}$ are likely to be comparable. This is because in general $\hat{A}$ is contructed using simulated events which are as close to a desired signal as possible. The importance of this will be made clear later. We now have:

$$A_{ij}\frac{x_j}{x_{ini,j}} = b_i \tag{6.4.10}$$

Another change is to introduce the detector response matrix $\hat{R}$ in favor of the probability matrix $\hat{A}$:

$$R_{ij} \equiv A_{ij}x_{ini,j} \tag{6.4.11}$$

so that we end up with:

$$\hat{R}\vec{w} = \vec{b} \tag{6.4.12}$$

Note that the elements $R_{ij} \in \hat{R}$ now list the actual number of events that were in bin j of the truth distribution that end up in bin i of the measured distribution instead of a probability. The reason for this has to do with statistics. As we shall see later the matrices $\hat{A}$ and $\hat{R}$ are mostly diagonal (the ATLAS detector would be not much good if this were not the case) but there are some far off-diagonal elements. These elements are tiny and statistical outliers.

Such an element is interpreted differently in $\hat{R}$ and $\hat{A}$. In the case of $\hat{R}$ the element $R_{ij} \in \hat{R}$ is some tiny number, say 1, to denote that a single event from the initial simulated sample found in the truth distribution in bin j ended up in bin i after detector simulation. But this element is tiny compared to the elements found along the diagonal of $\hat{R}$ where the bulk of the events end up. Therefore, this far off-diagonal element will carry very little weight in the procedure outlined below. Only elements that were

visited a lot in the simulation, and therefore have a low statistical uncertainty, will contribute a lot to the final result. In contrast, in the same situation, a far off-diagonal element $A_{ij} \in \hat{A}$ that has been filled only once during the initial simulation will have a probability value of 1, which is highest value to be found in the entire matrix! This element therefore carries a huge weight, which is undeserved since most of the events lay along the diagonal. Thus, $\hat{R}$ is more robust again statistical errors than $\hat{A}$ is[†].

Turning away from finding the exact inverse of the detector response matrix, one can also cast the problem of unfolding in a weighted least-square problem where one tries to minimize the following expression:

$$\sum_i \sum_j \left(\frac{R_{ij} w_j - b_i}{\Delta b_i} \right)^2 \tag{6.4.13}$$

where Δb_i is the error on the contents of bin i of $\vec{b}$. This expression can be rewritten into a matrix equation where one wants to minimize:

$$\left(\hat{R}\vec{w} - \vec{b} \right)^T \hat{B}^{-1} \left(\hat{R}\vec{w} - \vec{b} \right) \tag{6.4.14}$$

where $\hat{B}^{-1}$ is the inverse of the covariance matrix $\hat{B}$ of the measured distribution $\vec{b}$. The shape of $\hat{B}$ in our case is trivial since we can treat each bin as a separate measurement and so $\hat{B}$ ends up with only diagonal elements and which are the squares of the errors on elements of $\vec{b}$:

$$\hat{B} \equiv \begin{pmatrix} (\Delta b_1)^2 & 0 & 0 & 0 & \dots \\ 0 & (\Delta b_2)^2 & 0 & 0 & \dots \\ 0 & 0 & (\Delta b_3)^2 & 0 & \dots \\ 0 & 0 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \tag{6.4.15}$$

The procedure starts by performing an SVD on $\hat{B}$ which in our case is diagonal[‡] and therefore the two rotation matrices are equal:

$$\hat{B} = Q \Sigma Q^T \tag{6.4.16}$$

In the special case where $\hat{B}$ is arranged in such a way that its diagonal elements form a descending sequence the matrix Q will simply be the identity matrix and Σ will be equal to $\hat{B}$. In other cases Σ is equal to a copy of $\hat{B}$ where the elements are ordered in a descending fashion and Q is appropriately changed to rotate Σ into $\hat{B}$, since we have to adhere to our convention of ordering the singular values in a descending way. The inverse of $\hat{B}$ is then given by:

$$\hat{B}^{-1} = Q \Sigma^{-1} Q^T \tag{6.4.17}$$

[†]Note that the procedure outlined in this chapter does not take into account and statistical uncertainties on the detector matrix, whether it be $\hat{A}$ or $\hat{R}$. This makes it even more prudent to use $\hat{R}$. It should be mentioned that it is shown in [90] that the statistical uncertainties on $\hat{R}$ can be safely ignored if there are at least a hundred times more events in the initial simulation sample (used to determine $\hat{R}$) than are found in $\vec{b}$. In that case the statistical uncertainty of the end result will be dominated by the statistical uncertainties of $\vec{b}$. This is definitely the case in our application.

[‡]The procedure put forward in [88] is more general, and also deals with the non-diagonal case for which this decomposition is slightly more involved.

Inputting this into expression 6.4.14 we have:

$$\left(\hat{R}\vec{w}-\vec{b}\right)^T Q\Sigma^{-1}Q^T\left(\hat{R}\vec{w}-\vec{b}\right) \tag{6.4.18}$$

Which we can rewrite to something more pleasing to the eye by moving some terms around:

$$\begin{aligned}\left(\hat{R}\vec{w}-\vec{b}\right)^T Q\Sigma^{-1}Q^T\left(\hat{R}\vec{w}-\vec{b}\right) &= \left(Q^T\hat{R}\vec{w}-Q^T\vec{b}\right)^T\Sigma^{-1} \\ &\quad\cdot\left(Q^T\hat{R}\vec{w}-Q^T\vec{b}\right)\end{aligned} \tag{6.4.19}$$

Then one can split Σ^{-1} into two parts (note Σ and Σ^{-1} are diagonal) by saying:

$$\Sigma^{-1}=\Sigma^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}} \tag{6.4.20}$$

where:

$$\Sigma_{ij}^{-\frac{1}{2}}=\frac{1}{\sqrt{\Sigma_{ij}}} \tag{6.4.21}$$

And define:

$$\tilde{R} = \Sigma^{-\frac{1}{2}}Q^T\hat{R} \tag{6.4.22}$$

$$\tilde{b} = \Sigma^{-\frac{1}{2}}Q^T\vec{b} \tag{6.4.23}$$

So the expression to minimize is now:

$$\left(\tilde{R}\vec{w}-\tilde{b}\right)^T\left(\tilde{R}\vec{w}-\tilde{b}\right) \tag{6.4.24}$$

Note that we now have a least-squares problem without weights. We have thus essentially rescaled everything so every equation in this matrix minimization problem has the same importance, although this is simply semantics, the main aim was the reordering of the errors.

The exact solution to the minimization problem again involves finding the matrix inverse, in this case of $\tilde{R}$, which will run into the same stability problems as before. The authors of [88] propose to fix this by adding a stabilization term to the expression to suppress the basically random input from statistical uncertainties. They propose we change to minimizing this expression instead:

$$\left(\tilde{R}\vec{w}-\tilde{b}\right)^T\left(\tilde{R}\vec{w}-\tilde{b}\right)+\tau\cdot\left(C\vec{w}\right)^T C\vec{w} \tag{6.4.25}$$

where C is called the stabilization matrix whose dimensionality is equal to $n\times n$ where n is the dimensionality of $\vec{b}$, and τ is a stabilization parameter which governs the strength of the stabilization term. If τ is very large then minimizing the expression is dominated by trying to minimize the stabilization term whereas if it is very small it has only a limited effect. Therefore, the choice of τ is important and the authors' advice on choosing it is discussed later. The matrix C imposes a condition on $\vec{w}$ which, if it is deviated from too much, will make the stabilization term large. This way large variations caused by random fluctuations are killed. The choice of C is also very important, and as an

example, consider the case where C is the identity matrix. The stabilization term then becomes:

$$\begin{aligned}
\tau \cdot (C\vec{w})^T C\vec{w} &= \tau \sum_m \sum_i \sum_j w_m C_{mi} C_{ij} w_j && (6.4.26)\\
&= \tau \sum_m \sum_i \sum_j w_m \delta_{mi} \delta_{ij} w_j && (6.4.27)\\
&= \tau \sum_i w_i w_i && (6.4.28)\\
&= \tau \left|\vec{w}\right| && (6.4.29)
\end{aligned}$$

so that solutions with a large $\vec{w}$ norm are discriminated against. One should make a choice for C that is physically motivated. The authors propose the following form for C:

$$C = \begin{pmatrix}
-1 & 1 & 0 & 0 & ... & \\
1 & -2 & 1 & 0 & ... & \\
0 & 1 & -2 & 1 & ... & \\
 & ... & & & ... & \\
 & ... & & 1 & -2 & 1 \\
 & ... & & & 1 & -1
\end{pmatrix} \qquad (6.4.30)$$

In order to see what effect this particular choice has on the stabilization term we will analyze:

$$\sum_j C_{ij} w_j \qquad (6.4.31)$$

From the definition of C it follows that:

$$\sum_j C_{ij} w_j = \begin{cases}
-w_i + w_{i+1} & \text{if } i = 0 \\
w_{i-1} - w_i & \text{if } i = n \\
w_{i-1} - 2w_i + w_{i+1} & \text{otherwise}
\end{cases} \qquad (6.4.32)$$

where n is the length of $\vec{w}$. Since C is symmetric the same holds for $\sum_m w_m C_{mi}$ and so the combination yields:

$$\left(\sum_m w_m C_{mi}\right)\left(\sum_j C_{ij} w_j\right) = \begin{cases}
(w_{i+1} - w_i)^2 & \text{if } i = 0 \\
(w_{i-1} - w_i)^2 & \text{if } i = n \\
(w_{i-1} - 2w_i + w_{i+1})^2 & \text{otherwise}
\end{cases} \qquad (6.4.33)$$

which translates to:

$$\tau \sum_i \left[(w_{i+1} - w_i) - (w_i - w_{i-1})\right]^2 \text{ where } w_{-1} \equiv w_0 \text{ and } w_{n+1} \equiv w_n \qquad (6.4.34)$$

This appears to be a sum over the squares of the second derivatives of the elements of $\vec{w}$ with the end-point held fixed. The authors of [88] call this the "curvature" of the discrete distribution $\vec{w}$. The reasoning is that since we defined $\vec{w}$ to be the deviation from the initial simulated sample (which should be physics motivated), it should be rather smooth with little bin-to-bin variation, and this choice of C enforces that belief by killing

off any contribution to $\vec{w}$ which has large curvature and is therefore presumably caused by random statistical influences. Unfortunately this choice of C has a determinant of zero and is therefore not inversable. This can be easily fixed though by introducing a parameter called ξ that can be taken to be small (our choice shall be $\xi = 10^{-4}$) and substituting:

$$C_{ij} \rightarrow C_{ij} + \delta_{ij}\xi \tag{6.4.35}$$

which adds the tiny parameter to the diagonal values of C which makes the matrix non-degenerate[†], since ξ is small it will not change any numerical outcome, but it will make the calculation easier[‡].

We need to minimize expression 6.4.25 for an arbitrary value of τ. For this a method called *damped least squares* [90] is used, whereby first a solution is found for $\tau = 0$ which can then be trivially extended to a solution for an arbitrary value of τ. Note, however, that the solution for $\tau = 0$ cannot be used directly as an unfolding technique since this removes the regularization of the stabilization term in expression 6.4.25. Instead the aim is to find a solution for an arbirary value of τ and then establish the best value to use.

The first step is to make sure both terms deal with $\vec{w}$ in the same way by multiplying by CC^{-1}:

$$\left(\tilde{R}C^{-1}\left(C\vec{w}\right) - \tilde{b}\right)^{T}\left(\tilde{R}C^{-1}\left(C\vec{w}\right) - \tilde{b}\right) + \tau \cdot \left(C\vec{w}\right)^{T} C\vec{w} \tag{6.4.36}$$

For $\tau = 0$ this is tantamount to solving:

$$\tilde{R}C^{-1}\left(C\vec{w}\right) = \tilde{b} \tag{6.4.37}$$

Which can be solved by using SVD again:

$$\tilde{R}C^{-1} = U\,S\,V^{T} \tag{6.4.38}$$

where the same convention as always holds for S; its diagonal elements are ordered in a descending sequence. Inserting this into equation 6.4.37 we obtain:

$$U\,S\,V^{T}\left(C\vec{w}\right) = \tilde{b} \tag{6.4.39}$$

and by changing the order:

$$S\,V^{T}\left(C\vec{w}\right) \quad = \quad U^{T}\tilde{b} \tag{6.4.40}$$

One can now define two new vectors:

$$\vec{d} \quad \equiv \quad U^{T}\tilde{b} \tag{6.4.41}$$

$$\vec{z} \quad \equiv \quad V^{T}C\vec{w} \tag{6.4.42}$$

This simplifies equation 6.4.40 to:

$$S\vec{z} = \vec{d} \tag{6.4.43}$$

[†]One might wonder that if it is this easy to make a matrix inversable, why did we not apply that to $\hat{A}$? The reason is that unlike C the matrix $\hat{A}$ is not fixed and has statistical components. The ξ parameter will enhance these components of $\hat{A}$ and do nothing to solve the underlying problem to get rid of these.

[‡]If you carefully keep track and do all calculations analytically you will never need C^{-1}, but we need a numerically stable solution instead and for this ξ works perfectly.

which is a nicely diagonal system (S is by construction diagonal), and the solution to find $\vec{w}$ basically becomes a search for $\vec{z}$ which is trivial to find due to the before-mentioned diagonality of S:

$$z_i = \frac{d_i}{S_{ii}} \tag{6.4.44}$$

which via equation 6.4.42 means $\vec{w}$ can be written as:

$$\vec{w} = C^{-1} V \vec{z} \tag{6.4.45}$$

At this point one may wonder why we went through this seemingly complicated maneuver, could we not have inverted $\tilde{R}$ like before? The answer is yes, but there is a crucial difference; as is detailed in [90] the solution for a general τ can now simply be found by rescaling $\vec{d}$:

$$d_i^{(\tau)} \equiv d_i \frac{S_{ii}^2}{S_{ii}^2 + \tau} \tag{6.4.46}$$

From this the modified $\vec{w}$ can then be found:

$$\vec{w}^{(\tau)} = C^{-1} V z^{(\tau)} \tag{6.4.47}$$

$$z_i^{(\tau)} = \frac{d_i S_{ii}}{S_{ii}^2 + \tau} \tag{6.4.48}$$

which forms the solution to our unfolding problem with an appropriately chosen τ, since from $\vec{w}$ the needed $\vec{x}$ can be retrieved using expression 6.4.9. The covariance matrix corresponding to $\vec{w}$ can also be readily calculated as can be found in [88].

The remaining issue now is how to choose τ. The authors of [88] mention that a lot of information can be gleaned by scrutinizing the distribution[†] of $\log|d_i|$ versus i. To see why we investigate what $\vec{d}$ actually means. It is defined as:

$$\vec{d} \equiv U^T \tilde{b} \tag{6.4.49}$$

Where we can substitute in expression 6.4.23:

$$\vec{d} = U^T \Sigma^{-\frac{1}{2}} Q^T \vec{b} \tag{6.4.50}$$

We also know from expression 6.4.38 that:

$$\tilde{R} C^{-1} = U\, S\, V^T \tag{6.4.51}$$

From which we can extract U^T:

$$U^T \tilde{R} C^{-1} = S\, V^T \tag{6.4.52}$$

$$U^T = S\, V^T C\, \tilde{R}^{-1} \tag{6.4.53}$$

From expression 6.4.22 we can find an expression for $\tilde{R}^{-1}$:

$$\tilde{R}^{-1} = \hat{R}^{-1} Q\, \Sigma^{\frac{1}{2}} \tag{6.4.54}$$

Which substituting back into equation 6.4.53 gives:

$$U^T = S\, V^T C\, \hat{R}^{-1} Q\, \Sigma^{\frac{1}{2}} \tag{6.4.55}$$

[†]Note that this is the unregulated $\vec{d}$ and not the regulated $\vec{d}^{(\tau)}$.

Putting this into equation 6.4.50:

$$\vec{d} = S\,V^T C\,\hat{R}^{-1} Q\,\Sigma^{\frac{1}{2}}\Sigma^{-\frac{1}{2}} Q^T \vec{b} \tag{6.4.56}$$

Which we can simplify to:

$$\vec{d} = S\,V^T C\,\hat{R}^{-1}\vec{b} \tag{6.4.57}$$

And finally using expression 6.4.12 we find:

$$\vec{d} = S\,V^T C\,\vec{w} \tag{6.4.58}$$

To interpret this, recall that we know from our discussion when C was introduced that $C\,\vec{w}$ has the form given by expression 6.4.32. It is a a vector that gives a measure of the amount of bin-to-bin fluctuation around each bin. If the detector is perfect, and the simulation used to generate $\hat{R}$ agrees perfectly with the measured results than this vector has only elements that are zero. If this is not the case, but the detector is still pretty decent and the simulation well-chosen, the elements are still expected to be distributed around zero.

The forms of S and V^T are dependent upon $\hat{R}$, but in essence they enhance the elements of $C\,\vec{w}$ for which the error on the bin is smallest, and for which elements the corresponding element in $\hat{R}$ is largest. And since the elements are ordered (by the rotation matrices U and V) in a descending order, the magnitudes of last few terms of $\vec{d}$ are mostly statistical fluctuations around zero with a variance of one[†]. This all depends on $\vec{w}$ being smooth, which incidentally, is a reason for choosing to move to using $\vec{w}$ over $\vec{x}$.

Thus, from at $\log|d_i|$ versus i we can examine the extend to which we are dominated by statistical fluctuations. The authors of [88] recommend choosing τ by determining at which point the distribution of $\log|d_i|$ versus i transitions from significant to statistics dominated. If this should be at bin j then τ should be chosen as the square of the j'th diagonal element of $\tilde{R}C^{-1}$:

$$\tau = r_{jj}^2,\; r_{jj} \in \tilde{R}C^{-1} \tag{6.4.59}$$

This number j shall be called the unfolding rank. We shall see how we apply this in our analysis in the next section.

6.4.2 Application to the analysis

The first step to apply the singular value unfolding to our analysis is to determine the detector response matrices, $\tilde{R}$. The most straightforward way of doing this is to take a simulated sample of $b\bar{b}\gamma\gamma$ events within the fiducial volume we are focusing on, and put this sample through the analysis chain, recording for each observable and each event what the truth value was and where (and if) it ends up according to the analysis. This then naturally yields $\vec{R}$ and $\vec{x}_{ini}$. If an event is rejected by the analysis chain it does not end up in $\vec{R}$ but it does in $\vec{x}_{ini}$ which then accounts for any less than perfect acceptance of the analysis cuts, detector reconstruction inefficiencies, as well as trigger inefficiencies. Unfortunately such a sample at 8 TeV was not available and at the time of this analysis it was impossible to request. This was partly because the sample would need to undergo a full simulation of the ATLAS detector and reconstruction software

[†]This is because of the rescaling of $\vec{b}$ by the error on its elements, Δb_i.

in order to obtain a correct detector response matrix, which is a very computationally intensive task[†].

Our signal samples are inclusive and do not contain a large number of genuine $b\bar{b}\gamma\gamma$ events. However, all is not lost; our signal samples do have a large enough sample of $\gamma\gamma jj$ events and so we will unfold the cutflow in the pretag stage and then derive the acceptance of the final b-tags in a different way. We are bolstered in this approach by the conclusion reached in section 2.5; namely that the truth distributions for $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$ are very similar.

We start by deriving the detector response matrices in the pretag stage. Then we input the distributions obtained in section 6.3.2 into the unfolding process as though they were $\gamma\gamma jj$ events. This means the unfolding does not take into account any acceptance of the b-tagging cuts, but this is already taken care of in the procedure described in section 6.3 because this is essentially the b-tagging efficiency. It also means we do not take into account detector response related effects on the shapes of various observables that might be different between b-jets and other jets. This is unfortunately unavoidable, and to quantify the severity of these effects we studied the differences between unfolding on $\gamma\gamma jj$ and $\gamma\gamma bj$, because we do have a reasonable amount of $\gamma\gamma bj$ events. By $\gamma\gamma bj$ we mean the single tag stage where at least one of the two leading jets is b-tagged. A discussion of this can be found further down.

The truth selection needs mimic our fiducial volume as well as possible. Therefore, we apply the fiducial volume defined in section 2.5 as cuts at the truth level on the simulation samples. To make sure our events are prompt $jj\gamma\gamma$ events we also require that the two leading photons originate from the same vertex. The two jets are required to come from within 25 mm of this vertex. This distance was chosen because an investigation showed that this distance was enough to capture all spurious $jj\gamma\gamma$ events whilst rejecting virtually all pile-up in the simulation. We varied this distance as a systematic as mentioned below.

First we analyze the differences between our two signal samples. We examine the differences between PYTHIA and SHERPA in figure 6.4.1. These show how well events from both samples (required to pass the truth cuts mentioned above) do when subjected to the full analysis cutflow. We also find that the total acceptance of PYTHIA is much smaller than SHERPA being 0.026 compared to 0.13. The causes for this can be glimpsed from figure 6.4.1; PYTHIA performs worse in passing the preselection requirements, but also in the photon isolation, having the photons come from the most energetic primary vertex (goodPVindex), and in reconstructing the jets. SHERPA also loses acceptance at all these points, but PYTHIA loses more.

PYTHIA also performs significantly worse than SHERPA and the MADGRAPH5 NLO predictions (the latter we discussed in section 2.5) in the description of variables depending on the angular distribution of jets. This is most evident in variables directly related to the jet η as can be seen in figure 6.4.2. In the left plot in this figure we see that PYTHIA underestimates central jets and overestimates jets with a high η. Good jets need to have $|\eta| < 2.5$ so if one of the two leading jets fails the entire event fails. PYTHIA contains more such events as we see in the plot on the right side of figure 6.4.2, and this leads

[†]Faster simulations do exist but in order to achieve this they cut corners with the simulation of the detector. Whilst useful for other purposes, this fast sample can not be used for detector unfolding since the detector description used in generating it is not optimal.

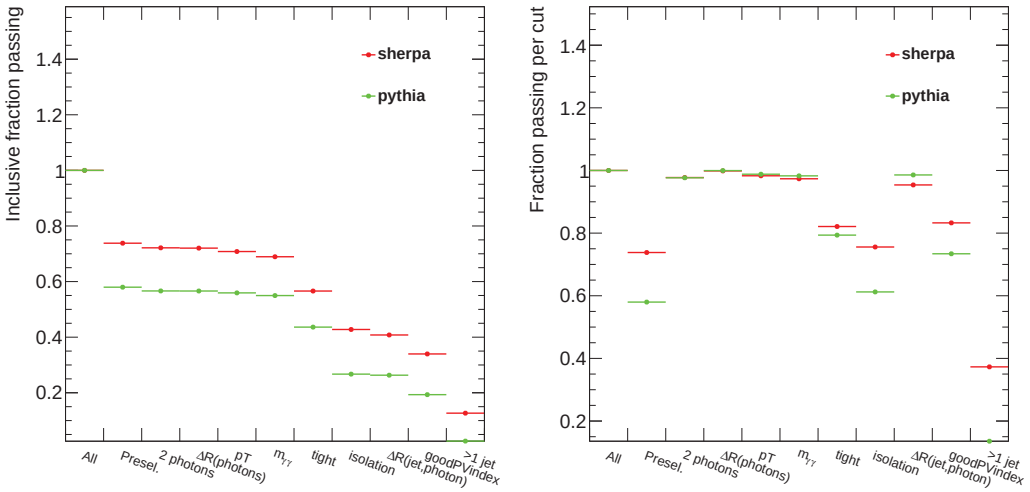

Figure 6.4.1: *Cutflow for simulation events passing the truth selection. On the left the inclusive numbers, on the right the numbers per bin (i.e. each bin denotes which fraction of events passing the previous cut remain after the current one).*

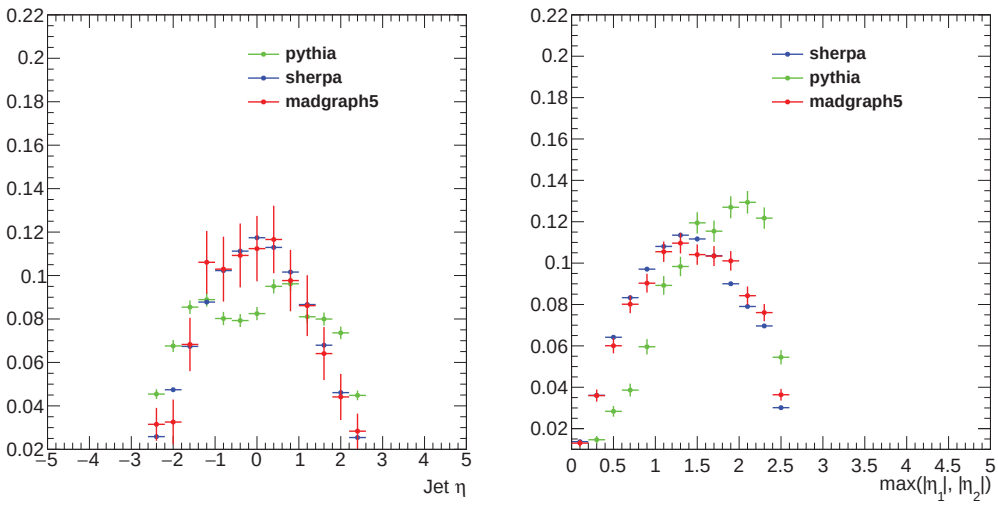

Figure 6.4.2: *Normalized comparison of the two full simulation samples,* PYTHIA *and* SHERPA*, and the theory predictions obtained using* MADGRAPH5*. On the left the pseudorapidity of the two leading jets. On the right the maximum of the absolute pseudorapidity of the two leading jets in the event.*

to the terrible acceptance of PYTHIA since we can observe in figure 6.4.1 that a lot of events fail the jet requirement.

We trace these differences to the fact that PYTHIA only contains $2 \to 1$ and $2 \to 2$ processes. This means that $pp \to jj\gamma\gamma$ (or $pp \to b\bar{b}\gamma\gamma$ for that matter) is a $2 \to 4$ process, so technically the PYTHIA sample does not contain our signal at all. The $jj\gamma\gamma$ events inside the sample contain mostly photons radiated off jets, jets coming from initial and final state radiation, or there are two separate pp interactions in the event creating the combined $jj\gamma\gamma$ system. The SHERPA sample does, however, contain $2 \to 4$ processes. As such we will only consider SHERPA in our unfolding process. We investigated using PYTHIA as a systematic variation but the low acceptance lead to huge systematic uncertainties. Because PYTHIA does not describe the signal well we left it out completely. Unfortunately this leaves us without a way to estimate unfolding uncertainties originating from the quality of the simulation.

The detector response matrices for our observables can be found in figures 6.4.3 to 6.4.5.
As can be seen in those figures the matrices are largely diagonal which means the reconstruction is competent. It can also be observed that the response matrices of observables which depend on more than one kinematic variable (such as the angular separations, which depend both on the p_T and η of two different objects) tend to be less diagonal than those based on only one such variable. This is expected; there is a greater opportunity for something to go wrong if an observable depends on the correct reconstruction of multiple aspects of an event as opposed to only one.

In figures 6.4.6 to 6.4.8 we show the $\vec{x}_{ini}$ and $\vec{b}_{ini}$ distributions used. The $\vec{x}_{ini}$ distribution describes all simulation events passing the truth selection. $\vec{b}_{ini}$ is the distribution after the reconstruction and analysis cuts. In these plots we can therefore clearly see how the reconstruction and analysis affect the original distributions. Major differences can be seen at low end of the p_T spectrum for the photons and jets; the low end is affected by the p_T cuts applied in the analysis. A large difference can also be seen in $m_{jj\gamma\gamma}$ where events at the low end fail the analysis cuts.

Note that the binning there is much finer than what we will actually use in the end; this is done mostly to convey more information to the reader, we are not actually sensitive to much of the finer details in these distributions. The distributions have also been normalized to emphasize the shape differences. Also included are the NLO theory predictions obtained using MADGRAPH5 from section 2.5. These closely follow the $\vec{b}_{ini}$ distributions, showing that the analysis cuts correctly select $b\bar{b}\gamma\gamma$ events. The error bars on the theory predictions have been dropped in the plots to increase clarity.

Unlike the simulation events used in the rest of the analysis the events are not corrected for various effects mentioned in section 5.6. Because otherwise the unfolding would become sensitive to these affects, and we would be essentially elevating simulation inaccuracies to the status of new physics. This is because both data and simulation go through the same reconstruction cycle, so any corrections made to the simulations are intended to mend problems in the actual truth generation of these simulations. These corrections are not physical, and so they should not enter the response matrices. In any case it is prudent to mention that the overall normalization of the simulation sample plays no role in the unfolding.

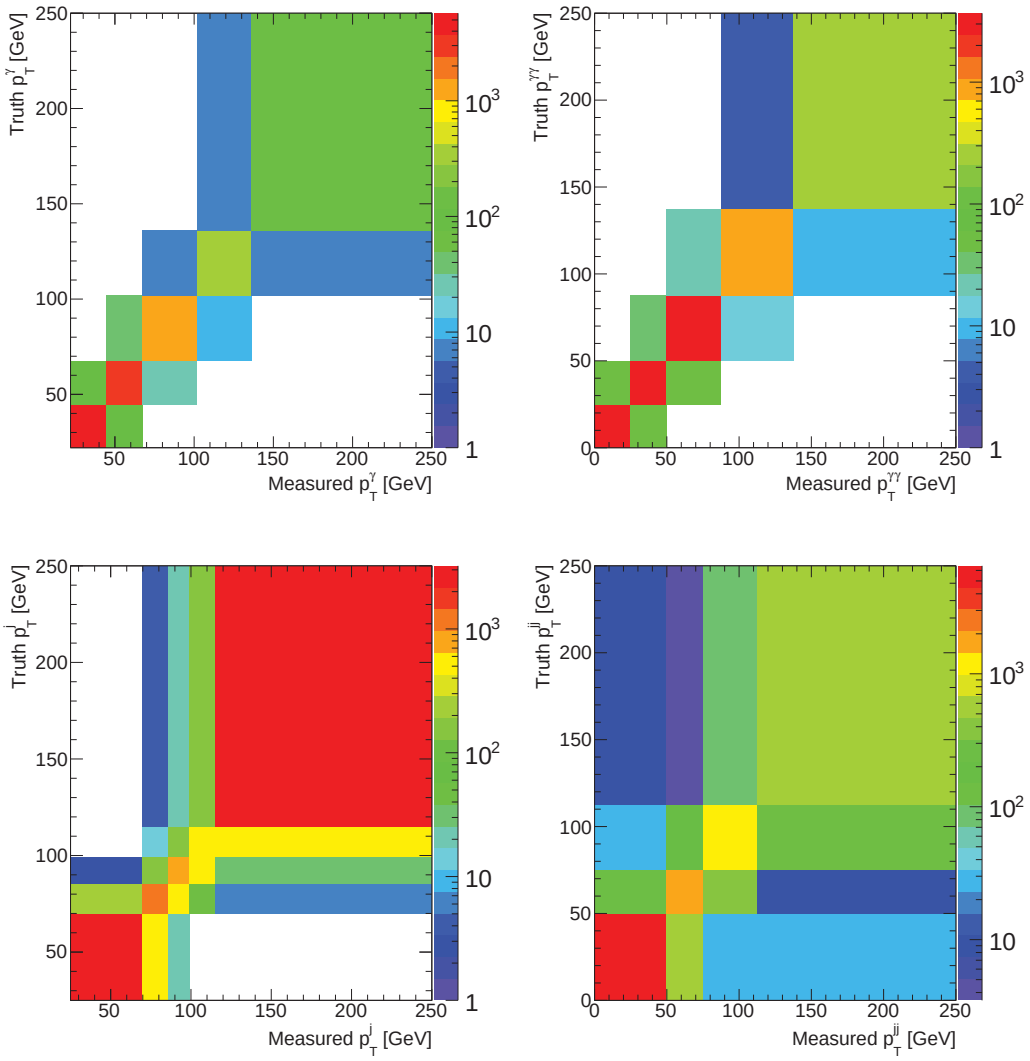

Figure 6.4.3: *The detector response matrices of the transverse momentum of the leading photon (top left), leading jet (bottom left), leading photon pair (top right), and leading jet pair (bottom right).*

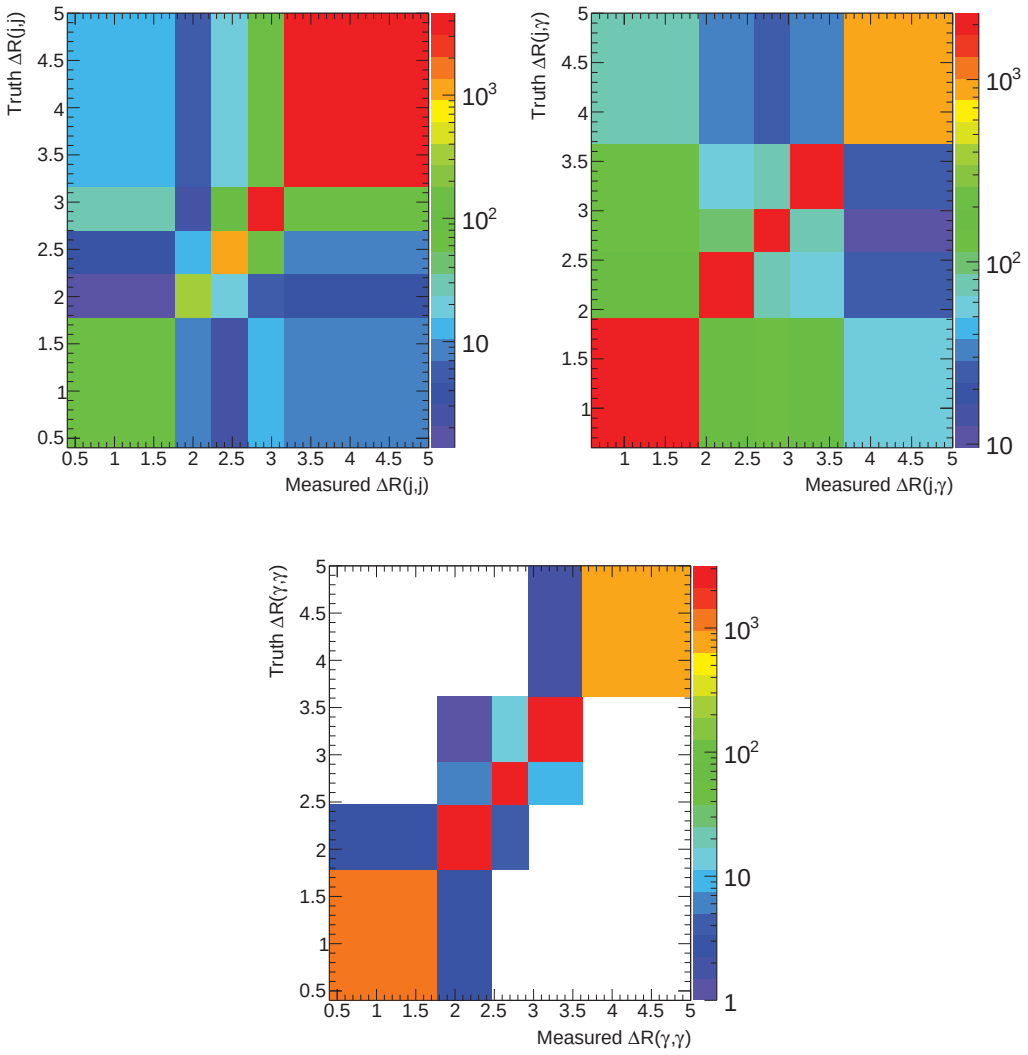

Figure 6.4.4: *The detector response matrices of the angular separation between the two leading jet (top left), leading jet and leading photon (top right), and the two leading photons (bottom).*

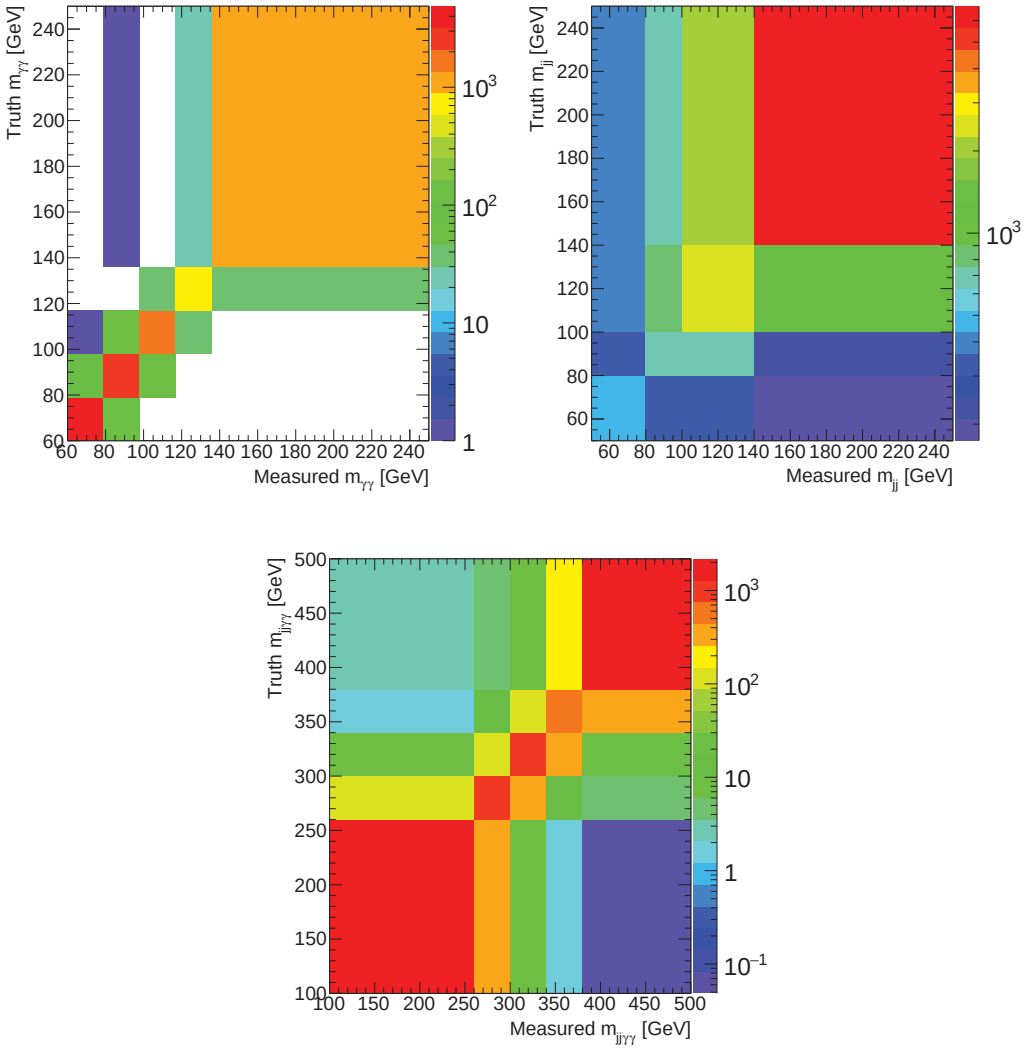

Figure 6.4.5: *The detector response matrix of the invariant mass of the leading photon pair (top left), leading jet pair (top right), and of the leading two jets and the leading two photons (bottom).*

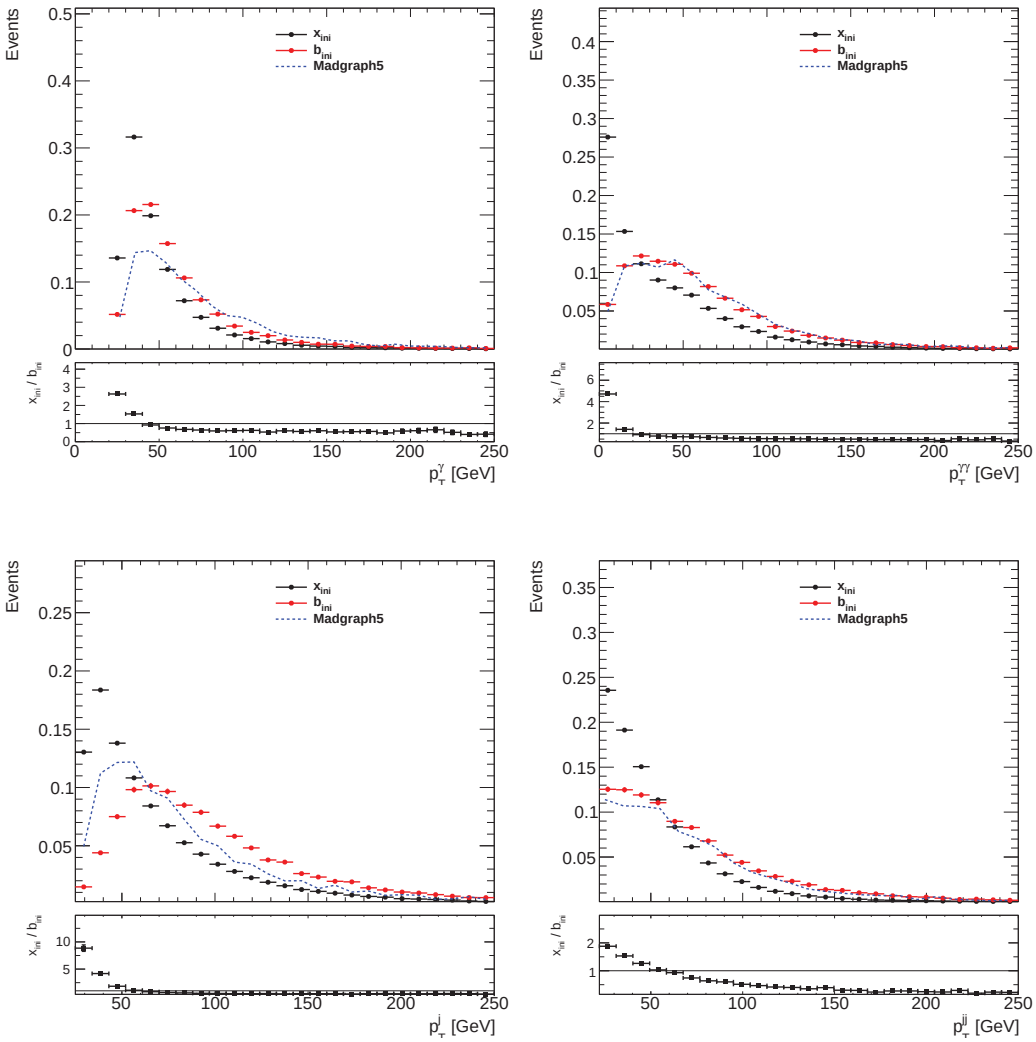

Figure 6.4.6: *Normalized distributions of $\vec{x}_{ini}$ and $\vec{b}_{ini}$ for the transverse momenta p_T^{γ} (top left) and $p_T^{\gamma\gamma}$ (top right), p_T^{j} (bottom left), and p_T^{jj} (bottom right). Also included are the* MADGRAPH5 *NLO predictions.*

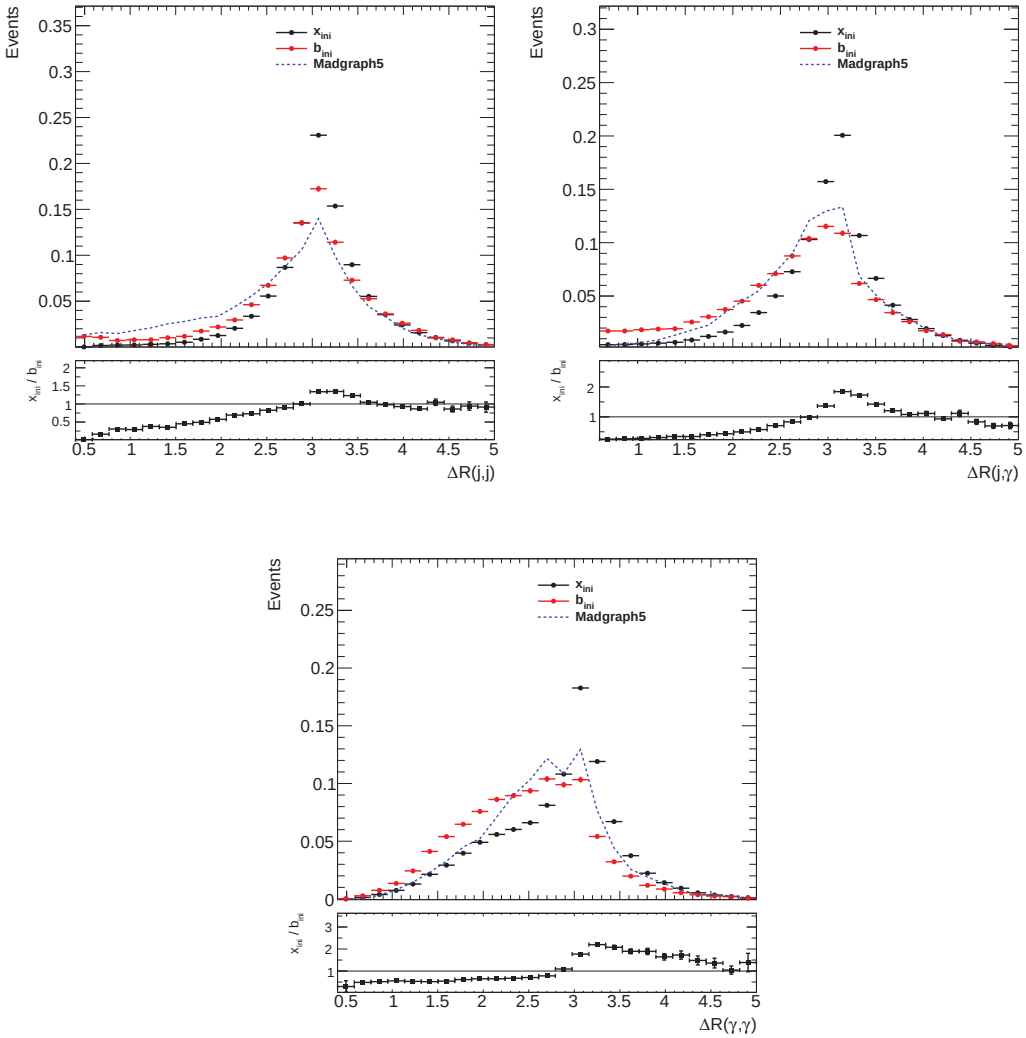

Figure 6.4.7: *Normalized distributions of $\vec{x}_{ini}$ and $\vec{b}_{ini}$ for the angular separations $\Delta R(j,j)$ (top left) and $\Delta R(j,\gamma)$ (top right), and $\Delta R(\gamma,\gamma)$ (bottom). Also included are the* MADGRAPH5 *NLO predictions.*

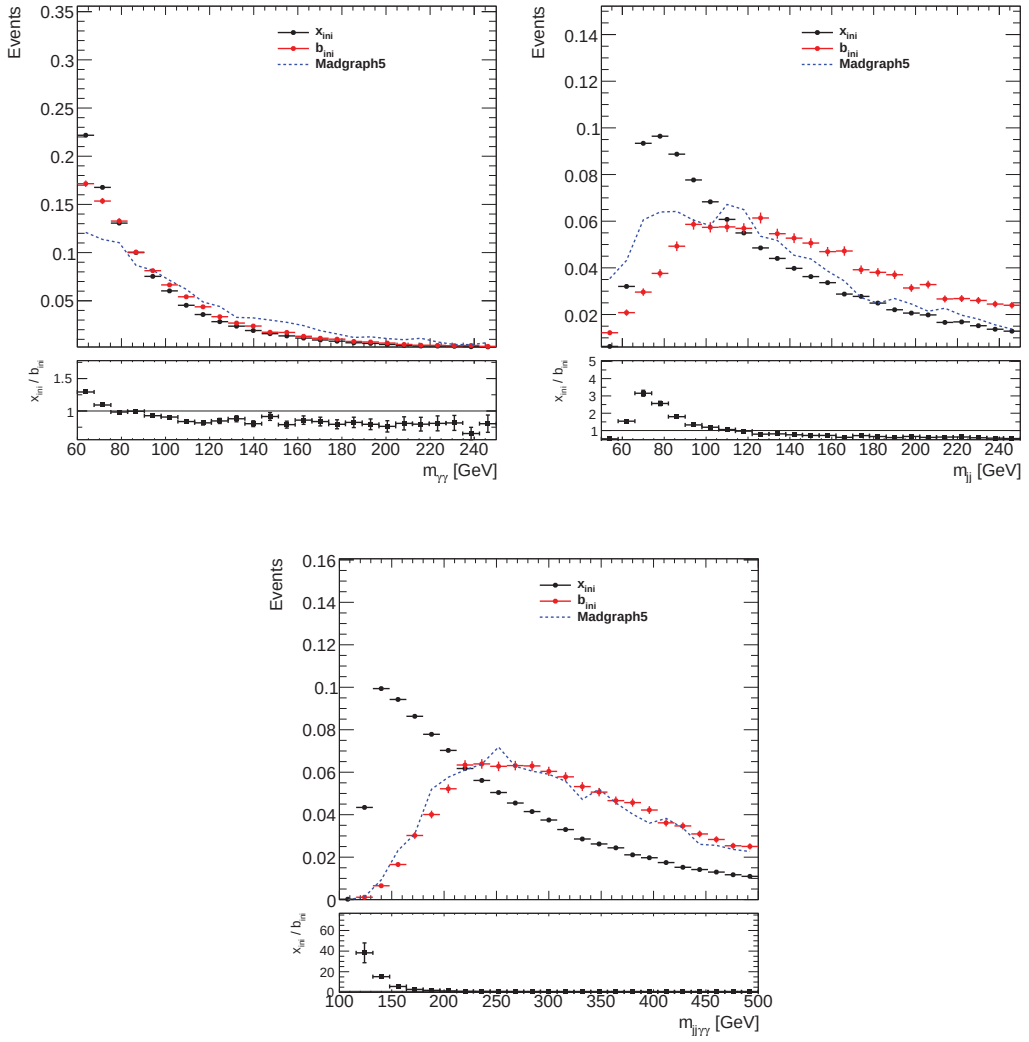

Figure 6.4.8: *Normalized distributions of $\vec{x}_{ini}$ and $\vec{b}_{ini}$ for the invariant masses $m_{\gamma\gamma}$ (top left) and m_{jj} (top right), and $m_{jj\gamma\gamma}$ (bottom). Also included are the* MADGRAPH5 *NLO predictions.*

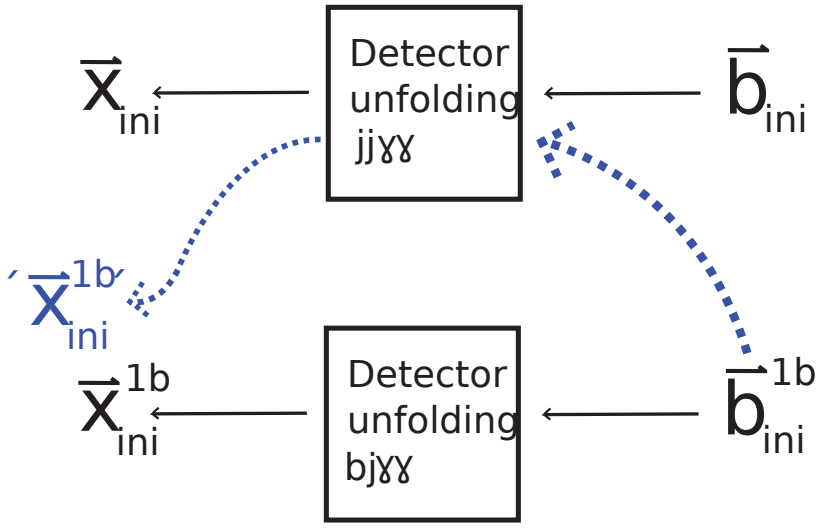

Figure 6.4.9: *Graphical description of how '$\vec{x}^{1b}_{ini}$' is defined. It is obtained by unfolding $\vec{b}^{1b}_{ini}$ using the response matrix obtained from $pp \to jj\gamma\gamma$ events. If the $jj\gamma\gamma$ response matrix is very similar to the $bj\gamma\gamma$ response matrix the resultant $\vec{x}$ should be very close to the $\vec{x}^{1b}_{ini}$ used in the creation of $\vec{b}^{1b}_{ini}$ because by construction if one would use the $bj\gamma\gamma$ response matrix to unfold $\vec{b}^{1b}_{ini}$ we would get $\vec{x}^{1b}_{ini}$.*

As mentioned before the effects of unfolding on $\gamma\gamma jj$ instead of $\gamma\gamma bb$ was assessed by examining $\gamma\gamma bj$ unfolding and comparing it to $\gamma\gamma jj$. The response matrices were regenerated adding to the truth selection the requirement of at least a single b-jet, and requiring that the event also passes the analysis cutflow up and including the single b-tag stage. If we had a proper $b\bar{b}\gamma\gamma$ signal sample we could create the response matrices using the full selection; requiring two b-jets in the truth selection and also two b-tagged jets in the analysis cuts. Unfortunately we do not have the statistics to do this in the simulation samples. But we can assess the effects of the b-tags by comparing $\gamma\gamma jj$ to $\gamma\gamma jb$.

To do so we take $\vec{b}^{1b}_{ini}$, which is $\vec{b}_{ini}$ obtained from the generation of the response matrix for $\gamma\gamma jb$, and unfold it using the response matrix for $\gamma\gamma jj$[†]. The '$\vec{x}^{1b}_{ini}$' thus obtained can be compared with the original $\vec{x}^{1b}_{ini}$ used to generate the response matrix for $\gamma\gamma jb$. Basically this tests how good the $\gamma\gamma jj$ response matrix works with $\gamma\gamma jb$ data. A schematic description of this method is found in figure 6.4.9. The results of this comparison for our observables can be found in figures 6.4.10 to 6.4.12. The distributions have been normalized because the response matrices have different normalizations due to a difference in acceptance[‡].

It can be seen in the figures that within the statistical uncertainties we have an agreement. This increases our confidence in using $\gamma\gamma jj$ unfolding, especially considering the coarse binning we have. Nonetheless it is prudent to add a systematic to capture the uncertainty inherent to this method. To do so we take the shape ratios (i.e. the ratios found below each plot) between $\vec{x}^{1b}_{ini}$ and '$\vec{x}^{1b}_{ini}$' and vary the nominal result we get from

[†]Unfolding it using the response matrix for $\gamma\gamma jb$ will of course trivially yield the $\vec{x}_{ini}$ used to create that response matrix, $\vec{x}^{1b}_{ini}$.

[‡]This acceptance is taken into account in the nominal analysis via the b-tagging efficiencies.

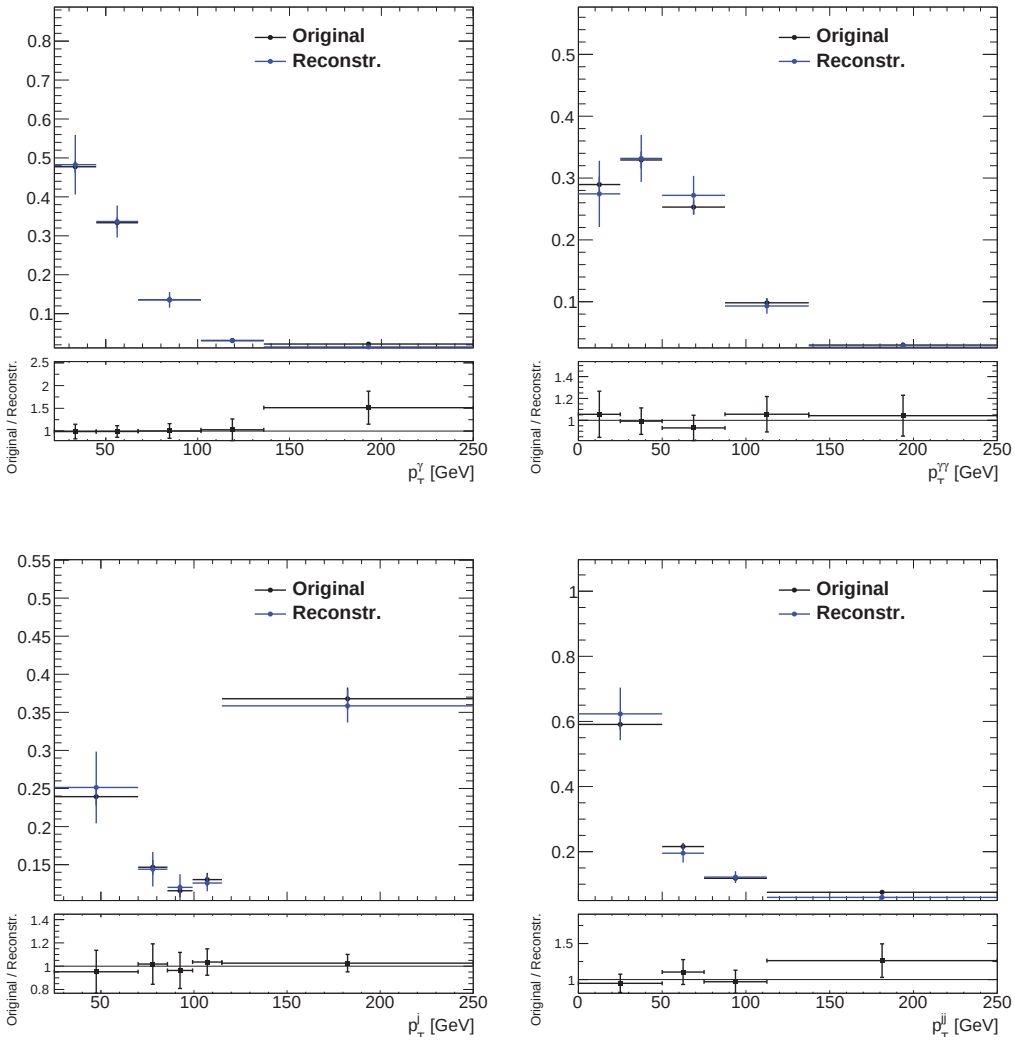

Figure 6.4.10: *Comparison of $\gamma\gamma jj$-based unfolding and $\gamma\gamma jb$-based unfolding for p_T^{γ} (top left), $p_T^{\gamma\gamma}$ (top right), p_T^{j} (bottom left), and p_T^{jj} (bottom right). We compare $\vec{x}_{ini}^{1b}$ (called original) from $\gamma\gamma jb$ with '$\vec{x}_{ini}^{1b}$' (called reconstructed) which was obtained by unfolding $\vec{b}_{ini}^{1b}$ using $\gamma\gamma jj$.*

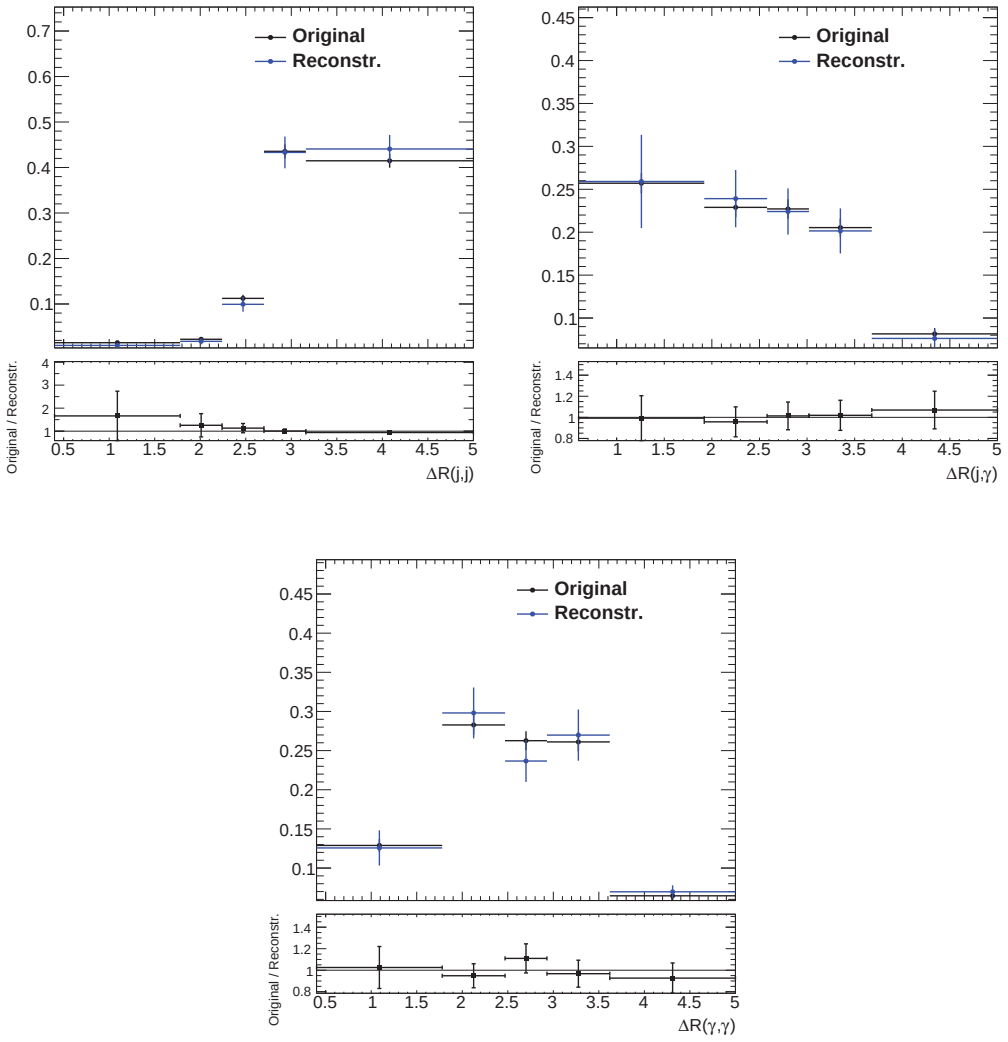

Figure 6.4.11: *Comparison of $\gamma\gamma jj$-based unfolding and $\gamma\gamma jb$-based unfolding for $\Delta R(j,j)$ (top left), $\Delta R(j,\gamma)$ (top right), and $\Delta R(\gamma,\gamma)$ (bottom). We compare $\vec{x}^{1b}_{ini}$ (called original) from $\gamma\gamma jb$ with '$\vec{x}^{1b}_{ini}$' (called reconstructed) which was obtained by unfolding $\vec{b}^{1b}_{ini}$ using $\gamma\gamma jj$.*

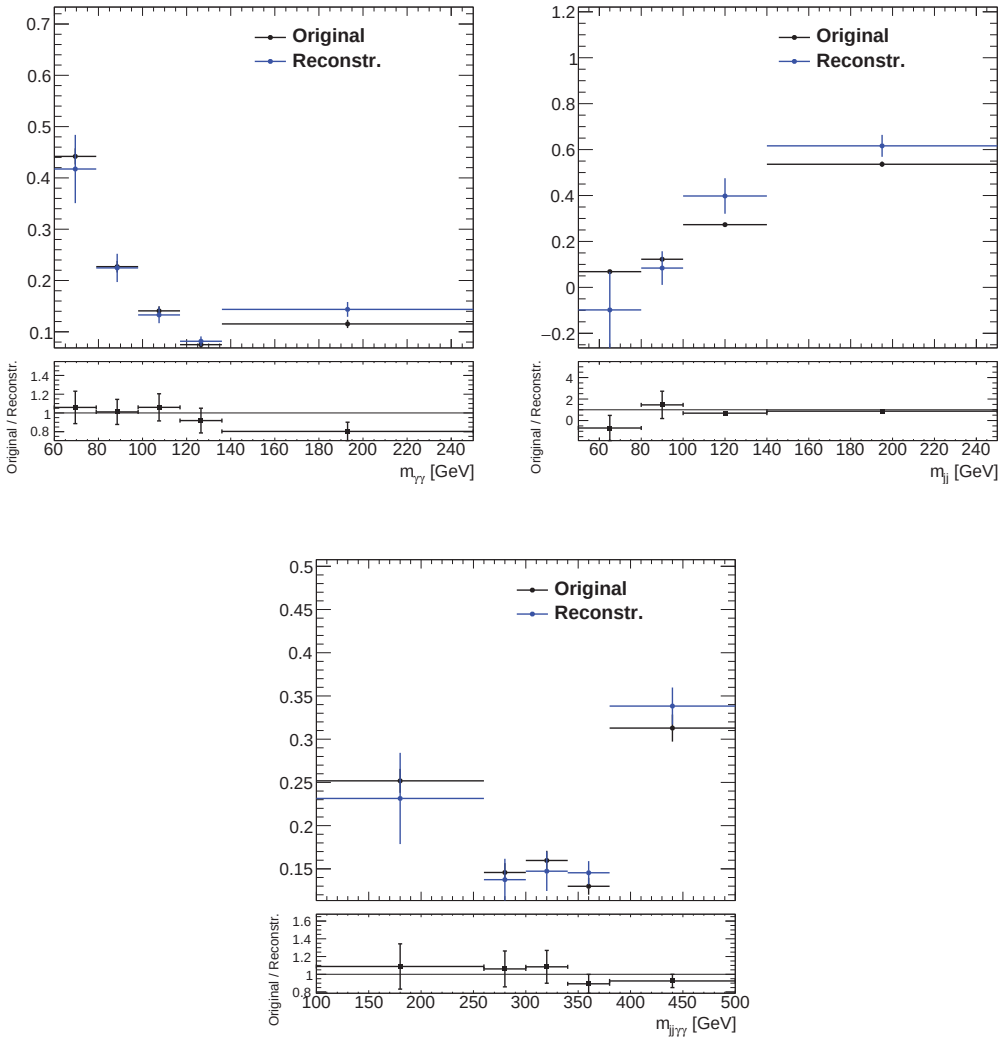

Figure 6.4.12: *Comparison of $\gamma\gamma jj$-based unfolding and $\gamma\gamma jb$-based unfolding for $m_{\gamma\gamma}$ (top left), m_{jj} (top right), and $m_{jj\gamma\gamma}$ (bottom). We compare $\vec{x}^{1b}_{ini}$ (called original) from $\gamma\gamma jb$ with '$\vec{x}^{1b}_{ini}$' (called reconstructed) which was obtained by unfolding $\vec{b}^{1b}_{ini}$ using $\gamma\gamma jj$.*

the unfolding up and down with this factor as a systematic uncertainty.

The value chosen for the bin number of $\vec{d}$ (i.e. the unfolding rank) to determine τ for each of the observables was three. This was also the unfolding rank used in the investigation of using $\gamma\gamma jb$ response matrices mentioned before. The distributions of $\vec{d}$ in terms of our observables were found to be very similar to each other leading us to the equal unfolding rank. A breakdown of the systematics inclusively for all events can be found in table 6.4.1. After the unfolding we estimate there to be approximately 540 $b\bar{b}\gamma\gamma$ events in the 2012 ATLAS dataset. The statistical uncertainties have been propagated through the unfolding procedure and also contain any statistical uncertainties contained in the response matrices. For the unfolding procedure the following systematics were considered:

Jet truth vertex distance In the truth selection the jets were accepted only if they were within 25 mm of the diphoton vertex. The sensitivity to this choice was assessed by removing the requirement for the up variation, and lowering it to 10 mm for the down variation. It also assesses the influence of pile-up in the creation of the detector response matrix.

Unfolding rank The effects of a particular choice for the unfolding rank are assessed by moving it up and down by one and quoting the difference with the nominal case. This systematic was not applied to the inclusive result since we are dealing with only a single bin in that case.

$jj\gamma\gamma$ **unfolding** The unfolding results were varied up and down by multiplying with the ratios found when comparing the $\gamma\gamma jb$ and $\gamma\gamma jj$ response matrices as mentioned before.

The systematics considered in other parts of the analysis have been propagated to this stage and are also found in table 6.4.1. The b-tagging MV1 template uncertainty is the largest systematic.

It can be noted that the number of events found in the various bins of our observables do not match the total found in the inclusive tally. This is because each bin is fitted separately which causes a bin-to-bin variation which in turn causes the mismatch with the total. The uncertainties are smaller in the inclusive case and therefore the total count is more trustworthy there. The final distributions of our observables converted to differential cross-sections can be found in the next section.

Events after unfolding	
Nominal value	539.8 ± 100.4
(unfolding) Jet truth vertex distance up [%]	2.89
(unfolding) Jet truth vertex distance down [%]	-0.96
(non-b-tag) Template uncertainties up [%]	-9.47
(non-b-tag) Template uncertainties down [%]	4.87
(electron) $f_{e\to\gamma}$ sideband shift up [%]	-0.99
(electron) $f_{e\to\gamma}$ sideband shift down [%]	1.86
(jet) Using E_T^{iso} for ϵ_I[%]	± 4.97
(jet) Using PYTHIA for ϵ_T[%]	± 0.59
(jet) ϵ_{I1}background boundary down [%]	-0.58
(jet) ϵ_{I2}background boundary up [%]	0.62
(jet) ϵ_{I2}background boundary down [%]	-0.82
(jet) ϕ_1 up [%]	-1.54
(jet) ϕ_1 down [%]	1.23
(jet) ϕ_2 up [%]	-2.16
(jet) ϕ_2 down [%]	1.68
(general) JVF cut up [%]	1.20
(general) JVF cut down [%]	1.65
Total systematic error [%]	± 11.74

Table 6.4.1: *Inclusive number of events after the unfolding procedure. The total systematic uncertainty was calculated by taking maximum of the up/down values of each systematic and adding them in quadrature. The word in parentheses is to remind the reader of the general source of the uncertainty. Only systematics with an effect of 0.5% or more are mentioned.*

6.5 Results

THE event rates obtained by the analysis can be converted into a cross-section determination via by using the total integrated luminosity that the used data represents. The integrated luminosity for the 2012 8 TeV proton-proton dataset useful for physics analyses is estimated in a recent ATLAS publication [54, 55] to be 20.24 fb^{-1} with an uncertainty of 1.9 %.

Using this integrated luminosity we can now present the differential cross-sections for the various observables and compare them with the theoretical calculations obtained in section 2.5. For the NLO predictions factorization and renormalization scale uncertainties were included by halving or doubling the scales set automatically by MADGRAPH5. In addition effects of the choice of PDF were assessed by using CTEQ6L [37] PDFs instead of the nominal NN23NLO [49] ones and symmetrizing the differences as a systematic. The scale uncertainty was calculated for the LO predictions in a similar way but the scales were varied separately. The PDF uncertainty was obtained by switching the nominal CT10 [48] PDFs for CTEQ6L.

Within the analysis' fiducial volume we find that the inclusive $pp \to jj\gamma\gamma$ production cross-section at $\sqrt{s} = 8$ TeV is:

$$5.43 \pm 0.036\,(\text{stat}) \pm 1.20\,(\text{syst})\,\text{pb}$$

The LO theory prediction is:

$$2.53 \pm 0.058\,(\text{stat}) \pm 0.37\,(\text{syst})\,\text{pb}$$

and for the NLO theory prediction we find:

$$3.43 \pm 0.052\,(\text{stat}) \pm 0.76\,(\text{syst})\,\text{pb}$$

By far the largest systematic uncertainty on the measurement originates from the JVF cut value which was varied up and down. Most jets fall into the low p_T range and therefore the JVF requirement of a good jet has a large effect on the acceptance of the analysis. Another reason for the heightened sensitivity to the JVF cut value is that this cut enforces low p_T jets to originate from the same primary vertex as the two leading photons. This is a rather strict requirement. We find that the analysis results agree better with the NLO predictions than the LO ones, but not within 1σ.

Within the fiducial volume we find that the inclusive $pp \to b\bar{b}\gamma\gamma$ production cross-section at $\sqrt{s} = 8$ TeV is:

$$0.0267 \pm 0.00496\,(\text{stat}) \pm 0.0032\,(\text{syst})\,\text{pb}$$

The theory predicts the following LO cross-section:

$$0.0138 \pm 3.04 \cdot 10^{-5}\,(\text{stat}) \pm 0.0022\,(\text{syst})\,\text{pb}$$

and also a NLO cross-section of:

$$0.0347 \pm 0.00124\,(\text{stat}) \pm 0.0092\,(\text{syst})\,\text{pb}$$

The measurement result falls in between the LO and NLO theory predictions. As mentioned in section 2.5 there are large k-factors between the LO and NLO theory cross-sections so this result is reasonable. The analysis agrees within 1σ with the NLO prediction. For $b\bar{b}\gamma\gamma$ the JVF cut variation systematic has only a small effect. This is because of the additional b-tagging requirements which enforce the JVF cut on their own. The dominating systematic here is the uncertainty in the MV1 templates used to estimate the flavor fractions. This uncertainty enters the analysis as a direct multiplication so it has large influence on the final result. Nonetheless the statistical uncertainty is larger than the systematic one.

All relevant systematics have been propagated fully throughout the analysis and a list with the largest ones can be found in table 6.5.1 for data and table 6.5.2 for simulation. The same tables also contain the numbers for $pp \to jj\gamma\gamma$. Plots describing the differential cross-section with respect to the observables can be found in figures 6.5.1 to 6.5.5. Because the background subtractions were performed bin-by-bin there are some fluctations bin to bin. Due to the same reason the integrals over these plots do not agree directly with the inclusive result.

In terms of shape the $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$ distributions agree well. They also mostly agree with the theory predictions apart from a normalization factor. In some cases the LO and NLO theory predictions diverge and we find that here the data follows the NLO predictions. Note for instance the first bin of $\Delta R(j,j)$ in the top left of figure 6.5.3 where the LO prediction overestimates the data. In the case of $pp \to jj\gamma\gamma$ we find that low end of the m_{jj} spectrum is not well described by the theory. Our analysis was, however, not optimized for the $jj\gamma\gamma$ channel directly; notably the cut requiring the diphotons to originate from the most energetic primary vertex might lead to this discrepancy since this aspect is not modelled by the MADGRAPH5 simulations.

In some bins the number of events was too small to assess all systematics properly inside that bin. In these cases the relevant systematic uncertainty on the nominal result was used instead. This affects mainly the MV1 template systematic. In the case of m_{jj} the $\gamma\gamma jb$ versus $\gamma\gamma jj$ systematic was found to inflate severely. This has been hidden from the picture to maintain clarity. This effect can be seen in the top right comparison plot in figure 6.4.12. The large statistical uncertainty on the bins of m_{jj} can be similarly traced to the unfolding procedure. Unfortunately there was no time left to analyze this issue further.

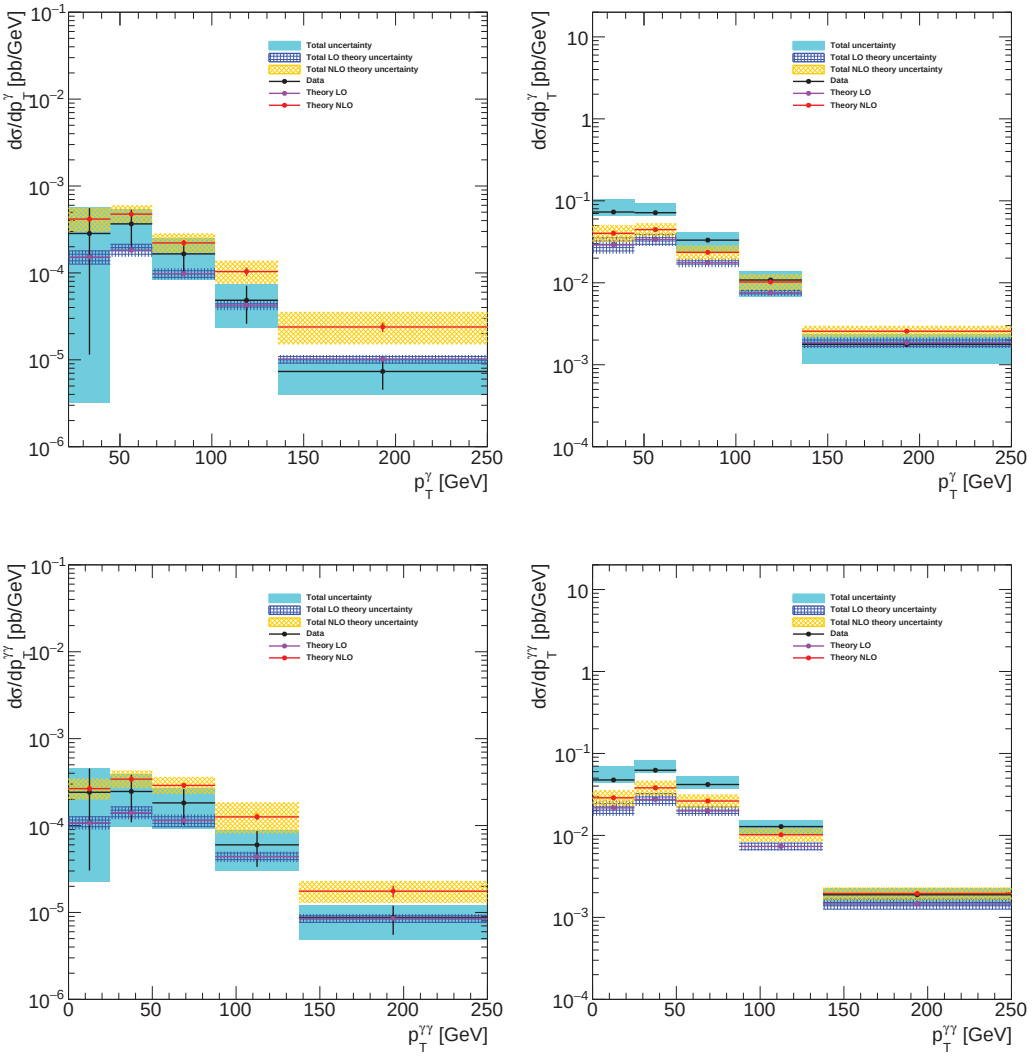

Figure 6.5.1: *On the left the distribution of the differential production cross-section of* $pp \to b\bar{b}\gamma\gamma$ *as obtained from this analysis compared to the theory predictions in terms of* p_T^{γ} *(top) and* $p_T^{\gamma\gamma}$ *(bottom). On the right the same, but for* $pp \to jj\gamma\gamma$.

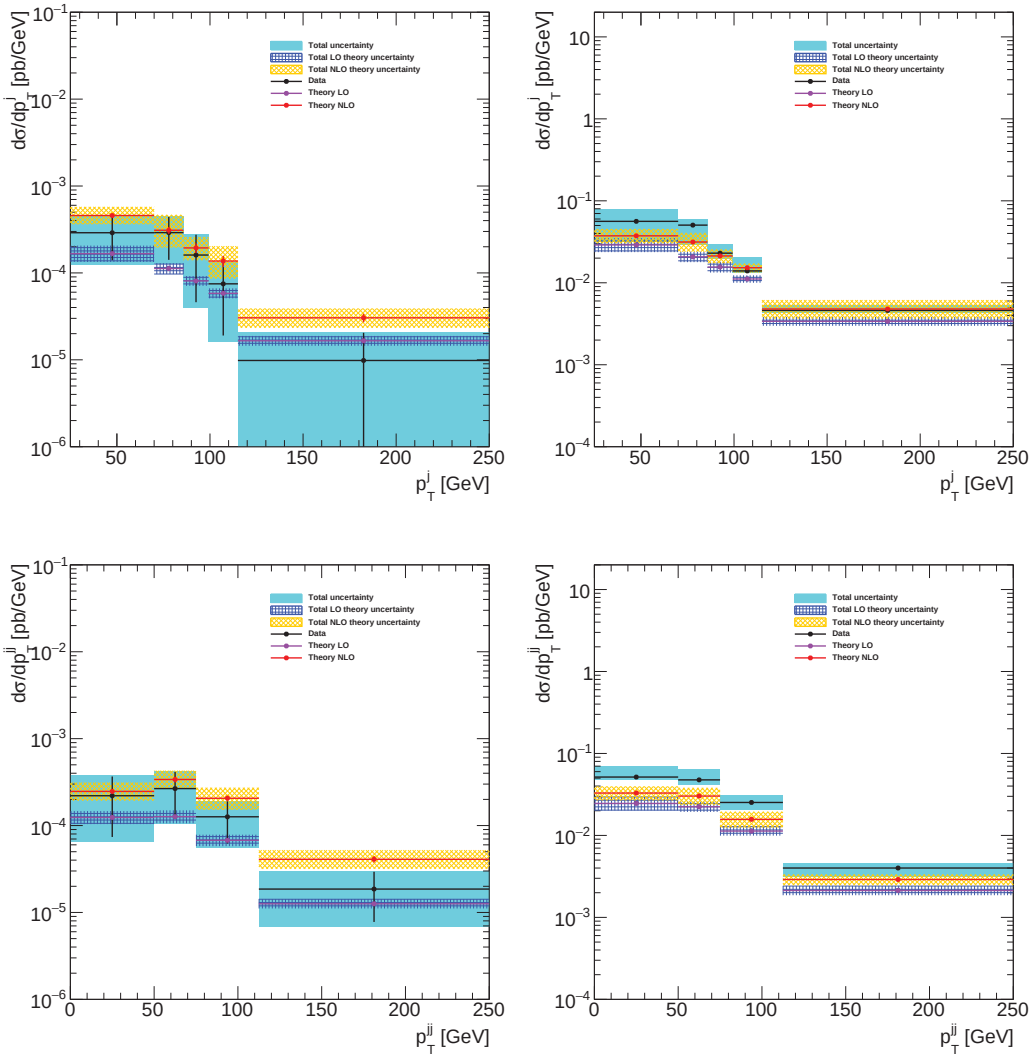

Figure 6.5.2: *On the left the distribution of the differential production cross-section of* $pp \to b\bar{b}\gamma\gamma$ *as obtained from this analysis compared to the theory predictions in terms of* p_T^j *(top) and* p_T^{jj} *(bottom). On the right the same, but for* $pp \to jj\gamma\gamma$.

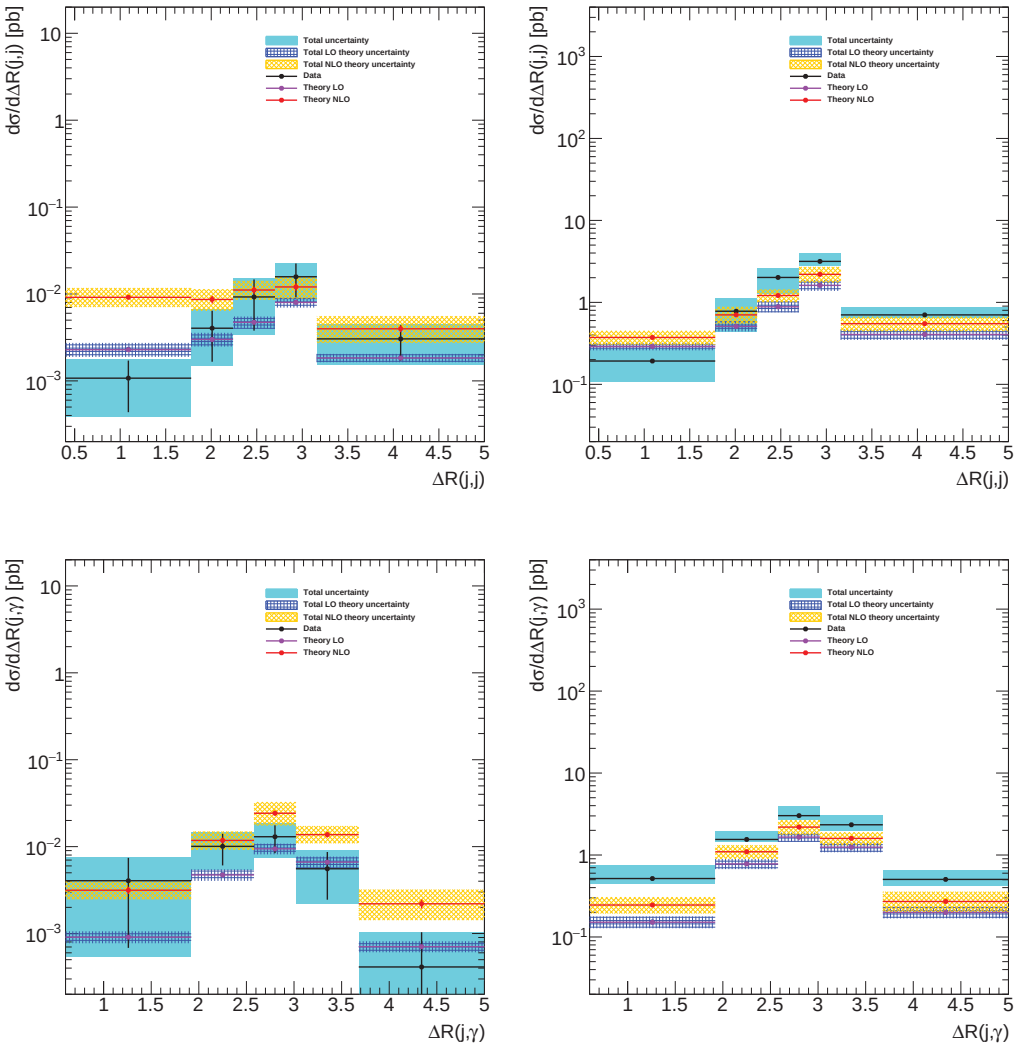

Figure 6.5.3: *On the left the distribution of the differential production cross-section of* $pp \to b\bar{b}\gamma\gamma$ *as obtained from this analysis compared to the theory predictions in terms of* $\Delta R(j,j)$ *(top) and* $\Delta R(j,\gamma)$ *(bottom). On the right the same, but for* $pp \to jj\gamma\gamma$.

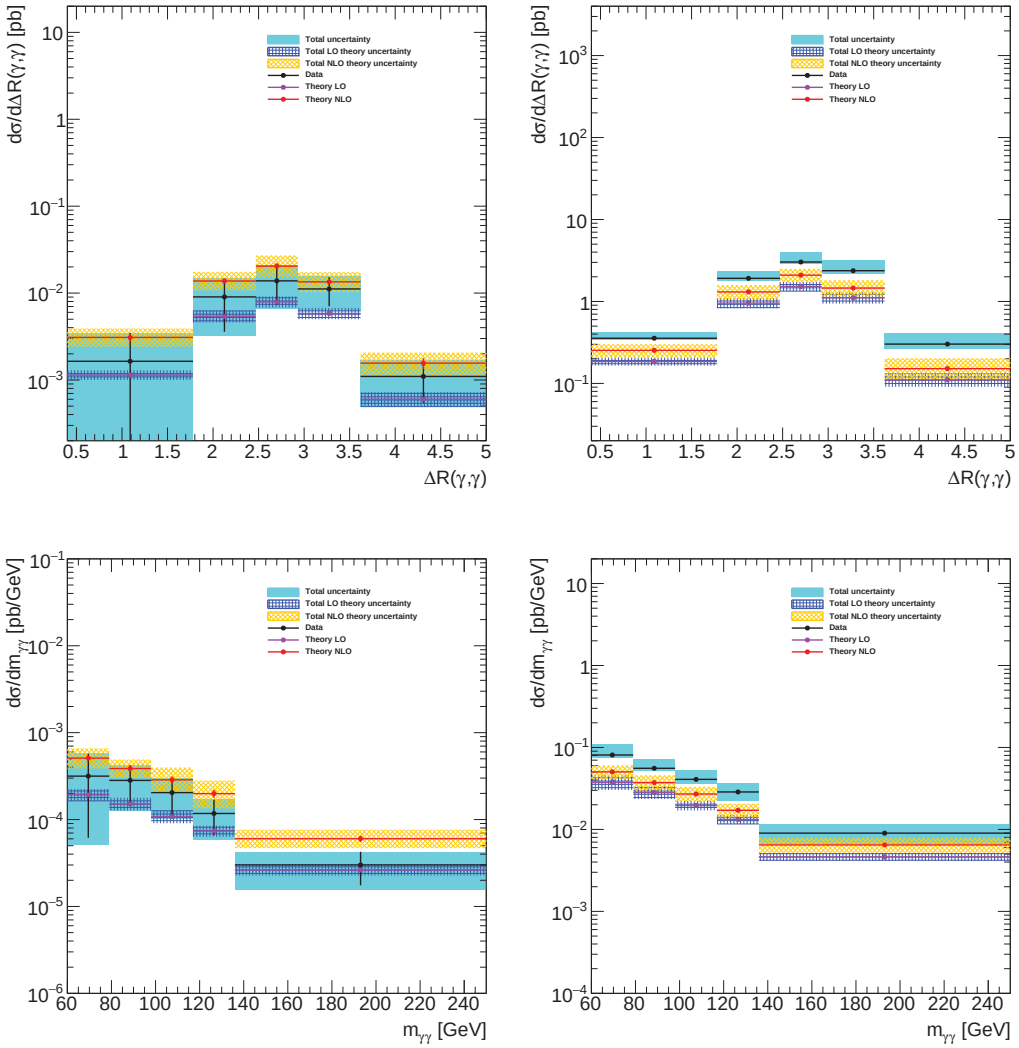

Figure 6.5.4: *On the left the distribution of the differential production cross-section of* $pp \to b\bar{b}\gamma\gamma$ *as obtained from this analysis compared to the theory predictions in terms of* $\Delta R(\gamma,\gamma)$ *(top) and* $m_{\gamma\gamma}$ *(bottom). On the right the same, but for* $pp \to jj\gamma\gamma$.

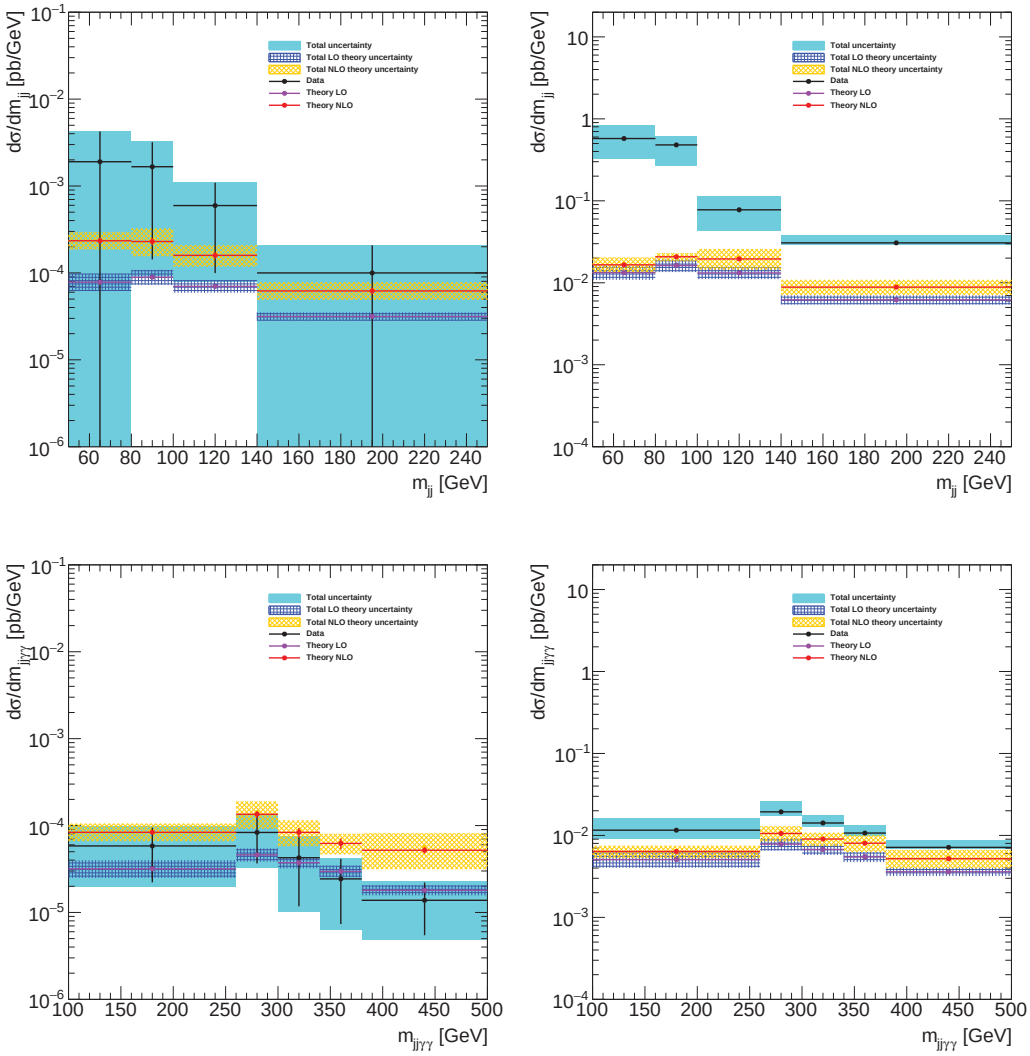

Figure 6.5.5: *On the left the distribution of the differential production cross-section of* $pp \to b\bar{b}\gamma\gamma$ *as obtained from this analysis compared to the theory predictions in terms of* m_{jj} *(top) and* $m_{jj\gamma\gamma}$ *(bottom). On the right the same, but for* $pp \to jj\gamma\gamma$.

	$\sigma\left(pp \to b\bar{b}\gamma\gamma\right)$[pb]	$\sigma\left(pp \to jj\gamma\gamma\right)$[pb]
Nominal value	0.0267 ± 0.0050	5.43 ± 0.036
(final) Luminosity up [%]	-1.86	-1.86
(final) Luminosity down [%]	1.94	1.94
(unfolding) Jet truth vertex distance up [%]	2.88	2.88
(unfolding) Jet truth vertex distance down [%]	-0.96	-0.96
(non-b-tag) Template uncertainties up [%]	-9.47	< 0.7
(non-b-tag) Template uncertainties down [%]	4.87	< 0.7
(electron) $f_{e\to\gamma}$ sideband shift up [%]	-0.99	< 0.7
(electron) $f_{e\to\gamma}$ sideband shift down [%]	1.87	< 0.7
(jet) Using E_T^{iso} for ϵ_I[%]	± 4.97	± 5.38
(jet) Using PYTHIA for ϵ_T[%]	< 0.7	± 0.70
(jet) ϵ_{I2}background boundary up [%]	< 0.7	0.70
(jet) ϵ_{I2}background boundary down [%]	-0.82	-0.95
(jet) ϕ_1 up [%]	-1.54	-1.24
(jet) ϕ_1 down [%]	1.23	1.00
(jet) ϕ_2 up [%]	-2.16	-2.09
(jet) ϕ_2 down [%]	1.68	1.61
(general) Lenient good PV cut [%]	< 0.7	0.92
(general) JVF cut up [%]	1.20	20.5
(general) JVF cut down [%]	1.65	20.9
Total systematic error [%]	± 11.9	± 22.1

Table 6.5.1: *Cross-section in the analysis' fiducial volume for $pp \to b\bar{b}\gamma\gamma$ (left) and $pp \to jj\gamma\gamma$ (right) as measured in data. The total systematic uncertainty was calculated by taking maximum of the up/down values of each systematic and adding them in quadrature. The word in parentheses is to remind the reader of the general source of the uncertainty. Only systematics with an effect of 0.7% or more are listed.*

	$\sigma\left(pp \to b\bar{b}\gamma\gamma\right)$[pb]		$\sigma\left(pp \to jj\gamma\gamma\right)$[pb]	
Simulation order	**LO**	**NLO**	**LO**	**NLO**
Nominal value	$0.0138 \pm 3.04 \cdot 10^{-5}$	0.0347 ± 0.00124	2.53 ± 0.0058	3.43 ± 0.0518
PDF [%]	± 11.2	± 10.0	± 10.2	± 20.5
Scale uncertainty [%]	-	± 24.4	-	± 8.59
Renormalization scale uncertainty [%]	± 11.0	-	± 10.2	-
Factorization scale uncertainty [%]	± 1.37	-	± 2.0	-
Total systematic error [%]	± 15.8	± 26.4	± 14.6	± 22.2

Table 6.5.2: *Cross-section in the analysis' fiducial volume for $pp \to b\bar{b}\gamma\gamma$ and $pp \to jj\gamma\gamma$ as found in the LO and NLO simulations. The total systematic uncertainty was calculated by adding each systematic in quadrature.*

CHAPTER 7

Conclusions and outlook

"We have a closed circle of consistency here: the laws of physics produce complex systems, and these complex systems lead to consciousness, which then produces mathematics, which can then encode in a succinct and inspiring way the very underlying laws of physics that gave rise to it."

- Roger Penrose

We have presented a first analysis using the ATLAS detector of the $pp \to b\bar{b}\gamma\gamma$ cross-section as a function of several prominent observables. We discussed how the main background to this analysis is formed by wrongly reconstructed photons and jets that fail to be b-tagged. We have also shown how to quantify these backgrounds and how to correct for them. Finally we have unfolded our results back to the parton level so they can be readily compared to theory predictions. Due to the low number of events recorded in the 8 TeV dataset of this kind we were only able to offer a coarse binning. Nonetheless we have showed that the results are largely consistent with NLO theoretical calculations and fall neatly within the LO and NLO predictions.

We should also realize the shortcomings of this analysis. A major one is the absence of a dedicated $pp \to b\bar{b}\gamma\gamma$ simulation sample fit for use in this 8 TeV analysis. We did our best to make due with the samples that were available, but it is clear that major improvements could be made with a correct signal sample. These would include the ability to do a proper unfolding based on signal events, and a better estimation of the tagging and photon identification efficiencies which are taken from Monte Carlo.

Another improvement is to use tracker-based jets instead of calorimeter-based jets. The former can better cope with pile-up than the latter which is expected to become much more important at higher collision energies. A calibration of b-tagging on such jets was performed by the author and details may be found in appendix A. The increased energy of 13 TeV used in the 2015 data will also improve the sensitivity to λ_{3h} as we saw in

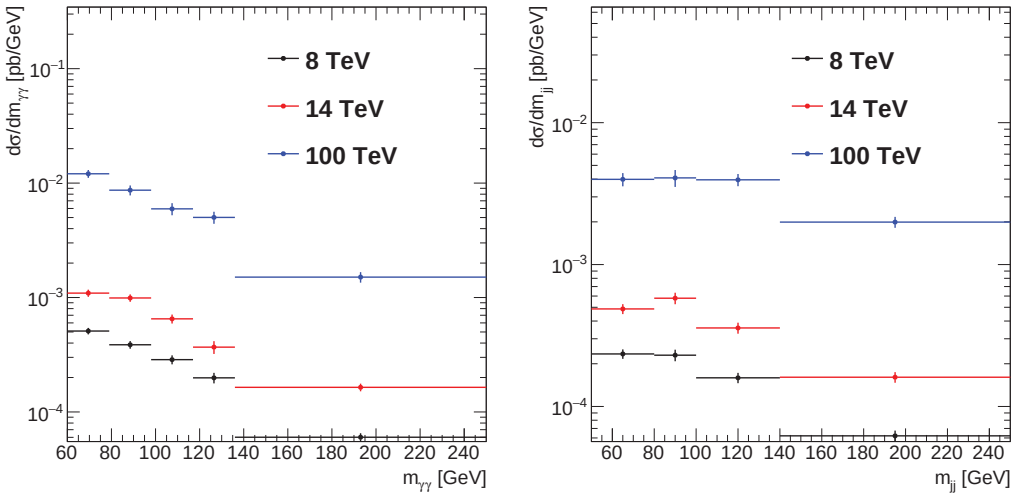

Figure 7.0.1: *Behavior of $pp \to b\bar{b}\gamma\gamma$ at various $\sqrt{s}$ energies for the diphoton invariant mass (left) and dijet invariant mass (right). No scale uncertainties are included and the statistical error is based on 50.000 events.*

section 2.4. A measurement of λ_{3h} is nevertheless out of the question at the LHC unless we collect a large amount of data. This might happen with the HL-LHC.

Also promising is a future 100 TeV hadron collider like the *Future Circular Collider* (FCC). We can briefly examine the behavior of $pp \to b\bar{b}\gamma\gamma$ at these energies using NLO MADGRAPH5 calculations. We show $m_{\gamma\gamma}$ and m_{jj} in figure 7.0.1 since these are important discriminants to examine and select Higgs bosons in a λ_{3h} measurement. We see that we can gain more than an order of magnitude in cross-section moving to 100 TeV. However, the present analysis is mainly constrained by fakes and the quality of b-tagging algorithms. Which means that the increase in energy will only be useful if the high performance of the reconstruction is maintained or surpassed. It will remain to be seen how well a future collider and detector will perform in this aspect. Another problem might be the increased cross-section of background channels. For example, a recent study found that the signal to background ratio for double Higgs analyses in $hh \to b\bar{b}b\bar{b}$ in the FCC is expected to be worse than for the HL-LHC [91]. This underlies the importance to obtain a good understanding of these backgrounds. An analysis might also need to combine several channels such as $W^+W^-b\bar{b}$, $b\bar{b}\gamma\gamma$, and $b\bar{b}\tau^+\tau^-$ to achieve enough combined sensitivity.

An e^+e^- linear collider, such as the *International Linear Collider* (ILC), might also offer an interesting opportunity. Since electrons are not composite objects the interaction energy of an e^+e^- collision can be carefully tuned to values around the Higgs mass. We have shown that these energies are the most sensitive to λ_{3h}. Of course a linear collider will not offer as much integrated luminosity as a hadron collider, but the events will be cleaner and the backgrounds better controllable. A dedicated study needs to be undertaken to see if the potential benefits outweigh this downside. Some preliminary studies are already underway [92]. However, in an e^+e^- collider the largest double Higgs production channels are $e^+e^- \to Z^* \to Zhh$ and $e^+e^- \to WW^*\nu\bar{\nu} \to hh\nu\bar{\nu}$. Both the ILC and FCC might also shed light on the role of the Higgs boson in baryogenesis

and dark matter [23].

Measuring λ_{3h} is of paramount importance as it plays an important role in many areas. The Higgs boson is the only elementary scalar particle we have discovered, and its role in the electroweak symmetry breaking should be clarified and experimentally verified. The exact value of λ_{3h} is also very sensitive to new physics and will likely lead to new clues in the areas of dark matter, cosmology, and yet-undiscovered particles. In this thesis we have shown that measuring the trilinear coupling is difficult since the signal is very tiny and buried under large backgrounds. The quantification and study of backgrounds, such as this one, will be very important and therefore we stress that analyses like the one presented here should be repeated at higher energies and at different colliders.

APPENDIX **A**

Calibration of the MV1 b-tagging algorithm for tracker-based jets

"I think nature's imagination is so much greater than man's, she's never going to let us relax."

- Richard Feynman

In this chapter we describe the calibration of a b-tagging algorithm on track-jets. This was part[†] of the ATLAS qualification task[‡] of the author. First we will discuss the reasons for wanting to do b-tagging on track-jets, then we move on to a description of the analysis setup and the simulation samples used, before moving on to the results.

A.1 The case for track-jet based b-tagging

With track-based b-tagging a cluster of tracks is used to define a region of interest instead of a calorimeter cluster[⊙]. This has a couple of advantages:

[†]The other part involved rewriting the configuration code of the ATLAS b-tagging software. This part will not be discussing within the confines of this thesis since it is very technical and requires specialized knowledge about the internal workings of the ATLAS software making it unsuitable for inclusion here.

[‡]A qualification task is typically performed during your first year of your PhD and is meant both as a way to train new members of ATLAS to use its software, and as a way to help them support ATLAS directly by doing things that are generally very important but unpopular and therefore suffer from manpower problems. Successful completion of the qualification tasks qualifies you as an ATLAS author, which basically means you get your name on all the ATLAS papers.

[⊙]Note that the actual tagging always happens on track-objects in both cases.

1. It is shown in Monte Carlo that the b-tagging efficiency drops rapidly if the angle between the b-hadron and the jet center increases. Track-based jets have centers that align better with the b-hadrons than the calorimeter-based ones. Therefore, track-jets should be better at b-tagging.
2. Since the tracks that make up the track-jet can be traced to a specific primary vertex a track-jet is less affected by higher pile-up conditions.
3. The number of calibrations that need to be performed by the b-tagging group can be greatly reduced by calibrating on track-jets and then allow end-users to match the track-jets to a calo-jet collection of their choosing. This is possible with, for example, ghost matching[†].
4. Since track-jets (and hence b-tagging) can be calibrated separately from calo-jets the b-tagging will not suffer from systematics that only affect calorimeter-based jets, such as jet energy resolution and scale uncertainties. In addition this allows one to optimize track-jet finding for b-tagging whilst optimizing the calorimeter-based jet finding for other things that may be of interest.

Because of these advantages it was deemed interesting by the ATLAS collaboration to investigate how well b-tagging performs on track-jets, which is the aim of this study.

A.2 Setup

THE efficiency with which a given b-tagging algorithm identifies b-jets is not the same in data and Monte Carlo, or at least we cannot assume it to be so. Therefore, when comparing data to simulated events one needs to add a correction for this fact. This correction depends on the ratio between the efficiency in data and Monte Carlo. The efficiency in Monte Carlo is trivial to measure; you have access to the knowledge of whether there was an actual b-hadron in the event and so you can check how often this hadron is correctly identified by a given algorithm. For data the situation is a bit more complex because you do not know what particles constituted an event, you only have access to reconstructed objects. To measure the efficiency in data it is therefore required that one identifies, to a large degree of certainty, a b-hadron without actually using the algorithm for which one is trying to determine the b-tagging efficiency[‡]. With the efficiencies in data and Monte Carlo in hand these can be divided to obtain a scale-factor with which the b-tagging rates in simulated data can be corrected. Because the efficiencies vary with jet radius and with the MV1 working point they have to measured using various jet radii and working points. The points selected where:

- ► $R = 0.4$ track-jets using the 70% working point. This was chosen because $R = 0.4$ calo-jets using the 70% working point was used a lot by ATLAS during Run 1, and it allows for a comparison to those.

[†]Ghost matching is a way of matching a track to a jet. They way it works is by creating a fake particle without any energy (the 'ghost') but with the direction of the track and injecting it into the set of particles that needs to be clustered in jets. If the ghost ends up in a jet the track is said to be matched to that jet. Ghost matching is an alternative to simply doing a ΔR cut, and is more robust for dense and boosted topologies.

[‡]Otherwise one falls into circular reasoning.

- $R = 0.3$ track-jets using a flat 0.7 cut on the MV1 output which was requested by a heavy resonance analysis in the exotics group within ATLAS.
- $R = 0.2$ track-jets, again using the 70% working point. The selection cuts for the track-jets is slightly modified here because track-jets with a small radius such as this can run into the problem where a b-jet with low transverse momentum is split into two smaller track-jets by the clustering algorithm which makes the association with the b-hadron ambiguous. Therefore, the transverse momentum cut is increased by a small amount.

The analysis for all these track radii, but for brevity we will in this thesis only show the results for the $R = 0.4$ track-jets, and refer the curious reader to the actual ATLAS internal note [93] for more information about the others.

To measure the efficiency the goal is to try and create a sample of events with a very high number of b-jets in them. The so-called b-purity should be as large as possible. The methodology and event topology employed here follows largely a previous calibration of the MV1 algorithm within ATLAS [94]. This calibration employed the so-called "tag and probe" method.

This involves selecting for $t\bar{t}$ events, because t-quarks frequently decay to b-quarks and so $t\bar{t}$ events have a large b-purity. The actual b-purity will not be 100% and so needs to be corrected for, which will happen later on. To further reduce backgrounds, events that have simultaneously a single electron and a single muon of opposite charge are selected. This suppresses many backgrounds that create a lepton pair of the same flavor. We will also require exactly two track-jets, which are going to be the b-jet candidates.

At this point one could envision attempting to tag the jets, see how often this works, correct for non-b-jets and arrive at a b-tagging efficiency. However, the b-purity is not high enough for this. Instead one can tag on the two jets using the MV1 algorithm. The jet is then known as the tag-jet. Tagging one of the jets results in events with a high degree of $b\bar{b}$ purity and so the other jet, called the probe-jet, has a large probability of also being a b-jet. One can then test MV1's performance on this probe-jet[†]. It is important to note that there are situations where both jets pass MV1. In such cases an event is considered twice since each jet is both a tag-jet and a probe-jet at the same time. See figure A.2.1 for a pictorial representation of the selected event topology.

The main backgrounds to the selected events include processes that generate two prompt leptons, such as diboson production with both bosons decaying leptonically, single top quark production together with a W boson, and events where a Z boson is produced with additional jets and where the Z boson decays to two tauons that then subsequently decay into an electron muon pair.

Failures to correctly reconstruct the required muon or electron leads to further backgrounds. This includes cases where one or both of the leptons are non-prompt, jets wrongly reconstructed as electrons, electrons coming from photons that converted in

[†]One might at this point wonder whether the inclusion of a tagged-jet would not necessitate the use of the very b-tagging scale factors we are trying to determine. This is a fair point. The correct way of proceeding is to calculate the scale factors and then reinsert them into the calibration (for use on the tag-jets) and rederive them again. Then continue this cycle until the values stabilize. We have not performed a full study of this due to time constraints but a quick test showed stabilization already after the first cycle thus increasing trust in the values recovered.

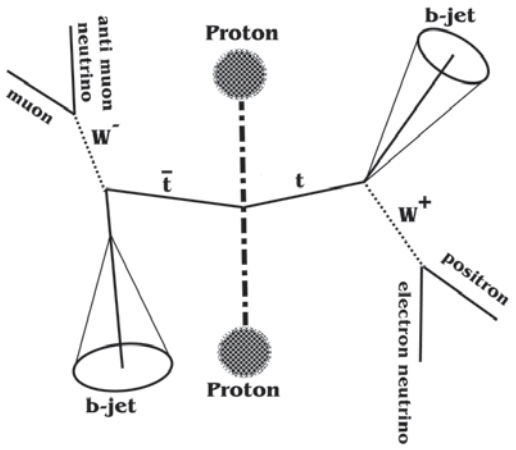

Figure A.2.1: $t\bar{t} \rightarrow b\bar{b}e^{-}\mu^{+}\bar{\nu}_{e}\nu_{\mu}$ *event topology. The charge-reversed situation where the muon comes from the* W^{+} *and the electron from the* W^{-} *is analogous.*

the detector material, and muons coming from pion or kaon decay. A large portion of this kind of background comes from $t\bar{t}$ events where one of the top quarks decays leptonically instead of decaying to a coveted b-quark. Other sources include Z bosons decaying to an electron or muon pair, W boson decay along with additional jets, and single top quark production via the s- or t-channel. The $W + jets$ contribution was investigated at but its effects on the calibration were found to be completely negligible and it was subsequently dropped from the analysis.

A.2.1 Samples

The data used corresponds to the full set of proton-proton collision data the ATLAS detector collected in 2012 at a center-of-mass energy of 8 TeV. The simulated data samples were the standard ATLAS Monte Carlo samples [95] produced in 2012 and which were subject to the same reconstruction algorithms as the data samples. The full ATLAS detector simulation [70] which is based on GEANT4 [69] was used for most of them, but some used a less accurate, but faster, simulation known as ATLFAST2. Minimum bias events generated using PYTHIA8 [63] were added to all simulated samples in order to simulate pile-up based backgrounds. The Monte Carlo samples were reweighted so that the average number of interactions per bunch crossing were the same in data and simulation. All samples were normalized to best available theoretical cross-sections and K-factors recommended by the TOP group in ATLAS.

The main $t\bar{t}$ sample was generated using POWHEG [79] interfaced with PYTHIA6 [62] and using the Perugia 2011C tune [96] and the CT10 parton density functions [48]. To assess systematics originating from MC modeling uncertainties an additional $t\bar{t}$ sample was used that was produced using MC@NLO [47] interfaced with HERWIG [97] and JIMMY [77] and using the ATLAS AUET2 [82] tune as well as the aforementioned CT10 parton distribution functions.

Two $t\bar{t}$ samples using the fast ATLFAST2 simulation generated with the ACERMC generator [98] interfaced to PYTHIA6 and using the CTEQ61 [37] parton distribution functions were used to investigate the effects of initial and final state radiation. The two samples use two tunes that are compatible with the variations found in an ATLAS study [99] of heightened jet activity in $t\bar{t}$ events. Systematics related to different methods of handling hadronization in simulations were studied using a fast simulation sample generated using POWHEG interfaced to HERWIG and using the AUET2 tune and CT10 parton distribution functions.

The s-channel single top quark background was simulated using POWHEG interfaced to PYTHIA6. The t-channel was simulated using ACERMC, also interfaced to PYTHIA6. The Wt single top quark background was generated using POWHEG, which again was interfaced to PYTHIA6. All these samples use the CT10 (NLO) PDFs and the Perugia 2011C tune.

The $Z \rightarrow \tau\tau$ with jets background was generated using SHERPA [64] and using the AUET2 tune and CTEQ6L1 PDFs. The samples were split into many different datasets using various criteria including the Z boson decay type, by the flavor of the additional jets, and into bins of the Z boson transverse momentum. The reasoning is getting a good amount of systematics across all of phase-space. Comparable SHERPA samples were used for the $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ backgrounds.

The diboson samples (WW, WZ, and ZZ) were obtained using ALPGEN [81] interfaced to HERWIG and JIMMY with the AUET2 tune and CTEQ6L1 PDFs. The different samples contained between 0 and 3 additional light jets.

A.2.2 Object selection

Electrons

Electrons were required to have a $p_T > 25\,GeV$ and an $|\eta| < 2.47$ where the calorimeter transition zone ($1.37 < |\eta| < 1.52$) is excluded. The TIGHT++ identification was used, and in addition two isolation requirements[†] are made: $ETcone20 < 6\,GeV$ and $pTcone30 < 6\,GeV$. Muons faking electrons due to catastrophic energy loss in the calorimeters or brehmsstrahlung are suppressed by rejecting any electron that shares a track with a reconstructed muon. If there was a jet within $\Delta R < 0.4$ of the electron the electron was also rejected. Note however, that jets within $\Delta R < 0.2$ of an electron are removed (see below).

Muons

Muons were reconstructed using both the inner detector and the muon spectrometer with the MUID algorithm. They were required to have a $p_T > 25\,GeV$ and an $|\eta| < 2.47$. Furthermore the impact parameter with respect to the primary vertex in the longitudinal direction needed to satisfy $z_0 < 2\,mm$. Muon production from heavy flavor decay in jets was suppressed by removing muons within $\Delta R < 0.4$ of a jet. Further isolation is obtained by requiring $I^l_{mini} < 0.05$, where I^l_{mini} is the mini-isolation [100] which is the ratio of the total p_T of all tracks in a variable-sized cone around the muon to the muon p_T.

Track-based jets

Track-jets were reconstructed with radii R equal to 0.2, 0.3, and 0.4 and using the anti-k_t jet clustering algorithm [72]. They were required to have a $p_T > 7\,GeV$ and $|\eta| < 2.5$. For $R = 0.2$ track-jets the p_T cut was raised to $10\,GeV$. A track-jet was required to be made from at least two tracks. The tracks that were used as input to the clustering algorithm were required to have a $p_T > 0.5\,GeV$ and $|\eta| < 2.5$. There has to be at least a single hit in a pixel detector, and at least five hits in the SCT detector[‡]. Further requirements are made to the impact parameters: $|d_0| \leq 1.5\,mm$ and $|z_0 \sin(\theta)| < 1.5\,mm$, where both are measured with respect to the primary vertex. These cuts greatly reduce pile-up uncertainties by making sure the tracks are tied to the primary interaction vertex. A track-jet is considered matched to a calo-jet if it is within a ΔR of 0.4 from a calo-jet. If there are multiple such jets it is matched to the closest one.

[†] $ETcone20$ quantifies the energy deposited in the calorimeter within a radius of $\Delta R < 0.2$ around the electron where the electron's own energy deposition is disregarded. Lower values therefore are a measure of isolation. $pTcone30$ measures the total p_T of all tracks within a radius of $\Delta R < 0.3$ where the electron's own track is excluded. This also gives a measure of isolation.

[‡] Note that if a track passes over a known faulty pixel sensor it is assumed to be a hit. This happens for both the pixel detector and the SCT detector.

Calorimeter-based jets

Calo-jets were reconstructed with a radius parameter of $R = 0.4$ and using the anti-k_t clustering algorithm. They were seeded using topological clusters that were calibrated using the local cluster weighting method and were corrected for pile-up using the jet area method.

b-tagging

The MV1 algorithm was used to determine whether a jet was b-tagged or not. The cut values varied per jet radius. For $R = 0.4$ the 70% working point was chosen, which results in a cut value of 0.8644. For $R = 0.3$ a cut of 0.7 was used, and for $R = 0.2$ the cut value was 0.8780, which is the 70% working point for the $t\bar{t}$ sample that was used. For studies involving truth data a jet was considered a b-jet if it was within a ΔR of 0.3 of a b-hadron.

A.2.3 Event selection

All selected events were required to fulfill some basic quality criteria including that all detector components were functioning at the time of recording. The muon data stream was used and events had to satisfy the logical OR of two triggers: EF_MU24I_TIGHT and EF_MU36_TIGHT[†]. The total integrated luminosity of the analyzed data sample was $20.4\,\text{fb}^{-1}$ with a total uncertainty of 2.8% which was calculated in a similar way to [54].

In addition at least one interaction vertex was required to be reconstructed in the inner tracker with at least three tracks associated to it. If a calo-jet was present with $p_T > 20$ GeV and which failed the BADLOOSEMINUS jet quality cut the event was not considered in this analysis. Likewise events were vetoed if they were corrupted, incomplete, or contained a LAr noise burst or corrupted calorimeter tile data[‡].

Events were then required to have exactly one electron and one muon with the characteristics detailed before. The muon was required to match to the before-mentioned triggers in order to increase the chances that this was the muon that triggered the event. Because an muon can be misidentified as an electron if it loses a lot of energy in the calorimeter or undergoes brehmsstrahlung events are rejected if the ΔR between the selected electron and selected muon is less than 0.15.

The samples after the brehmsstrahlung cut, but before any cut on the jets, form the preselection.

It was investigated whether it is prudent to add an azimuthal separation cut to reduce the effects of initial and final state radiation, like was done in an earlier calorimeter-based calibration. It was found that the cut did not significantly affect the b-tagging efficiency for track-jets, and the cut was dropped in order to increase statistics.

[†]The first of these triggers selects isolated muons with a transverse momentum of at least 24 GeV, the second one raises the momentum cut to 36 GeV but removes the isolation requirement but boosts the efficiency compared to the first one.

[‡]These are standard ATLAS practices which are detailed here:
twiki.cern.ch / twiki / bin / view / Atlas / DataPreparationCheckListForPhysicsAnalysis

A.2.4 Same-sign and opposite-sign events

The selected events were subdivided into two regions. In one of them the charges of the two selected leptons were required to have the opposite sign. As mentioned before this selection increases the b-purity of the sample. This sample will be known as the opposite-sign (OS) sample. Another region was defined where the charges were required to be of the same sign. This is the same-sign (SS) sample, which is used to assess the impact of fake leptons.

In an earlier b-tagging tag and probe calibration on calo-jets [101] it was used to gauge the fake and non-prompt lepton contribution to the OS sample. The reasoning being that the processes that contribute to fake leptons do not depend on the charge and so fake leptons contribute in similar fashions to both samples, but because of the inverted charge cut the SS contains virtually only fake lepton events. To investigate the kinematics of fake lepton events one can therefore take the SS distributions and subtract off the SS distribution from the OS distributions to remove the influences of fakes (and avoid double counting).

Sample	*data*	$t\bar{t}$	$Z + jets$	*single top*	*diboson*	*fake leptons*
Initial	$1.781 \cdot 10^8$	$2.785 \cdot 10^6$	$7.797 \cdot 10^7$	$1.068 \cdot 10^6$	$6.793 \cdot 10^5$	$1.781 \cdot 10^8$
GRL	$1.701 \cdot 10^8$	$2.785 \cdot 10^6$	$7.797 \cdot 10^7$	$1.068 \cdot 10^6$	$6.793 \cdot 10^5$	$1.701 \cdot 10^8$
Vertex	$1.699 \cdot 10^8$	$2.784 \cdot 10^6$	$7.733 \cdot 10^7$	$1.067 \cdot 10^6$	$6.781 \cdot 10^5$	$1.699 \cdot 10^8$
BadLoose	$1.698 \cdot 10^8$	$2.782 \cdot 10^6$	$7.725 \cdot 10^7$	$1.066 \cdot 10^6$	$6.774 \cdot 10^5$	$1.698 \cdot 10^8$
Cleaning	$1.694 \cdot 10^8$	$2.782 \cdot 10^6$	$7.725 \cdot 10^7$	$1.066 \cdot 10^6$	$6.774 \cdot 10^5$	$1.694 \cdot 10^8$
Trigger	$6.172 \cdot 10^7$	$6.523 \cdot 10^5$	$1.694 \cdot 10^7$	$1.733 \cdot 10^5$	$1.496 \cdot 10^5$	$6.172 \cdot 10^7$
Dilepton	$4.91 \cdot 10^4$	$2.733 \cdot 10^4$	$1.028 \cdot 10^4$	2716	6043	$4.91 \cdot 10^4$
Charges	$4.714 \cdot 10^4$	$2.708 \cdot 10^4$	$1.001 \cdot 10^4$	2684	5515	1956
Brehmsstrahlung	$4.714 \cdot 10^4$	$2.708 \cdot 10^4$	$1.001 \cdot 10^4$	2684	5515	1956
Dijet	$1.333 \cdot 10^4$	8500	953.9	861	780.3	338
b**-jet**	7798	7115	70.98	616.9	45.43	74

Table A.3.1: *Quantity of events passing the various cuts for $R = 0.4$ track-jets. Note that the fake lepton contribution consists of the full SS data sample.*

A.3 Data and simulation comparisons

HERE we will investigate how well the simulation describes the data in the preselection and after the full selection.

A.3.1 Preselection

The number of events passing each step in the event selection can be found in table A.3.1. The MC samples have all been scaled to the same integrated luminosity as the data. For the MC samples the SS events were subtracted off the OS events. The fake lepton sample consists of the SS data, which is why it follows the data in the cutflow up until the charge cut. In figure A.3.1 one can examine the track-jet kinematics and specifically the track-jet kinematics for jets that are b-tagged. In these plots, and those that will follow, the different $t\bar{t}$ samples are all shown. The nominal POWHEG+PYTHIA sample is part of the stack, and the other two feature as a green or red line. The other simulation samples have been added to these. All distributions are reasonably well described by the simulations.

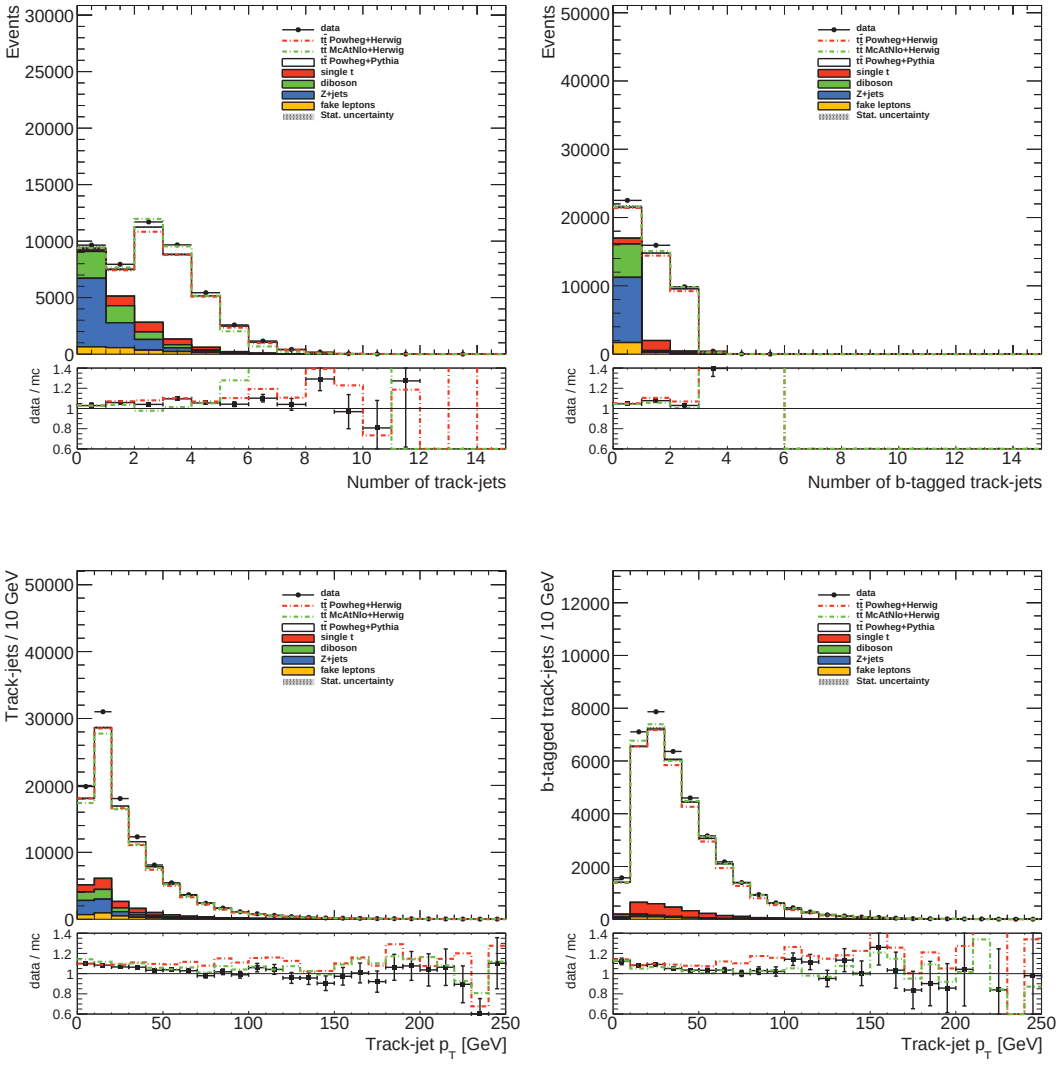

Figure A.3.1: *Shown for all $R = 0.4$ track-jets (left) and for only the ones that are b-tagged (right): their number (top) and their transverse momentum (bottom).*

Sample	Preselection	%	One b-tag	%	Two b-tags	%
$t\bar{t}$	27080	57.3	4147.4	86.3	2915.3	95.5
single top	2684	5.7	498.3	10.4	112.3	3.7
$Z + jets$	10010	21.2	65.2	1.4	3.9	0.1
Diboson	5515	11.7	38.6	0.8	1.0	0.0
fake leptons	1956	4.1	54.0	1.1	20.0	0.7
Total	47245	100	4803.5	100	3052.4	100
data	47145		4916		2882	

Table A.3.2: *Number of events in the preselection and the number of events with a single or double b-tag for the $R = 0.4$ track-jets. The last two categories are mutually exclusive.*

A.3.2 Full selection

The full selection adds the requirement that there be exactly two track-jets with the characteristics described earlier. In addition at least one of the must be b-tagged. In table A.3.2 one can find a subdivision in events with one or two b-tagged jets. As mentioned before, doubly b-tagged events enter the analysis twice since both jets can serve as both a probe and a tag jet. The correspondence between data and simulation is good, particularly after requiring a b-tag, boosting confidence that we have a good description of the events in simulation. However, significant variations between the various $t\bar{t}$ samples remain. These will later end up in the systematic errors. After the full selection the main background is formed by single t events, but the sample is mostly filled by $t\bar{t}$ events.

Figure A.3.2 shows the $R = 0.4$ track-jet kinematics for probe-jets and of jets that were tagged. Figure A.3.3 shows the flavor fractions as a function of probe-jet p_T. A summary of the flavor fractions and the number of probe-jets available after the full selection is found in table A.3.3.

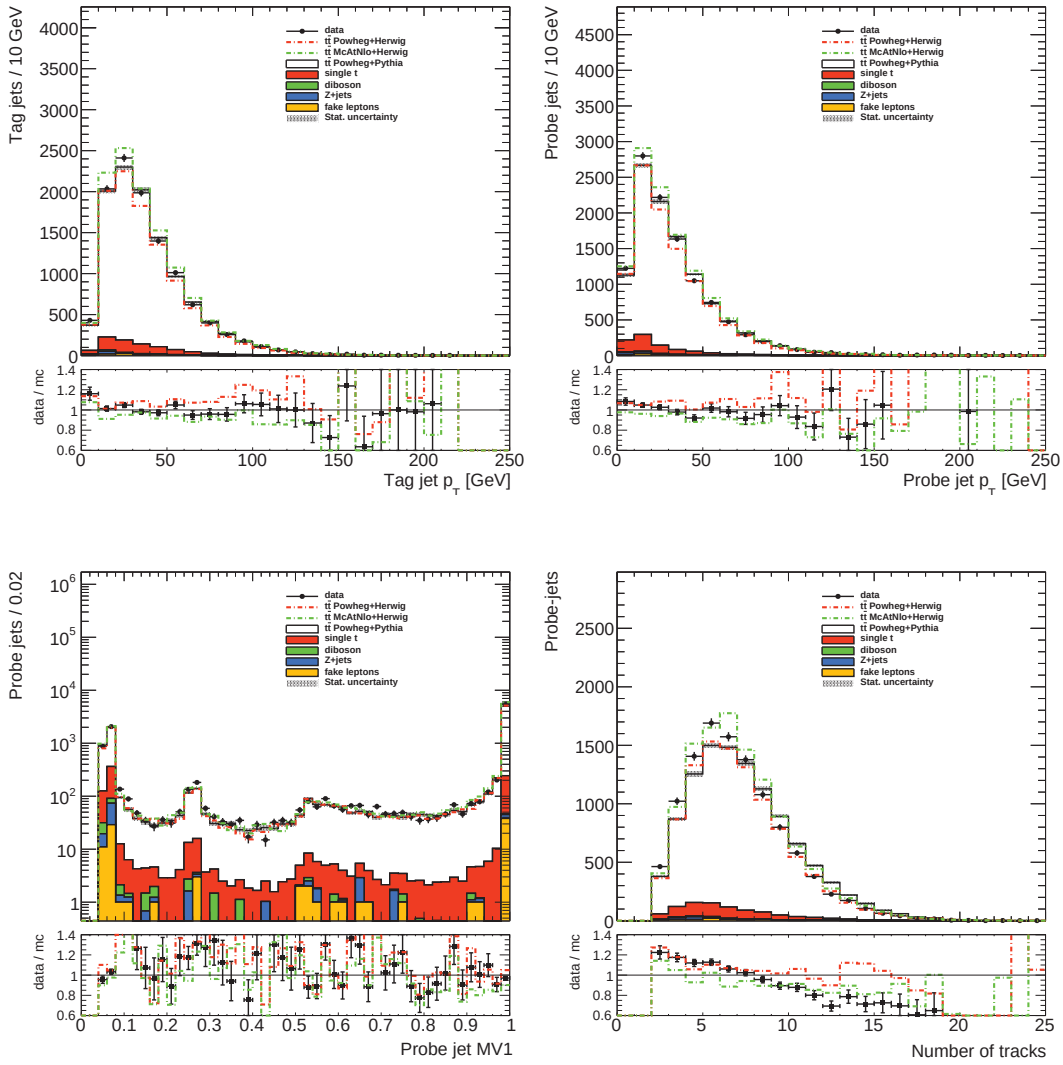

Figure A.3.2: *Shown are for the $R = 0.4$ track-jets after the full selection: the transverse momentum for tag-jets (top left) and for probe-jets (top right), and the probe-jet MV1 discriminant distribution (bottom left). Also shown is the distribution of the number of track-jets per event (bottom right).*

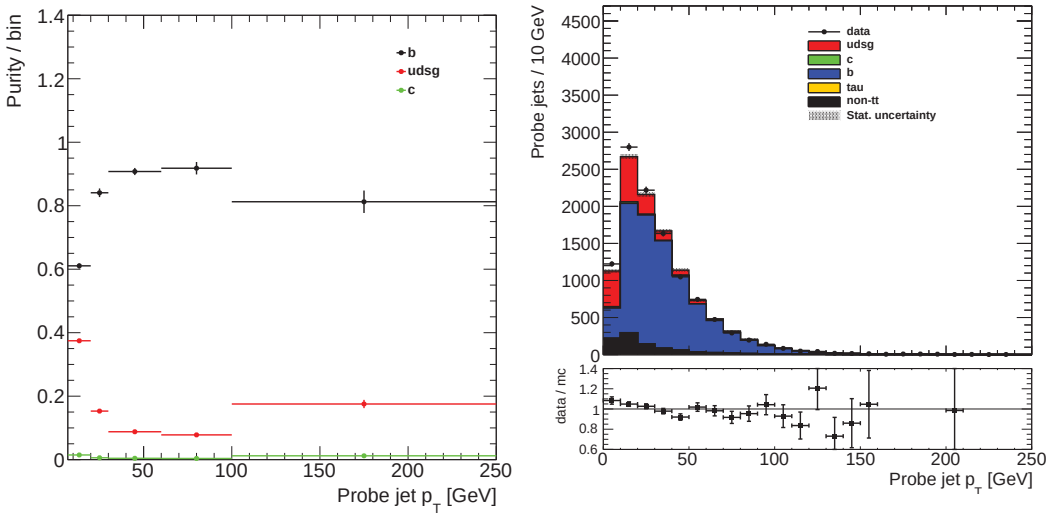

Figure A.3.3: *For the* POWHEG+PYTHIA $t\bar{t}$ *sample and also including all the background simulation samples for* $R = 0.4$ *track-jets we show: the flavor fraction as a function of probe-jet* p_T *(left) and the flavor labels in absolute numbers as a function of* p_T *(right).*

Sample	Events	%	Probe-jets	%	udsg [%]	c [%]	b [%]
non-$t\bar{t}$ backgrounds	733.3	9.3	836.3	7.7	57.0	2.5	42.0
$t\bar{t}$	7115.0	90.7	9978.0	92.3	17.2	0.7	82.7
Total simulation	7848.3	100.0	10814.3	100.0	20.3	0.8	79.5
Fake leptons	74.0		94.0				
Total	7922.3		10908.3				
Data	7798.3		10680.0				

Table A.3.3: *The number of events passing the full selection, including a single* b*-tag, and the number of probe-jets along with the flavor composition for* $R = 0.4$ *track-jets. The* $t\bar{t}$ *sample was generated using* POWHEG+PYTHIA*. Since the fake lepton contribution is extracted from data it not included in the total simulation sample.*

A.4 Extracting the b-tagging efficiency

CONFIDENT that the simulation provides a good-enough description of the data we press on and attempt to extract the b-tagging efficiency from the data. In addition we will discuss the systematic uncertainties.

A.4.1 Correcting for non-b-jet contamination and non-$t\bar{t}$-events

Since we set out to measure the b-tagging efficiency in data and simulation for $t\bar{t}$ events we need to subtract off the non-$t\bar{t}$ contributions. In addition, even in $t\bar{t}$ events there will still be light jets and a very small number of c-jets remaining. We need to correct for this less than perfect $b\bar{b}$-purity. If we call the number of tagged jets we expect to see in data, $N_{tagged,data}$, we can express it in the following way:

$$N_{tagged,data} = \epsilon_{b,data} \cdot N_{b,data} + \epsilon_{\neg b,data} \cdot N_{\neg b,data} \tag{A.4.1}$$

where $\epsilon_{b,data}$ is the tagging efficiency of actual b-jets, $\epsilon_{\neg b,data}$ is the tagging efficiency of light and charm jets, and $N_{b,data}$ and $N_{\neg b,data}$ are the number of actual b- and non-b jets in data. The tagging efficiency for light-jets and c-jets is slightly different, but because the contribution of c-jets is tiny we merge them with the light jets. Of course, using $f_{b,data}$ as the b-purity in data we can simplify this a bit:

$$N_{tagged,data} = N_{data}\left(\epsilon_{b,data} \cdot f_{b,data} + \epsilon_{\neg b,data}(1 - f_{b,data})\right) \tag{A.4.2}$$

where N_{data} represents the total number of data events. The measured b-tagging efficiency (which is $N_{tagged,data}$ divided by N_{data}) is then:

$$\epsilon_{data} = \epsilon_{b,data} \cdot f_{b,data} + \epsilon_{\neg b,data}(1 - f_{b,data}) \tag{A.4.3}$$

But what we are interested in is $\epsilon_{b,data}$, so solving for it gives:

$$\epsilon_{b,data} = \frac{\epsilon_{data} - \epsilon_{\neg b,data}(1 - f_{b,data})}{f_{b,data}} \tag{A.4.4}$$

$\epsilon_{\neg b,data}$ and $f_{b,data}$ are taken to be equal to their simulation equivalents, leaving ϵ_{data} as the only measured quantity:

$$\epsilon_{b,data} = \frac{\epsilon_{data} - \epsilon_{\neg b,MC}(1 - f_{b,MC})}{f_{b,MC}} \tag{A.4.5}$$

There systematic uncertainties associated with using the simulation values are discussed in section A.4.3 on page 204. We can express ϵ_{data} in the following way:

$$\epsilon_{data} = \frac{N_{data}^{pass}}{N_{data}^{total}} \tag{A.4.6}$$

where N_{data}^{total} represents the total number of probe-jets passing the full selection in the data sample, and N_{data}^{pass} the same probe-jets, but only those that also pass the MV1

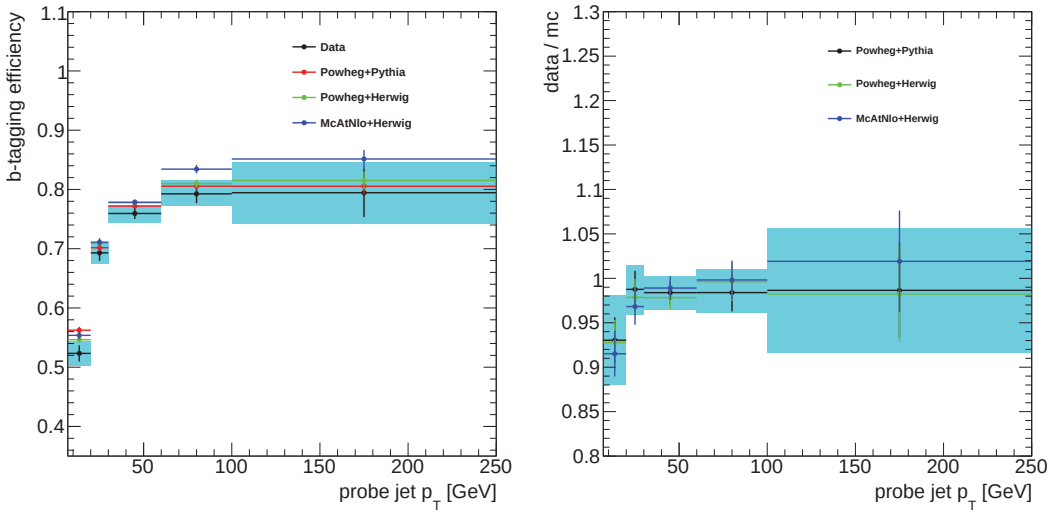

Figure A.4.1: *Measured b-tagging efficiencies and scale factors for $R = 0.4$ track-jets. The error bars denote the statistical error, the shaded area denotes the total uncertainty which consists of the statistical and systematic errors added in quadrature. The errors are only shown for the nominal* POWHEG+PYTHIA $t\bar{t}$ *sample.*

cut. We can then easily subtract off the non-$t\bar{t}$ contribution using the non-$t\bar{t}$ simulation sample:

$$\epsilon_{data} = \frac{N_{data}^{pass} - N_{\neg t\bar{t}}^{pass}}{N_{data}^{total} - N_{\neg t\bar{t}}^{total}} \tag{A.4.7}$$

Leading us to the following expression to obtain the corrected b-tagging efficiency:

$$\epsilon_{corr} = \frac{\left(\frac{N_{data}^{pass} - N_{\neg t\bar{t}}^{pass}}{N_{data}^{total} - N_{\neg t\bar{t}}^{total}}\right) - \epsilon_{\neg b,MC}^{t\bar{t}}(1 - f_{b,MC}^{t\bar{t}})}{f_{b,MC}^{t\bar{t}}} \tag{A.4.8}$$

where $\epsilon_{\neg b}$ and f_b are taken specifically from the $t\bar{t}$ simulation sample. This computation is performed per bin of the desired observable (for example p_T).

A.4.2 The measured b-tagging efficiencies and scale factors

The b-tagging efficiencies have been measured both as a function of probe-jet p_T and probe-jet $|\eta|$, but we shall only show the p_T here. The systematic uncertainties are described in the next section. The measured b-tagging efficiencies for the MV1 algorithm, corrected for non-b-jet and non-$t\bar{t}$-event contributions, for data and simulation, along with the ratio that forms the scale factor versus the probe-jet p_T are displayed in figure A.4.1 and listed in table A.4.1. The scale factors are close to unity which tells us that the simulations model the data quite well with regard to b-tagging.

p_T(GeV)	Data (%)	MC (%)	Scale factor	Total uncertainty
$7-20$	$52.2 \pm 1.4 \pm 1.6$	56.2 ± 0.6	$0.928 \pm 0.026 \pm 0.044$	0.051
$20-30$	$69.0 \pm 1.4 \pm 1.1$	70.2 ± 0.6	$0.984 \pm 0.021 \pm 0.019$	0.028
$30-60$	$76.2 \pm 0.9 \pm 1.2$	77.2 ± 0.4	$0.988 \pm 0.013 \pm 0.014$	0.019
$60-100$	$79.6 \pm 1.6 \pm 1.5$	80.6 ± 0.6	$0.989 \pm 0.021 \pm 0.012$	0.025
$100-250$	$80.0 \pm 4.2 \pm 3.2$	80.5 ± 1.4	$0.994 \pm 0.055 \pm 0.045$	0.071

Table A.4.1: *The MV1 b-tagging efficiency for the 70% working point for $R = 0.4$ track-jets in bins of probe-jet p_T in data and simulation (MC). Statistical errors are listed for both the data and the simulation samples. Systematics errors are listed for the data sample. The total uncertainty consists of the statistical error and systematic error added in quadrature.*

A.4.3 Systematic uncertainties

A large number of systematic uncertainties were investigated. For each also the non-b-tagging efficiency, and the b-purity, which were taken from simulation, are recalculated along with the number of probe-jets passing the cutflow. The total systematic uncertainty was obtained by adding together all individual contributions in quadrature. The exact effect of each systematic on the b-tagging efficiency as a function of probe-jet p_T can be found in table A.4.2. The same for the scale factors can be found in table A.4.3. In the tables any systematic that had an effect less of than 0.5 % in all bins was suppressed for readability. As a cross-check the analysis was repeated for $R = 0.4$ track-jets but only considering track-jets matched to calorimeter-jets within $\Delta R < 0.4$ so the b-tagging efficiency could be more directly compared to earlier $R = 0.4$ calorimeter-jet calibrations. The results were found to be consistent with earlier analyses. The systematics considered thus were:

Generation of $t\bar{t}$ events There are several different Monte Carlo generators available to generate collision events. Because we extract various quantities from these simulated events such as the non-b-tagging efficiency and the b-purity we should investigate how these quantities depend on the exact generator used.
To assess this a POWHEG+HERWIG $t\bar{t}$ sample was compared to a MC@NLO+HERWIG $t\bar{t}$ sample. The difference in the resulting b-tagging efficiencies was then symmetrized around the nominal POWHEG+PYTHIA value as a systematic.

Generation of Wt events Because Wt is our largest background we should also investigate the effects of different simulation techniques of this channel. It enters into the analysis when we subtract off the non-$t\bar{t}$ background from the data sample. A MC@NLO+HERWIG Wt sample was used to compare to the nominal POWHEG+PYTHIA sample. The difference was symmetrized.

Cross-section of $t\bar{t}$ and Wt The relative cross-sections between $t\bar{t}$ and Wt affect the flavor composition of the final selection. However, because we subtract off the

non-$t\bar{t}$ events from the data and compare directly to $t\bar{t}$ events the actual normalization of $t\bar{t}$ has no effect on the final results[†]. To assess the uncertainty on the Wt cross-section it was varied according to the uncertainties of the corresponding theoretical cross-section [102, 103]. This comes down to a $\pm 6.8\%$ variation.

Hadronization Uncertainty in the treatment of hadronization was assessed by comparing the nominal POWHEG+HERWIG sample where the events were hadronized by PYTHIA to the POWHEG+HERWIG sample where the hadronization was handled by HERWIG. The difference was taken as the systematic.

Initial and final state radiation The effects of the handling of the initial and final state radiation on $t\bar{t}$ and Wt (the largest background) samples was assessed by taking ACERMC+PYTHIA samples for these channels but with two different tunes. The tunes were chosen as to explore the full range of additional jet activity [99]. The difference in final results was symmetrized around the nominal result and taken as a systematic. The effects on $t\bar{t}$ and Wt were examined at the same time and the uncertainties between them are therefore treated as fully correlated.

Track-jet track p_T-scale and resolution The inner detector has uncertainties on the momentum scale and resolution with respect to the tracks that are measured, which we need to take into account since these tracks form the basis of the track-jets we analyze. The transverse momentum of the tracks out of which a track-jet consists were varied within the bounds of the systematic uncertainties of muon tracks in the inner detector as reported by a $Z \to \mu\bar{\mu}$ analysis [104]. All tracks were varied by the same amount at the same time; we thus assume full correlation here. To assess the resolution smearing the p_T of each track is then separately subjected to Gaussian smearing. After all this the kinematics of the track-jet in question are recomputed.

Track-jet track efficiency The efficiency of picking up a track is not the same in data as in simulation [105]. The systematic arising from this mismatch is modeled by randomly removing tracks from a given track-jet according to the tracking efficiency measured in [105] and recalculating the kinematics of the track-jet. The track efficiency is different inside the core of very hard ($p_T > 400$ GeV) calo-jets. The same method is therefore separately applied to such calo-jets. If a track thus removed is shared with a track-jet then it is removed from the track-jet as well. Both jets have their kinematics recalculated in this case.

Lepton uncertainties Uncertainties with leptons have been investigated before [106] and have been incorporated in this analysis via the use of scale factors and their associated uncertainties. This includes electron and muon reconstruction and trigger efficiencies as well as p_T-scale and resolution. It should be mentioned that, since leptons are only used for the preselection and do not change the flavor composition of the final selection much, the contribution of these uncertainties is negligible on the final b-tagging efficiency.

Fake lepton contribution The fake lepton contribution can affect the flavor composition of the final selection and therefore affect the measured b-tagging efficiencies.

[†]The normalization will only affect the simulation samples. And thus it will only affect the b-tagging efficiency in Monte Carlo. However, this efficiency is the number of probe-jets passing MV1 divided by all of them. The cross-section enters as a factor in both the nominator and the denominator and so cancels out.

Uncertainty [%]	$7-20$	$20-30$	$30-60$	$60-100$	$100-250$
Hadronization	-0.3	-0.9	-0.6	1.3	-0.4
$t\bar{t}$ generator up	-1.3	-1.1	1.1	0.2	3.7
$t\bar{t}$ generator down	1.3	1..1	-1.1	-0.2	-3.7
Wt generator up	2.0	-0.3	0.0	-0.2	-1.2
Wt generator down	-2.0	0.3	-0.0	0.2	1.2
ISR/FSR up	-1.4	-0.9	-0.9	-1.6	-0.4
ISR/FSR down	1.4	0.9	0.9	1.6	0.4
Top p_T correction up	-0.1	-0.0	-0.0	0.1	0.7
Top p_T correction down	0.1	0.0	0.0	-0.1	-0.6
Track efficiency up	-1.2	-0.4	0.2	0.2	0.2
Track efficiency down	1.2	0.4	-0.2	-0.2	-0.2

Table A.4.2: *Composition of the systematic uncertainty on the b-tagging efficiency for $R = 0.4$ track-jets in bins of transverse momentum [GeV]. Systematics with an effect of less than 0.5 % in all bins were suppressed.*

Because the small number of fake lepton events passing the selection cuts the effects are very small. The normalization was varied $\pm 50\%$ to gauge the systematic as was done in a previous calibration [94].

Pile-up All simulation samples were reweighted so that the mean number of proton-proton collisions per bunch crossing matches the data. To assess any uncertainties arising from this the samples were reweighted to vary the mean number of collisions by $\pm 10\%$ with respect to the data.

Top p_T spectrum It was seen in previous studies [107, 108] that the top p_T spectrum is mismodelled by simulations. We follow [109] in reweighing events to match the p_T spectrum in data. Half the correction is taken as the systematic uncertainty.

A.5 Conclusion

In this calibration we have derived scale factors that can be used to correct simulation samples to data with respect to b-tagging efficiencies. The scale factors are close to unity, and often equal to unity within the uncertainty. This shows the ATLAS simulation seems to do a good job at describing and predicting the data for b-tagging. This was the first calibration of track-jets, which offer some interesting advantages with respect to orthodox calorimeter-jets. The results of this calibration have been used within other analyses within ATLAS where bottom quarks play an important role.

Uncertainty [%]	$7-20$	$20-30$	$30-60$	$60-100$	$100-250$
Hadronization	-0.3	-0.9	-0.6	1.3	-0.4
$t\bar{t}$ generator up	-1.3	-1.1	1.1	0.2	3.7
$t\bar{t}$ generator down	1.3	1.1	-1.1	-0.2	-3.7
Wt generator up	2.0	-0.3	0.0	-0.2	-1.2
Wt generator down	-2.0	0.3	-0.0	0.2	1.2
ISR/FSR up	-3.2	-1.0	-0.7	-0.7	-2.1
ISR/FSR down	3.2	1.0	0.7	0.7	2.1
Top p_T correction up	-0.1	-0.0	-0.0	0.0	0.7
Top p_T correction down	0.1	0.0	0.0	-0.0	-0.6
Track efficiency up	-2.4	-0.9	-0.2	-0.1	0.0
Track efficiency down	2.4	0.9	0.2	0.1	-0.0

Table A.4.3: *Composition of the systematic uncertainty on the b-tagging scale factor for $R = 0.4$ track-jets in bins of transverse momentum [GeV]. Systematics with an effect of less than 0.5 % in all bins were suppressed.*

Bibliography

[1] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716: 1–29, 2012. *arXiv:1207.7214*.

[2] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012. *arXiv:1207.7235*.

[3] D. Hanneke, S. Fogwell, and G. Gabrielse. New Measurement of the Electron Magnetic Moment and the Fine Structure Constant. *Phys. Rev. Lett.*, 100:120801, 2008. *arXiv:physics.atom-ph/0801.1134*.

[4] UA2 Collaboration. An improved determination of the ratio of W and Z masses at the CERN pp collider. *Phys. Lett*, B3:354–364, 1992.

[5] T. Song and H. Berrehrah. The hadronization time of heavy quark in nuclear matter. *Phys. Rev. C*, 94:034901, 2016. *arXiv:nucl-th/1601.04449*.

[6] D0 Collaboration. Evidence for production of single top quarks. *Phys. Rev.*, D78: 012005, 2008. *arXiv:hep-ex/0803.0739*.

[7] S. Weinberg. *The Quantum Theory of Fields: Volume I*. Cambridge University Press, 1995. ISBN978-0521670531.

[8] C. Patrignani et al. (Particle Data Group). 2016 Review of Particle Physics. *Chin. Phys. C*, 40:100001, 2016.

[9] E. W. Kolb and M. S. Turner. *The Early Universe*. Perseus Press, 1994. ISBN978-0201626742.

[10] N. Fornengo. Dark matter overview. *XXV ECRS Proceedings*, 2016. *arXiv:astro-ph.HE/1701.00119*.

[11] A. Joyce, L. Lombriser, and F. Schmidt. Dark Energy vs. Modified Gravity. *Annu. Rev. Nucl. Part. Sci.*, 66:95–122, 2016. *arXiv:astro-ph.CP/1601.06133*.

[12] A. Strumia and F. Vissani. Implications of neutrino data circa 2005. *Nucl. Phys. B*, 726:294–316, 2005.

[13] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia. Higgs mass and vacuum stability in the Standard Model at NNLO. *JHEP*, 08:098, 2012. *arXiv:hep-ph/1205.6497*.

[14] P. Grange, J. F. Mathiot, B. Mutet, and E. Werner. Aspects of fine-tuning of the Higgs mass with finite field theories. *Phys. Rev. D*, 88:125015, 2013.

[15] R. H. P. Kleiss. Pictures, Paths, Particles and Processes. Lecture notes of the Theoretical Foundations of Elementary Particle Physics course at the Radboud University Nijmegen, November 2013.

[16] J. Lykken and M. Spiropulu. The future of the Higgs boson. *Phys. Today*, 66:28, 2013. *http://dx.doi.org/10.1063/PT.3.2212*.

[17] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.

[18] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[19] A. Kobakhidze and J. Yue. Electroweak Baryogenesis with Anomalous Higgs Couplings. *JHEP*, 04:011, 2016. *arXiv:hep-ph/1512.08922*.

[20] A. Kusenko, L. Pearce, and L. Yang. Postinflationary Higgs Relaxation and the Origin of Matter-Antimatter Asymmetry. *Phys. Rev. Lett.*, 114:061302, 2015.

[21] M. Tegmark. The Interpretation of Quantum Mechanics: Many Worlds or Many Words? *Fortsch. Phys.*, 46:855–862, 1997. *arXiv:quant-ph/9709032*.

[22] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov. On anomalous electroweak baryon-number non-conservation in the early universe. *Phys. Lett.*, B155:36–42, 1985. doi: 10.1016/0370-2693(85)91028-7.

[23] K. Assamagan, C. Chen, J. P. Chou, D. Curtin, M. A. Fedderke, et al. The Higgs Portal and Cosmology. *ACFI-T16-10*, 2016. *arXiv:hep-ph/1604.05324*.

[24] D. Maity. Minimal Higgs inflation. 2016. *arXiv:hep-ph/1606.08179*.

[25] X. Zhang. Operator Analysis for the Higgs Potential and Cosmological Bound on the Higgs-Boson Mass. *Phys. Rev.*, D47:3065–3067, 1993. *arXiv:hep-ph/9301277*.

[26] C. Grojean, G. Servant, and J. Wells. First-Order Electroweak Phase Transition in the Standard Model with a Low Cutoff. *Phys. Rev.*, D71:036001, 2005. *arXiv:hep-ph/0407019*.

[27] S. Coleman and E. Weinberg. Radiative Corrections as the Origin of Spontaneous Symmetry Breaking. *Phys. Rev.*, D7:1888–1910, 1973. doi: 10.1103/PhysRevD.7.1888.

[28] M. Slawinska, W. van den Wollenberg, B. van Eijk and S. Bentvelsen. Phenomenology of the trilinear Higgs coupling at proton-proton colliders. *arXiv:1408.5010*.

[29] S. Dawson, S. Dittmaier, and M. Spira. Neutral Higgs boson pair production at hadron colliders: QCD corrections. *Phys. Rev.*, D58:115012, 1998. *arXiv:9805244*.

[30] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke. Higgs boson pair production in gluon fusion at NLO with full top-quark mass dependence. *Phys. Rev. Lett.*, 117:012001, 2016. *arXiv:hep-ph/1604.06447*.

[31] D. de Florian and J. Mazzitelli. Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD. *Phys. Rev. Lett.*, 111:201801, 2013. *arXiv:1309.6594*.

[32] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *arXiv:1101.0593*.

[33] J. Baglio, A. Djouadi, R. Grober, M. Muhlleitner, J. Quevillon, et al. The measurement of the Higgs self-coupling at the LHC: theoretical status. *JHEP*, 1304:151. *arXiv:1212.5581*.

[34] E. N. Glover and J. van der Bij. Higgs boson pair production via gluon fusion. *Nucl. Phys.*, B309:282, 1988.

[35] T. Plehn, M. Spira and P. Zerwas. Pair production of neutral Higgs particles in gluon-gluon collisions. *Nucl. Phys*, B479:46–64, 1996. *arXiv:9603205*.

[36] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 1407:079, 2014. *arXiv:1405.0301*.

[37] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002. *arXiv:hep-ph/0201195*.

[38] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, et. al. Higgs pair production at the LHC with NLO and parton-shower effects. *Phys. Lett.*, B732: 142–149, 2014. *arXiv:1401.7340*.

[39] ATLAS Collaboration. Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 94:052002, 2016. *arXiv:1606.04782*.

[40] ATLAS Collaboration. Search For Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector. *Phys. Rev. Lett.*, 114:081802, 2015. *arXiv:1406.5053*.

[41] ATLAS Collaboration. Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Eur. Phys. J.*, 75: 412, 2015. *arXiv:1506.00285*.

[42] ATLAS Collaboration. Searches for Higgs boson pair production in the $hh \to bb\tau\tau$, $\gamma\gamma WW^*$, $\gamma\gamma bb$, $bbbb$ channels with the ATLAS detector. *Phys. Lett. D*, 92:092004, 2015. *arXiv:1509.04670*.

[43] CMS Collaboration. Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 8 TeV. *Phys. Lett. B*, 749:560, 2015. *arXiv:1503.04114*.

[44] U. Baur, T. Plehn, and D. L. Rainwater. Probing the Higgs selfcoupling at hadron colliders using rare decays. *Phys. Rev.*, D69:053004, 2004. *arXiv:0310056*.

[45] M. J. Dolan, C. Englert, and M. Spannowsky. Higgs self-coupling measurements at the LHC. *JHEP*, 1210:112, 2012. *arXiv:1206.5001*.

[46] M. Spira. HPAIR program. http://tiger.web.psi.ch/proglist.html.

[47] S. Frixione and B. Webber. Matching NLO QCD calculations and parton shower simulations. *JHEP*, 06:029, 2002. *arXiv:hep-ph/0204244*.

[48] H. H. Lai et al. New parton distribution functions for collider physics. *Phys. Rev.*, D82:074024, 2010. *arXiv:1007.2241*.

[49] NNPDF Collaboration, R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, and J. Rojo. Parton distribution with QED corrections. *Nucl. Phys.*, B877:290–320, 2013. *arXiv:hep-ph/1308.0598*.

[50] ATLAS Collaboration. Measurement of the diphoton in association with jet cross sections with the ATLAS detector at $\sqrt{s} = 8\,TeV$. *STDM-2016-12*, 2016.

[51] B. Humpert and R. Odorico. Multi-parton scattering and QCD radiation as sources of four-jet events. *Phys. Rev. Lett. B*, 154:211–214, 1985.

[52] ATLAS Collaboration. Study of hard double-parton scattering in four-jet events in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment. *JHEP*, 11:110, 2016. *arXiv:hep-ex/1608.01857*.

[53] ATLAS Collaboration. ATLAS Public Luminosity Results Run I. 2016. *https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults*.

[54] ATLAS Collaboration. Improved luminosity determination in pp collisions at $\sqrt{s} = 7\,TeV$ using the ATLAS detector at the LHC. *submitted to Eur. Phys. J.*, . *arXiv:1302.4393*.

[55] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC. . *arXiv:1608.03953*.

[56] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3:S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.

[57] Joao Pequenao and Paul Schaffner. An computer generated image representing how ATLAS detects particles. *CERN-EX-1301009*, 2013. *https://cds.cern.ch/record/1505342*.

[58] M. Lyman and H. R. Varian. How Much Information. 2010. *http://www2.sims.berkeley.edu/research/projects/how-much-info/how-much-info.pdf*.

[59] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne and R. Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010. *https://cds.cern.ch/record/1291633*.

[60] S. Eidelman et al. Review of Particle Physics. *Physics Letters B*, 592:1+, 2004. *http://pdg.lbl.gov*.

[61] R. Placakyte. Parton Distribution Functions. 2011. *arXiv:hep-ph/1111.5452*.

[62] P. Skands S. Mrenna, T. Sjostrand. PYTHIA 6.4 physics and manual. *JHEP*, 05: 0265, 2006. *arXiv:hep-ph/0603175*.

[63] P. Skands T. Sjostrand, S. Mrenna. A brief introduction to PYTHIA 8.1. *Comp. Phys. Comm.*, A178:852, 2008. *arXiv:0710.3820*.

[64] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert, J. Winter. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009. *arXiv:0811.4622*.

[65] T. Sjostrand and M. van Zijl. A multiple-interaction model for the event structure in hadron collisions. *Phys. Rev.*, D36:2019, 1987.

[66] B. R. Webber. A QCD model for jet fragmentation including soft gluon interference. *Nucl. Phys. B*, 238:492, 1984.

[67] Talk given by T. Sjostrand at IPPP Durham. Old ideas in hadronization: The Lund String. 2009. *http://home.thep.lu.se/ torbjorn/talks/durham09.pdf*.

[68] D. Amati and G. Veneziano. Preconfinement as a Property of Perturbative QCD. *Phys. Lett.*, B83:87–92, 1979.

[69] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instr. Meth.*, A506:250, 2003.

[70] ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70: 823, 2010. *arXiv:1005.4568*.

[71] ATLAS Collaboration. Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at sqrts = 7 TeV in 2011. *ATLAS-CONF-2012-048*, 2012. *https://cds.cern.ch/record/1450089*.

[72] M. Cacciari, G. P. Salam, and G. Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008. *arXiv:0802.1189*.

[73] M. Cacciari and G. P. Salam. Dispelling the N^3 myth for the Kt jet-finder. *arXiv:0512210*.

[74] ATLAS Collaboration. Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data. *ATLAS-CONF-2011-102*, . *https://cdsweb.cern.ch/record/1369219*.

[75] ATLAS Collaboration. Performance of b-Jet Identification in the ATLAS Experiment. *JINST*, 11:P04008, 2016. *arXiv:hep-ex/1512.01094*.

[76] ATLAS Collaboration. Expected photon performance in the ATLAS experiment. *ATL-PHYS-PUB-2011-007*, . *https://cds.cern.ch/record/1345329*.

[77] J. R. Forshaw J. M. Butterworth and M. H. Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys.*, C72:637, 1996. *arXiv:hep-ph/9601371*.

[78] ATLAS Collaboration. Further ATLAS tunes of PYTHIA6 and Pythia8. *ATL-PHYS-PUB-2011-014*, . *http://cds.cern.ch/record/1400677*.

[79] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 0711:070, 2007. *arXiv:hep-ph/0709.2092*.

[80] ATLAS Collaboration. Summary of ATLAS Pythia 8 tunes. *ATL-PHYS-PUB-2012-003*, . *https://cds.cern.ch/record/1474107*.

[81] M. L. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003. *arXiv:hep-ex/0206293*.

[82] ATLAS Collaboration. New ATLAS event generator tunes to 2010 data. *ATLAS-PHYS-PUB-2011-008*, . *https://cdsweb.cern.ch/record/1345343*.

[83] ATLAS Collaboration. Measurement of the isolated di-photon cross section in pp collisions at $\sqrt{s} = 7\,TeV$ with the ATLAS detector. *ATL-COM-PHYS-2012-1078*, . *https://cds.cern.ch/record/1463021*.

[84] A. Schwartzman. Jet energy calibration at the LHC. *arXiv:1509.05459*.

[85] ATLAS Collaboration. Signal studies in $H \to \gamma\gamma$ search with 8TeV data. *ATLAS-COM-PHYS-2012-755*, . *https://cds.cern.ch/record/1453770*.

[86] Anulli, Fabio and Bieniek, S and Jansen, E and Konstantinidis, N and Salvatore, Daniela and Verducci, Monica. Measurement of the $b\bar{b}$-dijet cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. Technical Report ATL-COM-PHYS-2013-618, CERN, Geneva, May 2013. *https://cds.cern.ch/record/1547076*.

[87] V. Blobel. Unfolding Methods in High-energy Physics Experiments. *DESY-84-118*, 1984. *inspirehep.net/record/207948*.

[88] A. Hocker and V. Kartvelishvili. SVD Approach to Data Unfolding. 1995. *arXiv:hep-ph/9509307v2*.

[89] Forsythe, George E. and Malcolm, Michael A. and Moler, Cleve B. *Computer Methods for Mathematical Computations*. Prentice Hall Professional Technical Reference, 1977. ISBN 0131653326.

[90] C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems*. Series in Automatic Computation. Prentice-Hall, Englewood Cliffs, NJ 07632, USA, 1974.

[91] N. P. Hartland. Double Higgs production from $HH \to (b\bar{b})(b\bar{b})$ at a 100 TeV hadron collider. *arXiv:1609.00254*.

[92] N. Haba, K. Kaneta, Y. Mimura, and E. Tsedenbaljir. Higgs Pair Production at the LHC and ILC from general potential. *Phys. Rev. D*, 89:015018, 2014. *arXiv:hep-ph/1311.0067*.

[93] W. van den Wollenberg, M. Kagan, E. Thompson, and L. Zhou. Measurement of the Track-Jet b-Tagging Efficiency of the MV1 algorithm in pp Collisions at $\sqrt{s} = 8\,TeV$ using $e\mu$ Dilepton $t\bar{t}$ events. *ATL-COM-PHYS-2015-009*. *https://cds.cern.ch/record/1981533*.

[94] R. Hawkings. Measurement of the b-tagging efficiency of the MV1 algorithm in pp collisions at $\sqrt{s} = 8\,TeV$ using $e\mu$ dilepton $t\bar{t}$ events. *ATLAS-CONF-2013-381*. *https://cdsweb.cern.ch/record/1537694*.

[95] ATLAS Collaboration. ATLAS Top 2012 Monte Carlo samples. 2013. *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopMC12*.

[96] P. Z. Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys. Rev.*, D82:074018, 2010. *arXiv:1005.3457*.

[97] G. Corcella et al. HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 101:010, 2011. *arXiv:hep-ph/0011363*.

[98] B. P. Kersevan and E. Richter-Wąs. The Monte Carlo event generator AcerMC version 2.0 with inferfaces to PYTHIA 6.2 and HERWIG 6.5. *arXiv:hep-ph/0405247*.

[99] ATLAS Collaboration. Measurement of $t\bar{t}$ production with a veto on additional central jet activity in pp collisions at $\sqrt{s} = 7\,TeV$ using the ATLAS detector. *Eur. Phys. J.*, C72:2043, 2013. *arXiv:1203.5015*.

[100] ATLAS Collaboration. Object selection and calibration, background estimations and MC samples for the winter 2013 top quark analyses with 2012 data. *ATLAS-COM-PHYS-2013-088*, . *https://cdsweb.cern.ch/record/1509562*.

[101] ATLAS Collaboration. Calibration of b-tagging using di-leptonic top pair events and a combinatorial likelihood approach. *ATLAS-CONF-2013-395*, . *https://cdsweb.cern.ch/record/1567111*.

[102] M. Aliev et al. HATHOR - Hadronic top and heavy quark cross section calculator. *Comp. Phys. Comm.*, A182:1034, 2011. *arXiv:1007.1327*.

[103] N. Kidonakis. Two-loop anomalous dimensions for single top quark associated production with a W^- or H^-. *Phys. Rev.*, D82:054018, 2010. *arXiv:1005.4451*.

[104] ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data. *Eur. Phys. J.*, C74:3130, 2014. *arXiv:1407.3935*.

[105] ATLAS Collaboration. Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC. *New J. Phys.*, 13:053033, 2011. *arXiv:1012.5104*.

[106] ATLAS Collaboration. Object selection and calibration, background estimations and MC samples for the winter 2013 top quark analyses with 2012 data. *ATLAS-COM-PHYS-2013-088*, . *https://cdsweb.cern.ch/record/1509562*.

[107] ATLAS Collaboration. Measurement of top quark pair differential cross-section in the l+jets channel in pp collisions at $\sqrt{s} = 7\,TeV$ using the ATLAS detector. *ATLAS-CONF-2013-099*, . *https://cdsweb.cern.ch/record/1600778*.

[108] ATLAS Collaboration. Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum using the ATLAS detector in $\sqrt{s} = 8\,TeV$ proton-proton collisions. *ATLAS-CONF-2014-057*, . *https://cdsweb.cern.ch/record/1951328*.

[109] ATLAS Collaboration. Search for the Standard Model Higgs boson in associated production with a vector boson and decaying to bottom quarks with the ATLAS detector. *ATLAS-COM-PHYS-2013-465*, . *https://cdsweb.cern.ch/record/1543067*.

English summary

"Greed has poisoned men's souls, has barricaded the world with hate, has goose-stepped us into misery and bloodshed. We have developed speed, but we have shut ourselves in. Machinery that gives abundance has left us in want. Our knowledge has made us cynical. Our cleverness, hard and unkind. We think too much and feel too little. More than machinery we need humanity. More than cleverness we need kindness and gentleness. Without these qualities, life will be violent and all will be lost...."

- Charlie Chaplin, The Great Dictator

WITH this appendix the author helpfully provides a summary of the work described in this thesis for the layperson. We will here and there gloss over details or take some liberties with certain facts in order to provide a coherent intuitive story. Anyone desiring to learn more may find additional information in any standard textbook on the subject. It is hoped that after reading this summary the layperson will have a clearer picture of the author's work.

Nature and the Higgs mechanism

ALL things around us, chairs, tables, aeroplanes, animals, and even you and me, are made up of molecules, which in turn are made up of atoms. An atom is a collection of even tinier particles; neutrons and protons clumped together in the nucleus surrounded by a whirling cloud of electrons. As far as we know an electron does not consist of smaller particles; it is as we call it an elementary particle. Protons and neutrons do consist of smaller particles, the quarks, which are, in turn, elementary particles again, see figure A.5.1. Outside of these there are other elementary matter particles, all of which are either unstable[†] or difficult to detect so we do not see them much in

[†]This means that they rapidly fall apart into a few smaller lighter particles. Note that this does not mean the mother particle consisted out of the daughter particles; rather the mother particle changes into the daughter ones in the moment of decay. Particles can typically decay in several different sets of daughter particles and the set is chosen at random at decay.

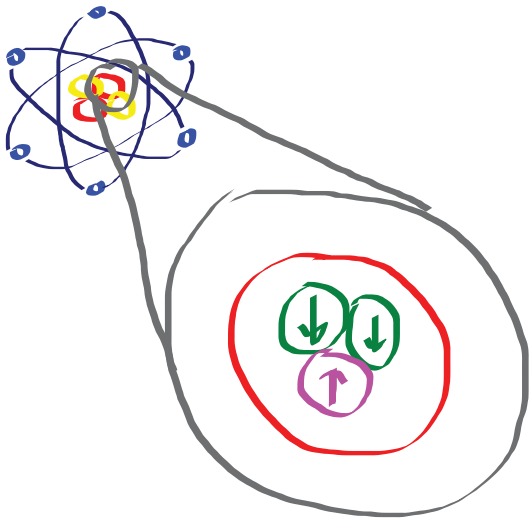

Figure A.5.1: *Author's schematic drawing showing an atom in the upper left with yellow protons and red neutrons in the core, surrounded by blue electrons. The zoomed-in picture in the lower right shows the interior of a neutron containing two down-quarks and one up-quark. The drawing is not to scale. If it were the atomic nucleus would be invisible compared to the orbit of the electrons.*

everyday life. Of these only the b-quark warrants any attention from us. It is a particle quite similar to the quarks found inside the protons and neutrons except heavier and unstable, and is therefore only created in high energy particle collisions.

In addition to this we should have a brief look at the way fundamental forces are treated in modern particle physics. Here forces are ascribed elementary messenger particles, which in a sense *carry* a force. For example the electromagnetic force, responsible for things like magnetism, electricity, light, but also the entire field of chemistry, is conveyed via the use of the photon. The photon is a stable massless particle that interacts with elementary particles that carry an electric charge. A large number of these photons interacting with charges then gives rise to the phenomena of electromagnetism. Physics has identified four such fundamental forces: electromagnetism, gravity, and the weak and strong nuclear forces. All other forces may be derived from these. The weak nuclear force is related to radioactivity and particle decay. Its messenger particles are heavy and unstable themselves and are known as the W and Z bosons. The strong nuclear force is mediated by the massless gluon, and which is responsible for keeping the neutrons and protons inside an atomic nucleus together. It does this by acting on the quarks found in these.

Incidentally it is the huge binding energy of the atomic nucleus that finds its origin in the strong nuclear force which is unleashed in atomic explosions and exploited in nuclear reactors. This binding energy also provides a large fraction of the mass of a proton or neutron. A proton for example consists of three quarks[†], but these only contribute a small part of the total mass of the proton. The rest is supplied by the

[†]More accurately these three quarks are known as the valence quarks. In truth a proton is a complicated composite object consisting not only of the three valence quarks but also of a virtual sea of other quarks and gluons. We will neglect these details here.

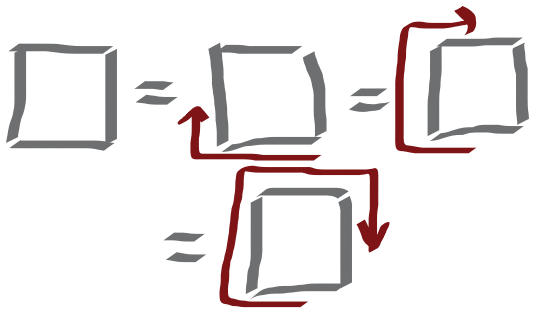

Figure A.5.2: *Author's schematic drawing illustrating how a square is symmetric (looks the same) under 90 degree rotations.*

binding energy caused by the gluon-quark interactions, which by Einstein's famous $E = mc^2$, manifests itself as an additional mass.

The physics of elementary forces are described by the so-called Standard Model. The mathematical underpinnings of this model are built upon symmetries. A symmetry is a property of an object that remains the same under some kind of operation. For example, consider the square drawn in figure A.5.2. The square appears (assuming the square is perfect, and not the poorly drawn copy like in the figure), the same when we rotate it 90 degrees. It looks the same still when rotated around 180 degrees, or in fact any multiple of 90 degrees. We say that the square is *invariant under* 90 degree rotations. The same principle applies to the Laws of Nature themselves. For example, the Laws of Nature are the same everywhere in the Universe. This means the laws are *invariant under translation*.

Physicists realized that symmetries of the laws of nature can actually be used to explain some of these laws. For example it can be shown that the translational invariance mentioned before leads to the conservation of (linear) momentum. In the same way the conservation of energy can be derived from the requirement that the Laws of Nature are static in time. On top of this it is the case that some more esoteric symmetries can be used to derive the four fundamental forces. It is unfortunately beyond the scope of this meager summary to show how and why, save to say that from these symmetries one can derive the properties of the messenger particles, like the photon and the gluon, and see how they interact with ordinary matter. The application of symmetries in this way is very powerful; one replaces a large number of properties that the fundamental forces have, by a very small statement of symmetry.

As powerful and successful this symmetry method turned out to be, it was soon realized that there was a small problem. And this is the problem of mass. We know from measurements that quite a large number of elementary particles have a non-zero restmass. The issue with mass is related to the symmetry that gives rise to the weak nuclear force. This symmetry forbids the inclusion of massive particles. Or, stated the other way around, massive particles preclude the weak nuclear force symmetry. Since we do see that particles have mass this means that the weak nuclear force is not describable using the symmetry method.

A way out was discovered by several people, most notably François Englert, Peter Higgs,

and Robert Brout[†], in what was to become known as the Higgs mechanism. This mechanism predicts the existence of a new elementary particle, the Higgs boson, and a pervasive Higgs field. To see how this works we must first introduce the concept of a particle field.

The Standard Model is a model of a particular kind of *quantum field theory* (QFT). In QFT a particle is nothing more than an excitation of an underlying field. For example, the electron is a ripple in, or distortion of, something called the electron field. This field is supposed to be stretched out across the Universe. It exists everywhere, will forever exist, and has existed forever[‡]. This is comparable to the classical notion of a magnetic field; it exists everywhere, it is just that close to a magnet it is perturbed and warped. In fact a magnetic field is a form of electromagnetic field, which is basically just the photon field.

Just like for the classical case of magnetism the field itself can contain energy. The energy is related to field strength, and this relation is called the potential. For most fields, and certainly for the electromagnetic field, the energy inside the field amounts to zero if the field strength is zero. This makes intuitive sense; a zero magnetic field is like having no magnetic field, and no magnetic field means no field in which to store energy.

In the early Universe, shortly after the Big Bang, the Universe was very hot and very compact. In these cramped conditions a lot of energy is stored in the various fields. However, as the Universe expanded and cooled this energy will diminish until minima were reached. These minima correspond to a disappearing field strength, see the left side of figure A.5.3. This is comparable to a magnetic field that has lost all energy and thus the field strength becomes zero.

The Higgs mechanism introduces a new field, the Higgs field, whose excitations are called Higgs bosons. This field does not follow the normal rule of having zero energy in the case of zero field strength. It's potential does have energy minima, and it is in one of these points that the Universe will settle after cooling. These minima, however, do not correspond to a field strength of zero. This means that even at the current era the Universe is bathed in a non-zero Higgs field, see the right side of figure A.5.3.

As certain particles travel through this field they interact with this non-zero Higgs field. Just like the binding energy of the gluons holding together quarks in protons and neutrons there is an energy associated with these interactions. Again via Einstein's $E = mc^2$ this energy manifests itself as a mass. Thus, all the particles the Higgs field interacts with gain a mass that is proportional to the strength of this interaction. Different particles have different Higgs interaction strengths and therefore have different masses. Some particles, like the photon, do not directly interact with the Higgs field and therefore remain massless.

At this juncture one might accuse the physicists of just reinterpreting particles' masses and that it does not really seem to solve the issue with the weak nuclear force symmetry. Surely with these added masses the symmetry is still as violated as before? Technically this is true, but what the Higgs mechanism provides for is a way for this symmetry to be broken spontaneously. The Laws of Nature are now symmetric under the required

[†]The latter unfortunately missed out on the eventual Nobel prize because he died in 2011, too early for the 2013 award and also for the 2012 discovery of the Higgs boson at the LHC.

[‡]Of course the Universe as we know it seems to have a beginning in the Big Bang, and it might also have a similar end, but to simplify the mathematics it is often easier to remove these *end-points*.

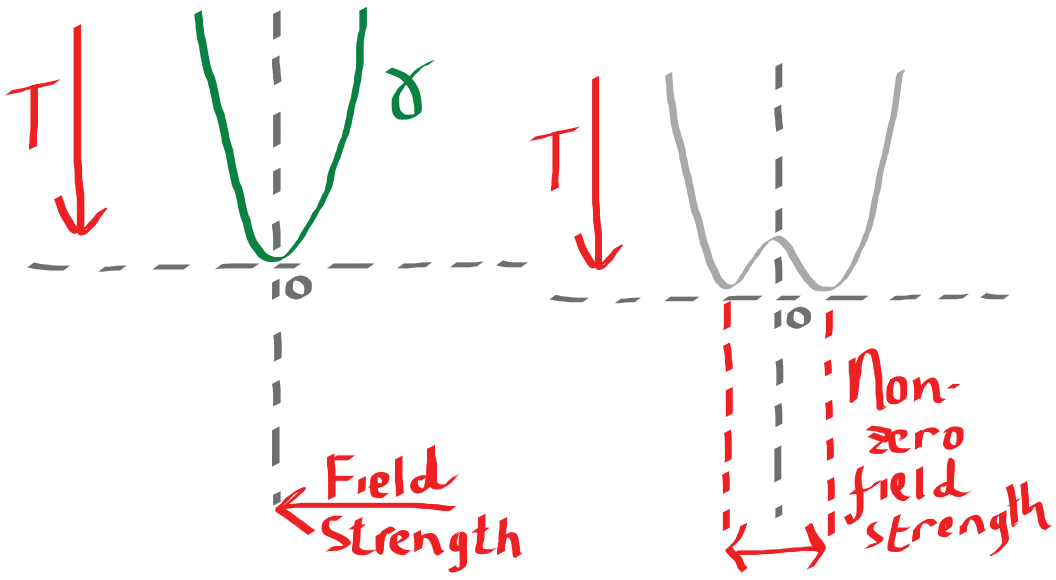

Figure A.5.3: *Author's schematic drawings illustrating the photon (electromagnetic) potential on the left, which is a function of the field strength in terms of the temperature in the Universe. The field strength is zero at low temperature and energy. On the right the Higgs potential whose field strength does not go to zero in the energy minima. Note also that a choice must be made for either minimum when the energy drops.*

symmetry, it is just that the groundstate, i.e. the lowest energy state, is not. So when the Universe cools the symmetry breaks at some point, and only then do the particles get their mass. This allows us to keep on use the symmetry principle and also have particles that have mass.

We have strong indications that this picture of reality is right. In 2012 the Higgs boson was discovered at the *Large Hadron Collider* (LHC) at CERN, Geneva, by two different experiments: the *Compact Muon Solenoid* (CMS) and *A Toroidal LHC ApparatuS* (ATLAS). At the LHC protons are accelerated to very high energies and then made to smash into each other, creating many new particles in the debris. This discovery shows that the Higgs field exists. It does not however establish that this Higgs field has the required properties for it to generate a mass for the other particles. For this we need to measure the potential of the Higgs field and show that it has minima with a non-zero field strength.

Phenomenology and motivation

MEASURING the Higgs potential turns out to be tricky. The crucial parameter needed to discover the shape of it is the so-called Higgs trilinear coupling, λ_{3h}, and it governs how strongly the Higgs boson interacts with itself. The difficulty lies in that in order to measure it we not only have to create a Higgs boson in the LHC (a feat on its own), but we need that Higgs boson to emit another Higgs boson as well. Such an event is exceedingly rare, and we need for it to happen enough times so that sufficient

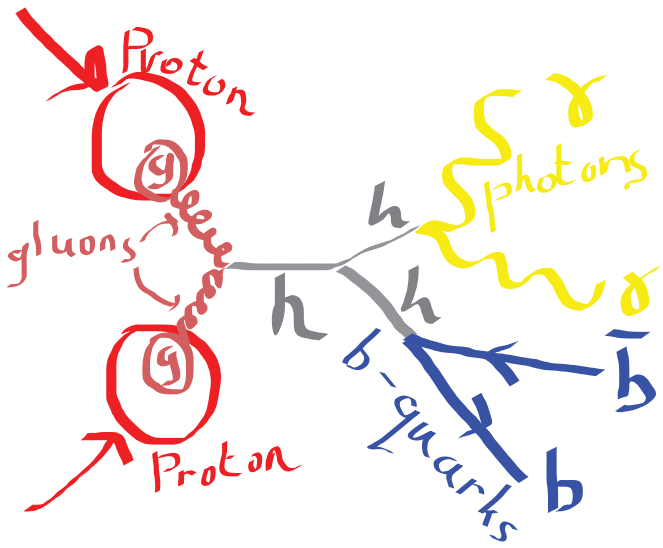

Figure A.5.4: *Author's schematic drawing of the $pp \to hh \to b\bar{b}\gamma\gamma$ process. On the left two protons collide in which, for example, two gluons interact to form a Higgs boson (h in the picture). The Higgs boson then splits into two, which is the event that involves λ_{3h}, before one Higgs boson decays to a photon pair, and the other one to a b-quark pair.*

statistics may be collected. In addition there are many processes that can be generated by the LHC[†] that might appear like such an event, but are not. One will need ways to reduce the effects of these so-called background processes in the analysis.

In this thesis we first analyzed which processes are the most promising ones to examine with respect to their sensitivity to λ_{3h}. We need to do this because the Higgs boson cannot be directly observed. It is an unstable particle and will therefore decay into different other particles before reaching the detector. We have two Higgs bosons in our process, one original one, and another one emitted by the first one, both of which decay. We find that a process in which one of the two Higgs bosons decays to a b-quark pair and the other one into a photon pair is the most sensitive and has the least backgrounds. This process is known as $pp \to hh \to b\bar{b}\gamma\gamma$, where the p stands for a proton, h for a Higgs boson, γ for a photon, and b for a b-quark. $\bar{b}$ stands for the anti-matter partner of the b-quark but that detail is not really relevant for this summary. A visualization of this process can be found in figure A.5.4. In the same study we also found that the sensitivity for λ_{3h} is greatest at energies comparable to the Higgs boson's mass.

Unfortunately the coveted process was found to be too rare to detect in the dataset collected by ATLAS (which is the experiment with which the author was associated) at the present time. However, it was also found that the main background to any λ_{3h} measurement in $pp \to hh \to b\bar{b}\gamma\gamma$, which is $pp \to b\bar{b}\gamma\gamma$, has large theoretical uncertainties[‡]. The second process has the same outgoing particles, which is why it is a background, but no

[†]Note that physicists cannot choose what happens when two protons collide. Quantum mechanics teaches us that this is totally random. We may only say that certain events will happen with certain probabilities. We cannot tell the machine to produce only the event we are interested in and nothing else.

[‡]That it has large theoretical uncertainties means that the prediction the Standard Model gives for how often the process should occur has significant uncertainties in its calculation. Often these calculations are very complex and certain approximations had to be made.

longer requires those particles to originate from two Higgs bosons. This process is rare, but not as rare as the λ_{3h} one. Any future λ_{3h} measurement will need to deal with this irreducible background and so it was judged prudent to investigate this channel and measure the rate at which it occurs at the LHC.

The main analysis

OUR analysis thus aims at determining the rate with which the $pp \to b\bar{b}\gamma\gamma$ process occurs at the LHC. The main objects of interest are thus photons and jets. Jets are a large collection of particles that enter the detector in a cascade-like fashion and are more easily treated as a single composite object. The origins of jets are often gluons or quarks. These interact with the strong nuclear force, which has a property making it so that gluons and quarks decay into a shower of particles. The traits of any particular jet can be analyzed to determine what kind of quark created it. In our case we are interested in finding jets originating from b-quarks (these are known as b-jets) and the process of identifying these jets is known as b-tagging.

The ATLAS detector and software sometimes make mistakes when reconstructing a proton-proton collision, and these so-called 'fakes' form the main background to the analysis. We have electrons being wrongly reconstructed as photons and also jets masquerading as photons. In addition to these there is a background coming from non-b-jets that erroneously received a b-tag.

The electrons were dealt with by examining a different process where electrons dominate and then seeing how often these end up being regarded as photons by the detector. Armed with this knowledge one can write down several equations predicting how many events we should see in our own process with two photons, two electrons, and with one photon and one electron. Solving this system of equations then leads to the size of the electron background, which was found to be around 5.3%.

Dealing with the jet background was slightly more complicated. Similar to the three event categories used in the electron background we define sixteen categories by defining four different ones for each individual photon in the event. Since we have two photons this means we have $4 \times 4 = 16$ event categories. The four different photon categories were generated by asking the photon two questions. One is whether the photon was isolated, i.e. whether there were a lot of other particles detected near the photon, and the second is whether the photon passes certain quality cuts made to the way the photon interacted with the detector material.

For each of these sixteen categories one can again write down an equation predicting how many events should be there based on the total number of events measured, and on how often a jet is wrongly reconstructed as a photon (and some more things). This system of equations is too large to solve directly, so a computer algorithm finds the best possible solution and from this we estimate the jet background to be on the order of 10%.

The contribution of non-b-jets was gauged in the following way. We divided events into four categories based on the types of jets found in the event. One of these was the bb category, which is our signal, the others contained either only one b-jet or none at

all. The output of applying b-tagging on a jet is a number that can be interpreted as the probability is a true b-jet or not. One can study the distribution of this output number on simulated events which are fed through the same ATLAS event reconstruction software as the normal data was. One can also combine the distribution for the two jets into one single distribution and obtain shapes for each of the four different event categories. This we did, and afterwards this distribution was measured in data and the four different shapes were fit to it. The relative amplitude with which each shape is found in the data distribution corresponds to the fraction of events of the shape's type in data. Thus, the amplitude of the bb shape in the data distribution gives the signal contribution, and it was found to be around 54%.

After all the backgrounds are taken care of one needs to perform a so-called 'unfolding'. This is to revert any effects imposed upon the data by the detector itself[†]. If we do not do this it is impossible to compare the analysis result to any theoretical calculation. The unfolding itself is a technical story which we shall not go into here.

Finally after all this work one is rewarded by nice plots containing the measurement and theory predictions together. In the end we found that the LHC produced approximately 540 genuine $pp \to b\bar{b}\gamma\gamma$ events in 2012, which is a small number since there were about a 100 million proton-proton collisions generated per second in that year on average. We also found that the measurement and the theory predictions agree with each other within the various uncertainties. A plot comparing reality with theory can be found in figure A.5.5.

[†]For example the detector might incorrectly estimate the energy of photons by a certain factor. One then needs to divide all photon energies by this factor.

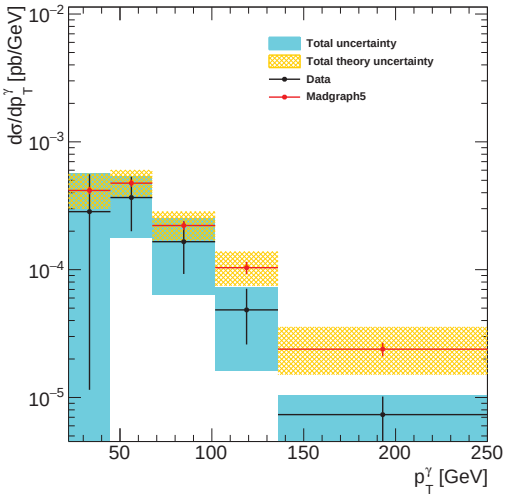

Figure A.5.5: *The distribution of the differential production cross-section (related to the production probability) of* $pp \to b\bar{b}\gamma\gamma$ *as a function of the leading photon transverse momentum as obtained from this analysis compared to the theory predictions in terms of the transverse momentum of the leading photon. The transverse momentum is related to the momentum, but excludes the direction along which the protons collide. The leading photon is the photon with the highest such momentum in the event. The* MADGRAPH5 *line refers to the theory prediction.*

Nederlandse samenvatting

"Alles wat enige status heeft, heeft namelijk ook een vorm van macht en macht corrumpeert altijd. Het moet geridiculiseerd kunnen worden, als dat niet meer kan, dan krijg je enge toestanden, een dictatuur of zoiets. En daarom moet dat gebeuren."

- Hans Teeuwen

Met dit hoofdstuk wil de auteur hulpvaardig een samenvatting geven voor een leek van hetgeen dat in deze dissertatie beschreven is. We zullen hier en daar wat technische details weglaten en af en toe wat vrijheden nemen met feiten zodat we een coherent begrijpbaar verhaal kunnen leveren. Eenieder die meer informatie wenst zal dit kunnen vinden in elk willekeurig tekstboek over moderne deeltjesfysica of kwantumveldentheorie. Het wordt gehoopt dat de lezer, na het lezen van deze samenvatting, zich een beter beeld kan vormen van het werk dat de auteur verricht heeft.

De natuur en het higgsmechanisme

Alle dingen rondom ons, stoelen, tafels, vliegtuigen, dieren, en zelfs jij en ik, zijn opgebouwd uit moleculen. Deze moleculen bestaan op hun beurt weer uit atomen. Een atoom is een verzameling van nog kleinere deeltjes: neutronen en protonen die samengeklonterd de atoomkern vormen en omringt zijn door een zwerm van elektronen. Zover wij weten bestaat het elektron niet uit kleinere deeltjes; het is een zogenaamd elementair deeltje. Protonen en neutronen bestaan wel uit kleinere deeltjes, de quarks. Quarks zijn wel elementaire deeltjes, zie figuur A.5.6. Naast deze deeltjes bestaan er ook nog andere elementaire materiedeeltjes. Deze zijn allemaal of instabiel[†], of moeilijk waar te nemen, en daarom komen we ze in het dagelijks leven niet vaak tegen. Van deze deeltjes vereist enkel het b-quark enige aandacht van ons. Dit deeltje is verwant aan de quarks die men kan vinden in het binnenste van protonen en neutronen met de uitzondering dat het veel zwaarder en instabiel is. Het kan daarom alleen gevonden worden in hoogenergetische deeltjesbotsingen.

[†]Dit houdt in dat het deeltje zeer snel uit elkaar valt in een aantal lichtere deeltjes.

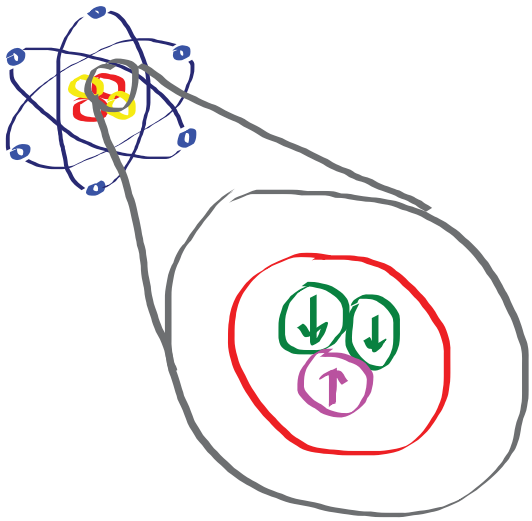

Figure A.5.6: *Schematische tekening door de auteur. Linksboven wordt een atoom weergegeven, met gele protonen en rode neutronen in de kern omcirkelt door blauwe elektronen. Rechtsbeneden wordt een neutron vergroot weergegeven. Zichtbaar zijn de twee down-quarks en de ene up-quark die het samen het neutron vormen. De tekening is niet op schaal. Als dat wel het geval zou zijn, was de atoomkern totaal onzichtbaar geweest ten opzichte van de grootte van de elektronbanen.*

Hiernaast zullen we ook een snelle blik moeten werpen op de manier waarop de fundamentele natuurkrachten behandeld worden in de moderne deeltjesfysica. Daar worden natuurkrachten boodschapperdeeltjes toebedeeld, welke de natuurkracht *dragen*. Neem als voorbeeld de elektromagnetische kracht, verantwoordelijk voor dingen zoals magnetisme, elektriciteit, licht, maar ook alle scheikunde. Deze kracht wordt overgedragen door het foton. Dit is een stabiel massaloos deeltje welke koppelt aan deeltjes die een elektrische lading hebben. Twee objecten met een tegengestelde elektrische lading trekken elkaar, bijvoorbeeld, aan. Dit gebeurt omdat deze twee objecten fotonen uitwisselen.

In de natuurkunde kunnen we vier van deze fundamentele natuurkrachten identificeren: elektromagnetisme, zwaartekracht, en de zwakke en sterke kernkrachten. Alle andere krachten der natuur kunnen uit deze vier afgeleid worden. De zwakke kernkracht is gerelateerd aan radioactiviteit en deeltjesverval, en wordt overgedragen door massieve instabiele deeltjes genaamd de W- en Z-bosonen. De sterke kernkracht heeft als boodschapperdeeltje het massaloze gluon en is verantwoordelijk voor het samenhouden van de neutronen en protonen in de atoomkern. Het doet zijn werk door interactie aan te gaan met de quarks die zich in deze deeltjes bevinden.

De sterke kernkracht leidt ook tot de enorme bindingsenergie[†] gevonden in een atoomkern die wordt ontketend in de explosie van een atoomwapen, of wordt uitgebuit in een nucleaire reactor. Deze bindingsenergie levert ook een groot deel van de massa van

[†]Als deeltjes een wisselwerking hebben kunnen ze een binding vormen (denk wederom aan objecten met een tegengestelde lading die elkaar aantrekken). Om deze binding te breken is energie nodig, de bindingsenergie.

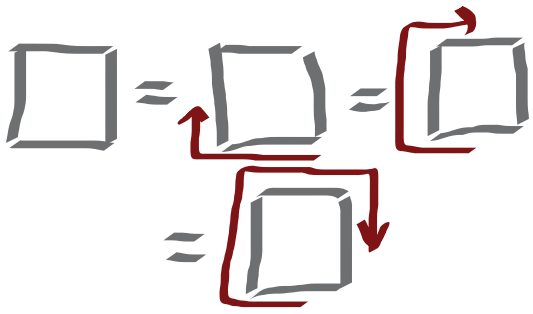

Figure A.5.7: *Schematische tekening door de auteur die laat zien hoe een vierkant symmetrisch is (er hetzelfde uitziet) onder rotaties van negentig graden.*

een proton of neutron. Een proton bijvoorbeeld bestaat uit drie quarks[†], maar deze zijn slechts verantwoordelijk voor een kleine fractie van de totale massa van het proton. De rest wordt geleverd door de bindingsenergie die voortkomt uit de quark-gluon interacties. Deze energie vertaalt zich via Einstein's bekende formule, $E = mc^2$, in extra massa. Deze formule laat zien dat massa niets anders is dan een verschijningsvorm van energie.

De regels van de fundamentele natuurkrachten worden beschreven door het zogenaamde Standaardmodel. De achterliggende wiskunde van dit model is gebaseerd op symmetrieën. Een symmetrie is een eigenschap van een object dat hetzelfde blijft onder een bepaalde operatie. Beschouw als voorbeeld het vierkant getekend in figuur A.5.7. Het vierkant ziet er (aangenomen dat het vierkant perfect is, en niet een slecht getekende kopie zoals in dit figuur) hetzelfde uit wanneer we het 90 graden draaien. Het zit er na een 180 graden rotatie nog steeds hetzelfde uit, en evenzo na 270 graden, of inderdaad een willekeurig aantal van 90 graden draaiingen. We zeggen dat het vierkant *invariant is onder* rotaties van 90 graden. Hetzelfde principe kan worden toegepast op de wetten der natuur. De wetten zijn bijvoorbeeld overal in het universum hetzelfde. Dit houdt in dat de wetten *invariant zijn onder translatie*.

Natuurkundigen realiseerden zich in de jaren zestig dat de symmetrieën van de natuurwetten zelfs gebruikt kunnen worden om enkele van deze wetten te verklaren. Het kan bijvoorbeeld worden aangetoond dat translatie invariantie leidt tot de wet van het behoud van (lineaire) impuls. Op dezelfde manier kan behoud van energie afgeleid worden uit het gegeven dat de natuurwetten niet veranderen in de tijd. Met behulp van deze symmetrieën zijn ook de eigenschappen van de boodschapperdeeltjes, zoals het foton en het gluon, af te leiden. En ook hoe deze deeltjes interactie hebben met de materiedeeltjes. Dit gebruik van symmetrieën is krachtig; we vervangen een groot aantal eigenschappen van de natuurkrachten door een handjevol symmetrieën.

Helaas werd er vrij snel gerealiseerd dat zich een addertje onder het gras bevond welke op het punt stond roet in het eten te gooien. Dit is het probleem van massa. We weten van metingen die gedaan zijn aan elementaire deeltjes dat een groot aantal van hen een massa heeft. Het probleem nu is dat de symmetrie waaruit de zwakke kernkracht zou volgen niet toestaat dat deze deeltjes een massa hebben. Of anders gezegd: deeltjes

[†]Om iets preciezer te zijn zullen we vermelden dat deze drie quarks bekend staan als de valentiequarks. Het proton is een zeer gecompliceerd object die niet enkel uit drie valentiequarks bestaat, maar ook uit een virtuele soep van andere quarks en gluonen. Op deze details zullen we hier verder niet ingaan.

met massa sluiten de zwakke kernkracht symmetrie uit. Omdat we wel degelijk deeltjes zien die een massa hebben, zou dit dus betekenen dat de zwakke kernkracht niet uit te leggen is zijn met behulp van een symmetrie.

Gelukkig werd er op den duur een oplossing bedacht door diverse mensen. De meest bekende onder dezen zijn François Englert, Peter Higgs, en Robert Brout[†], en de oplossing zou bekend worden onder de naam van het higgsmechanisme. Dit mechanisme voorspeld het bestaan van een nieuw elementair deeltje, het higgsdeeltje, en een allesdoordringend higgsveld. Alle andere deeltjes danken hun massa dan aan dit veld. Om te zien hoe dit werkt zullen we eerst het concept van een deeltjesveld moeten introduceren.

Het Standaardmodel is een vorm van kwantumveldentheorie (KVT). In KVT is een deeltje niets meer dan een excitatie van een onderliggend veld. Het elektron bijvoorbeeld is een golf in, of een verstoring van, iets dat een elektronveld heet. Dit veld strekt zich uit over het gehele universum. Deze stand van zaken is te vergelijken met het klassieke concept van een magnetisch veld; het bestaat overal, het is alleen dat het in de buurt van een magneet verstoord en vervormt wordt. Feitelijk is het magnetisch veld een vorm van het elektromagnetisch veld, wat uiteindelijk gewoon het fotonveld is.

Net zoals in het klassieke voorbeeld der magnetisme is het zo dat het veld zelf energie kan bevatten. Deze energie is gerelateerd aan de veldsterkte, en deze relatie heet de potentiaal. Voor de meeste velden, en zeker zo voor het elektromagnetisch veld, is de energie in een veld nul zodra de veldsterkte ook nul is. Dit klinkt zeer redelijk; een magnetisch veld met veldsterkte nul is zo'n beetje hetzelfde als de afwezigheid van het veld, en zonder veld is er ook niets om de energie in op te slaan.

In het vroege heelal, vlak na de oerknal, was het universum zeer heet en heel erg compact. In deze krappe omstandigheden zat er een hoop energie opgeslagen in de diverse velden die het universum rijk is. Echter, terwijl het heelal uitdijde en afkoelde verdeween deze energie totdat de minimale energie werd bereikt die elk veld kan bevatten. Deze minima corresponderen op hun beurt met een verdwijnende veldsterkte, zie de linkerzijde van figuur A.5.8. Dit is vergelijkbaar met een magnetisch veld dat al zijn energie verloren is en nu een veldsterkte van nul heeft.

Het higgsmechanisme introduceert een nieuw veld, het higgsveld, wiens excitaties de higgsdeeltjes zijn. Dit veld volgt niet de normale regel waar de veldsterkte en de energie in het veld simultaan nul zijn. De higgspotentiaal heeft wel energie minima, en het heelal zal zich in één van deze punten bevinden na afkoeling, maar deze minima corresponderen niet een veldsterkte van nul. Dit betekent dat zelfs in ons huidige tijdperk het heelal nog steeds doorgeven is van een niet-nul higgsveld, zie de rechterzijde van figuur A.5.8.

Als bepaalde deeltjes door dit veld reizen zullen ze interageren met het niet-nul higgsveld en met deze interactie gaat een bindingsenergie gepaard. Wederom zal deze energie zich via Einstein's $E = mc^2$ manifesteren als een massa. Hierdoor krijgen alle deeltjes die een interactie hebben met het higgsveld een massa die proportioneel is aan de sterkte van deze interactie. Voor de verschillende deeltjes is de sterkte van deze interactie verschillend waardoor deze deeltjes allemaal een andere massa hebben. Sommige

[†]Robert Brout heeft helaas, in tegenstelling tot de twee anderen, zijn Nobelprijs nooit ontvangen. Hij overleed in 2011, te vroeg voor de uitreiking die in 2013 plaatsvond en ook voor de ontdekking van het higgsdeeltje in 2012.

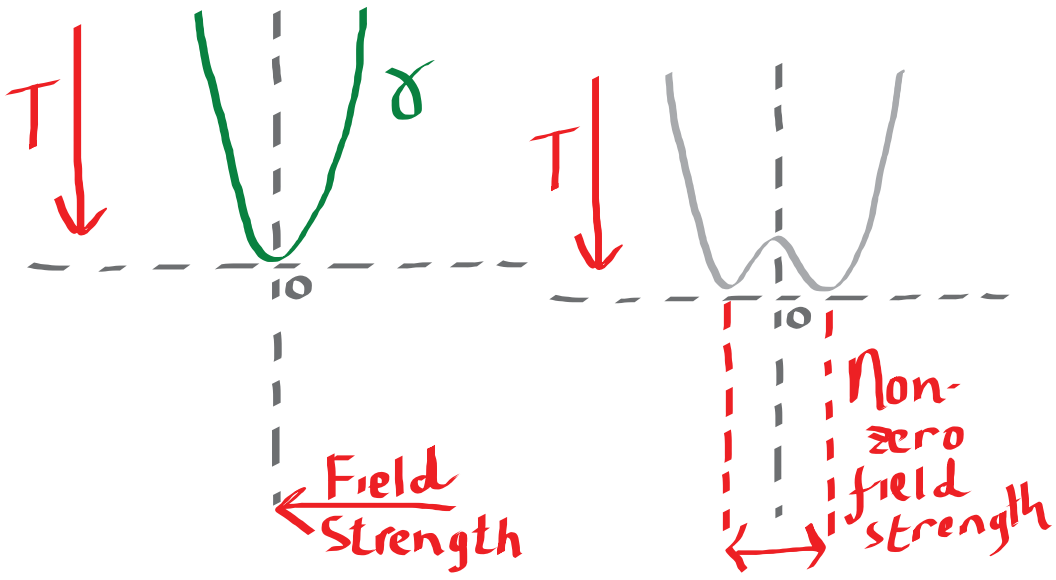

Figure A.5.8: *Schematische tekening door de auteur die links de foton (elektromagnetisch) potentiaal weergeeft, welke een functie is van de veldsterkte in termen van de temperatuur van het heelal. De veldsterkte gaat naar nul bij de minimale temperatuur en energie. Rechts word de Higgs potentiaal weergegeven. Hier is de veldsterkte niet nul bij de minimale energie. Merk op dat een keuze moet worden gemaakt voor één van de twee minima zodra de energie laag genoeg wordt.*

deeltjes, zoals het foton, hebben geen directe interactie met het higgsveld, en blijven daarom massaloos.

Op dit moment zou de lezer de natuurkundigen kunnen beschuldigen van valsspelen. Herinterpreteren zij niet gewoon de massa's van deeltjes? Met deze extra massa wordt de zwakke kernkracht symmetrie nog steeds onmogelijk gemaakt te bestaan! Het lijkt alsof het Higgs mechanisme het oorspronkelijke probleem helemaal niet oplost! Technisch gezien is dit waar, maar wat het Higgs mechanisme ons brengt is de mogelijkheid om deze symmetrie spontaan te breken. De natuurwetten zijn symmetrisch onder de gevraagde symmetrie, het is alleen zo dat de grondtoestand, d.w.z. de laagste energietoestand, dat niet is. Dus zodra het heelal voldoende is afgekoeld zal de symmetrie breken en op dat moment krijgen deeltjes hun massa. Dit zorgt ervoor dat we in staat zijn zowel het symmetrieprincipe te behouden alsmede deeltjes een massa te geven.

We hebben sterke indicaties dat dit beeld van de werkelijkheid juist is. In 2012 werd het higgsboson ontdekt in de *Large Hadron Collider* (LHC) op CERN nabij Genève, door twee verschillende experimenten: de *Compact Muon Solenoid* (CMS) and *A Large Toroidal LHC ApparatuS* (ATLAS). De LHC is een deeltjesversneller waar protonen tot hele hoge energie worden versnelt alvorens ze op elkaar te laten botsen waarbij veel nieuwe deeltjes gecreëerd worden. Deze ontdekking laat zien dat het higgsveld bestaat, maar het laat niet zien dat die higgsveld de eigenschappen bezit die nodig zijn om een massa te geven aan andere deeltjes. Om dit laatste wel aan te tonen zullen we de vorm van de higgspotentiaal moeten meten en laten zien dat het minima heeft waarbij de veldsterkte niet nul is.

Fenomenologie en motivatie

HET meten van de higgspotentiaal is helaas niet zo eenvoudig. De cruciale parameter die gemeten dient te worden om de vorm van de potentiaal vast te stellen is de zogenaamde *'Higgs trilineaire koppelingsconstante'*, λ_{3h}. Deze constante bepaalt hoe sterk het higgsdeeltje interageert met zichzelf. De moeilijkheid zit hem in het gegeven dat om λ_{3h} te meten we niet alleen een higgsdeeltje moeten maken in de LHC (een kunst op zich), maar dat dat higgsdeeltje vervolgens een ander higgsdeeltje zal moeten uitzenden. Zo'n gebeurtenis is heel erg zeldzaam, en we zullen het ook nog eens vaak genoeg moeten doen om voldoende data te verzamelen. Daarnaast is het ook nog eens zo dat vele processen die door de LHC gegenereerd kunnen worden er uit kunnen zien als zo'n gebeurtenis maar het niet zijn. Men zal manieren moeten vinden om deze zogeheten achtergronden te kunnen filteren.

In deze dissertatie hebben we eerst gekeken naar welke botsingen het meest veelbelovend zijn om naar te kijken als we λ_{3h} zouden willen meten. We moeten dit doen omdat we het higgsdeeltjes niet direct kunnen waarnemen. Het is een instabiel deeltjes dat vervalt in andere deeltjes voordat het de detector bereikt. We hebben twee higgsdeeltjes nodig, één origineel, en een andere die aangemaakt wordt door de eerste. Deze zullen beiden vervallen, en niet noodzakelijk op dezelfde manier. We hebben gevonden dat een proces waarbij één van de twee deeltjes vervalt naar een b-quark paar, en het andere naar een fotonpaar, de minste achtergronden heeft. Dit houdt in dat er maar weinig processes zijn die er hetzelfde uitzien voor een deeltjesdetector waardoor het eenvoudiger te onderscheiden is.

Dit proces kan geschreven worden als $pp \to hh \to b\bar{b}\gamma\gamma$, waar p staat voor een proton, h voor een higgsdeeltje, γ voor een foton, en b voor een b-quark. $\bar{b}$ staat voor het antimaterie partnerdeeltje van het b-quark, maar dat detail is niet echt belangrijk voor deze samenvatting. Een schematische weergave van dit proces kan gevonden worden in figuur A.5.9. We hebben ook gevonden dat de gevoeligheid voor λ_{3h} het grootst is als de interactie-energie gelijk is aan de massa van het higgsdeeltje.

Helaas bleek dit proces te zeldzaam om gevonden te kunnen worden in de data die door de ATLAS-detector (dit was het experiment met welke de auteur verbonden was) is opgenomen. Maar het bleek ook dat de grootste achtergrond voor elke λ_{3h} meting in $pp \to hh \to b\bar{b}\gamma\gamma$ grote theoretische onzekerheden heeft. Dit wil zeggen dat de voorspelling die het Standaardmodel geeft voor hoe vaak dit proces zou moeten voorkomen geplaagd wordt door significante onzekerheden in de berekeningen. Vaak zijn zulke berekening erg complex en moeten soms benaderingen worden gebruikt. Deze achtergrond, $pp \to b\bar{b}\gamma\gamma$, heeft leidt tot dezelfde deeltjes als het originele proces (daarom is het een achtergrond), maar in dit geval is het niet noodzakelijk dat deze deeltjes afkomstig zijn van het verval van twee higgsdeeltjes. Dit proces is ook zeldzaam, maar niet zo zeldzaam als die voor het λ_{3h} proces. Elke toekomstige meting aan λ_{3h} zal met deze achtergrond te maken krijgen, en dus werd het noodzakelijk geacht dit proces goed te bestuderen en te meten precies hoe vaak het voorkomt in de LHC.

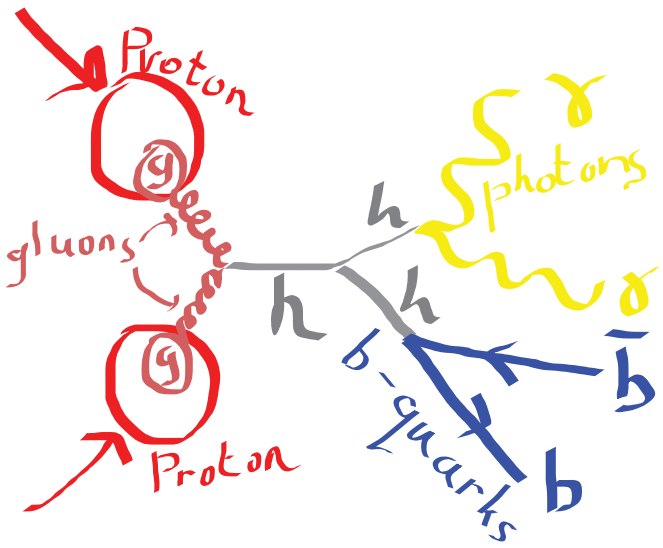

Figure A.5.9: *Schematische tekening door de auteur van het $pp \rightarrow hh \rightarrow b\bar{b}\gamma\gamma$ proces. Links botsen twee protonen op elkaar waarbij, bijvoorbeeld, twee gluonen samensmelten tot een Higgs deeltje (h in het plaatje). Het Higgs deeltje splitst dan in twee, en dit is het onderdeel dat gevoelig is voor λ_{3h}. Daarna vervalt één Higgs deeltje naar een fotonpaar (photons), en het andere naar een b-quark paar.*

De eigenlijke analyse

Onze analyse probeert te meten hoe vaak het $pp \rightarrow b\bar{b}\gamma\gamma$ proces zich voordoet in de LHC. De belangrijkste objecten in onze analyse zijn dus fotonen en jets. Jets zijn een grote collectie deeltjes die de detector binnentreden op een lawine-achtige manier en zijn gemakkelijker te beschrijven als een singulier composiet object dan als individuele deeltjes. De eigenschappen van een gegeven jet kunnen worden onderzocht om er achter te komen wat voor een soort deeltje de oorzaak ervan was. In ons geval zijn we geïnteresseerd in het vinden van jets die van b-quarks afkomen. Deze jets worden b-jets genoemd, en de onderzoeksmethode om deze te vinden wordt b-stempeling genoemd.

De ATLAS-detector en -software maken soms fouten tijdens de reconstructie van een proton-proton botsing, en deze fouten vormen de voornaamste achtergrond[†] tot onze analyse. We hebben te maken met elektronen die verkeerd gereconstrueerd worden als fotonen, en ook jets die zich voordoen als fotonen. Hier bovenop bevindt zich ook een achtergrond van niet-b-jets die toch een b-stempel krijgen.

De elektronen werden onder de loep genomen in een ander proces waarbij er veel meer elektronen dan fotonen voorkomen en dan te kijken hoe vaak deze elektronen eruit zien als fotonen voor de detector. Gewapend met deze kennis konden we een aantal vergelijkingen opstellen die voorspellen hoeveel botsingen we zouden moeten zien met ons proces waarbij zich twee fotonen vormen, of twee elektronen, of een elektron en een

[†]Een achtergrond is een signaal dat lijkt op hetgene we willen meten maar het eigenlijk niet is. Het is zaak deze tot een minimum te beperken om zo verwarring te voorkomen.

foton. Uit dit stelsel van vergelijkingen kan dan de grootte van de elektronachtergrond worden afgeleid. Deze kwam neer op 5.3%.

De achtergrond afkomstig van jets was iets moeilijker om mee om te gaan. Vergelijkbaar met de elektronachtergrond met zijn drie botsingscategorieën hebben we zestien categorieën gedefinieerd door elke van de twee fotonen in ons proces in te delen in één van vier categorieën. Omdat we twee fotonen hebben leidt dit tot $4 \times 4 = 16$ botsingscategorieën. De vier fotoncategorieën werden aangemaakt door twee vragen te stellen over de fotonen. De eerste was of het foton geïsoleerd was, d.w.z. of er andere deeltjes vlakbij het foton gevonden werden, en de tweede was of de interactie tussen het foton en detectormateriaal aan bepaalde kwaliteitseisen voldeed.

Voor elk van deze zestien categorieën schreven we opnieuw een vergelijking op die voorspelde hoeveel botsingen we zouden moeten vinden in die categorie aangenomen dat een bepaald aantal daadwerkelijk plaatsvond, alsmede bepaalde andere parameters zoals hoe vaak een jet eruit ziet als een foton etc. Dit systeem van vergelijkingen was te groot om direct op te lossen, en dus hebben we een computeralgoritme gebruikt om de beste oplossing te vinden en hiermee schatten we de totale contributie van deze achtergrond op zo'n 10%.

De non-b-jet achtergrond werd geschat op de volgende manier. We hebben botsingen onderverdeeld in vier categorieën gebaseerd op het type jets dat in de botsing gevonden werd. Eén van deze is de bb-categorie, wat ons signaal vormt, en de botsingen in de andere categorieën miste ten minste één b-stempel. Bij het toepassen van een b-bestempeling krijgt men een getal wat geïnterpreteerd kan worden als de waarschijnlijkheid dat een gegeven jet een b-jet is. De distributie van dit getal over vele botsingen kan bestudeerd worden op computersimulaties die door dezelfde ATLAS reconstructiesoftware gehaald zijn als de echte data. De distributies voor de twee jets in ons proces kunnen ook gecombineerd worden tot een enkele distributie en de vorm van deze distributie kan afgeleid worden voor elk van de vier categorieën met dezelfde computersimulaties. Daarna kan dezelfde distributie gemeten worden in de data en vergeleken worden met de vorm van de vier distributies van de vier categorieën uit de simulaties. Via een computeralgoritme kan dan bepaald worden met welke fracties elke categorie zich voordoet in de data. Hiermee vonden we dat slechts 54% van alle botsingen aan de bb-categorie voldoet.

Nadat alle achtergrond verwijderd zijn, moet men een zogeheten 'unfolding' doen. Dit is om effecten die door de detector aan de gemeten data worden opgelegd ongedaan te maken[†]. Als we dit niet zouden doen is het onmogelijk onze meting te vergelijken met een theoretische voorspelling.

Uiteindelijk na al dit werk wordt men beloont met een paar mooie uitziende grafiekjes weermee de meting en de theoretische voorspelling met elkaar vergeleken kunnen worden. We vonden dat de LHC ongeveer 540 echte $pp \to b\bar{b}\gamma\gamma$ botsingen geproduceerd heeft in 2012, wat een klein aantal is gezien er 100 miljoen proton-proton botsingen plaatsvonden per seconde gemiddeld door het jaar heen. We vonden ook dat de theoretische voorspelling en de gemeten waarden overeenkomen binnen de diverse onzekerheden. Een plaatje die de theorie met de werkelijkheid vergelijkt kan gevonden worden in figuur A.5.10.

[†]De detector kan, bijvoorbeeld, de energie van foton met een bepaalde factor verkeerd meten. Tijdens de unfolding deelt men dan door die factor om zo de ware energie terug te krijgen.

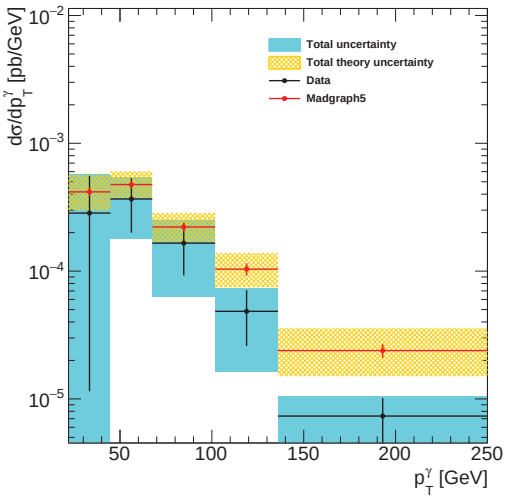

Figure A.5.10: *De distributie van de differentiële botsingsdoorsnede (gerelateerd aan de productiewaarschijnlijkheid) van $pp \rightarrow b\bar{b}\gamma\gamma$ ten opzichte van de transversale impuls van het leidende foton zoals gemeten in onze analyse vergeleken met de theoretische voorspelling (*MADGRAPH5*). De transversale impuls is de impuls waar de bijdrage van de richting waarlangs de protonen botsen is weggelaten. Het leidende foton is het foton met de grootste transversale impuls.*

Acknowledgements

"First they came for the Socialists, and I did not speak out— Because I was not a Socialist.

Then they came for the Trade Unionists, and I did not speak out— Because I was not a Trade Unionist.

Then they came for the Jews, and I did not speak out— Because I was not a Jew.

Then they came for me—and there was no one left to speak for me."

- Pastor Martin Niemöller

HERE I would like to thank various people without whom this thesis might not have existed. First and foremost I thank my supervisor Bob, and my co-supervisor Magdalena. Over the years we have had many fruitful discussions and it hardly needs mentioning that this thesis is better for it. Bob, your ability to identify problems with a simple glance to a plot is unparalleled. You always managed to help me focus on the important things, and were always available for a nice chat. Magdalena, I really enjoyed exploring experimental physics together with you, and gained much from our very frequent dialogues. You were also always ready to explain the theoretical underpinnings of whatever we were working on. To both of you I also apologize for any indicidental stubbornness on my part. I feel I would do a disservice not to mention Michael Kagan in a similar vein. Michael, whilst we just worked on a 'service task' for my first two years, I can say without a doubt that I learned much from this experience and your expert guidance. You were always ready, willing, and able, to explain new concepts in a very illuminating manner, which were very useful to me when performing the analysis described here. I thank all three of you deeply. I would also like to thank my promotion committee for reading, commenting on, and judging my work.

Apart from this I would also like to thank my other ATLAS collegues with which I've had beneficial and friendly interactions. Daniël, your help in bending the computers systems to my will has been invaluable. Lydia, your Sinterklaas parties are legendary, I hope you will continue to organize them (and invite me). Gabriele, you were the best office

mate I could have wished for at CERN. Together we were put in the backwater ATLAS office, but our frequent pleasant talks about politics were more than worth it. The only sad point is that I have to concede you were right; you predicted Trump would win, and he did. I hope we will all be alright. Vince, you and I went on fantastical journeys together, from the cold Spine of the World, to different timelines and planes. We slew monsters both big and small, but let us forget about that one incident involving an orc. Antonia, I enjoyed our year living together in St. Genis, and before that our regular dinner+wine evenings; you are an awesome cook. Marie, like me you did your ATLAS service task within in the b-tagging group and I took much pleasure in our discussions surrounding that subject. You are also a fierce friend, and certaintly brightened my stay at CERN. Our frequent dinners are fond memories. Koen, my room mate for the last year in Amsterdam. We did not see each other as often as perhaps we should have, but I did like living with you in the Langestraat nonetheless. Nadezhda, you arrived quite late on the scene, around halfway through the last year, but we had lots of fun together. I recall with much joy our many evenings watching movies and discussing at lengths the intricate affairs of the couple across the street from my place in Amsterdam. Of course there are many more people within ATLAS; thank you Pieter, Lars, Hartger, Stefan, Tim, Matteo, Jörn, Ingrid, Ivan, Pierfrancesco, Lucrezia, Robin, Rosemarie, Naghmeh, Marco, Hannah, Nika, Rogier, Priscilla, Remko, Luca, Christian, and Sarah. Also the new batch of future physicists who joined me in the PhD factory: Kees, Marc, Pepijn, Daan, Milo, Martijn, and Terry; I thank you for keeping me company in my last year, and wish you a bright future.

No acknowledgement section would be complete without mentioning the group of people I befriended during my year long stay at CERN. I think I can, without any controversy, say that there was a hardcore group of friends formed by Tarjei, Sergio, Ellis, Romy, Flo, and myself. And, even though she arrived late, I feel I can also include Camille *'Romy 2.0'* in that tight group. You made my year at CERN a spectacular adventure! I could relate some anecdotes here, but truly there are too many great ones to choose from, and I am sure many new ones will be formed in the future.

To my friends from Armenia: Vard, Narine, and Armine, I say shnorhakalut'yun! Thank you for being my friends and teaching me about your culture and language, and introducing me to the very best your country has to offer.

I also thank my friends from the Netherlands who supported me during my PhD and studies in various ways: Remco, Elena, Guus, Resie, Han, Martijn, Frank, Laura, Guy, Sjoerd, Larissa, Bas, Frederik, and Michael. And of course those I saw often in the last few years for dinner and games: Eveline, Jeroen, Ans, and Jorrit. I hope we will continue to meet regularly for a long time to come! Maaike, I thank you for keeping me company during the long train trips between Amsterdam and Nijmegen, and for the endless 'what's your favorite...?' questions.

As for my relatives, I thank my mother for raising me and giving me the opportunity to pursue my dreams, my brother for keeping me company on the road to maturity, and my uncle Frans for instilling in me a sense of wonder and curiousity. In addition I like to thank Amber for keeping me company and being my mainstay during the first years of my PhD adventure. Finally I thank my soulmate Ani for coming down from her planet to join me on this mundane ball of earth and rock. Let us continue to sing a song of magic and love together!

Ձիշտ ասողի գլխարկը ծակ է:

"The cosmos is all that is, or ever was, or ever will be. Our comtemplations of the cosmos stirs. There is a tingling in the spine, a catch in the voice, a faint sensation, as if a distant memory of falling from a great height. We know we are approching the grandest of mysteries. The size and age of the cosmos are beyond ordinary human understanding, lost somewhere between immensity and eternity is our tiny planetary home, the Earth.

For the first time we have the power to decide the fate of our planet and ourselves. This is a time of great danger. But our species is young and curious and brave, it shows much promise. In the last few millenia we have made the most astonishing and unexpected discoveries about the cosmos and our place within it. I believe our future depends powerfully on how well we understand this cosmos, in which we float like a mote of dust in the morning sky.

We wish to pursue the truth no matter where it leads. But to find the truth we need imagination and skepticism both. We will not be afraid to speculate, but we will be careful to distinguish speculation from fact. The cosmos is full beyond measure of elegant truths, of exquisite interrelationships, of the awesome machinery of Nature.

The surface of the Earth is the shore of the cosmic ocean. On this shore we have learned most of what we know. Recently we have waded a little way out, maybe ankle-deep, and the water seems inviting. Some part of our being knows, this is where we came from. We long to return. And we can, because the cosmos is also within us. We are made of star stuff. We are a way for the cosmos to know itself."

Carl Sagan, Cosmos: A Personal Voyage