Measuring the gluon transverse momentum dependent parton distribution function f_1^g with a $J/\psi + \gamma$ final state using ATLAS data.

Alina Isobel Hagan, MSc

Department of Physics

Lancaster University

A thesis submitted for the degree of

Doctor of Philosophy

May, 2025

Measuring the gluon transverse momentum dependent parton distribution function f_1^g with a $J/\psi + \gamma$ final state using ATLAS data.

Alina Isobel Hagan, MSc.

Department of Physics, Lancaster University

A thesis submitted for the degree of *Doctor of Philosophy*. May, 2025.

Abstract

This thesis presents a measurement of the unpolarised, transverse-momentum-dependent (TMD) parton distribution function (PDF) f_1^g using a $J/\psi+\gamma$ final state with the ATLAS detector at the Large Hadron Collider (LHC). 2.6 fb^{-1} of data from proton-proton collisions collected during 2015 as a part of LHC Run 2 is used. The final state of a J/ψ meson produced in association with an isolated photon, initiated from gluon-gluon fusion, is used to give insight into the internal transverse dynamics of the proton. The transverse momentum distribution of the $J/\psi + \gamma$ system is reflective of the transverse momentum of the colliding gluons. By measuring the distribution of the gluon transverse momentum and combining this with a model of the gluon distribution in the proton and a collinear PDF, a full description of a TMD is possible. A series of techniques were utilised to extract this gluon transverse momentum distribution, including statistical tools such as boosted decision trees and binned template profile likelihood fits. The mean transverse momentum of the gluons in a Gaussian parameterisation was found to be $\langle k_T^2 \rangle^{\frac{1}{2}} = 1.61 \pm 0.28$ GeV.

Acknowledgements

Thank you to Vato. Your expertise and patience made this entire endeavour possible. The work herein and the completion of this PhD was not possible without you, your guidance, or patience.

Thanks to my mother and my father, your support and help over these four years has been invaluable. You made this possible.

To Victoria, Charlotte, Joe, Lorenzo, Nathan, Chris, Maya, and Moth. Your company, friendship, conversation and fellowship at CERN, in Lancaster, and down south, has been irreplacable. You have made me a different person in the best possible ways.

Thank you to Katie, Kerry, Blair, Leo, Courteny. Your constant encouragement and support from our undergraduate days and now has been instrumental throughout my studies. Your belief in me has been a source of strength, and without you I would not have made it this far.

To my brothers, David, Matthew, and Tom, for your unfaltering understanding throughout my journey. Your belief in me, without question, has meant more than words can express, and I wouldn't have reached this milestone without you by my side.

Thanks to my collected friends and more from across the country: Violet, Emma, Rain, Sammy, Emily, Lissa, Luci, Jon, Bil, Col, George, Elliot, Eleanor, Joseph, Amar, Emily, Giovanna and Aodhan.

To Neža. Your feedback and companionship pushed me through this whole thing. Thanks for keeping me together, buddy.

For Danny, Laura, Jemma, Ea, Willow, Emily, Claire, and Becca. You kept me afloat with your love and friendship in my hardest time and I can never fully repay you.

I love you all. Thank you.

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography. A rough estimate of the word count is: 33000.

Alina Isobel Hagan

Glossary

TMD Transverse-momentum-dependent parton distribution function.

PDF Parton distribution function. A function mapping the fraction of a nucleons collinear momentum to the probability of finding a quark or gluon of a particular flavour.

QCD Quantum Chromodynamics. The quantum field theory describing the strong nuclear interaction.

Hadron Either a Baryon or a Meson: bound states of matter, held together with QCD.

Meson A bound state of a quark and an anti-quark.

Baryon A bound state of three quarks.

Exclusive A precise and defined physical process where all the products are observed and measured.

Inclusive A physical process with some of the final state products unmeasured.

Prompt production When a final state, or component of it, is produced directly from the initiating particles

Non-prompt production When a final state, or component of it, is produced through an intermediate decay, such as that of a B-Hadron.

Trigger Hardware or software used to identify events interesting to physics analysis.

Contents

List of Tables

List of Figures

1. Introduction

Since the beginning of the twentieth century scientific inquiry has provided continually more detailed insight into the structure of nature, starting from the first atomic models in the 1890s. In the 1910s, via foil and radiation, the basic structure of the atom was revealed: a dense positively-charged core surrounded by orbiting negative charges, the nucleus and electrons. Then, quantum mechanics began its life in the 1920s enhancing the description of 'orbits' to that of region with a defined probability. Further hypotheses were made in the 1920s, verified by experiments in the '30s, that allowed the dense positive core of the atom to be decomposed further. The nucleus now consisted of the positive proton and the uncharged neutron. It was a further 17 years before the next paradigm shift in understanding the nucleus came about, the development of the nuclear shell model in 1949. This gave the nucleons (protons and neutrons) a quantum mechanical model much like that of the electrons in the fields surrounding them.

In the backdrop of this nuclear scene, quantum mechanics birthed a hopeful but troubled child: quantum field theory. By the end of the 1920s, progress had been made, including relativity, creating free-field theories of various particle types, and (after another round of quantisation) basic predictions of electromagnetism were starting to be made. However, all was not entirely well. In a number of different tests quantum field theory failed to produce physical answers; the theoretical process of perturbation theory predicted infinite values for physical quantities.

Solutions to these divergences did not arrive until the late '40s and early '50s, with the appearance of renormalization. Troublesome divergences could be tidied away using some clever mathematical tricks and accurate predictions

could be made. Following shortly after this, gauge theories as a guiding approach for modelling fundamental interactions were developed in the '50s. A fully rigorous definition of quantum field theory for electromagnetism followed in the next decade: quantum electrodynamics (QED).

While revolutions were being made on the front of quantum field theories, focus was not lost on the nature of the nucleon. Throughout the late '40s and '50s further particles had been discovered, but not explained: pions, kaons, lambdas and more. An effort was made to classify these new members of the 'particle zoo' in the early '60s, using similar ideas of symmetry to those found in gauge theories. If these discovered particles are not fundamental and instead composed of different types of constituents, inferring the rules for combining these constituents lets others be predicted. Predicted, and then discovered.

With the discovery of the omega-minus in '64 the first experimental hint for another layer 'lying within' appeared. The pieces that make up the particle zoo, the 'quarks', may be more than a mathematical convenience. And so the search continued, culminating in the establishment of the idea that nucleons and the subsequently discovered mesons and baryons were not fundamental objects. Deep inelastic scattering experiments of electrons off protons in '68 showed that the nucleon contained smaller fundamental constituents.

Shortly after, in the early '70s, gauge theories were put to work in an effort to create a theoretical description of the interactions between these quarks, or partons as they had become known. It was discovered, after some renormalization, that an accurate (albeit quite different to QED) description of the strong nuclear force could be made and properties could be predicted with the already-existing tools of perturbation theory. This framework that described the quarks and the carrier that bound them together, the gluon, became known as quantum chromodynamics (QCD). QCD along with QED and its recently unified cousin, the weak nuclear force, comprise the standard model (SM). This is a full quantum field theory that describes all fundamental physical forces apart from gravity.

The SM lies at the heart of modern particle physics, and has often been described to work 'embarrassingly well'. Experimental and theoretical efforts, put somewhat bluntly, are oriented towards finding a way to break it. Although impressively accurate, there remain more unanswered questions on topics such

as dark matter. Finding a point at which the SM starts to give way opens a door to expanding the model and expanding human understanding of nature.

The largest machine on the planet, the Large Hadron Collider (LHC), and the multiple attached experiments are engaged in this exploration. In order to create experimental data with which to test the SM , the Large Hadron Collider impacts protons together at extremely high energies. At this energy scale, protons do not interact, their partons do. More specifically, in the current energy regime of the Large Hadron Collider] 85% of interactions in collisions are between two gluons.

This is where the 'quite different' nature of QCD must be considered. While many predictions of its behaviour can be made using perturbation theory, some important properties are theoretically restricted. Perturbation theory relies on small effects that converge at higher orders to make predictions, the strength of the coupling between two particles being an example one of those effects. The QCD coupling is 'inverted', meaning that it is very strong at low energies and very weak at high energies. With this configuration, it becomes impossible to theoretically predict how partons are distributed when collected together as this is a low energy phenomenon. Using a process called factorization this unpredictable 'non-perturbative' component of QCD, that describes the probability of finding a parton with a specific momentum, can be replaced with the results of experimental measurement. This experimental measurement allows practical theoretical predictions in some kinematic areas to be made with QCD.

These non-perturbative parton distribution functions (PDFs) have been traditionally parameterised with a single variable, the fraction of collinear momentum a parton could be carrying. In an era where ever more precise tests of QCD and other phenomena are being explored, expansions have been made to this to include momentum in other dimensions. Expanded parton distribution functions, parameterised with transverse momentum, are known as transverse momentum dependent parton distribution functions (TMDs). The measurement of these not only provides opportunities to constrain and quantify effects of the transverse dynamics of partons on physical processes outside QCD, but also provides data to test the tools and methods that QCD provides, such as exploring

how TMDs evolve over different energy scales.

This thesis details an analysis that uses data collected with the ATLAS experiment at the Large Hadron Collider (LHC) and a final state of a J/ψ meson and a photon to quantify the transverse dynamics of the gluon within the proton. The text begins with a brief theoretical overview in Part I, Chapters 2 and 3 detail the SM and QCD, and the background of TMDs at high-energy experiments, respectively. Next, Part II first describes the experimental setup of the LHC, and the ATLAS detector in Chapter 4. The remainder of this part is Chapter 5, detailing the setup of this analysis and the techniques that are applied in the last phase of the text. Part III covers the processes used to extract the transverse dynamics within the gluon, and presents the parameterisation of the transverse momentum dependent parton distribution function across three Chapters 6-9. Chapter 10 is reserved for an outline of the findings.

Part I

Theoretical Foundation

2. The Standard Model & QCD

2.1 The Standard Model

2.1.1 Overview

The Standard Model (SM) [1] is a quantum field theory that described the fundamental components of nture and their interactions. As a theoretical framework it is used to predict the properties of all currently observed particles. It contains particles which divide into two categories: fermions and bosons. Fermions, particles with half-integer spin, follow Fermi-Dirac statistics and are bound by the Pauli Exclusion Principle. The Pauli Exclusion principle determines that no two fermions within the same system may carry the same set of quantum numbers. Twelve of the members of the Standard Model are fermions: the six quark flavours and the six leptons. Both quarks and leptons break down into three generations, each heavier than the last. Fermions are the building blocks of all currently understood matter in nature. Bosons, particles with integer spin, compose the remainder of the Standard Model. These obey Bose-Einstein statistics and can be further split into two more categories within the SM. The SM contains *vector* bosons: gauge bosons that operate as force carriers. The vector bosons are the Photon, 8 Gluons, the Z boson and the two W bosons. The remaining *scalar* component is the Higgs Boson. Fermions carry multiple other quantum numbers governing their coupling with the gauge bosons and consequentially their interactions.

2.1.2 Lagrangians and Symmetries

The SM (in brief) is represented by the following Lagrangian density [2]:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + \text{h.c.} + \psi_i y_{ij}\psi_j\phi + \text{h.c.} + |D_\mu\phi|^2 - V(\phi). \tag{2.1}$$

The first term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ describes the interactions of the gauge fields for the vector bosons. $F^{\mu\nu}$ is the field strength tensor and carries the kinetic and potential terms of the fields. Next, the $i\bar{\psi}\not{D}\psi$ term (and its hermitian conjugate h.c) describes the coupling of the force-carrying bosons in the covariant derivative $\not{D}$ to the fermions ψ. The terms $\psi_i y_{ij}\psi_j\phi$ and $|D_\mu\phi|^2$ describe the coupling of the fermions and bosons to the scalar Higgs field ϕ, and $-V(\phi)$ is the potential of this field.

Symmetries are *the* key concept that the SM is dependent on. The symmetries within this Lagrangian determine both conservation laws and the definition of all the force carriers [3, 4]. Traditional symmetries such as invariance during time and space rotations and translations yield the familiar kinematic conservation laws. If the fields in the SM are transformed with a spatial translation and the Lagrangian remains unchanged, Noethers theorem can be applied. As the Lagrangian is invariant under a continuous transformation Noethers theorem determines that there is an associated conservation law. In the case of spatial translations, the symmetry produces conservation of momentum. Separate from the readily familiar transformations, are gauge transformations. These are local transformations (dependent on space-time coordinates) that form a Lie group. Gauge, in this context, refers to how a specific field is transformed. Symmetry under gauge transformations yields further essential conservation laws. Requiring that the SM Lagrangian maintains certain symmetries provides insight into two things: first, into the conservation laws the SM maintains, through the application of Noethers theorem; secondly, information on the field structure of the SM, and hence the force carriers and the nature of SM couplings.

For example, the inclusion of the photon into the SM from the $U(1)$ gauge transform of a fermionic field [5]:

$$\psi \to \psi' = e^{ie\chi}\psi \qquad\qquad \bar{\psi} \to \bar{\psi}' = \bar{\psi}e^{-ie\chi}, \tag{2.2}$$

where the Lagrangian of this free fermionic field ψ is

$$\mathcal{L} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi. \tag{2.3}$$

The phase χ has dependence on space-time coordinates, making this a local gauge transformation, rather than a global one. In the global case $\partial_\mu\chi = 0$. For the first term of the free Lagrangian, the local gauge transformation produces the following:

$$i\bar{\psi}\gamma_\mu\partial^\mu\psi \rightarrow i\bar{\psi}'\gamma_\mu\partial^\mu\psi' = i\bar{\psi}\gamma_\mu\partial^\mu\psi + i(ie\partial^\mu\chi)\bar{\psi}\gamma_\mu\psi. \tag{2.4}$$

The additional rightmost term violates the invariance. To correct this, one must use the “minimal coupling prescription" and introduce the gauge covariant derivative $D_\mu = \partial_\mu + ieA_\mu$. If the newly introduced vector gauge field A^μ transforms as $A'_\mu = A_\mu - \partial_\mu\chi$, then the Lagrangian becomes invariant:

$$i\bar{\psi}'\gamma_\mu D'^\mu\psi' = i\bar{\psi}\gamma_\mu(\partial^\mu + ieA'^\mu + ie\partial^\mu\chi)\psi = i\bar{\psi}\gamma_\mu D^\mu\psi, \tag{2.5}$$

the introduction of this covariant derivative introduces a new interaction term coupling the quark to this gauge field $-e\bar{\psi}\gamma^\mu\psi A_\mu$. This leaves a larger Lagrangian:

$$\mathcal{L} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi - e\bar{\psi}\gamma^\mu\psi A_\mu. \tag{2.6}$$

For A^μ to correspond to a physical degree of freedom, kinetic and mass terms are required. The gauge invariant field strength tensor $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ can be constructed, which yields $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$. Any mass term would take the form $m_A^2 A_\mu A^\mu$, which would again violate gauge invariance. Setting $m_A = 0$ and including the final kinetic term, the Lagrangian becomes:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi - e\bar{\psi}\gamma^\mu\psi A_\mu \tag{2.7}$$

Recognise next that the relativistic description of Electromagnetism provides field strength tensors and currents that satisfy all the above relations. This would associate the gauge field A_μ with the photon. It can be seen that a Lagrangian with a description for electromagnetism can be extracted from applying the requirement of invariance under symmetry to the free fermion Lagrangian. This principle is continued with other symmetries of the SM. The SM maintains a $SU(3) \times SU(2) \times U(1)$ local gauge invariance. $SU(3)$ corresponds to the strong nuclear force and its associated carriers. The combination of $SU(2) \times U(1)$ produces a unified description of both the weak nuclear force and the electromagnetic force.

2.1.3 Interactions

The SM provides a description of three fundamental forces:

Electromagnetism. This is the force related to the $U(1)$ gauge symmetry introduced above. The gauge boson for this interaction, the photon, couples to the electric charge carried by all quarks and three of the leptons: the electron, the muon, and the tau.

The Weak Nuclear Force. Predominantly associated with the $SU(2)$ symmetry, the weak nuclear force is mediated by the $W^{\pm}$ and Z bosons, coupling to weak hypercharge and weak isospin. These quantum numbers are carried by all fermions and the W and Z bosons themselves.

The Strong Nuclear Force. This force is associated with the $SU(3)$ symmetry, mediated by gluons, coupling to color charge. The non-abelian nature of $SU(3)$ (as with $SU(2)$) determines that the color charge is carried by the gluons as well as the quarks.

When the theoretical neutral gauge boson of $SU(2)$ mixes with that of $U(1)$ two physical fields are produced as a result, the photon and the Z boson. With the existence of massive gauge bosons, gauge invariance is not preserved in electroweak interactions. The Higgs boson is included in the Standard Model to restore the invariance. This inclusion allows Electroweak unification, a combined $SU(2) \times U(1)$ gauge symmetry describing electromagnetism (EM) and Weak interactions simultaneously. There has not been a successful inclusion of gravity into the Standard Model. Several other open questions remain: there is no origin for the mass hierarchy present in the fermion masses, a sufficient explanation for dark matter, whether neutrinos are their own antiparticles, and so on.

2.2 QCD

2.2.1 Color & SU(3)

The fundamental charge that is conserved within QCD interactions is color charge [6]. Similarly to the case of EM, where charge determines the coupling of photons to objects, color charge determines the coupling of gluons to quarks. Experimental evidence leads to the conclusion that 3 colors exist for QCD,

subsequently named red, green and blue (r, g, b). A quark of arbitrary flavour can exist in one of these three color states and this triplet forms a fundamental representation of $SU(3)$, the color group. Transformations between these color charges can be represented by 3×3 complex unitary matrices. $SU(3)$ has eight generators, the Gell-Mann (λ) matrices. The eight 'rotation' angles in $SU(3)$ are represented by η. Considering this it is possible to have a common field for the quarks (and anti-quarks) which transform as:

$$\psi_n \rightarrow e^{\frac{1}{2}ig(\eta\cdot\lambda_{nk})}\psi_k \qquad\qquad \bar{\psi}_k \rightarrow \bar{\psi}_n e^{-\frac{1}{2}ig(\eta\cdot\lambda_{nk})}, \tag{2.8}$$

where $n, k = r, g, b$. g is the coupling constant. The Gell-Mann matrices, λ, are as follows:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

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

Under this transformation the derivative terms produce non-invariance, as with the $U(1)$ case. Using the previous prescription, a covariant derivative can be defined as:

$$D_\mu = \partial_\mu + \frac{1}{2}ig(G_\mu \cdot \lambda). \tag{2.9}$$

G_μ form 8 new gauge fields, one per η angle. The fields are gluons and the λ matrices describe their allowed color configurations. When requiring Lagrangian invariance, the following transformational properties of the gauge field are produced:

$$G^a_\mu \rightarrow G^{'a}_\mu = G^a_\mu - \partial_\mu \eta^a - g f^{abc} \eta^b G^c_\mu, \tag{2.10}$$

where f^{abc} describes the fully antisymmetric structure constants of $SU(3)$. The structure constants are found in the commutation relation defining the Lie algebra of $SU(3)$,

$$[\lambda_a, \lambda_b] = 2i f_{abc} \lambda_c. \tag{2.11}$$

This additional final term in the gauge field transform, Equation 2.10, is a consequence of the non-Abelian nature of $SU(3)$. The necessary kinetic term remains similar to the $U(1)$ example with $F^a_{\mu\nu}F^{\mu\nu\,a}$, though $F^{\mu\nu}$ now includes the structure constants to maintain gauge invariance:

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\nu - g f^{abc} G^b_\mu G^c_\nu. \tag{2.12}$$

This introduces the coupling of the gauge fields together. In other words, the gluons must hold color charge and must self-interact. Color charge is held by only the quarks and gluons, and no other components of the Standard Model.

2.2.2 Confinement & Asymptotic Freedom

Asymptotic freedom and confinement are two related, but distinct, concepts in QCD [6]. There is currently no solid quantitative theoretical explanation for the property of confinement, the observation that objects carrying color charge cannot be found 'bare'; at large (macroscopic) distances, color-carriers cannot exist freely, they are bound together into collections. All observed QCD objects are colour-neutral, built from color-carrying particles grouped together in a colorless combination. Both bosons and fermions can be found as these composite objects. Examples include mesons and baryons, quark-antiquark and 3-quark bound states, respectively.

Connected to this is the concept of asymptotic freedom. This is a grounded theoretical and experimental observation of the running behaviour of the QCD coupling constant is dependent on the energy scale.

In QED the evolution of the coupling as a function of energy is expressed as:

$$\alpha(q^2) = \frac{\alpha(\mu^2)}{1 - \frac{1}{3\pi}\alpha(\mu^2)\log\frac{q^2}{\mu^2}}, \tag{2.13}$$

where q^2 and μ^2 are the energy and the regularisation scale from the renormalisation process, respectively. By examination, as the energy scale increases $q^2 \gg \mu^2$ (and the length scale decreases) the coupling strength increases. However, unlike QED, the QCD coupling constant evolves in the opposite direction:

$$\alpha_s(q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{1}{12\pi}(11n - 2f)\alpha_s(\mu^2)\log\frac{q^2}{\mu^2}}. \tag{2.14}$$

Here, n and f are the number of colors and flavours, 3 & 6, respectively. So, as the energy scale increases, and length scales become shorter, the coupling decreases and QCD objects become loosely bound. When the energy scale decreases, the logarithm becomes negative, and the denominator shrinks. In this region color charged objects become tightly coupled. This can, intuitively, be seen as the emergence of confinement. However, although the coupling does appear to increase quickly at the confinement scale, there is no quantitative description of the dynamics past this point.

3. Transverse Momentum Dependent Parton Distribution Functions

3.1 Factorization, the Parton Model, and Parton Distribution Functions

Color confinement presents a challenge for both experimental and theoretical approaches to describing QCD. There is no method for direct physical observation of the bound components within a hadron. Neither can perturbative QCD alone describe the entire range of hadronic interactions. Insights on many topics like the quark-gluon structure inside hadrons, the nature of strong decays, hadron spectroscopy and low energy scattering are theoretically locked out[1]. Asymptotic freedom, however, offers something of an escape-hatch.

In order to produce a framework that can produce results using perturbation theory, the strongly coupled interactions in QCD must be factorised out. The remaining components would be short distance, high energy scale, and weakly coupled, calculable within a perturbative framework. Equation 2.14 is derived in this perturbative region and cannot provide reliable couplings in the region $q^2 \ll \mu^2$. This process of separating the short-distance and long-distance features is known as QCD factorisation [10]. The short-distance component makes

[1]The techniques of Lattice QCD [7] are beginning to push this boundary, with non-perturbative methods able to produce predictions of hadron masses[8]. Methods are steadily evolving for predicting PDFs [9], though these are techniques are still maturing.

up the process-dependent partonic cross-section. Then, the remaining long distance process-independent components of the cross-section are collected into non-perturbative distribution functions. These are probability distributions, parameterised by momentum, that describe the probability of finding a certain parton. They are known as parton distribution functions (PDFs). PDFs are phenomenological in nature, experimental data is collected from processes such as Deep Inelastic Scattering (DIS) that are used to create a description of the internal hadronic structure.

Parton distribution functions, optimistically, maintain universality: PDFs parameterised from data with one experimental method should be applicable to another, they are agnostic of the subprocess. However, PDFs are not inherently universal. When performing renormalization of the field theory definition of a PDF, dependence on the renormalization scale μ is included. This, by the necessity that different processes require different μ, means that there is no supposed process-independence. Thankfully, the perturbatively calculable DGLAP evolution equations [11] allow the dependence of PDFs on μ to be methodically evaluated. Though no *current* method exists to calculate the PDFs themselves, some promise exists for this to be done with Lattice QCD.

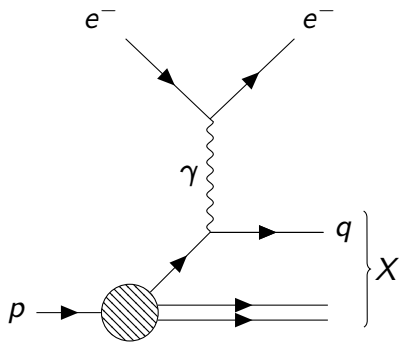


Figure 3.1: Feynman diagram for the Deep Inelastic Scattering process.

To illustrate, consider the process of applying the parton model to DIS [12, 13], shown in Figure 3.1, where it was originally verified. Here, a high-energy lepton (typically an electron) of 4-momentum l is used as an electromagnetic probe to a target proton of momentum p:

$$l + p \to l' + X. \tag{3.1}$$

The scattered outgoing electron with momentum l' is measured and the resultant momentum transfer $q = l - l'$ is inferred, with $q^2 = q \cdot q$ the invariant mass of a virtual photon. X represents the hadronic remnants. With a space-like photon with $-Q^2 = q^2 < 0$, consider this in the infinite momentum frame where $p \to \infty$, where the proton and photon are moving along the $z - axis$. In this frame, any incoming quark inside the proton would have a fraction of the proton's momentum x and an elastic scattering relation between the quark and the electron can be formed:

$$(xp + q)^2 = x^2 M_p^2 + 2xp \cdot q + q^2 \tag{3.2}$$

$$\underset{x^2 \sim 0}{\to} 2xp \cdot q - Q^2 = m_q^2 \tag{3.3}$$

$$\underset{m_q^2 \approx 0}{\to} x = \frac{Q^2}{2p \cdot q} \tag{3.4}$$

This x, often known as Bjorken-x for the eponymous scaling relation, can indicate the nature of the nucleon. Should p be a fundamental object with no internal constituents, then the scattering is elastic at all energies, meaning $x = 1$. However, as observed at SLAC[14, 15], the scattering is inelastic corresponding to $0 < x < 1$. Now, define the fraction of lost momentum from the lepton as $y = \frac{p \cdot q}{p \cdot l}$ and use the expression for $2 \to 2$ scattering [16], the cross-section for this electron-quark elastic scattering process becomes:

$$\frac{d\sigma}{dQ^2} = \frac{2\pi\alpha^2 e^2}{Q^4}[1 + (1 - y)^2], \tag{3.5}$$

where e is the charge of the scattered quark. As the proton is made of multiple partons, the cross-section of scattering off a particular parton a within the momentum range $x \to x + dx$ can be expressed as:

$$\frac{d^2\sigma_a}{dxdQ^2} = \frac{2\pi\alpha^2 e_a^2}{Q^4}[1 + (1 - y)^2] f_a(x), \tag{3.6}$$

where f_a is the probability of finding a parton of flavour a with momentum fraction x. Hence, the full lepton-proton scattering cross-section in the parton model can be expressed as a sum:

$$\frac{d^2\sigma_{lp}}{dxdQ^2} = \frac{2\pi\alpha^2}{Q^4}[1 + (1 - y)^2] \sum_a e_a^2 f_a(x) \tag{3.7}$$

A derivation of the inelastic lepton-proton scattering cross-section that does not use a parton model yields a result in terms of structure functions:

$$\frac{d^2\sigma_{lp}}{dxdQ^2} = \frac{4\pi\alpha^2}{Q^4}\left[\frac{(1 - y)}{x} F_2(x, Q^2) + y^2 F_1(x, Q^2)\right]. \tag{3.8}$$

Comparing Equations 3.7 and 3.8, the structure functions can be expressed as:

$$2xF_1(x,Q^2) = F_2(x,Q^2) = x\sum_a e_a f_a(x). \tag{3.9}$$

These are momentum-weighted and charge-weighted sums over the parton distribution functions f_a. The Callan-Gross Relation $F_2(x,Q^2) = 2xF_1(x,Q^2)$ must hold for this, and determines that the quarks must be spin-1/2. A compacted cross-section with the quark model can be written as:

$$\frac{d^2\sigma_{DIS}}{dxdQ^2} = \frac{2\pi\alpha^2}{Q^4}[1+(1-y)^2]F_2(x) \tag{3.10}$$

$$\rightarrow \sigma_{DIS}(x,Q^2) = \sum_a \int dx f_a(x)\sigma_a(x,Q^2). \tag{3.11}$$

Here in Equation 3.11, the hard component σ_a has been (naïvely) factored out from the softer and non-perturbative parton distribution f_a. This process works well for DIS, though in processes seen at the LHC and other hadron-hadron colliders, questions arise. Consider the case of the Drell-Yan (DY) process [17], shown in Figure 3.2.

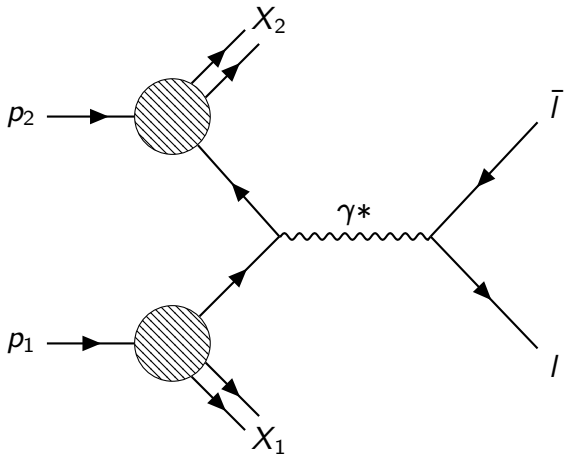


Figure 3.2: Feynman diagram for the Drell-Yan process [17].

Here $p_{1,2}$ are the protons, $X_{1,2}$ are the hadronic remnants, l and $\bar{l}$ are leptons and γ^* is the virtual photon. So far, the parton elucidated between Equations 3.4-3.11 has only considered quarks directly, with the gluons appearing consequentially as missing momenta. In a DY process there are two partons interacting electromagnetically and this partonic cross-section can be calculated perturbatively,

though as color-charged objects they may also interact via gluons, which this parton model does not explicitly account for. The question is whether this additional QCD interaction appears as a correction to the parton-antiparton interaction or is part of the softer nucleon interactions [18]. Some factorisation scale μ must be selected (dependent on the process being studied), allowing the corrections to be factored into renormalised parton distribution functions:

$$\sigma_{hh} = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu^2) f_b(x_2, \mu^2) \sigma_{a,b}(x_1, x_2, Q^2/\mu^2), \tag{3.12}$$

This cross-section describes hadronic collisions where the addition of and squaring of individual amplitudes has been replaced by a convolution of a partonic cross-section $\sigma_{a,b}$ and parton distribution functions f_a and f_b. There has been a factorisation of probabilities into parton distribution functions f_a, and f_b, where a and b need to be antiparticles of each other for photon production. Additionally, the partonic cross-section $\sigma_{a,b}$ is independently calculable through perturbation theory. This process of applying a factorisation scale, $\mu_F^2 = Q^2$ in the DY case, is known as QCD factorisation and leads to the QCD improved parton model.

3.2 Transverse Momentum Dependent Parton Distribution Functions

3.2.1 From PDFs to TMDs

So far, the entirety of the parton model and QCD factorization have been discussed only considering collinear momentum fractions [19]. Although the nucleon is not a one-dimensional object, this assumption is not invalid. At the high energies in experiment the partons within nucleons can be thought of as travelling essentially parallel to each other. In the context of large momentum transfers, where $Q \gg \Lambda_{QCD} \sim 200$ MeV, the contribution of transverse components is small but not negligible. Λ_{QCD} is the charecteristic scale of QCD. The internal dynamics of the nucleon, often known as the confined motion, can be represented in more than a single dimension. PDFs that are parameterised in additional dimensions, such as transverse momentum k_T, form a set known as Transverse Momentum Dependent PDFs, TMDs. A review of

TMDs and their associated theory can be found in [20]. Different sets of TMDs exist for quarks and gluons, the former of which have been measured in some detail [21, 22, 23, 24]. Although there is no generally agreed framework for the parameterization and evolution of TMDs, as with PDFs, their definitions have been settled in [25], following the most universal functions.

Intuitively, in the picture of the parton model previously explored, the extension to TMDs can be carried out simply by allowing the PDFs carry an extra parameter of $\boldsymbol{k}_T$ [20]:

$$\frac{d\sigma_{pp}}{dQ^2 dy d\boldsymbol{q}_T} \approx \sum_a \int\int d\boldsymbol{k}_{T1} d\boldsymbol{k}_{T2} f_a(x_1, \boldsymbol{k}_{T1}) f_a(x_2, \boldsymbol{k}_{T2}) \tag{3.13}$$

$$\times \delta(\boldsymbol{q}_T - \boldsymbol{k}_{T1} - \boldsymbol{k}_{T2}) \sigma_B(x_1, x_2, \boldsymbol{k}_{T1}, \boldsymbol{k}_{T2}, Q^2). \tag{3.14}$$

The dependence of the DY cross-section on the resultant lepton state's transverse momentum q_T makes the cross-section sensitive to measuring TMDs. This type of process is sometimes known as a two-scale probe. While there exists the high energy scale of $Q \gg \Lambda_{QCD}$, there also exists a second scale. Here, the observable on a lower scale q_T, with $q_T > \Lambda_{QCD}$, gives access to the confined motion inside the proton in additional dimensions. This does, however, introduce a constraint. For the TMD factorization to be applicable, $Q^2 \gg q_T \gg \Lambda_{QCD}$ must be enforced [26]. As well as kinematical constraints, QCD also presents some color-based conditions. If the produced final state does not consist of color singlets, the color flow in the final state cannot be isolated from the TMD. This prevents the use of subprocesses that may exhibit final state radiation for the measurement of TMDs. Furthermore, any state with more than two gluons interacting cannot be used. In this case, double parton scattering (DPS) smears out the information of the TMD in an ill-defined way.

Multiple subprocesses are capable of preserving TMD factorisation, such as the Higgs boson production [27]. However, with $gg \rightarrow H$ the presence of high backgrounds and limited statistics present significant hurdles. Another option could be the production of C-even quarkonia at leading order, though the acceptance limits of ATLAS combined with the low transverse momentum limits make this unviable. Preferable conditions for accessing gluon TMDs in gg fusion can be found with a subprocess of vector quarkonium produced in association with an isolated photon [26]. This state allows quarkonia to

be produced at leading order with a color-singlet final state and suppressed color-octet contributions that would break the TMD factorization. Additionally, the contribution of feed-down from C-even state decays is predicted to be suppressed [26].

3.2.2 Studying Gluon TMDs with a $J/\psi + \gamma$ final state

The cross-section for subprocesses including gluon TMDs, when initiated by gg scattering, is a convolution of the short-distance matrix elements and the gluon TMD correlators. The correlator and matrix elements are denoted as Φ and $\mathcal{M}$, respectively [26]:

$$d\sigma = \frac{(2\pi)^4}{8s^2} \int d^2\boldsymbol{k}_{1T} d^2\boldsymbol{k}_{2T} \delta^2(\boldsymbol{k}_{1T} + \boldsymbol{k}_{2T} - \boldsymbol{q}_T) \mathcal{M}_{\mu\rho} (\mathcal{M}_{\nu\sigma})^* \Phi_g^{\mu\nu}(x_1, \boldsymbol{k}_{1T}) \Phi_g^{\rho\sigma}(x_2, \boldsymbol{k}_{2T}) d\mathcal{R} \tag{3.15}$$

$\boldsymbol{q}_T$ is the transverse momentum of the gg system, and the Feynman diagram is shown in Figure 3.3, where the heavy quarkonium $\mathcal{Q}$ is a J/ψ in this analysis.

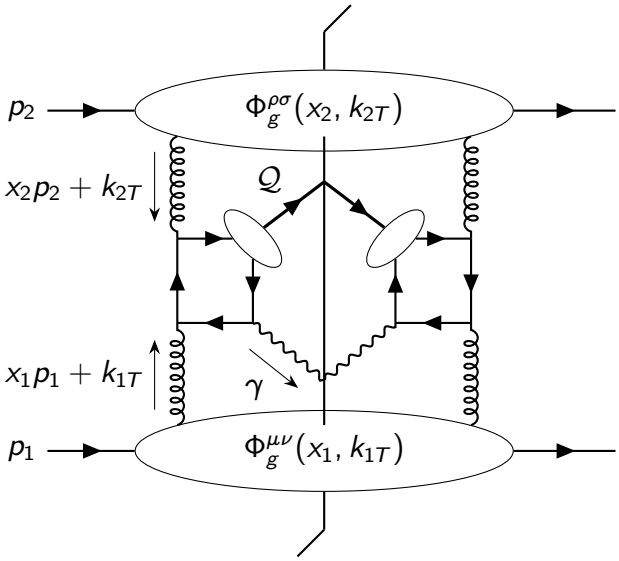


Figure 3.3: Feynman diagram for $pp \to \mathcal{Q} + \gamma + X$ with TMD factorization [26].

For an unpolarised proton, the gluon correlator expands to the following:

$$\Phi_g^{\mu\nu}(x, \boldsymbol{k}_T) = -\frac{1}{2x} \left\{ g_T^{\mu\nu} f_1^g - \left(\frac{k_T^\mu k_T^\nu}{m^2} + g_T^{\mu\nu} \frac{\boldsymbol{k}_T^2}{2m^2} \right) h_1^{\perp g} \right\}. \tag{3.16}$$

In this formulation, m is the proton mass, and k_T^μ are the gluon transverse momentum components. Two individual TMDs are introduced in Equation 3.16 [28].

$f_1^g(x, k_T)$ describes the probability of finding an unpolarised gluon with some collinear momentum fraction x and transverse momentum k_T, inside an unpolarised proton. Similarly, $h_1^{\perp g}(x, k_T)$ describes the probability of finding a linearly *polarised* gluon with some collinear momentum fraction x and transverse momentum k_T.

With this, the cross-section for the subprocess $gg \to J/\psi + \gamma$ is expressed as:

$$\frac{d\sigma}{dQdyd^2\boldsymbol{q}_T d\Omega_{cs}} = \frac{C_0(Q^2 - M_Q^2)}{sQ^3 D}\Bigg\{ F_1 \mathcal{C}\left[f_1^g f_1^g\right] + F_3\cos(2\phi_{cs})\mathcal{C}\left[w_3 f_1^g h_1^{\perp g} + x_1 \leftrightarrow x_2\right] + F_4\cos(4\phi_{cs})\mathcal{C}\left[w_4 h_1^{\perp g} h_1^{\perp g}\right]\Bigg\} + \mathcal{O}\left(\frac{q_T^2}{Q^2}\right). \quad (3.17)$$

Here $\mathcal{C}$ denotes a convolution of terms in square brackets, with corresponding weights w_i.

$$\mathcal{C}[w\, f\, g] \equiv \int d^2\boldsymbol{k}_{1T} \int d^2\boldsymbol{k}_{2T}\, \delta^2(\boldsymbol{k}_{1T} + \boldsymbol{k}_{2T} - \boldsymbol{q}_T) w(\boldsymbol{k}_{1T}, \boldsymbol{k}_{2T}) f(x_1, \boldsymbol{k}_{1T}^2) g(x_2, \boldsymbol{k}_{2T}^2), \quad (3.18)$$

with weight $w_1 = 1$ and w_3 and w_4 defined in equations 3.25 and 3.26. All present angular variables are in the Collins-Soper frame, detailed in Appendix A, with solid angle $d\Omega_{cs} = d\cos\theta_{cs} d\phi_{cs}$. λ is a dimensionless mass measure defined as $\lambda = \frac{Q^2}{M^2}$. Q and M are the invariant mass of the system, and the mass of the J/ψ meson, respectively. The remaining terms are defined as follows:

$$C_0 = 4\alpha_s^2 \alpha_{\text{em}} e_c^2 |R_0(0)|^2/(3M^3), \quad (3.19)$$

$$D = \left((\lambda+1)^2 - (\lambda-1)^2\cos^2\theta_{cs}\right)^2, \quad (3.20)$$

$$F_1 = 1 + 2\lambda + 9\lambda^2 + (6\lambda^2 - 2)\cos^2_{cs}\theta + (\lambda-1)^2\cos^4_{cs}\theta, \quad (3.21)$$

$$F_3 = 4\lambda\sin^2\theta_{cs}, \quad (3.22)$$

$$F_4 = (\lambda-1)^2\sin^4\theta_{cs}, \quad (3.23)$$

$$w_1 = 1 \quad (3.24)$$

$$w_3 = \frac{\boldsymbol{q}_T^2\boldsymbol{k}_{2T}^2 - 2(\boldsymbol{q}_T\cdot\boldsymbol{k}_{2T})^2}{2m^2\boldsymbol{q}_T^2}, \quad (3.25)$$

$$w_4 = 2\left[\frac{\boldsymbol{k}_{1T}\cdot\boldsymbol{k}_{2T}}{2m^2} - \frac{(\boldsymbol{k}_{1T}\cdot\boldsymbol{q}_T)(\boldsymbol{k}_{2T}\cdot\boldsymbol{q}_T)}{m^2\boldsymbol{q}_T^2}\right]^2 - \frac{\boldsymbol{k}_{1T}^2\boldsymbol{k}_{2T}^2}{4m^4}, \quad (3.26)$$

$R_0(0)$ is the quarkonium radial wave function, and $\alpha_{s,em}$ are the QCD and EM

coupling constants. TMD factorisation is expected to hold in regions with $Q^2 \gg k_T^2$ ($\lambda \gg 1$). With the dependence of the above terms on powers of λ, F_3 becomes negligible with respect to $F_{1,4}$. This allows Equation 3.17 to be reduced to the following form:

$$\frac{d\sigma}{dQdyd^2q_Td\Omega_{cs}} = \frac{C_0(Q^2-M^2)}{s\,Q^3D}F_1\mathcal{C}[f_1^g f_1^g]\left\{1+\frac{F_4}{F_1}\cos(4\phi_{cs})\frac{\mathcal{C}[w_4 h_1^{\perp g} h_1^{\perp g}]}{\mathcal{C}[f_1^g f_1^g]}\right\}. \quad (3.27)$$

The full transverse momentum of the resultant system produced by the $gg \to J/\psi + \gamma$ subprocess is present only in convolution terms. Dependence on the azimuthal Collins-Soper angle is present only in terms associated with the polarised gluon contribution $h_1^{\perp g}$. The polar Collins-Soper angle θ_{cs} appears in common factors D and F_1, and through the ratio F_4/F_1.

$$\frac{F_4}{F_1} = \frac{\sin^4\theta_{cs}}{9+6\cos^2\theta_{cs}+\cos^4\theta_{cs}} = \frac{(1-z^2)^2}{9+6z^2+z^4} \quad (3.28)$$

The terms associated with the identity at the start of the curly braces are related to the unpolarised gluon TMD contributions. These are present as the convolution $\mathcal{C}[f_1^g f_1^g]$. The remaining term in the braces encapsulates the polarised gluon effects, $\mathcal{C}[w_4 h_1^{\perp g} h_1^{\perp g}]$. Additional dependence on Collins-Soper angles is present in two forms: firstly, on the polar angle through $\cos(\theta_{cs}) = z$, from Equation 3.28; secondly, dependence on the azimuthal angle, ϕ_{cs}, in the form of the modulation $\cos 4\phi_{cs}$.

When integrating over the angular terms, the polarised convolution term in Equation 3.27 zeroes out, leaving the following:

$$\frac{d\sigma}{dQdyd^2q_T} = \frac{C_0(Q^2-M^2)}{sQ^3}F_1'\mathcal{C}[f_1^g f_1^g]. \quad (3.29)$$

Here, F_1' is the ratio F_1/D after their angular dependence is integrated out.

Typical modelling in literature [29] assumes a Gaussian structure for the unpolarised contribution f_1^g:

$$f_1^g(x,k_T^2) = \frac{G(x)}{\langle k_T^2\rangle\sqrt{2\pi}}\exp\Big(-\frac{k_T^2}{2\langle k_T^2\rangle}\Big) \quad (3.30)$$

$G(x)$ represents the collinear PDF for gluons, $\langle k_T^2\rangle$ is the mean gluon transverse momentum squared, with no assumed dependence on x. The convolution can be calculated analytically:

$$\mathcal{C}\,[f_1^g f_1^g] = \frac{G(x_1)G(x_2)}{2\langle k_T^2\rangle\sqrt{2\pi}}\exp\Big(-\frac{\bar{q_T}^2}{2(2\langle k_T^2\rangle)}\Big). \quad (3.31)$$

x_1 and x_2 are the collinear momentum fractions of the two gluons.

In cases with no polarised contributions ($h_1^{\perp g} = 0$), or when angular terms are integrated, Equations 3.31 and 3.27 allow direct measurement of the $\langle k_T^2 \rangle$ parameter. This $\langle k_T^2 \rangle$ can be determined from the cross-section of q_T^2. The minimum transverse momentum thresholds of the ATLAS detector for photons and muons obscure the measurement of q_T^2 by breaking the relation between $\langle k_T^2 \rangle$ and q_T^2. Any analysis process to extract $\langle k_T^2 \rangle$ undistorted by these acceptance cuts must be carefully designed.

3.2.3 Feasibility Study of Polarised Gluon Effects

With an ideal experimental setup, Equation 3.27 indicates that the cross-section with respect to the Collins-Soper azimuthal angle will exhibit modulation dependent on the size of $h_1^{\perp g}$. At ATLAS, the minimum transverse momentum (p_T) of a reconstructed muon is $p_T(\mu) > 4$ GeV. As this analysis concerns the $J/\psi \to \mu\mu$ channel, the prescence of this cut on both of the muons translates to $p_T(J/\psi) > 8$ GeV. To produce a system with a balanced transverse momentum requires a symmetrical cut on photons, $p_T(\gamma) > 8 - 9$ GeV. This results in a distorted ϕ_{cs} distribution. Any measurement of the modulation of $\cos 4\phi_{cs}$ becomes difficult, and consequentially it is difficult to analyse $h_1^{\perp g}$.

The production of the $J/\psi + \gamma$ system from gg fusion is simulated with PYTHIA 8[30]. No polarised gluon contribution was present in this simulation ($h_1^{\perp g} = 0$). To emulate the effects of non-zero $h_1^{\perp g}$, events are given a weight determined by the braced term in Equation 3.27. Restrictions on the parameterization of $h_1^{\perp g}$ are given by a model-independent positivity bound [28]:

$$k_T^2 |h_1^{\perp g}(x, k_T^2)| \leq 2M^2 f_1(x, k_T^2). \tag{3.32}$$

Two models are considered for simulating the gluon k_T effects. Model 1, Equation 3.33, is based on the Gaussian model in Equation 3.30.

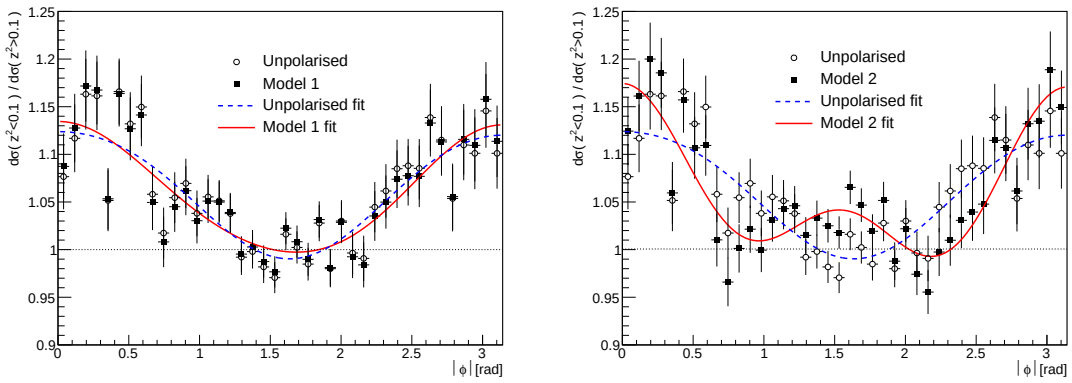
$$h_1^{\perp g}(x, k_T^2) = \frac{M^2 G(x)}{\pi \langle k_T^2 \rangle^2} \exp\left(1 - \frac{k_T^2}{r\langle k_T^2 \rangle}\right) \tag{3.33}$$

Model 2 saturates the bound in Equation 3.32.

Before reducing to only the unpolarised contribution, it can be seen $h_1^{\perp g}$ is more prevalent in the low-$|z|$ region, where $z = \cos\theta_{cs}$. By picking a working point

in z, the region where the azimuthal modulation is enhanced can be identified. Additionally, distortion from acceptance cuts is independent of z. The ratio of the distribution of ϕ_{cs} in the low-z region against that in the high-$|z|$ region can be used to remove the acceptance effects and retain sensitivity to measure h_1^g. Using a cut around $|z|^2 = 0.1$ halves the sample and is chosen for the working point.

Figure 3.4 compares the ϕ_{cs} modulation of the ratioed distributions. This shows



(a) (b)

Figure 3.4: The ratios of differential cross-sections for events with $z^2 < 0.1$ over the events with $z^2 > 0.1$, for Model 1 (left) and Model 2 (right) described in the text. On both plots, open points describe the unweighted distribution, corresponding to $h_1^{\perp g} = 0$, with the fit shown as dashed blue lines. The crosses describe the weighted distributions, with the fits shown as solid red lines.

both the models M1 and M2, the model with no simulated polarisation, and the fits to the models to extract the parameters. The fit applied is given by:

$$\frac{d\sigma(z^2 < 0.1)}{d\sigma(z^2 > 0.1)} = p_0 \cdot (1 + p_1 \cos\phi_{cs} + p_2 \cos 2\phi_{cs} + p_3 \cos 3\phi_{cs} + p_4 \cos 4\phi_{cs}) \,. \quad (3.34)$$

The statistics of this MC study are approximately equivalent to an integrated luminosity of 60 fb^{-1} at 13 TeV. No additional attempt to correct for acceptance effects is made here. From the fits in Figure 3.4, the parameters associated with the modulation in $\cos 4\phi_{cs}$ are not expected to change by a significant margin in the case of M1 ($\Delta p_4(M1) = (9 \pm 6)10^{-3}$). The case of M2 ($\Delta p_4(M2) =$

$(50 \pm 6)10^{-3}$) could allow a measurement with sufficient real-world statistics. The total effective integrated luminosity available for a `2mu4`-type trigger in ATLAS Run II data is on the order of 2.6 $\text{fb}^{-1}(\pm 1.13\%)$. `2mu4`-type triggers are those that select events where the two muons of a minimum transverse momentum of 4 GeV are detected. Sensitivity to $h_1^{\perp g}$ is not achievable with this dataset, though the possibility to measure f_1^g remains.

3.2.4 Effects of Minimal Acceptance Cuts

Considering Equations 3.27 and 3.29, the transverse momentum of the $J/\psi + \gamma$ system reflects the convolution of two f_1^g functions. With the model in Equation 3.31 the distribution of this momentum equates to $\langle k_T^2 \rangle$, and parameterises the TMD f_1^g. As with the $h_1^{\perp g}$ case, acceptance cuts prevent f_1^g parameterised directly from q_T. In addition to the minimal acceptance cuts on the transverse momentum of muons of $p_T(\mu) > 4$ GeV, photons also have a minimum threshold of $p_T > 5$ GeV. With the same sample as used in Section 3.2.3, the effects of these acceptance thresholds can be studied for the case of parameterising f_1^g.

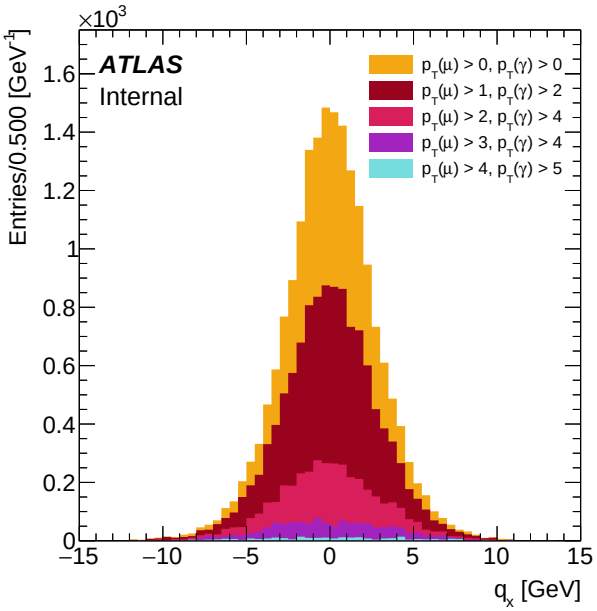


Figure 3.5: Distribution of the x-component of $J/\psi + \gamma$ system transverse momentum in the lab frame before minimal acceptance (yellow), and through progressively increasing cuts up to the minimum realistic cuts of $p_T(\mu) > 4$ GeV and $p_T(\gamma) > 5$ GeV (cyan). This distribution is in the lab frame, hence the y-component is identically distributed.

Take either of the lab frame components of q_T, $q_x = q_T \cos\phi_{\text{lab}}$, or $q_y = q_T \sin\phi_{\text{lab}}$. q_y will be symmetric with q_x. Figure 3.5 compares the distributions before and after acceptance cuts. After the cuts emulating the acceptance effects are applied, the distribution is no longer the expected Gaussian shape and has significantly broadened. Under acceptance cuts, the relationship between the observed q_T distribution and the gluon transverse momentum k_T is broken. The observed q_T distribution is no longer reflective of the underlying gluon transverse momentum, and hence any measurement of the features of this distribution should not be used to measure the TMD.

To recover an appropriate distribution, the basis in which the underlying transverse momentum is measured can be changed. In order to move out of the lab frame, q_T can be decomposed into two new components using the typically back-to-back configuration of the J/ψ and γ. First define $\epsilon = \pi - \Delta\phi$, and express q_T^2 in terms of $p_T(J/\psi)$ and $p_T(\gamma)$. With these starting points it is possible to transform to a new basis:

$$q_T^2 = p_T(J/\psi)^2 + p_T(\gamma)^2 + 2p_T(J/\psi)p_T(\gamma)\cos\Delta\phi \tag{3.35}$$
$$= p_T(J/\psi)^2 + p_T(\gamma)^2 - 2p_T(J/\psi)p_T(\gamma)\cos\epsilon \tag{3.36}$$
$$\simeq p_T(J/\psi)^2 + p_T(\gamma)^2 - 2p_T(J/\psi)p_T(\gamma)\left(1 - \frac{\epsilon^2}{2}\right) \tag{3.37}$$
$$= (p_T(J/\psi) - p_T(\gamma))^2 + p_T(J/\psi)p_T(\gamma)\sin^2\epsilon \tag{3.38}$$
$$= (q_T^A)^2 + (q_T^B)^2. \tag{3.39}$$
$$q_T^A = p_T(J/\psi) - p_T(\gamma) \tag{3.40}$$
$$q_T^B = \sqrt{p_T(J/\psi)p_T(\gamma)}\sin(\epsilon). \tag{3.41}$$

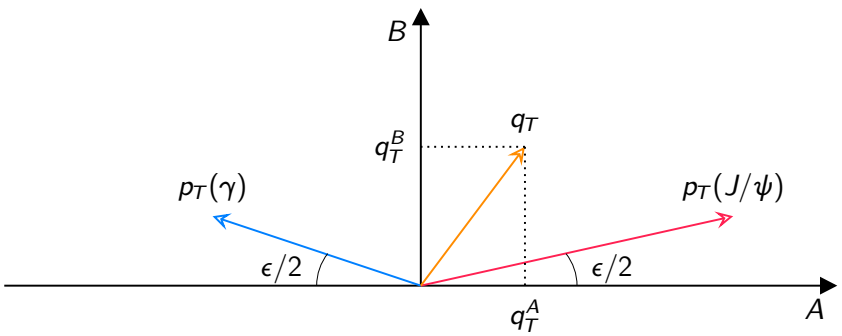


Figure 3.6: Diagram of the A-B frame and the vectors representing q_T^A and q_T^B.

q_T^A, and q_T^B, Figures 3.6, 3.7 show the $J/\psi + \gamma$ system transverse momentum components along two new axes, A and B, respectively. These axes are defined uniquely for each event. Each increase in acceptance cuts reduces the statistics available. For the $p_T(\mu) > 4$ GeV, $p_T(\gamma) > 5$ GeV case in Figure 3.7 there are almost no events left in the original Monte-Carlo (MC) sample. A second sample with raised generator level cuts of 3.5 GeV for muons and 4 GeV cuts for photons is used here. The effects of taking the distributions of the new variables

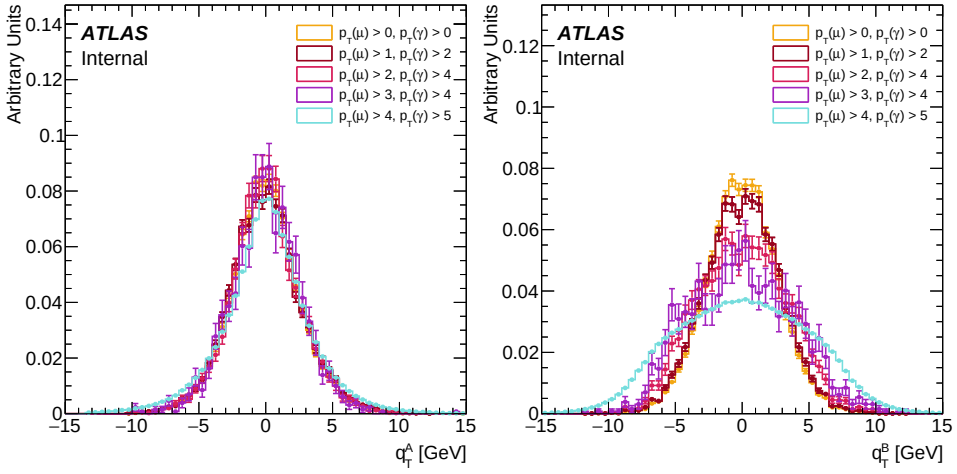


Figure 3.7: Normalised distributions of MC simulated (left) q_T^A and (right) q_T^B overlaid with progressively increasing acceptance cuts. Distribution recentred to zero. Acceptance cuts increase from none to ATLAS minimal. q_T^A distributions are recentred to manage asymmetric acceptance cuts.

and passing them through progressively increasing simulated acceptance cuts are seen in Figure 3.7. q_T^B can be seen to broaden and change shape as the acceptance cuts increase, absorbing the detrimental effects that were previously seen in q_x. As q_T^A is the scalar difference in J/ψ and γ transverse momenta, the mean will increase proportionally to the imbalance between the photon minimum acceptance cut and the effective J/ψ acceptance cut. This corresponds to the distribution of q_T^A shifting rightwards. For Figure 3.7b this shift has been removed by recentering the distribution. After correction, it can be seen that the q_T^A distribution is resistant to broadening as acceptance cuts increase. Using this new basis f_1^g can be parameterised through the variance of q_T^A, which is unaffected by the acceptance cuts, and still representative of the underlying k_T.

Part II

Experimental Setup & ATLAS Contributions

4. The ATLAS Experiment & The Large Hadron Collider

4.1 The Large Hadron Collider

The Large Hadron Collider (LHC)[31, 32] is a modern proton-proton synchrotron located at the CERN complex, outside Meyrin, Switzerland. At a circumference of approximately 27 km the LHC is currently the largest particle accelerator on the planet, and similarly produces the highest centre-of-mass energies for pp collisions, reaching $\sqrt{s} = 13$ TeV. The LHC was constructed from $1998 - 2008$ in the tunnel previously occupied by the Large Electron-Positron Collider (LEP). It extends from the CERN site, across the French border towards the Jura Mountains and round returning through Geneva and to Meyrin. It is positioned on a plane inclined at 1.41% between $45 - 175$ metres underground, a measure to protect both the experimental recordings from a large fraction of cosmic ray interference and the surface from any hazardous emissions from either the particle collisions or the synchrotron radiation.

The LHC is supported by a number of smaller accelerator systems designed to provide proton bunches of the correct spacing and energy [33], these are visible in Figure 4.1. The acceleration process begins with the injection of Hydrogen anions[1] into LINAC [34][2], which provides an acceleration to energies of 160 MeV before the beam is passed on to the Proton Synchrotron Booster (PSB) [35]. This continuous beam from LINAC is stripped of electrons and split consecutively

[1] Ions in Run 2.

[2] This was LINAC 2 until 2020, reaching 50 MeV. During Long Shutdown 2 (LS2) this was obsolesced by LINAC 4

amongst the four PSB rings, accelerated to 1.4 GeV, and injected into the Proton Synchrotron (PS). PS [36] produces bunches spaced by 25 ns, accelerates these to 25 GeV, and passes them over to the Super Proton Synchrotron (SPS) for the final energy increase up to 450 GeV [37]. Bunches inserted from SPS into the LHC

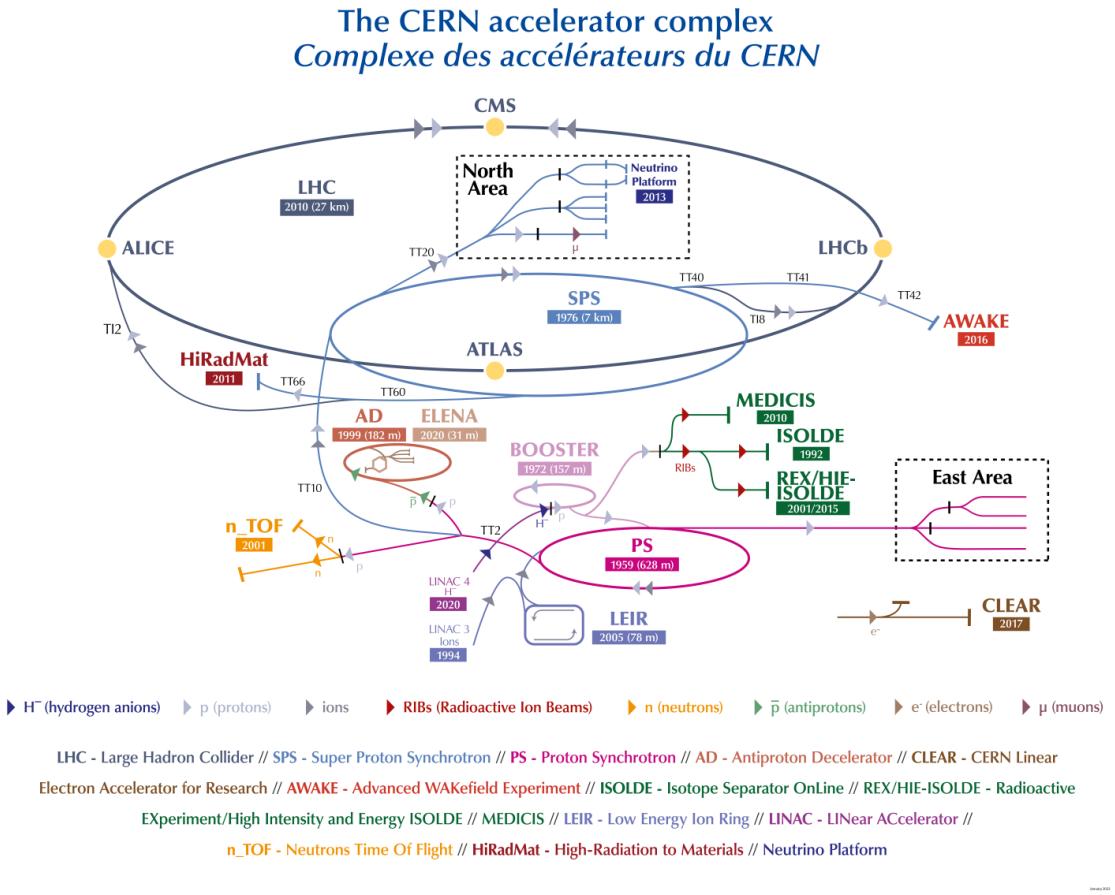


Figure 4.1: CERN Accelerator Complex schematic [38].

are brought from 450 GeV up to 6.8 TeV over a timespan of approximately 20 minutes [31]. This is driven by the 16 RF cavities around the LHC ring, oscillating at 400 MHz each providing up to 2 MV of acceleration. The LHC beam pipe is a dual-core design allowing bunches to travel in different directions allowing the Centre of Momentum (CoM) to coincide with the lab frame of a detector. For a single turn through the LHC, a proton bunch will pass through the magnetic field of 1232 superconducting dipole magnets [39]. The magnets are of Niobium-titanite construction, are held at 1.9 K, and produce a field of $\sim$ 8.3 T [40]. In order to amplify luminosity at the interaction point a focussing effect is produced by

the 474 quadrupole magnets, arranged into FODO lattices positioned throughout the circumference [41].

There are 8 experiments that are part of the LHC, four of which are positioned around the circumference. These 4 are ATLAS, CMS [42, 43], LHCb [44], and ALICE [45]. located at points 1, 5, 8, and 2, respectively. These experiments (and others) are collectively engaged in Run 3 of the LHC. After commissioning and a magnet quench incident in 2008, the LHC entered full operation on 20 November 2009 with first collisions at 7 TeV occurring on 30 March 2010. Run 1 concluded in early 2013, delivering a *pp* integrated luminosity of 23.3 fb^{-1} to ATLAS and CMS [46]. Run 2 began in 2015 and extended into 2018, raising the collision energy to 13 TeV, and delivering a total of 160 fb^{-1} [47]. After Long Shutdown 2 (LS2)[3] and the upgrades to both the LHC and detectors associated with this, the LHC began Run 3 operations with an increase in the collision energy, again, to 13.6 TeV, and is currently ongoing in 2024. After Run 3, the LHC enters Long Shutdown 3 (LS3) for a series of upgrades to prepare for the High-luminosity (HL) phase where there is a roughly eight-fold increase in peak luminosity planned, and the total integrated luminosity over the HL-LHC period from Run 4 to the end of Run 6 planned in 2038 aims to be 3000 fb^{-1}.

4.2 The ATLAS Detector

4.2.1 Overview

The ATLAS Detector [48] is a general-purpose detector with cylindrical geometry and a forwards-backwards symmetry. It has nearly complete coverage, approaching 4π, in solid angle around the interaction point (IP) which is located centrally within the detector layout. The detector is located at Point 1 on the LHC ring, next to the Swiss-French border, approximately 100 m below the surface, with the ATLAS control room and supporting infrastructure located directly above at the Meyrin site of CERN. There are a number of main subsystems within the ATLAS detector, a cross-section of the detector outlining their positions is shown in Figure 4.2. The detector is generally split into two regions, the barrel and endcap. Barrel regions can be described as the volume around the

[3] 2018-2022

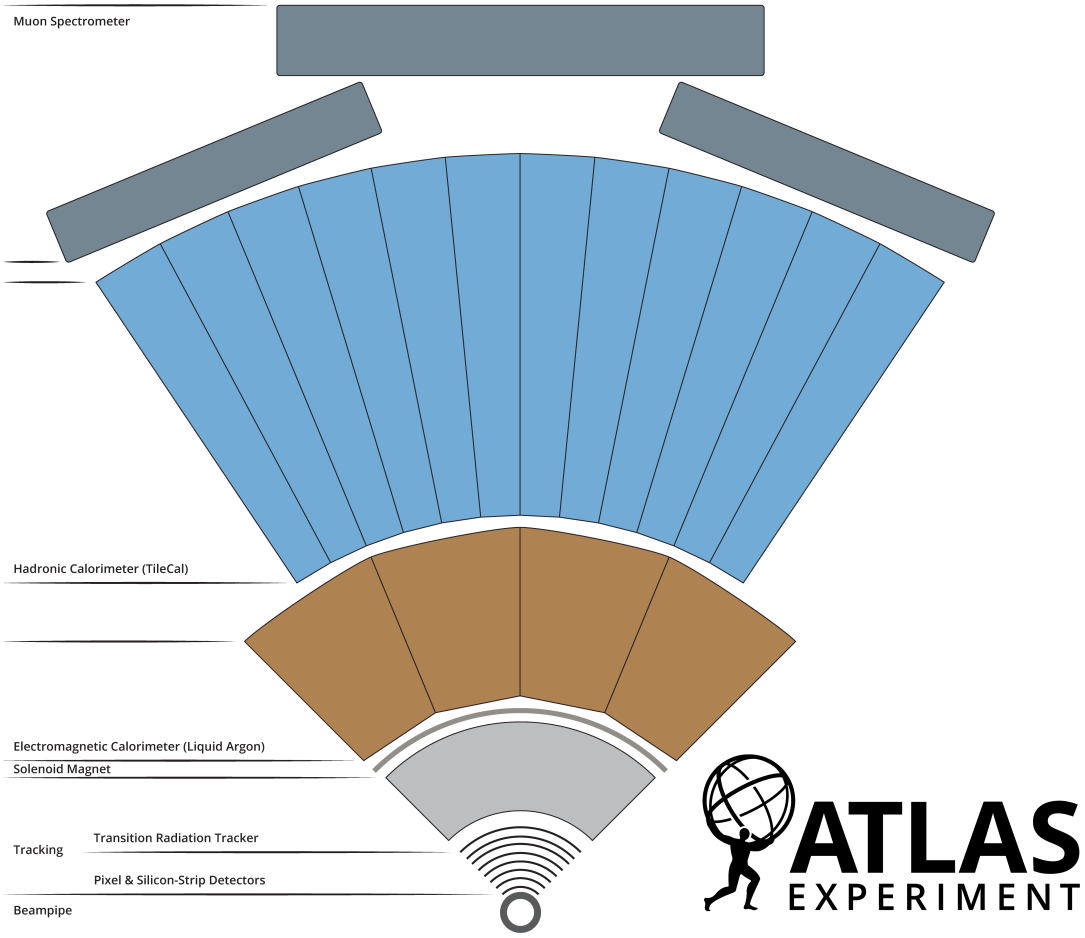


Figure 4.2: Cross-sectional slice of the ATLAS detector [49].

interaction point, high polar angle, symmetric around $\theta = \pi/2$. The endcap region is the cylindrical volume further down the beam line from the interaction point, at lower polar angles, symmetric around $\theta = 0$.

Starting centremost radially, closest to the interaction point there is the Inner Detector (ID) [50, 51] comprised of the Pixel Detector, the Semi-Conductor Tracker (SCT), and the Transition Radiation Tracker (TRT). This is used to provide tracking information for all charged particles created in an event. The ID is immersed within the central solenoidal magnet [52], producing a 2T field to generate a curvature in charged particle paths.

Outside this, the calorimeter systems begin with the Electromagnetic Calorimeter (ECal). The ATLAS ECal is a Liquid Argon (LAr) Calorimeter [53] in both the

barrel and endcap Regions, and is used for the precision measurement of the energy of electrons and photons. The Hadronic Calorimeter (HCal) systems follow the ECal, these are designed for the measurement of strongly interacting particles which typically pass through the ECal and form showers further out from the interaction point. Three main components form the HCal, the Tile Calorimeter (TileCal) [54] in the barrel region, and LAr calorimeters form the detector setup in the endcap and forward regions.

Last in the sequence is the Muon Spectrometer (MS) [55], placed in the outermost region to collect muons that typically penetrate through the earlier stages of the detector, or any other charged particles that may leak from the other inner stages. The MS is constructed from a multitude of technologies.

4.2.2 Coordinate system and common variables

To perform physics analysis at a pp experiment, several variables must be defined to accurately describe the physics objects that will be detected. The cylindrical geometry of ATLAS lends towards using a cylindrical coordinate system for analysing events and is defined from a Cartesian system centred around the interaction point. The x-axis points towards the centre of the LHC ring, the y-axis directly upwards, and the z-axis along the beamline. The Azimuthal angle ϕ in this cylindrical system begins at the x-axis, and the polar angle θ at the z-axis.

Energy E. The energy of particles that are produced in reactions. This can be assessed as the particle will deposit energy into the calorimeter materials as it travels outwards from the interaction point.

Transverse Momentum p_T. This is the component of a particles momentum in the plane transverse to the beam line. In ATLAS this is used in preference to overall momentum on the basis that it is a conserved variable and is more informative of the physics given the typical kinematic distribution of events that would be expected in a pp collision of the nature at the LHC.

Invariant mass m. Invariant mass can be used in the context of single particles or systems of them. Calculating the invariant mass of selected combinations of particles can yield information on the resonant particles within an event or interaction. This is typically calculated through combining

information collected on the momentum and energy of a particle, or collection of particles.

Rapidity y, Pseudorapidity η, and Angular Distance ΔR. These three variables help define a particle position after an interaction. Rapidity is an angular measure and relativistic measure of velocity defined as

$$y = -\frac{1}{2} \ln \frac{|E + p_z|}{|E - p_z|}. \tag{4.1}$$

Beneficially, rapidity differences are Lorentz invariant. When the particle masses are far smaller than the energy the rapidity can be approximated to the pseudorapidity, a more easily calculable and purely angular variable defined as

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{4.2}$$

When combined with the azimuthal angle ϕ, a Lorentz invariant distance measure appropriate for the cylindrical geometry can be defined:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{4.3}$$

Tracking Parameters z_0, d_0, and $\frac{q}{p}$. When combined with the beamspot position, θ and ϕ, these 5 variables are used to fully define a track in ATLAS. z_0 is the longitudinal impact parameter, measuring the closest approach of the track to the beamspot along the axis of the beam, d_0 is the distance of the closest approach of a track to the beamline in the transverse plane, and $\frac{q}{p}$ is the charge-momentum ratio of a reconstructed track.

Transverse displacement L_{xy}, and lifetime τ. L_{xy} is the transverse displacement of a point, typically the vertex of a particle decay, from the beamline. Due to the relativistic nature of particles and the fact L_{xy} is not a variable natural to the rest frame of any decaying particle, it is preferable to use the lifetime τ in many cases:

$$\tau = \frac{L_{xy} m_{parent}}{c \cdot p_T}. \tag{4.4}$$

In practice, τ becomes the pseudo-proper lifetime when the full parent hadron (e.g. B) mass and p_T are not reconstructed, and replaced with that of the child hadron (e.g. J/ψ), presuming the ratio is the same.

4.2.3 Inner Detector

The ATLAS Inner Detector (ID) [56, 57, 58], shown in Figure 4.3, is the first detector subsystem. It provides the backbone of tracking and vertexing information for events recorded during LHC runs. Starting at a radius of 3.3 cm the detector extends out to 1.05 m and is 6.2 m long. Innermost to this system, 3.3 cm from the beamline, is the pixel detector (PIX/IBL), consisting of 4 layers of pixels, the innermost Insertable B-Layer (IBL) was installed in 2014. The IBL has a smaller pixel size, of 50×250 μm, than the other pixel layers with 50×400 μm. It was inserted to increase the tracking resolution of the ID. PIX/IBL provides 92 million readout channels, 12 million from the IBL alone, with a total 1.9m^2 of silicon area.

The second subsystem of the ID is the semiconductor tracker (SCT) [59, 60, 61, 62]. As of Run 3 this is operating with 6.3 million readout channels consisting of silicon 'micro-strips', in 4088 modules arranged across 4 barrel layers (2112 modules) and 18 endcap disks, nine each side (988×2), covering an area of 60 m^2. Detector planes are offset by 40 mrad to assist in providing a readout accuracy as tight as 25 μm. All modules in the barrel region are identical, they consist of four sensors paired off and connected back to back. Across the four sensors in a module there are 3072 readout strips (768 per sensor), pitched identically at 80 μm. The endcap modules have a variable micro-strip pitch dependent on their position as a function of their wedge-like geometry, though still have 768 readout strips per sensor.

The outermost section of the ID consists of the Transition Radiation Tracker (TRT) [63]. This is constructed from 300,000 thin-walled drift tubes of 4 mm diameter with a total volume of 12 m^3. In the barrel region the straws are aligned parallel to the beam axis and are positioned from 560 mm to 1080 mm in radius. In the endcap region, the straws are aligned perpendicularly. Particle identification is performed with a combination of ionisation-loss curve analysis, and using the probability of certain particles to produce readout pulses in the TRT that cross certain low and high eV thresholds, LT and HT respectively. The HT threshold for example is used for electron-pion discrimination.

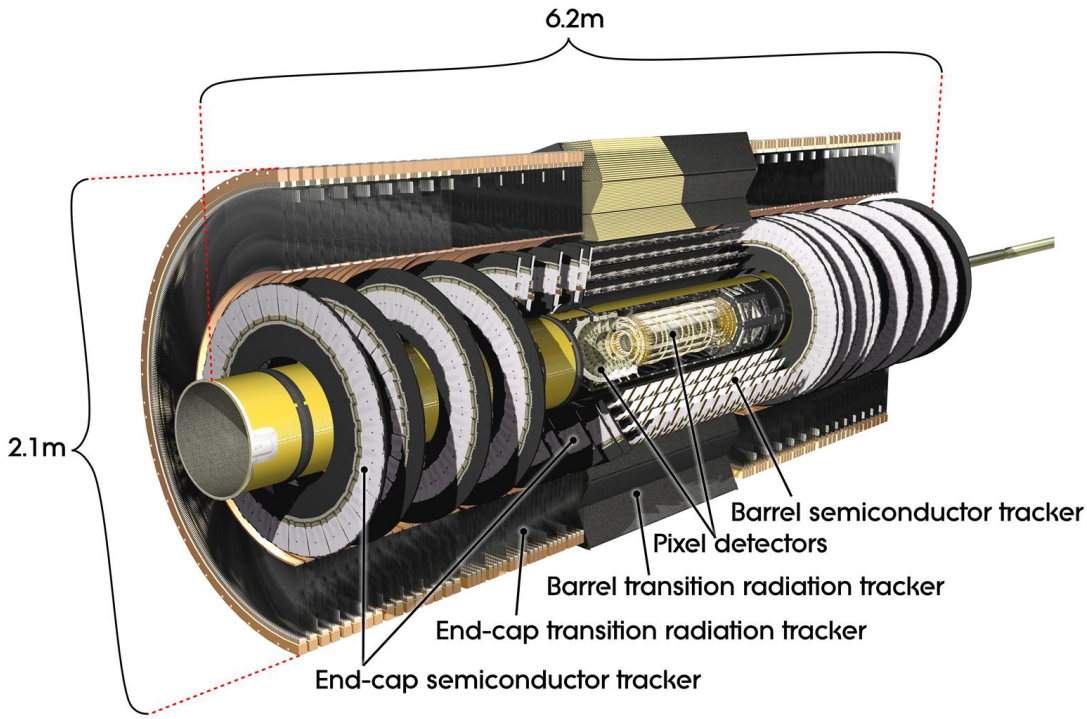


Figure 4.3: Cutaway of ATLAS Inner Detector systems [64].

4.2.4 Calorimetry

The ATLAS calorimeter system is composed of two detector types, each type predominantly assigned to the two different tasks of Electromagnetic (ECal) and Hadronic (HCal) energy measurements. The LAr Calorimeter system consists of four components, the LAr Forward Calorimeter (FCal) for both ECal and HCal purposes, the LAr hadronic end-cap (HEC), the LAr electromagnetic end-cap (EMEC), and the LAr electromagnetic barrel. HCal in the barrel region is performed by the Tile Calorimeter (TileCal).

4.2.4.1 Electromagnetic Calorimeter

The LAr ECals [53, 65, 66], shown in Figure 4.4, are designed to measure the energy of photons and electrons produced as a result of collisions at the interaction point. These are sampling calorimeters [4] with copper, lead, and tungsten

[4] Calorimeters with distinct mediums for initiating particle showers and measuring their energy.

absorbers, and the eponymous Liquid Argon active medium.

Electron and photon calorimetry is performed in the ECal systems by using the nature of EM showers. Produced electrons undergo bremsstrahlung emission of lower-energy photons inside the absorber material of the calorimeter. Photons, including those from the bremsstrahlung process, can undergo pair production of lower-energy electron-positron pairs. This process is initiated in the absorber material where the shower will evolve faster due to the shorter radiation length of the denser material, following this the shower will then ionise the active medium producing a current, this can be measured with readout electronics and used to produce an estimate of the energy of an incident photon or election.

The electromagnetic barrel calorimeter (EMB) [67] sits in the central region of the detector covering a rapidity range of $|\eta| < 1.475$. It is 6.8 m in length and extends through the radial distance 1.15-2.25 m. Absorber layers are arranged in an accordion-like structure with the Liquid Argon filling the spaces between, producing a hermetic design held at 89 K. The barrel is made from two sections, with a millimetres wide gap centred around $\eta = 0$. Each assembly contains 1024 absorbers constructed of stainless-steel coated lead. The ECal endcaps are disks 0.632 m thick, and extend to a radius of 2.077 m, covering a rapidity region from $1.52 < |\eta| < 3.2$. A crack region exists between $1.375 < |\eta| < 1.52$ for ID cooling.

The forward calorimeter doesn't fall cleanly into either category of EM or Hadronic calorimeter due to its dual-purpose design. It extends ECal and HCal out to cover a range of $3.1 < |\eta| < 4.9$, is constructed from three absorber layers, and again uses LAr as an active medium. The first absorber layer is copper, less dense than the later layers, designed to initiate and evolve photon and electron showers. The next two layers are denser tungsten, specifically to focus on hadronic depositions and measurement.

4.2.4.2 Hadronic Calorimeter

After punching through the EM calorimeter, hadronic particles deposit their remaining energy into the surrounding hadronic calorimeter. In the far forward region hadronic calorimetry is provided by the dual-purpose FCal. Coverage of the central region is provided by the Tile long barrel calorimeter (LB), extending through a region of $|\eta| < 1.0$ Further reach is provided by the Tile extended barrel

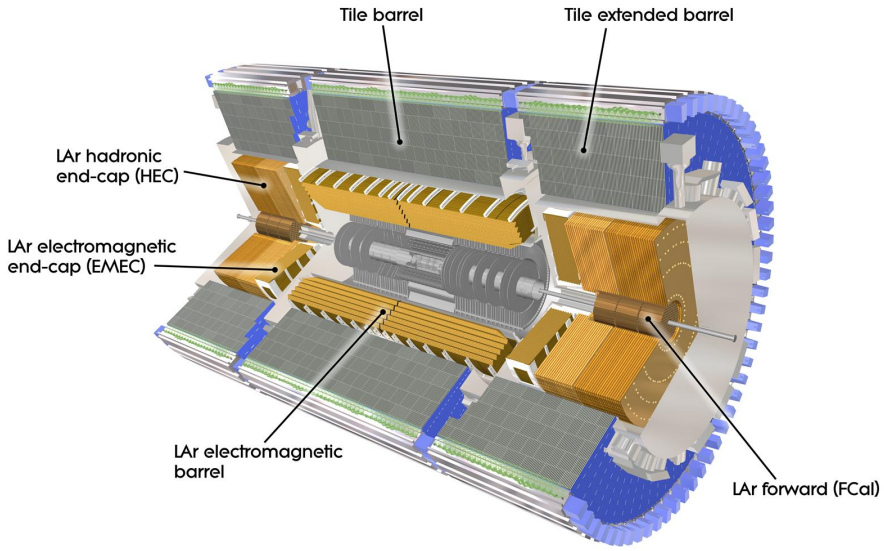


Figure 4.4: Cutaway of ATLAS Calorimeter System. [68].

(EB) calorimeters at each end of the LB, ranging from $0.8 < |\eta| < 1.7$ [69]. The central LB is 5.8 m in length and each EB is 2.6 m, both TileCals cover a radius from 2.28 m to 4.23 m. The gap region between these two systems is filled with scintillators to ensure complete coverage and reconstruction.

Like the LAr calorimeters, the Tile calorimeter is a sampling calorimeter [70]. The absorber in this case consists of steel with interspersed plastic providing a scintillating active medium. Scintillating layers, normal to the beam, are each connected to two wavelength shifting fibres. These, in turn, each relay detected signal to a dedicated photomultiplier tube (PMT) located outside the tile assembly.

There are three layers to the TileCal, A, BC, and D. In layers A and BC each tile is $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$ in size. Layer D has larger $\Delta\phi \times \Delta\eta = 0.1 \times 0.2$ sized tiles. There are 64 modules azimuthally, wedge shaped, providing a complete azimuthal enclosure. In total there are 5182 cells connected to 9852 PMTs [71].

There are also the hadronic endcap calorimeters (HEC) of LAr design [72, 73, 74]. The HEC cover a rapidity range of $1.5 < |\eta| < 3.2$ and are located within the same cryostat as the ECal endcaps. HEC is split into two wheels, HEC1 and HEC2, of length 0.82 m and 0.96 m. The inner HEC1 has copper plates of 25 mm thickness

and HEC2 a thicker 50 mm. Both wheels have 32 modules. This system has pads of size $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$ and 3072 readout channels.

4.2.5 Muon Spectrometer

The ATLAS muon spectrometer [75, 76, 77, 55] (MS), shown in Figure 4.5, lies on the outer layer of the detector and consists of five distinct technologies:

- Resistive Plate Chambers (RPC)
- Thin Gap Chambers (TGC)
- Monitored Drift Tubes (MDT)
- Cathode Strip Chambers (CSC)
- Micromegas (MM)

Each of these technologies serves a different purpose in the muon system. In general, the RPC and TGC detectors are used for their fast response time to allow proper triggering. MDTs are used for their precision tracking properties but operate on a larger timescale. The MS is positioned at the outer boundary of the ATLAS detector due to the behaviour of muons as minimally-interacting particles. In other words, muons should pass through the inner layers depositing only a small fraction of their energy. The MS consists of a barrel region and two end cap systems known as the 'wheels'. There are 10 total wheels, arranged into three stations, the intermediate of which is movable to provide access to the detector.

The barrel region of the MS is made from three stations and provides rapidity coverage $|\eta| < 1.05$. Innermost is an isolated MDT station followed by the intermediate station consisting of an MDT layered between two RPCs. Outermost in the barrel MS array is a consecutive MDT and RPC.

The first of the wheels are the two 'small wheels' positioned before the end-cap toroidal magnet. In Run 2 these small wheels consisted of TGCs layered on top of MDT arrays and a CSC in the most forward section. During Long Shutdown 2, these small wheels were replaced by the New Small Wheel 'NSW', designed to maintain efficiency as background increases throughout future runs. The New Small wheel utilises the new small strip TGCs (sTGCs) and the MM technologies. These new

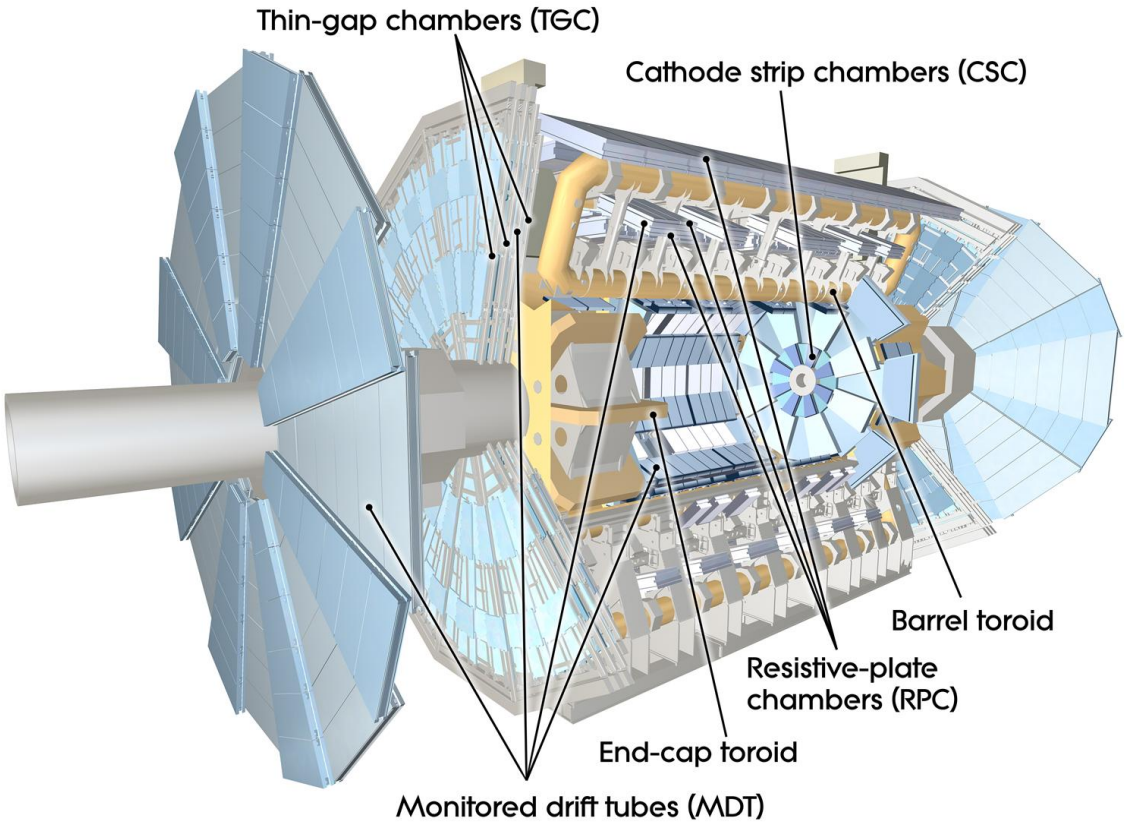


Figure 4.5: Cutaway of ATLAS Muon Subsystem. [78].

detectors are arranged into sectors built from four wedges. Two sTGC wedges sandwich two MM wedges, separated by a central spacer. The intermediate, movable, large wheel of range $1.0 < |\eta| < 2.7$ is constructed from an MDT layer and three TGC layers. A single outermost large muon wheel consist of MDTs, also extending as far as $|\eta| = 2.7$.

MDT in the barrel and endcap regions consists of 656 and 494 chambers, respectively. Each tube, 354384 of them, is 29.97 mm in diameter, constructed of aluminium and filled with a 0.93:0.07 Ar:CO_2 mixture at 3 bar. The wires within are Tungsten-Rhenium at a diameter of 50 μm. Holding a potential of 3080 V produces a maximum drift time of approximately 750 ns and a resolution of 75 μm.

CSC are between 7 and 8 m from the interaction point and around 1.5 m from the beamline. There are 32 wedge-shaped chambers, 16 small and 16 large, around 1.1 m in thickness. Each wheel consists of 8 of each chamber type,

alternating azimuthally. The chambers are angled at $11.59°$ off perpendicular from the beamline. T,, they contain a $0.8 : 0.2$ Ar:CO_2 mixture.

RPC use a 1.8 mm thick resistive plate made of a plastic laminate with an operating voltage of 4.9 kV. There are 1116 units in total, covering $|\eta| < 1.05$, giving a timing resolution of ~ 1 ns.

The TGCs in the intermediate muon wheel provide azimuthal information to assist in a building a complete picture of muon trajectories in coordination with the MDTs. They are filled with a $0.55 : 0.45$ CO_2:*n*-pentane mix, with 50 μm wires held at 2900 V. There are a total of 7 TGC layers, arranged into one triplet positioned before the intermediate MDT and two doublets spaced out after this. Triplets are used to assist in rejection of fake muons from various backgrounds. Over both sides, there are a total of 3588 TGC chambers.

4.3 ATLAS Contributions

Contributions to the ATLAS collaboration are a key part of maintaining an efficient scientific research endeavour. Not only does this allow the day-to-day operation, contributions to the collaboration make up the backbone of the technical advancement. Operations activities facilitate smooth data collection and allow a broad and complete physics program to be conducted. Technical tasks contributing to evolution of the detector has positive effects both during the data taking cycle and beyond. Often technical tasks contribute to advancing the Athena code base and the various tools available to analysis and CP groups. Along with this, a large fraction of this work is dedicated to producing the next generation of hardware and software for the detector and the HL-LHC era. Various contributions to the collaboration were conducted across both of these categories.

Authorship Qualification: BLS Trigger Validation Package. A technical task conducted as part of authorship qualification. This project involved taking the old B and Light States Trigger group's trigger validation package, and updating it to the new release of Athena for Run 3 Analyses. The updated validation package gives fine-grained evaluation of trigger efficiencies at all stages of the BLS trigger chain.

Muon+BPhys Trigger Signature Expertship. Operations activity as a Trigger Signature Expert for Muon+BPhys. Conducted in a remote capacity, this role involved evaluating the trigger information from the Muon and BPhysics trigger chains that were active during a data collection run. Signature experts are responsible for examining output from the multiple data streams and verifying their accuracy against reference information. This signature shifter position covers both Muon and BPhys due to the interlinked nature of the trigger stacks. BPhys triggers, being highly dependent on muon signatures, are built on top of the existing muon triggers, and perform additional advanced processing.

Run Control Desk Shifter. On-site operations activity. The Run Control (RC) role involved managing the ATLAS detector's transitions between certain states, starting and stopping data-taking runs and ensuring proper configuration and coordination between other operations desks that managed various subsystems. Some of these shifts were conducted during milestone weeks, where the various subsystem teams operate in unison to resolve issues identified in previous runs and integrate new technology.

5. Reconstruction, Event Selection & Statistical Techniques

This analysis performs the first dedicated measurement of the $g + g \rightarrow J/\psi + \gamma$, continuum (non-resonant) production of J/ψ in association with an isolated photon. Performing this measurement requires careful reconstruction and selection of events. Both real ATLAS data and MC simulation samples are used to this end. Multiple phases of cuts, reconstruction and trigger selection are applied to both types of sample. This chapter first details the choice and format of both the samples and reconstructed objects, then the format of selection cuts that are used for the entire analysis. Following this, a background on the statistical methods utilised to complete the analysis is presented. The analysis process seeks to extract the signal contribution within a *data* sample. Later in the process, this is performed using a binned template profile likelihood method.

5.1 Object Definitions, MC Samples and Datasets

5.1.1 Trigger

The trigger utilised for this analysis is `HLT_2mu4_bJpsimumu_noL2`, part of the ATLAS B-Physics trigger menu. There are several requirements that this trigger enforces for the selection of an event [79]. Firstly, a pair of opposite sign muons originating from the same production vertex are located. Secondly, the tracks from these muons must be successfully fit to a common vertex. After this the B-

physics trigger stack enforces an invariant mass window on the two muons [80]. The dimuon candidates produced from this fit are accepted if $2.5 \leq m_{\mu\mu} \leq 4.3$ GeV and $\chi^2/\text{d.o.f} \leq 20$. This χ^2 is the goodness-of-fit measure for the muon track-fitting procedure. Invariant mass windows are applied with respect to the desired hadron, in this case the window corresponds to the J/ψ peak at $m_{J/\psi} = 3.096$ GeV. This trigger was only active unprescaled during Run 2 of the LHC and only during 2015 [81]. A trigger prescale was introduced towards the end of 2015, though this trigger was unprescaled for most of the year [82]. The inner detector is used for identifying muon tracks for this trigger and as a consequence there is an implicit $|\eta| < 2.4$ cut. The trigger is also seeded at L1 (the initial event processing stage conducted in hardware) with the data from TGC muon chambers, with a similar coverage requirement. The common dimuon vertex fit is performed later at the High-Level Trigger (HLT) stage. Therefore, muon tracks from full reconstruction are used in this analysis, offering an efficiency improvement.

This analysis uses a segment of the ATLAS Run II dataset, specifically the data taken in 2015 at a centre of mass energy of $\sqrt{s} = 13$ TeV. 2015 offers the lowest prescales on the trigger and the lowest transverse momentum thresholds, $p_T(\mu) > 4$ GeV. Cases where q_T is much smaller than the $J/\psi+\gamma$ system invariant mass $\mathcal{Q}$ are most desirable. Conditions like this are most common in topologies where the J/ψ and γ are back-to-back in the lab frame. This configuration gives a minimum invariant mass of $\mathcal{Q} \simeq 2 \cdot (4 + 4) = 16$ GeV, with $q_T = 0$ GeV. To include the remaining Run II datasets from 2016 through to 2018 a trigger with a higher transverse momentum threshold on the muon must be used. This higher threshold is due to the fact that dimuon triggers with `2mu4` requirements are not available unprescaled for 2016-2018 [81]. Higher threshold triggers are available such as `2mu6`, as are higher multiplicity triggers such as `3mu4`. Both of these options include some compromise over `2mu4`, the detrimental effects of higher thresholds are apparent from the Section 3.2.4. Higher multiplicity triggers are available at L1, with topological constraints, and not as a part of the b-physics menu. Their inclusion in a future study will introduce significant challenges.

Beneficially, conditions in 2015 led to reduced pileup contributions. Trigger studies pertaining to this are conducted in [83]. Pileup corrections to efficiencies were small enough not to warrant further systematic studies in this case. This

analysis follows on from this conclusion, no systematic studies are conducted of pileup reweighting.

5.1.2 Reconstruction

After the trigger stage, the reconstruction stage applies further cuts and processing originating from the internal ATLAS BPHY14 derivation[1]. Muons obtained so far must meet 'tight' quality criteria [84] and dimuon candidates are restricted to a mass range of $2.6 \leq m_{\mu\mu} \leq 3.5$ GeV. Photons are required to satisfy 'tight' isolation criteria, a selection requirement using calorimeter and track-based isolation [85]. Isolation is applied in a ΔR cone around the γ direction and is defined as [85]:

$$E_T^{Iso}|_{\Delta R<0.4} < 0.022 \times E_T + 2.45 \text{ GeV} \tag{5.1}$$

$$p_T^{Iso}|_{\Delta R<0.2} < 0.05 \times E_T \text{ GeV} \tag{5.2}$$

The lower efficiency of the 'tight' working point is balanced out by using calorimetric photons. Tight isolation requirement greatly assists in enhancing the signal-to-background ratio.

The following ATLAS datasets were used in this analysis [86]:

- data15_13TeV.periodD.physics_Main.PhysCont.DAOD_BPHY14.grp15_v01_p3955
- data15_13TeV.periodE.physics_Main.PhysCont.DAOD_BPHY14.grp15_v01_p3955
- data15_13TeV.periodF.physics_Main.PhysCont.DAOD_BPHY14.grp15_v01_p3955
- data15_13TeV.periodG.physics_Main.PhysCont.DAOD_BPHY14.grp15_v01_p3955
- data15_13TeV.periodH.physics_Main.PhysCont.DAOD_BPHY14.grp15_v01_p3955
- data15_13TeV.periodJ.physics_Main.PhysCont.DAOD_BPHY14.grp15_v01_p3955

[1]ATLAS Derivations are a well-defined set of procedures applied to a dataset to prepare it for a specific analysis or group of analyses. Data unneccesary for an analysis is removed, and the dataset is augmented with relevant additional data. ATLAS maintains roughly 100 different derivation formats.

5.1.3 MC Simulation

Monte Carlo (MC) simulations of the signal and backgrounds of this analysis are created for multiple purposes. Two samples are used to assess the signal and background contributions to data and the third is used as a cross-check of background removal methods. Three background contributions are considered:

1. Uncorrelated J/ψ and γ pairs
2. Non-prompt J/ψ production from B-Hadrons, combined with stray photons
3. Contributions from the dimuon continuum

All samples are generated with PYTHIA 8.230 [30] event generator and using the CTEQ6L1 PDF set [87], in conjunction with `Photos` [88]. This is to improve the simulation of radiative corrections through EM showers for final state radiation (FSR).

The signal MC sample, known as *sign*, is a simulation of the $gg \rightarrow J/\psi + \gamma$ subprocess. Approximately 9.4m events are produced, with generator-level cuts of 3.5 GeV on muons, and 4 GeV on photons.

The first background MC sample, known as *bckg*, is responsible for modelling inclusive prompt J/ψ production from QCD in the hard-scatter primary interaction. This is the main background of uncorrelated J/ψ and γ contributions. Approximately 5m events are produced, with generator-level cuts of 2.5 GeV on muons.

The second background MC sample, known as *bbbg*, models inclusive non-prompt J/ψ production, through the creation of B-hadrons. Approximately 10m events are generated, with generator level cuts of 3.5 GeV on muons. *bbbg* is the cross-check sample. As a process for subtracting non-prompt contributions is applied to the data sample, this sample is used as an indicator of the efficacy of the method.

All generated samples are reconstructed using a full `GEANT4` model of the ATLAS detector [89] and have the same BPHY14 derivation as is used for data applied to them. The *sign* sample contains MC truth information for both muons and photons, whereas the *bckg* and *bbbg* samples do not contain any truth photons.

The following Monte Carlo derivation samples were used:

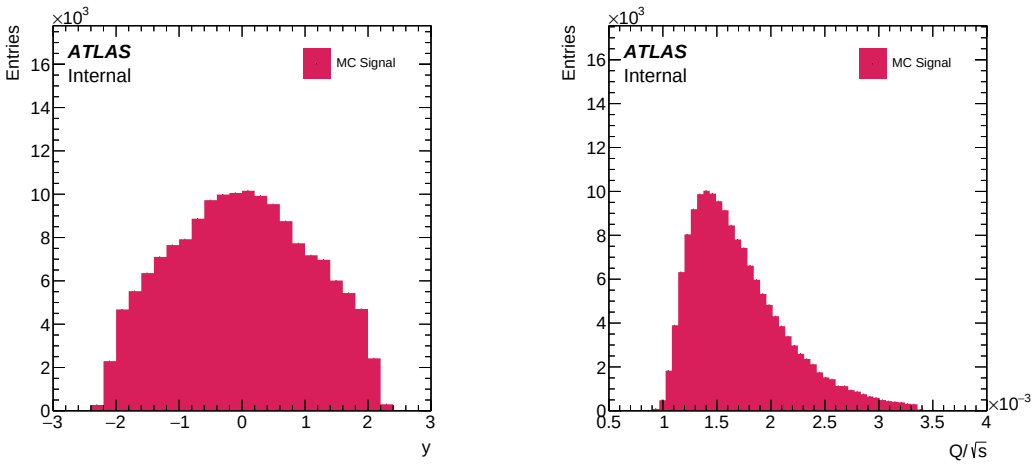
sign MC: mc16_13TeV.300011.Pythia8B_AU2_CTEQ6L1_pp_Jpsi_mu3p5mu3p5 _gamma4.deriv.DAOD_BPHY14.e6863_e5984_s3126_r9364_r9315_p3665

bckg MC: mc16_13TeV.300011.Pythia8B_AU2_CTEQ6L1_pp_Jpsi_mu3p5mu3p5 _gamma4.deriv.DAOD_BPHY14.e6863_e5984_s3126_r9364_r9315_p3665

bbbg MC: mc16_13TeV.300203.Pythia8BPhotospp_A14_CTEQ6L1_bb _Jpsimu3p5mu3p5.deriv.DAOD_BPHY14.e4889_a875_r9364_p3751

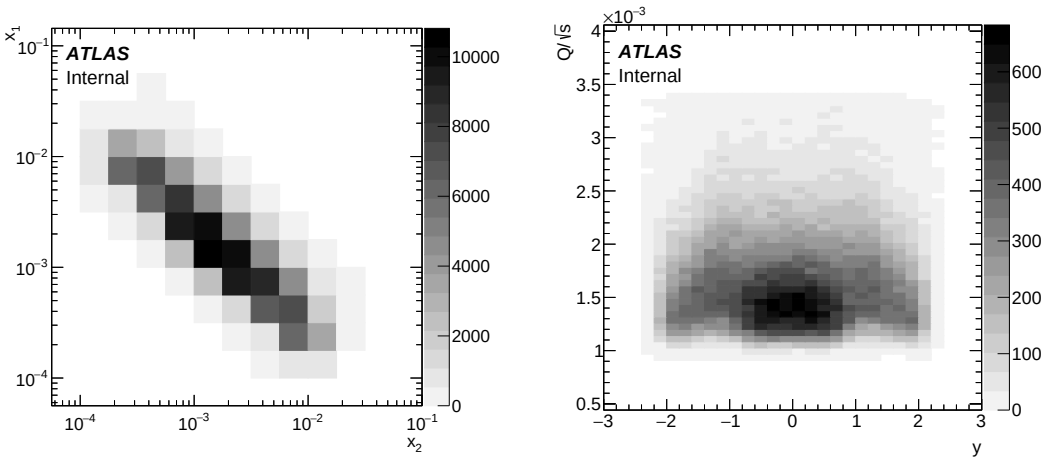
5.1.4 Accessible Kinematic region

The analysis seeks to extract the average gluon transverse momentum $\sqrt{\langle k_T^2 \rangle}$ to parameterise the Gluon TMD model. To have any hopes of evolving these PDFs, a full parameterisation with collinear momentum fraction x is required. Using the *sign* sample, limits on x can be estimated from $x_{1,2} = \frac{Q}{\sqrt{s}} e^{\pm y}$. The LHC provides a $\sqrt{s}$ of 13 TeV. The kinematic region associated with this is shown in Figures 5.1 and 5.2.



(a) (b)

Figure 5.1: Distributions of the kinematic variables in the *sign* sample. 5.1a shows the $J/\psi + \gamma$ system rapidity, and 5.1b shows the system invariant mass fraction $\frac{Q}{\sqrt{s}}$.



(a) (b)

Figure 5.2: Maps of 5.2a the x_1 and x_2 values accessible in the *sign* sample and 5.2b the $J/\psi + \gamma$ system rapidity y vs the system invariant mass fraction $\frac{Q}{\sqrt{s}}$.

5.2 Event Selections and Variables

5.2.1 Reconstructed Variable Distributions

Several variables are used throughout the analysis, Table 5.1 summarises those in use from the outset. Where appropriate, additional variables will be introduced later in the text. Most variables are defined within the lab frame, though some are in the Collins-Soper framewhich is detailed in Appendix A:

- $p_{\mathrm{T}}(\mu\mu)$: The transverse momentum of the dimuon system, measured in the lab frame. Also shown as $p_T(J/\psi)$ after procedure in Section 5.3 is applied.
- $p_{\mathrm{T}}(\gamma)$: The transverse momentum of the produced photon, measured in the lab frame.
- $\tau_{\mu\mu}$: The pseudo-proper lifetime of the dimuon system, measured with respect to the primary vertex that has the smallest 3D unsigned impact parameter to the $J\psi$ direction. Defined as

$$\tau_{\mu\mu} = \frac{L_{xy} m_{J\psi}}{c p_{\mathrm{T}}} \tag{5.3}$$

where $m_{J/\psi}$ is the known J/ψ mass [90], p_{T} is the transverse momentum of the J/ψ candidate dimuon system and L_{xy} is the 'decay length', the

distance between the primary pp interaction vertex and the $J/\psi \to \mu^+\mu^-$ decay vertex in the transverse plane. Measured in the lab frame. This lifetime is 'pseudo'-proper as the momentum of the parent system cannot be measued with certainty.

- $\Delta\phi(\mu\mu,\gamma)$: The difference in ϕ between the produced dimuon and photon, measured in the lab frame, and brought to the standard interval $-\pi < \Delta\phi \leq \pi$.

- $\Delta y(\mu\mu\ ,\gamma)$: The difference in rapidity of the produced dimuons and the pseudorapidity η of the produced photon, measured in the lab frame.

- λ: Defined as $\left(\frac{p_{\mu\mu}+p_\gamma}{m_{J/\psi}}\right)^2 \equiv \left(\frac{Q}{m_{J/\psi}}\right)^2$, the invariant mass of the $J/\psi+\gamma$ system, Q, divided by the known value of the J/ψ mass [90], all squared. Division by the $m_{J/\psi}$ provides a natural scale, and renders λ dimensionless. Measured in the lab frame.

- q_T^2: Defined as the vector sum squared $(\boldsymbol{p}_\mathrm{T}(\mu\mu)+\boldsymbol{p}_\mathrm{T}(\gamma))^2$, this is the square of the transverse momentum of the $J/\psi+\gamma$ system, measured in the lab frame.

- `HLT_2mu4_bJpsimumu_noL2`: the flag which is set if this trigger has fired for this event.

- $\cos\theta_{CS}$: cosine of the polar angle of J/ψ candidate production in the subprocess, measured in the Collins-Soper frame.

- ϕ_{CS}: azimuthal angle of J/ψ production in the subprocess, measured in the Collins-Soper frame.

- q_T^A: $J/\psi+\gamma$ system transverse momentum oriented along the A-axis. This is defined as the difference between the J/ψ and γ transverse momenta, see Section 3.2.4. Axes defined uniquely in each event.

- q_T^B: $J/\psi+\gamma$ system transverse momentum oriented along the B-axis. This is dDefined as $\sin(\epsilon)\sqrt{p_T(\mu\mu)p_T(\gamma)}$, see Section 3.2.4. Axes are defined uniquely in each event.

All variables are plotted for all four samples before selection cuts in Figures 5.3-5.12 These distributions are after the reconstruction process up to the BPHY14

derivation creation. An additional requirement that at least one photon exists in the event is included to allow all variables to be calculated.

Variable	Description
q_T^A	Component of q_T along A axis. Refer to Section 3.2.4
q_T^B	Component of q_T along B axis. Refer to Section 3.2.4
$\Delta\phi$	$\phi_{\mu^+\mu^-} - \phi_\gamma$ if $\Delta\phi > \pi \rightarrow \Delta\phi - 2\pi$ if $\Delta\phi < -\pi \rightarrow \Delta\phi + 2\pi$
Δy	Measured in lab frame $y_{\mu^+\mu^-} - \eta_\gamma$
$m_{\mu\mu}$	-
$p_{\mathrm{T}}(\mu\mu)$	-
$p_{\mathrm{T}}(\gamma)$	-
λ	$\frac{(p_{\mu\mu}+p_\gamma)^2}{m^2_{J/\psi}}$
$\tau_{\mu\mu}$	Measured in lab frame
$\cos\theta_{CS}$	Measured in the Collins-Soper frame
ϕ_{CS}	Measured in the Collins-Soper frame
`HLT_2mu4_bJpsimumu_noL2`	Trigger flag

Table 5.1: Table summarizing the main variables in the analysis.

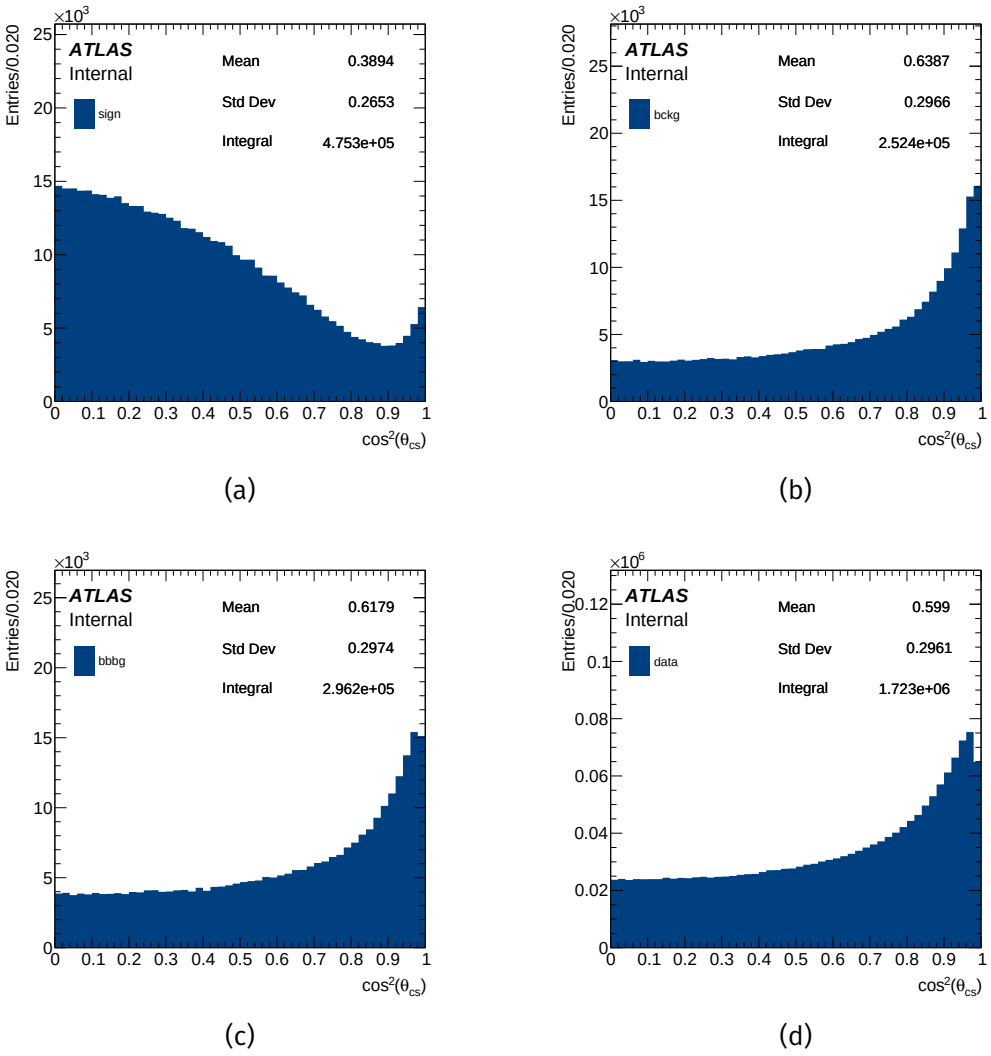


Figure 5.3: Distribution of reconstructed $\cos^2\theta_{CS}$, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

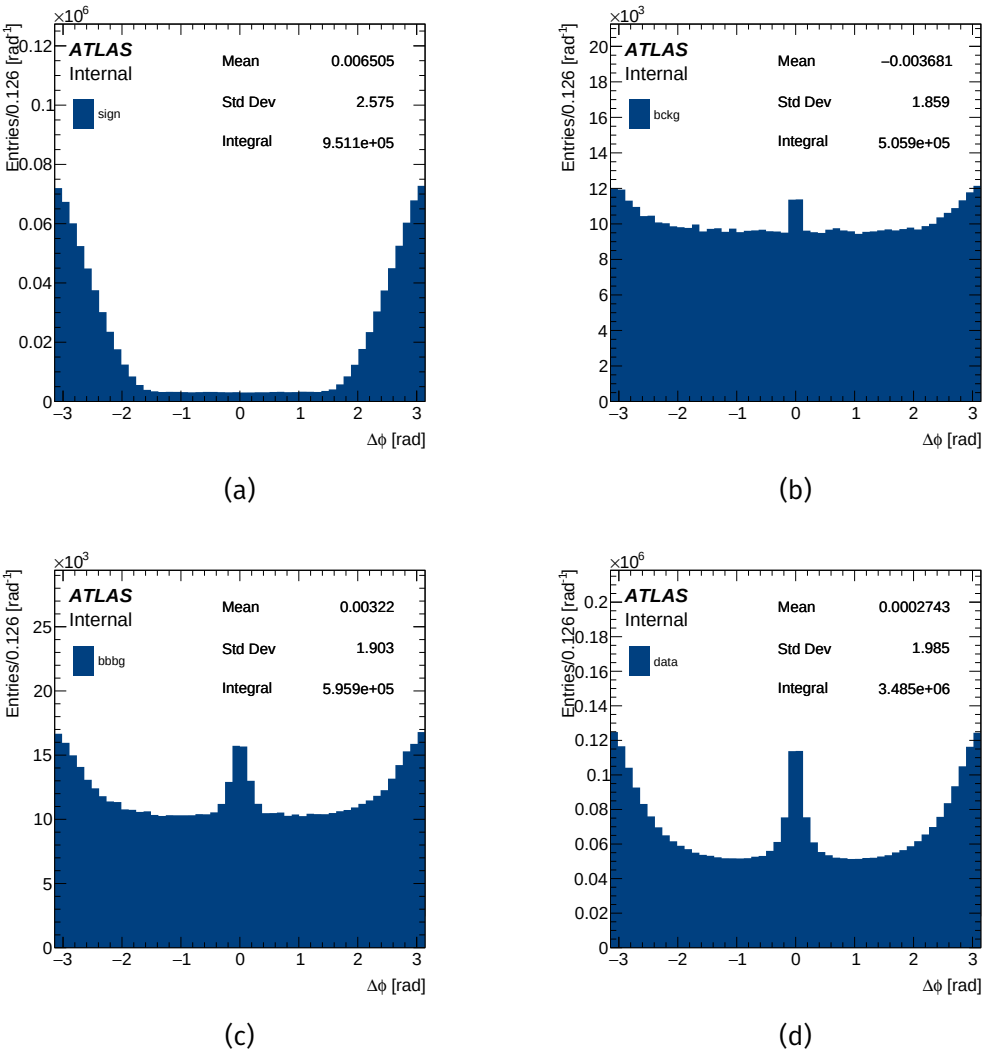


Figure 5.4: Distribution of reconstructed $\Delta\phi$, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

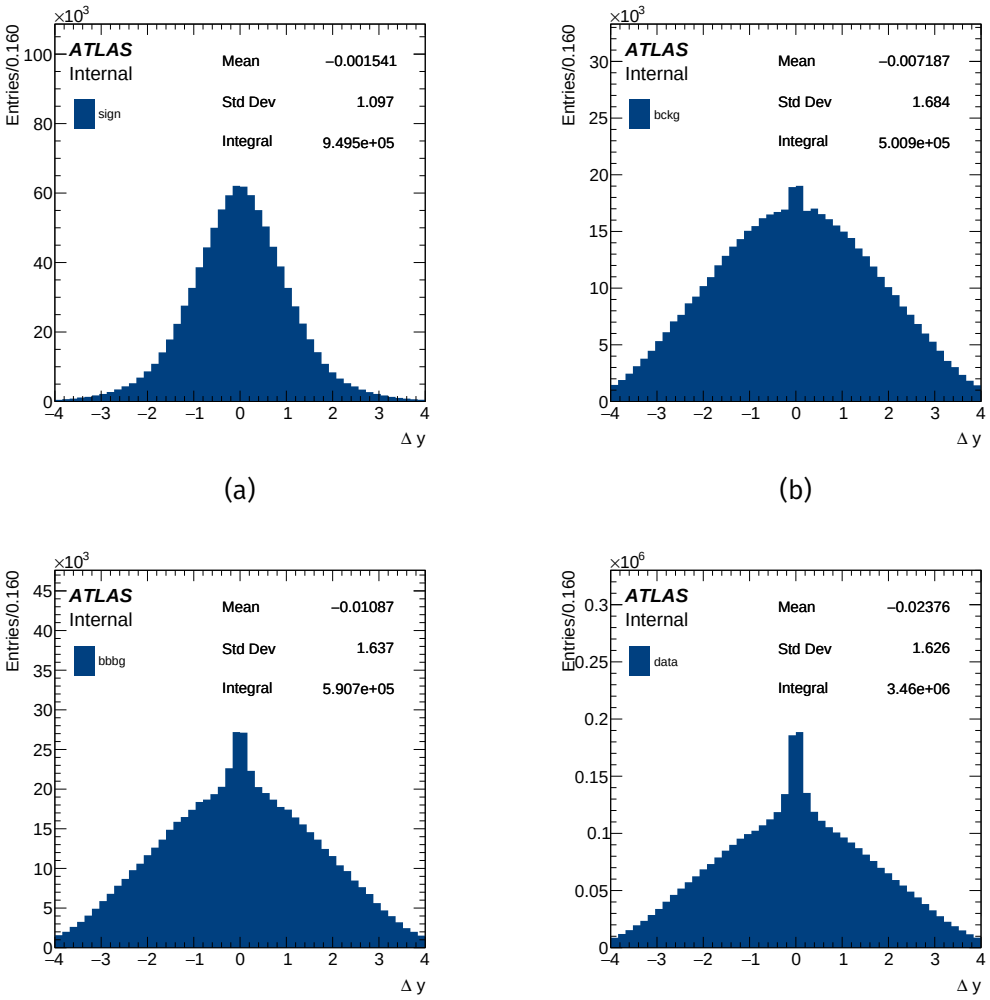


Figure 5.5: Distribution of reconstructed Δy, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

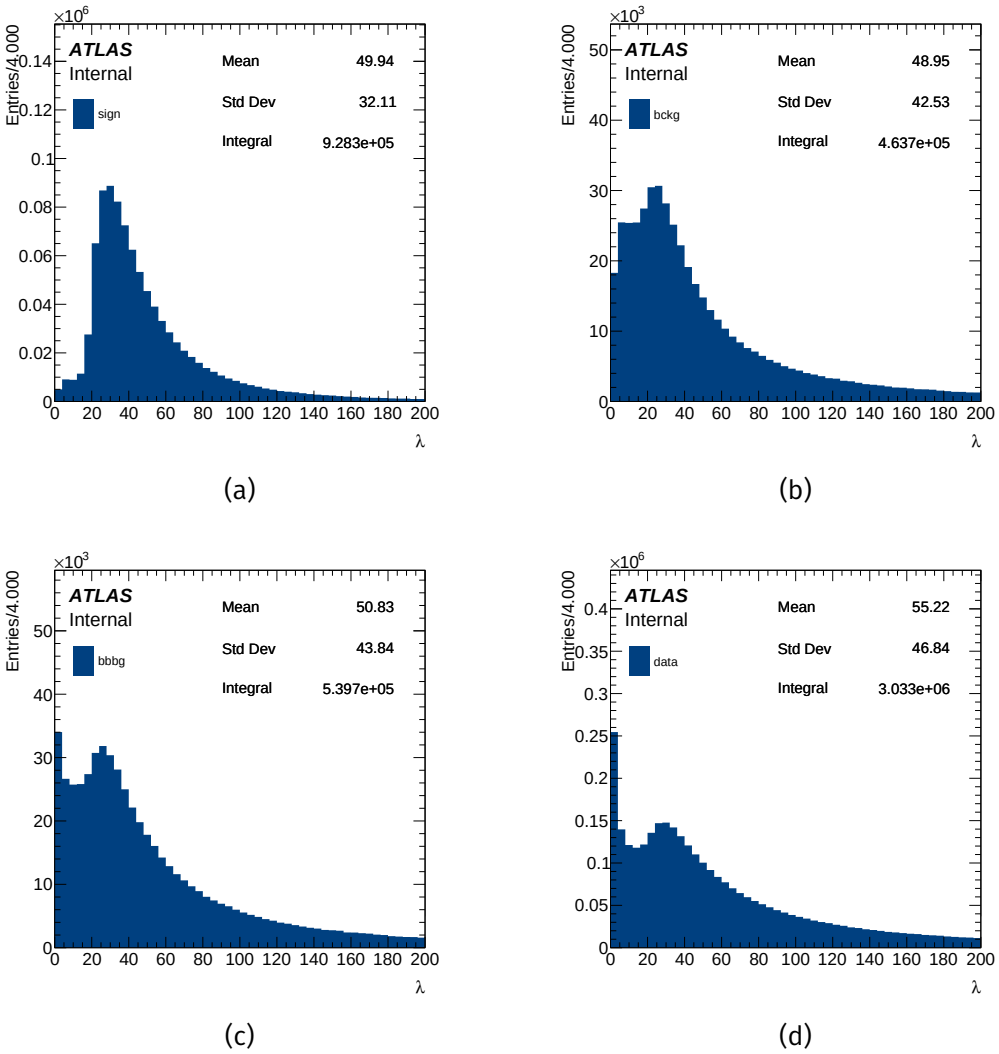


Figure 5.6: Distribution of reconstructed λ, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

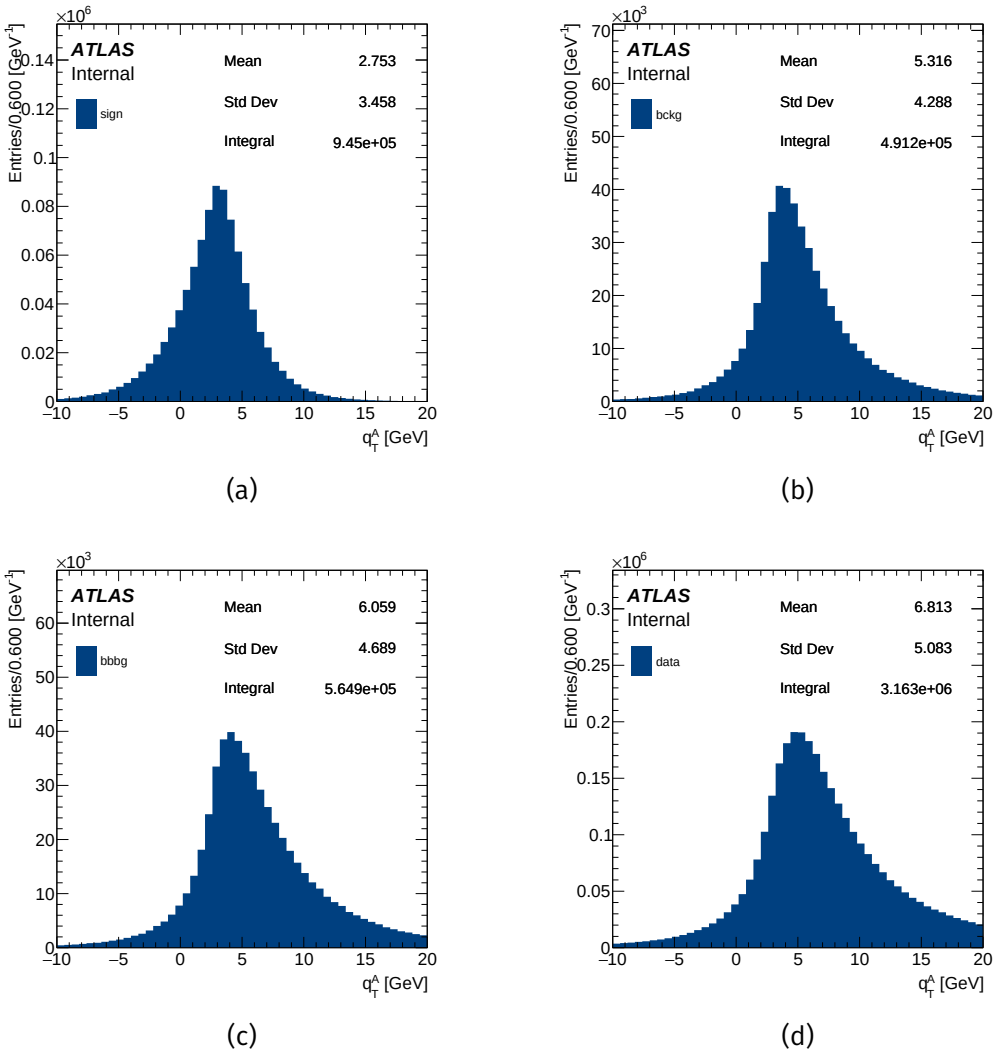


Figure 5.7: Distribution of reconstructed q_T^A, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

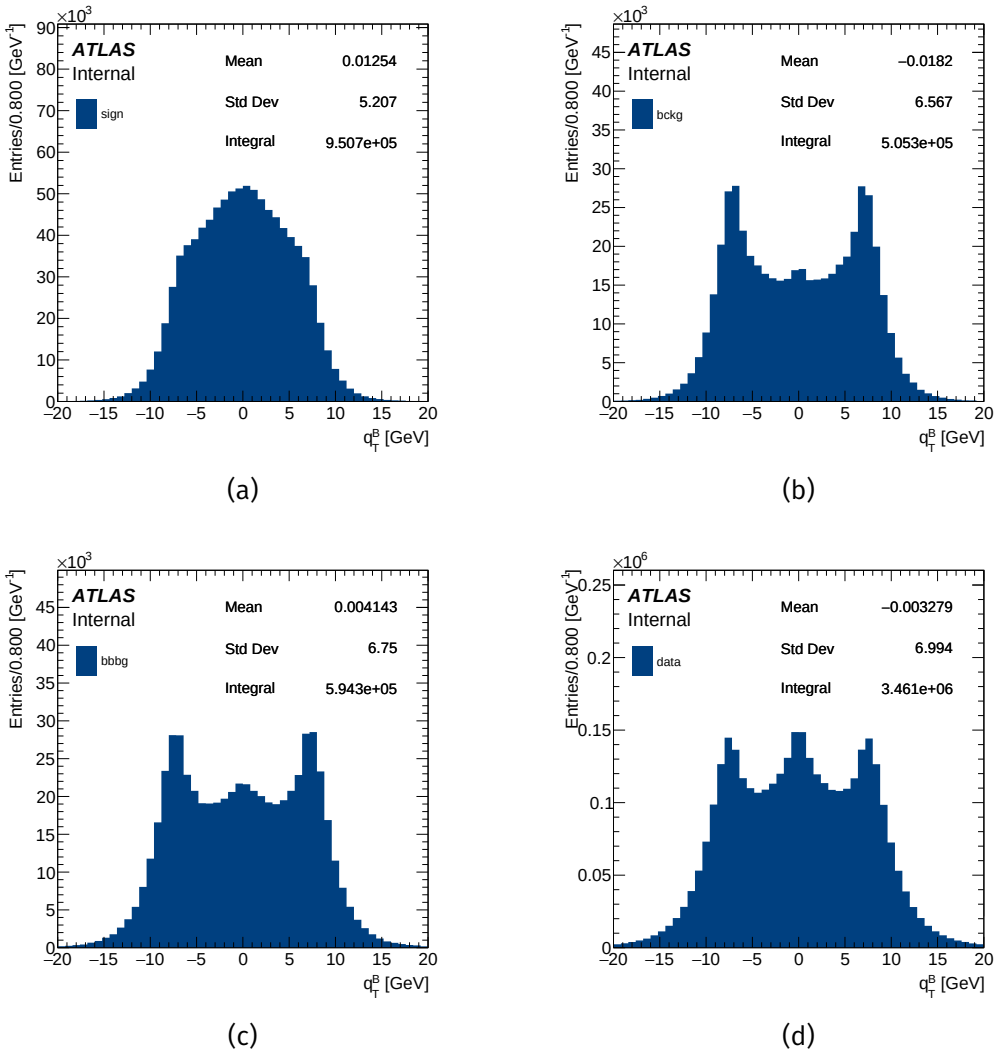


Figure 5.8: Distribution of reconstructed q_T^B, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

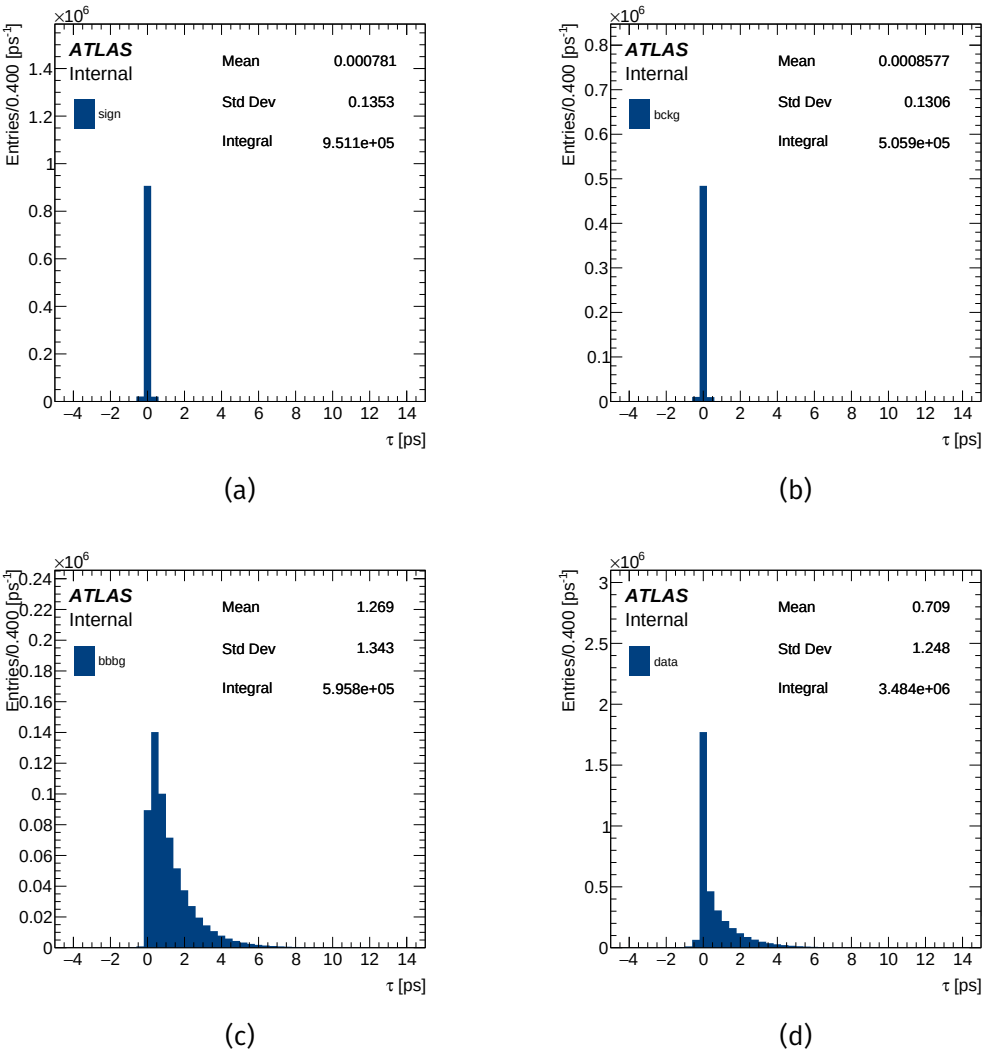


Figure 5.9: Distribution of reconstructed $\tau_{\mu\mu}$, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

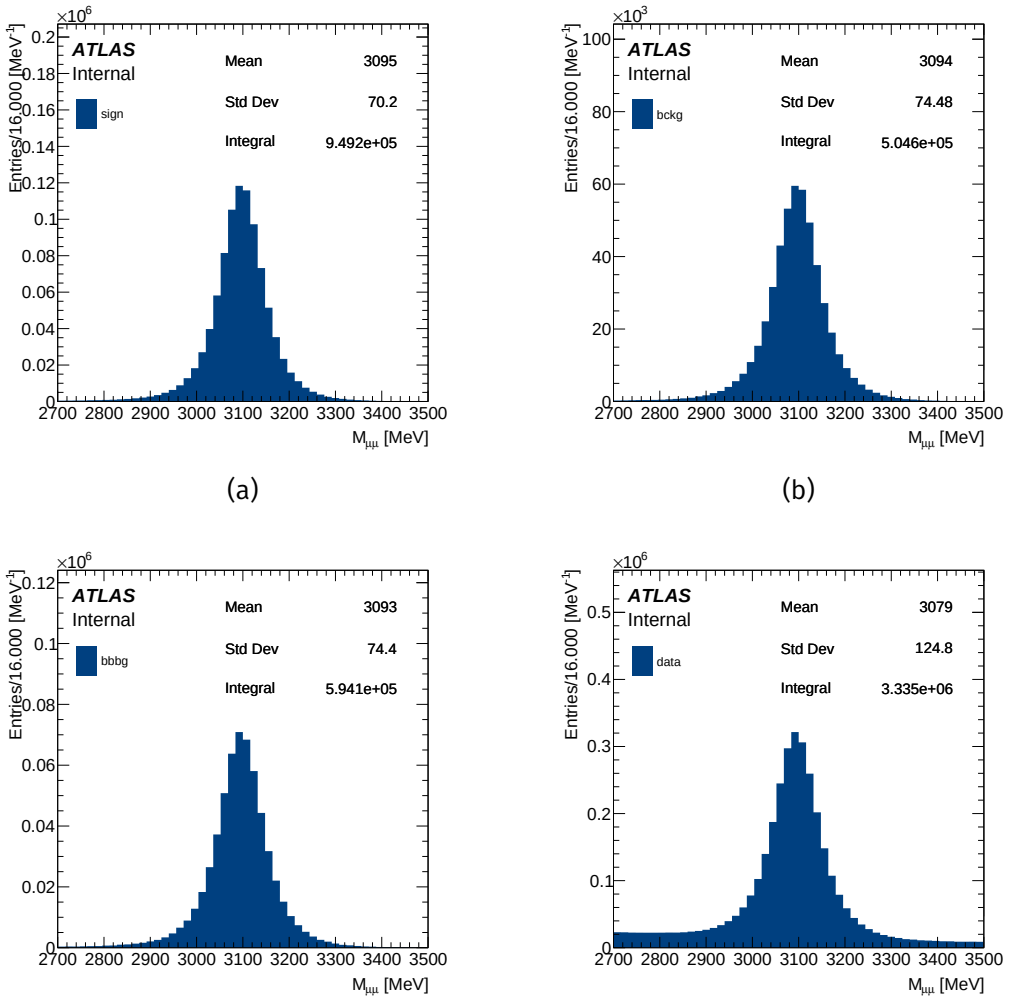


Figure 5.10: Distribution of reconstructed $m_{\mu\mu}$, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

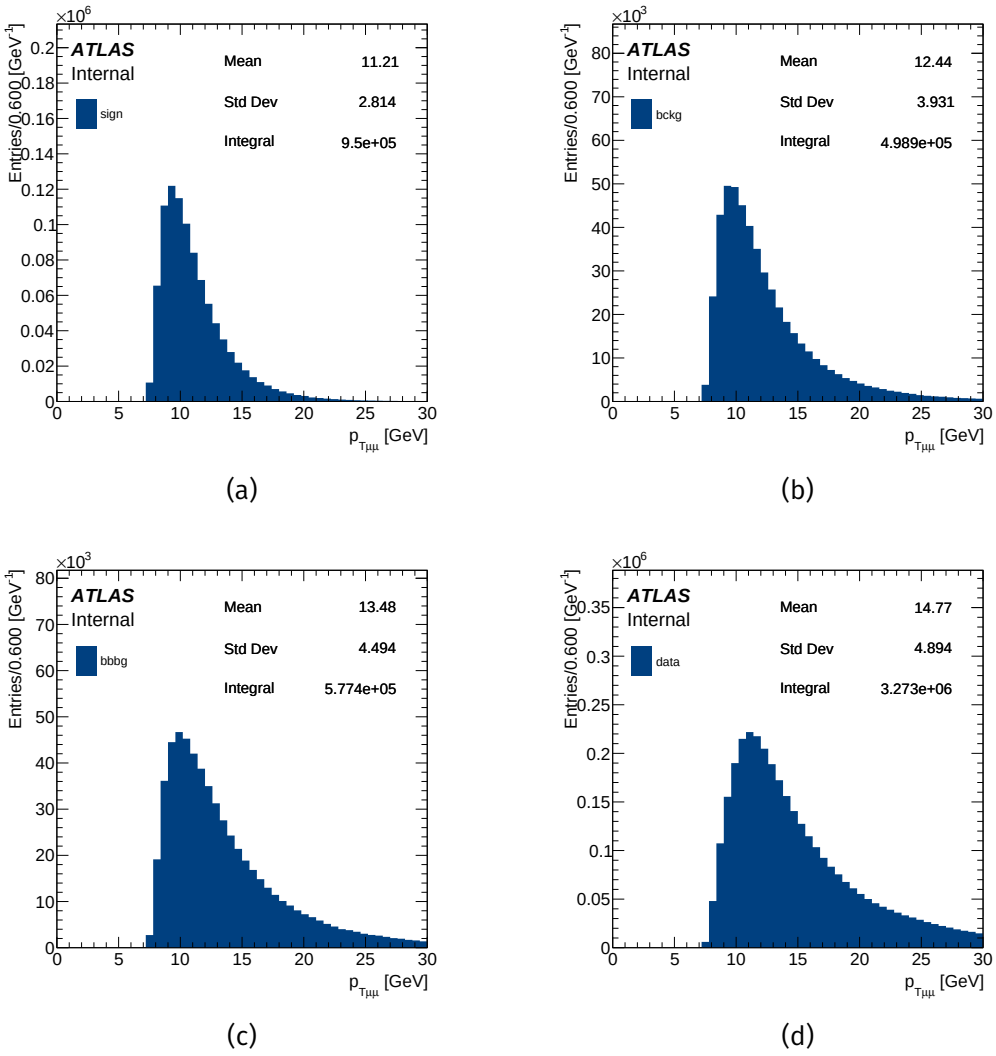


Figure 5.11: Distribution of reconstructed $p_T(\mu\mu)$, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

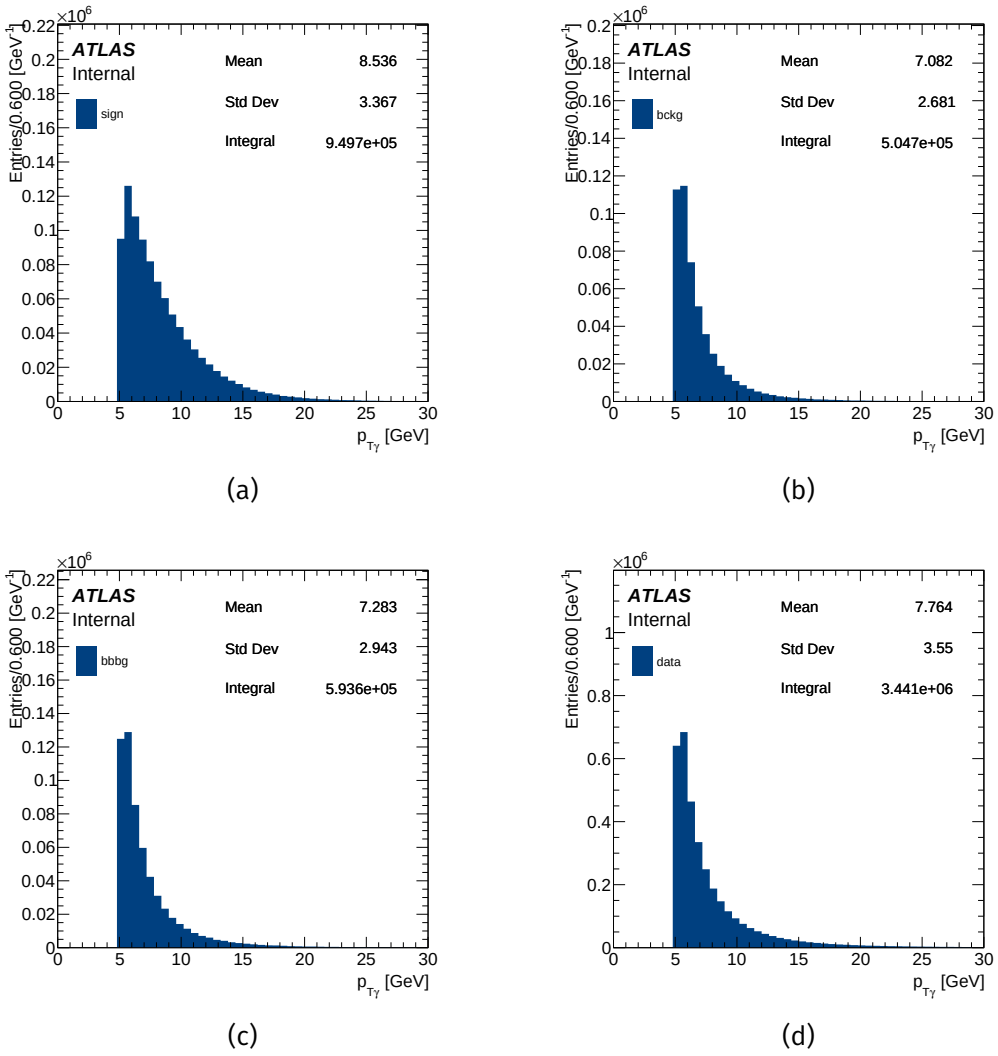


Figure 5.12: Distribution of reconstructed $p_T(\gamma)$, before selection cuts, with $N_{photons} > 0$ for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

5.2.2 Selection Cuts

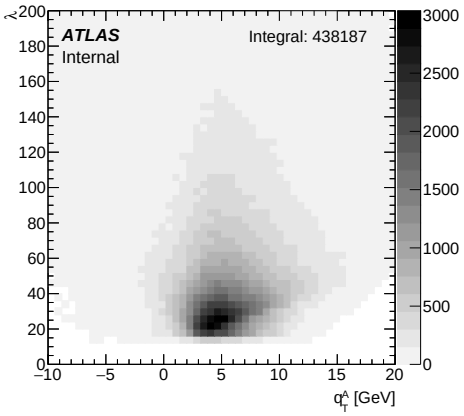
Following on from [26], TMD factorization is expected to hold in regions satisfying the relation $q_T^2 \ll \mathcal{Q}^2$. The ATLAS acceptance cuts present a problem in light of this requirement, with an implicit minimum transverse momentum threshold of $p_T(J/\psi) > 8$ GeV. . The $J/\psi + \gamma$ system is unbalanced when combined with a minimum acceptance photon of $p_T(\gamma) \geq 5$ GeV. The transverse momentum of

the J/ψ is computed from the vector sum of the two μ transverse momenta. In this configuration, the system will probably have a large transverse momentum. The minimum system transverse momentum of ~ 3 GeV is produced when the J/ψ and γ are back-to-back. This corresponds to $Q \simeq 13$ GeV. The alternative topology where the J/ψ and the γ are at a small angle also poses a problem. Here, the invariant mass can be reduced, but q_T^2 exceeds this. In different terms, it is possible to have access to small q_T values with unbalanced cuts. However, this is at much larger Q, where cross-sections are suppressed, and q_T^2 is biased as a result.

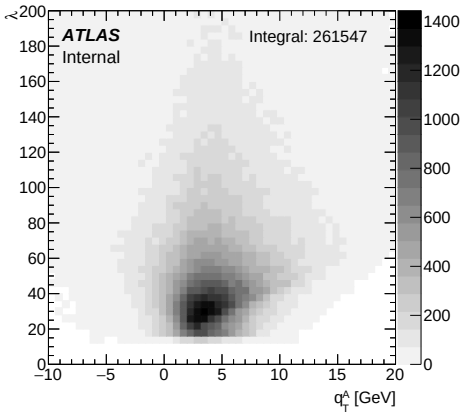
To find an arrangement with $q_T^2 \ll Q$ the cuts on photons must be increased to $p_T(\gamma) \geq 9$ GeV to the threshold of the implicit J/ψ cut. Hence, configurations of $q_T \approx 0$ at the smallest possible invariant mass can occur. However, this severely affects the the effective sample size and distorts the shape of measured distributions.

Problems with both the imbalanced and balanced cut scenarios are resolved with the introduction of q_T^A. However, this introduction forms part of a larger correction to the data, in comparison to using a detector with lower acceptance cuts. Using this axis of measurement is beneficial for back-to-back events. This event topology is more strongly associated with the $gg \to J/\psi + \gamma$ subprocess. The transverse momentum of the $J/\psi + \gamma$ system becomes the difference in momentum between the two particles. This is equivalent to the definition of $q_T^A = p_T(J/\psi) - p_T(\gamma)$. The unbalanced acceptance cuts alter measured distributions, increasing the mean value of q_T^A. However, this shift effect is not relevant to the parameterisation of the TMD as the width of the distribution of q_T^A stays unchanged. Therefore, balanced acceptance cuts can be dropped to achieve higher selection efficiency.

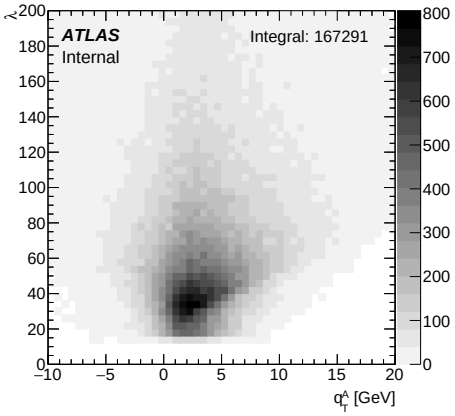
To further elucidate this point, maps of q_T^A and λ are given in Figure 5.13. Three changes are visible with the progressive acceptance cuts. First, the leftward shift in q_T^A, stemming from the balance between the cuts. Second, the progressive decrease in statistics, only 20% of events survive the increase to $p_T(\gamma) > 9$ GeV. Third, average λ increasing from $\simeq 25 \to 40$. This corresponds to an increase in Q from ~ 15 GeV to ~ 19 GeV.



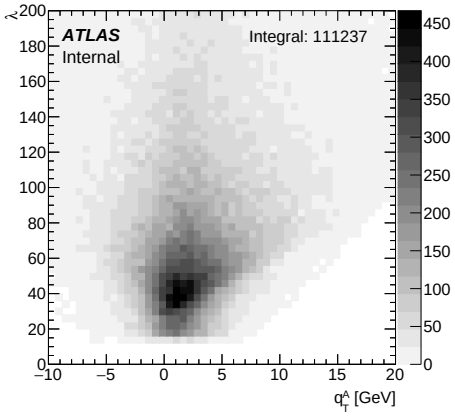
(a) $p_T(\mu) > 4$ GeV, $p_T(\gamma) > 5$ Gev



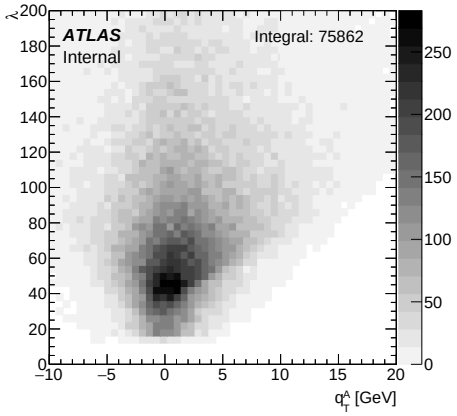
(b) $p_T(\mu) > 4$ GeV, $p_T(\gamma) > 6$ Gev



(c) $p_T(\mu) > 4$ GeV, $p_T(\gamma) > 7$ Gev



(d) $p_T(\mu) > 4$ GeV, $p_T(\gamma) > 8$ Gev



(e) $p_T(\mu) > 4$ GeV, $p_T(\gamma) > 9$ Gev

Figure 5.13: Distributions of λ vs. q_T^A at varying p_T cuts on the photon, increasing from 5 GeV (a) to 9 GeV (e) in steps of 1 GeV on the *data* sample. Along with the q_T^A distribution shifting leftwards with increasing $P_T(\gamma)$, the distribution lifts off the lower $\lambda > 25$ bound, at the cost of statistics.

Often there are events with more than one photon candidate, the normalised photon count per event for all four samples is presented in Figure 5.14. These candidates are ordered in terms of decreasing transverse momentum, and the first candidate (highest $p_T(\gamma)$) is selected. Muons were organised similarly, though the number of dimuon candidates per event is close to 1. The leading photon and the leading dimuon are used. Table 5.2 outlines the selection cuts utilised in this analysis, and the event counts at each stage.

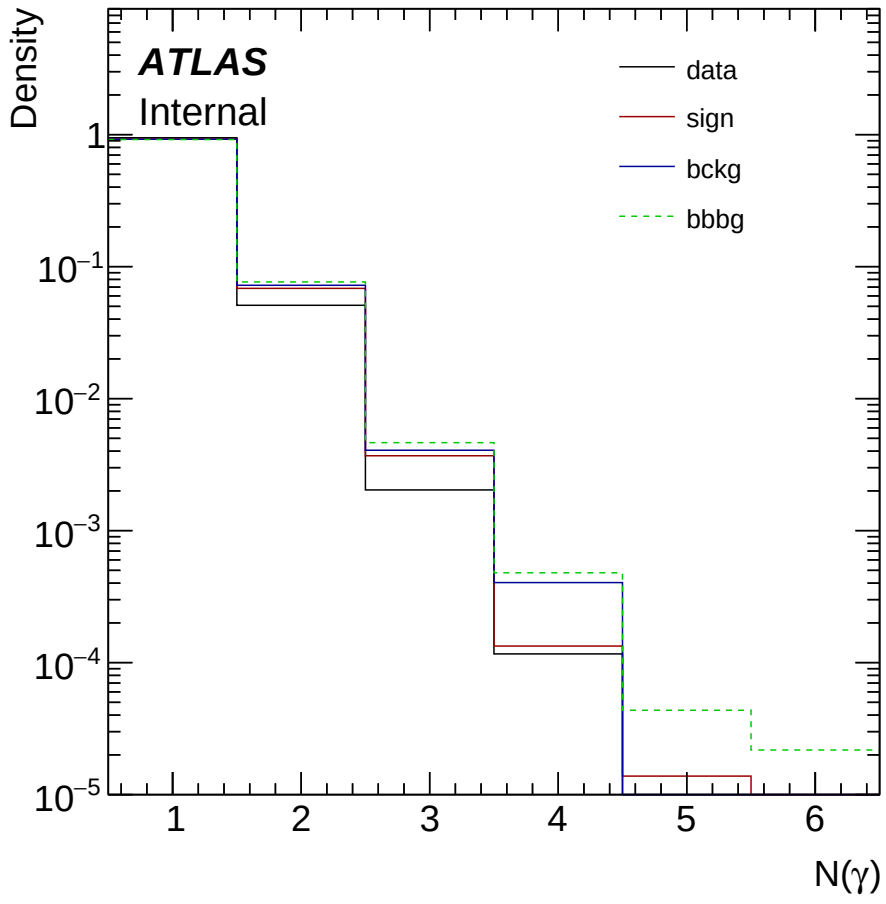


Figure 5.14: Normalised (by sample) count of photons present in each event.

Cut	Definition	sign	bckg	bbbg	data	Additional information
No cut	–	8399000 100.00%	10000000 100.00%	10000000 100.00%	13265752 100.00%	BPHY14 Reconstruction in place
1	$N_{\text{photon}} > 0$	11.32%	5.06%	5.95%	26.27%	Some reconstructed events do not contain a photon, these need to be removed from the analysis.
2	Trigger	5.25%	2.54%	3.16%	20.03%	Pass `HLT_2mu4_bJpsimumu_noL2` trigger
3	$p_{\text{T}}(\mu^+) > 4$ GeV, $p_{\text{T}}(\mu^-) > 4$ GeV	5.21%	2.45%	2.96%	18.61%	μ acceptance cut.
4	$15 \leq \lambda \leq 200$	4.92%	1.83%	2.17%	13.35%	$Q \lesssim 44$ GeV
5	$0 \leq \cos^2\theta_{CS} \leq 1$	4.92%	1.83%	2.17%	13.35%	Trivial
6	$0 \leq q_T^2 \leq 400$ GeV^2	4.87%	1.60%	1.84%	11.57%	$q_T \leq 20$ GeV
7	$0 \leq \lvert\phi_{CS}\rvert < \pi$ rad.	4.87%	1.60%	1.84%	11.57%	Trivial
8	2.7 GeV $\leq m_{\mu\mu} \leq 3.5$ GeV	4.86%	1.60%	1.84%	11.11%	Dimuon mass range
9	$-5\ ps \leq \tau_{\mu\mu} \leq 15$ ps	4.86%	1.60%	1.84%	11.11%	Dimuon lifetime range
10	$-10 < q_T^A < 20$ GeV	4.86%	1.60%	1.84%	11.11%	Trivial
11	$\lvert q_T^B\rvert < 15$ GeV	4.86%	1.60%	1.84%	11.11%	Trivial
12	$\lvert\eta(\gamma)\rvert < 2.5$ and $\lvert y(\mu)\rvert < 2.4$	4.80%	1.57%	1.81%	10.96%	Acceptance
13	$p_{\text{T}}(\gamma) > 5$ GeV	4.80%	1.57%	1.81%	10.96%	Photon Acceptance
14	Photon quality < 1	2.72%	0.40%	0.46%	3.03%	Tight photons only
15	$\Delta R < 0.12$	2.59%	–	–	–	Section 5.2.2.1
Overall	-	217240	39610	45919	438186	-

Table 5.2: Basic selection cuts for events deemed useful for this analysis and the event count present at every stage.

5.2.2.1 MC Signal Photon Matching

In an event, more than one photon or dimuon candidate can be present. Events with 0 γ's and $\mu\mu$'s inside the mass cut $2.7 \leq m_{\mu\mu} \leq 3.5$ are removed from all samples in the first stages of the analysis. In the remaining events, candidates are arranged in descending order by their p_T, the leading candidate in the 0^{th} position. When examining the *sign* sample the highest p_T reconstructed MC photon is not guaranteed to correspond to the truth MC photon from the subprocess. When considering the leading candidates, events with uncorrelated J/ψ's and γ's enter into the *sign* sample. This effect should be mitigated to provide a sample representing the pure signal. The distribution of ΔR between the truth and reconstruction after selection cuts is shown in Figure 5.15

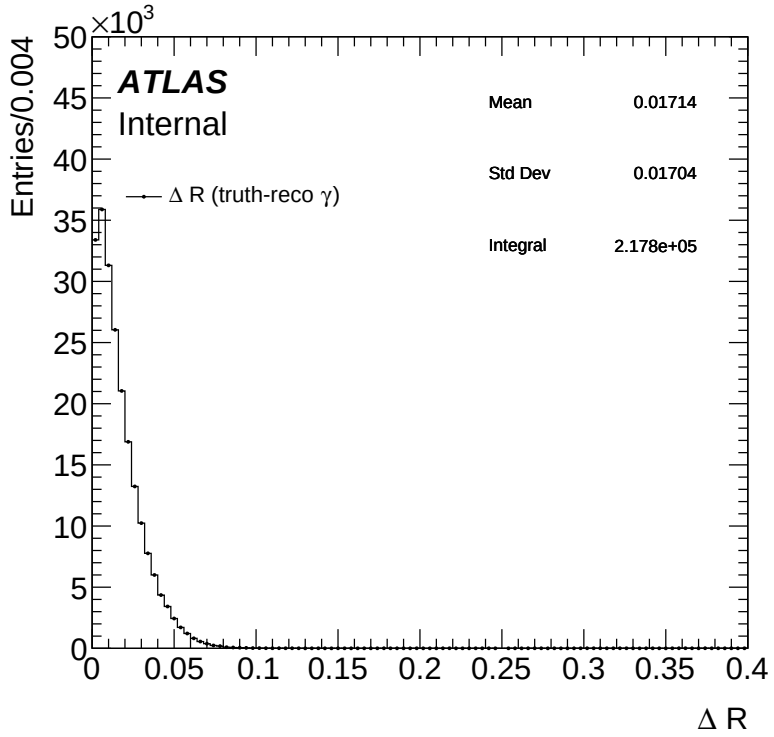


Figure 5.15: ΔR between truth and 0^{th} reconstructed photons. Based on this histogram a cut of $\Delta R < 0.12$ is sufficient to say a truth and reconstructed photon are matched. This cut is named cut 15.

Following from this, introduce a cut at $\Delta R < 0.12$ (only on *sign*) that removes any cases of unordered reconstruction. Approximately 95% of reconstructed events pass this cut, and the truth and reconstructed photons are considered matched. The remaining 5% of events are considered background and discarded.

5.2.2.2 Additional q_T^B cuts

Improved signal-to-background ratios can improve the fit quality. Hence, the phase space is explored for additional selection cuts. The 2D distribution of q_T^A and q_T^B in both the *sign* and *bckg* are shown in Figure 5.16.The difference in the maps is a consequence of the dependence on the azimuthal difference between J/ψ and γ, $\Delta\phi$.

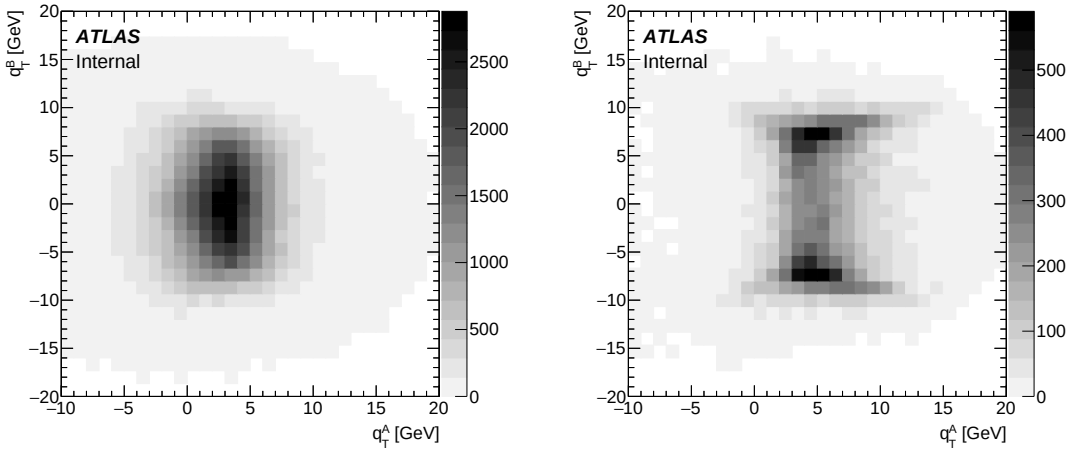


Figure 5.16: Map of q_T^A vs q_T^B for *sign* and *bckg*

In Figures 5.17 and 5.18, greater signal-to-background ratio regions can be visually identified. When a cut restricting events to $|q_T^B| \leq 6$ GeV is introduced the analysis region is limited to an area with high signal-to-background ratio. Further support for this change can be evidenced by the evolution of the shape of q_T^A when viewed in slices of $|q_T^B|$ in Figure 5.19. The $|q_T^B| < 6$ GeV cut is not applied until later in the analysis. This is to provide greater sample size in both *sign* and *bckg* for building a multivariate classifier.

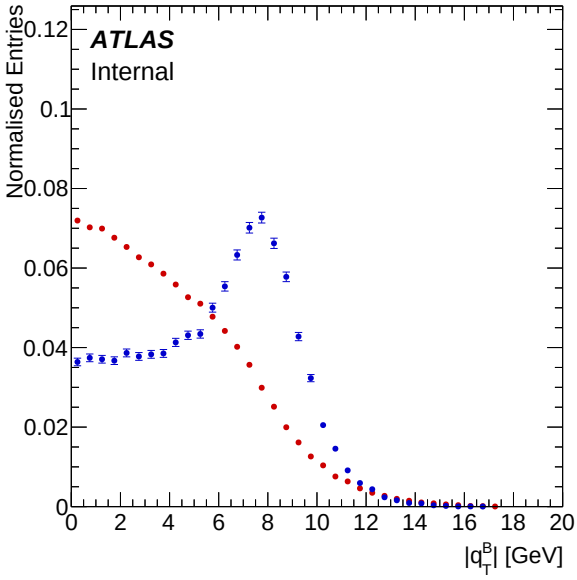


Figure 5.17: Distribution of $|q_T^B|$ in the *sign* sample (red) and *bckg* sample (blue) after selection cuts.

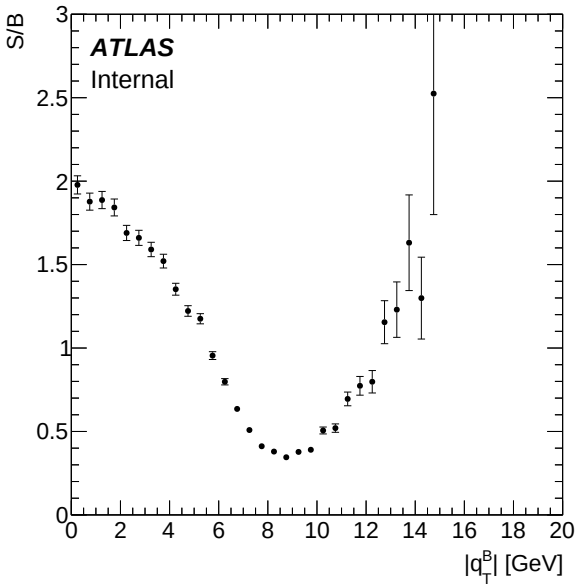


Figure 5.18: The signal-to-background ratio with respect to q_T^B.

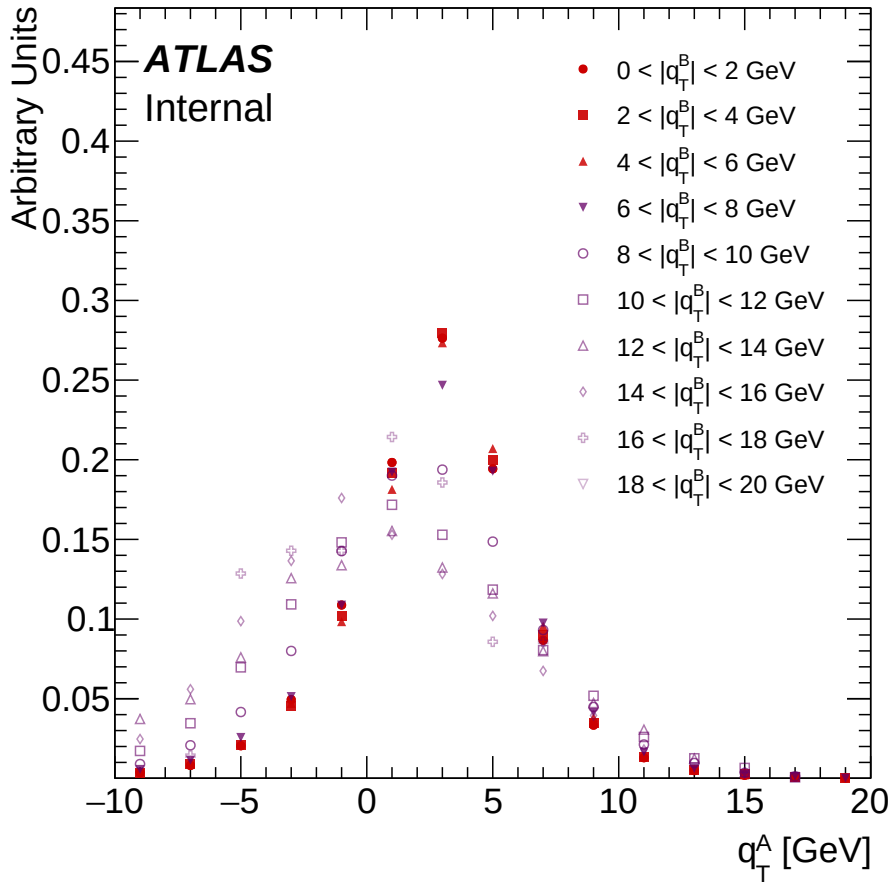


Figure 5.19: Normalised distributions of q_T^A in progressively increasing 2 GeV windows of q_T^B. Areas of smaller q_T^B, where the J/ψ and γ are more back-to-back, preserve the shape of q_T^A. The shape of q_T^A breaks down at regions of higher q_T^B. Distributions where $q_T^B < 6$ Gev are shown in red, distributions where $q_T^B > 6$ GeV are shown in purple.

5.2.3 Analysis Variable Distributions

The rest of the analysis continues with all the cuts in Table 5.2 present. Distributions of the variables used in the analysis and the Boosted Decision Tree (BDT) process for all samples are shown in Figures 5.20–5.30.

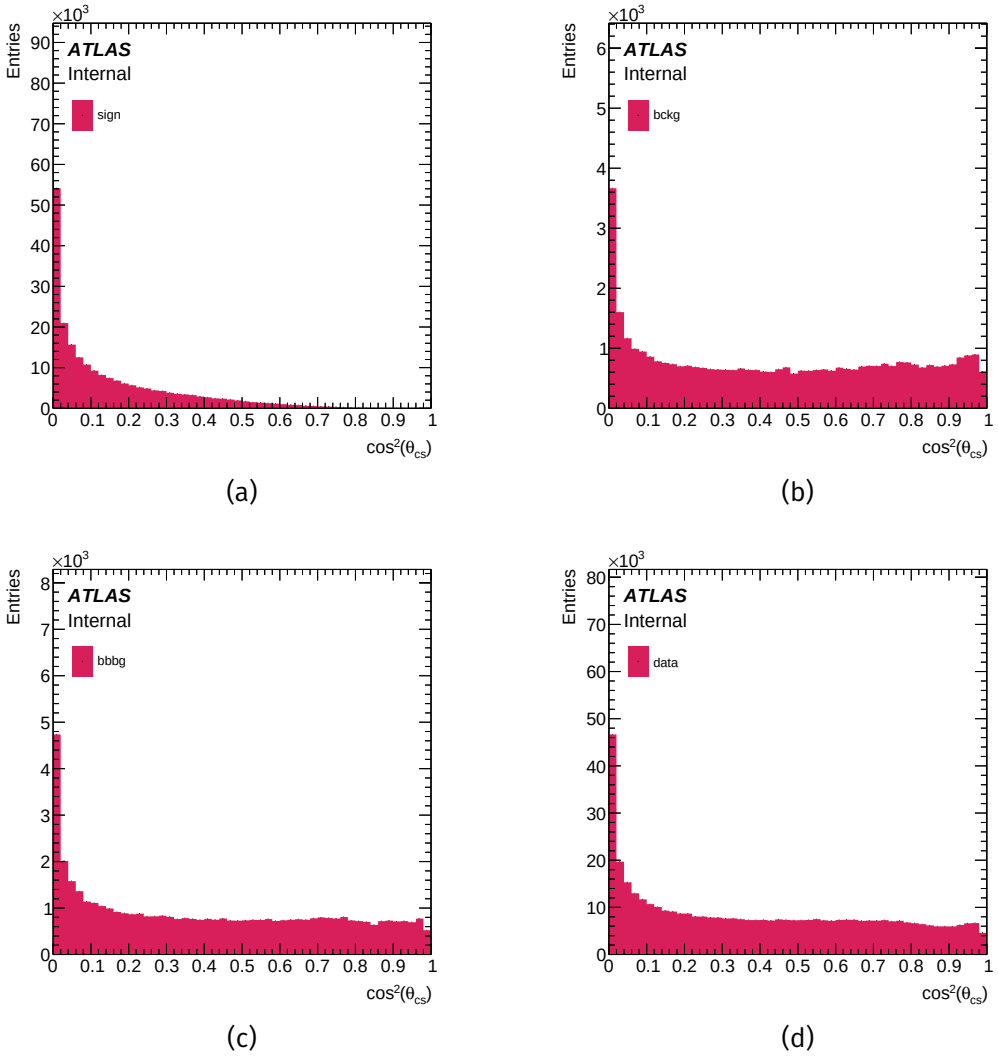


Figure 5.20: Distributions of the variable $\cos^2\theta_{CS}$ for (a) *sign*, (b) *bckg*, (c) *bbbg*, and (d) *data* samples after basic selection cuts.

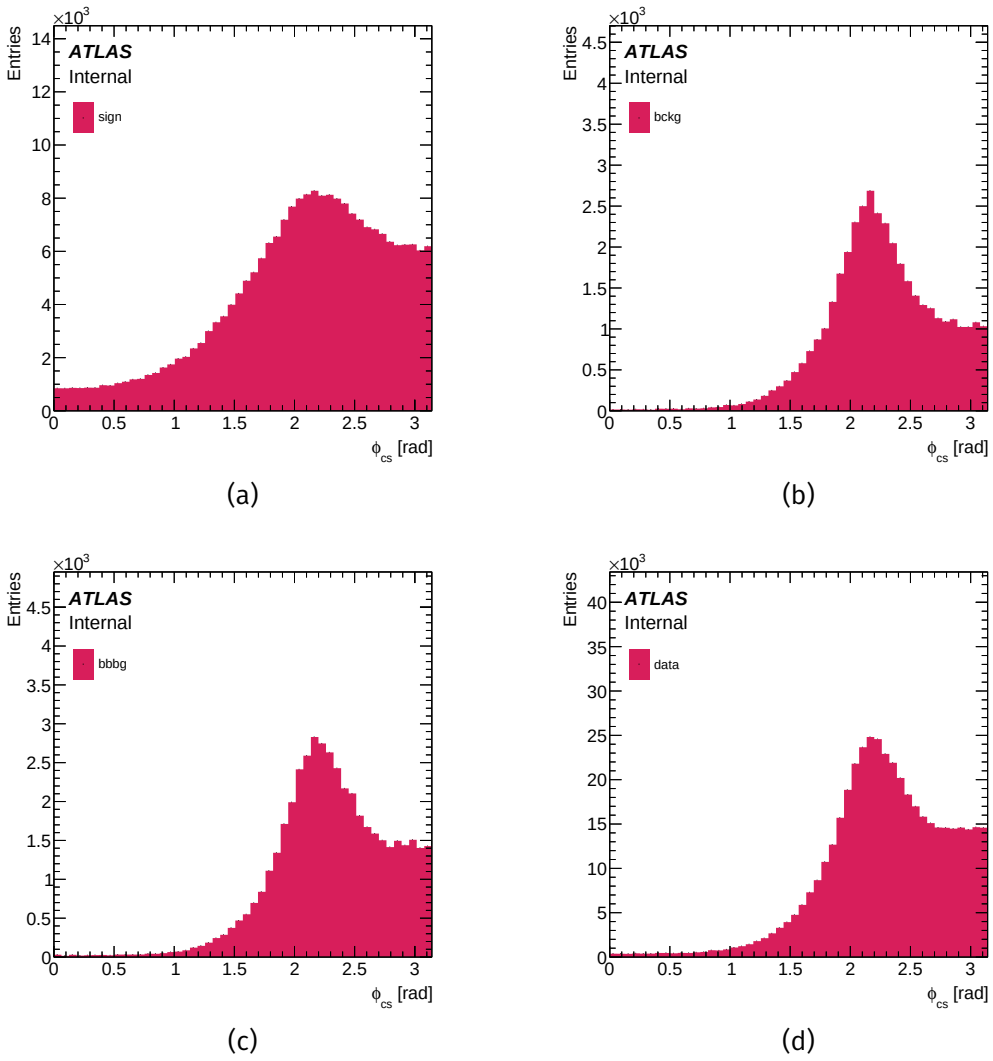


Figure 5.21: Distribution of reconstructed $|\phi_{cs}|$, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

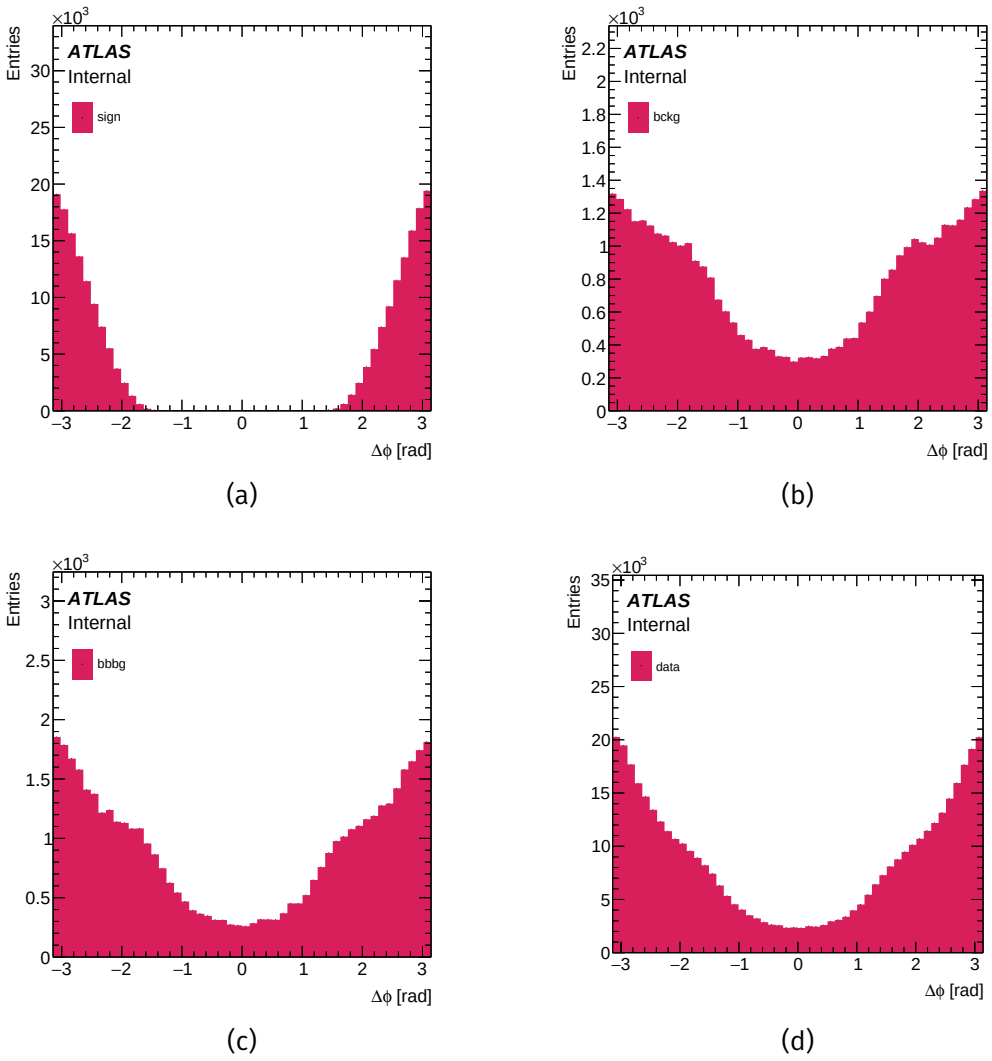


Figure 5.22: Distribution of reconstructed $\Delta\phi$, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

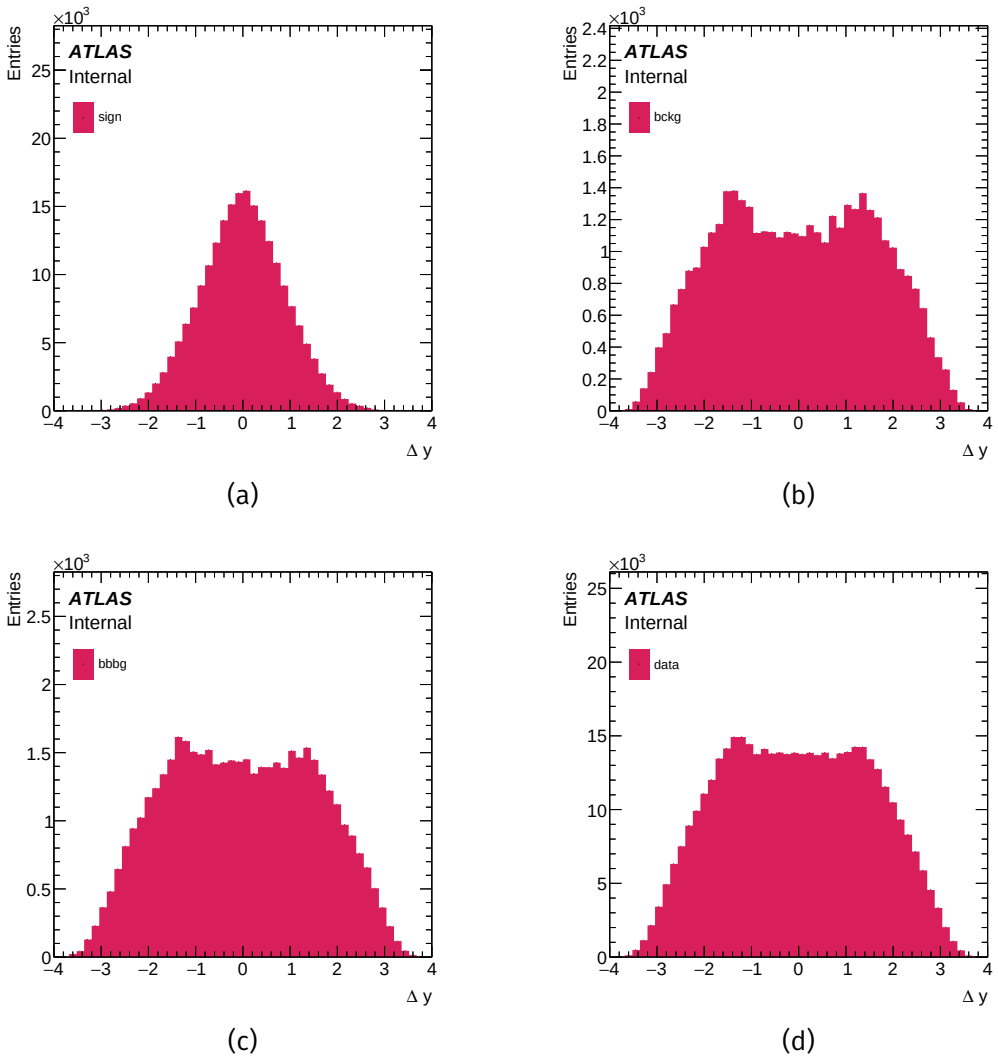


Figure 5.23: Distribution of reconstructed Δy, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

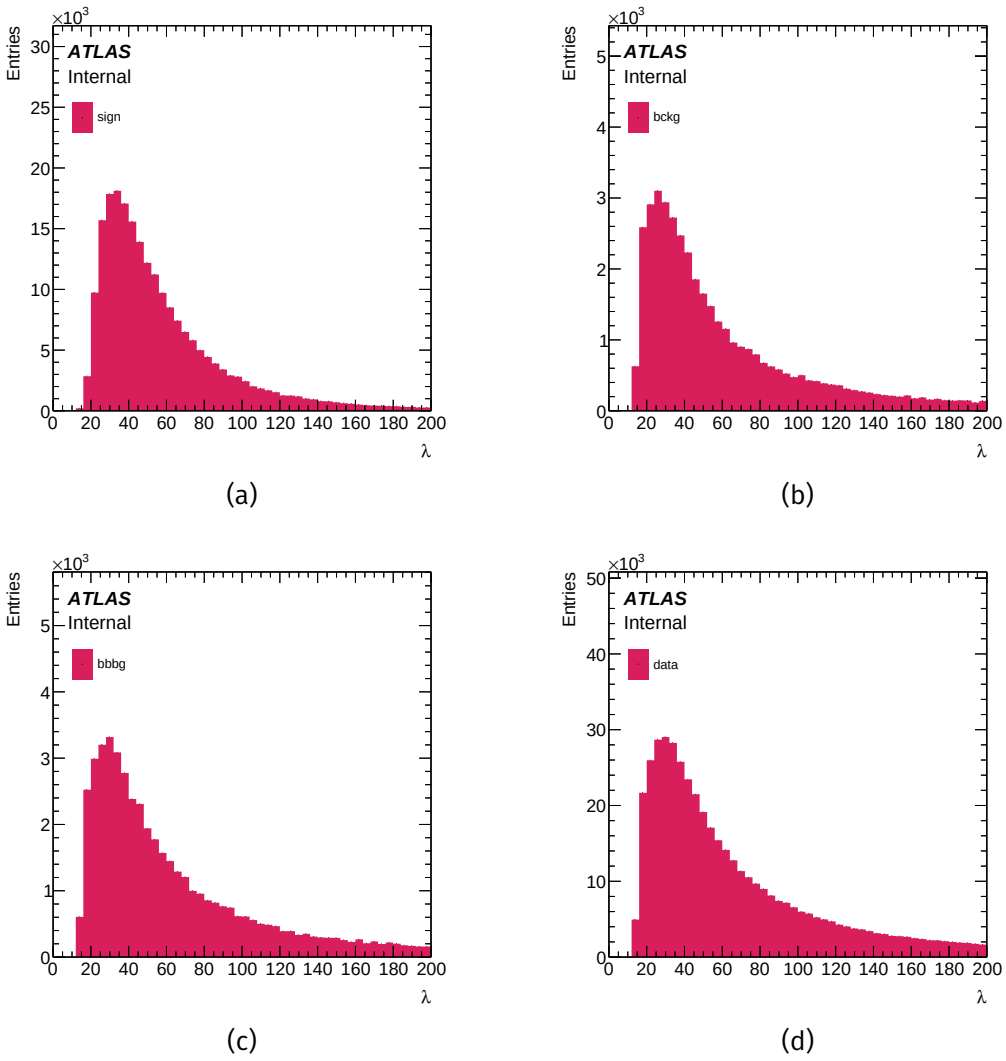


Figure 5.24: Distribution of reconstructed λ, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

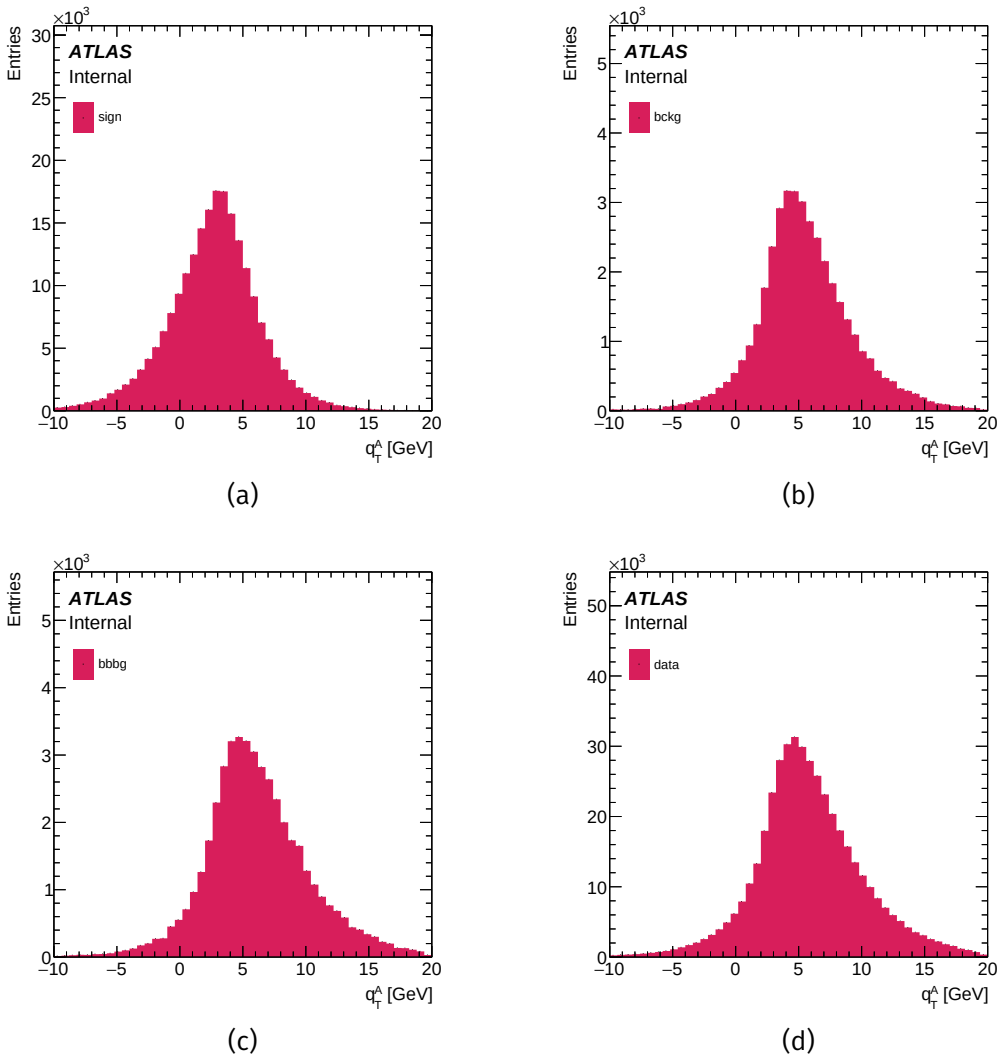


Figure 5.25: Distribution of reconstructed q_T^A, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

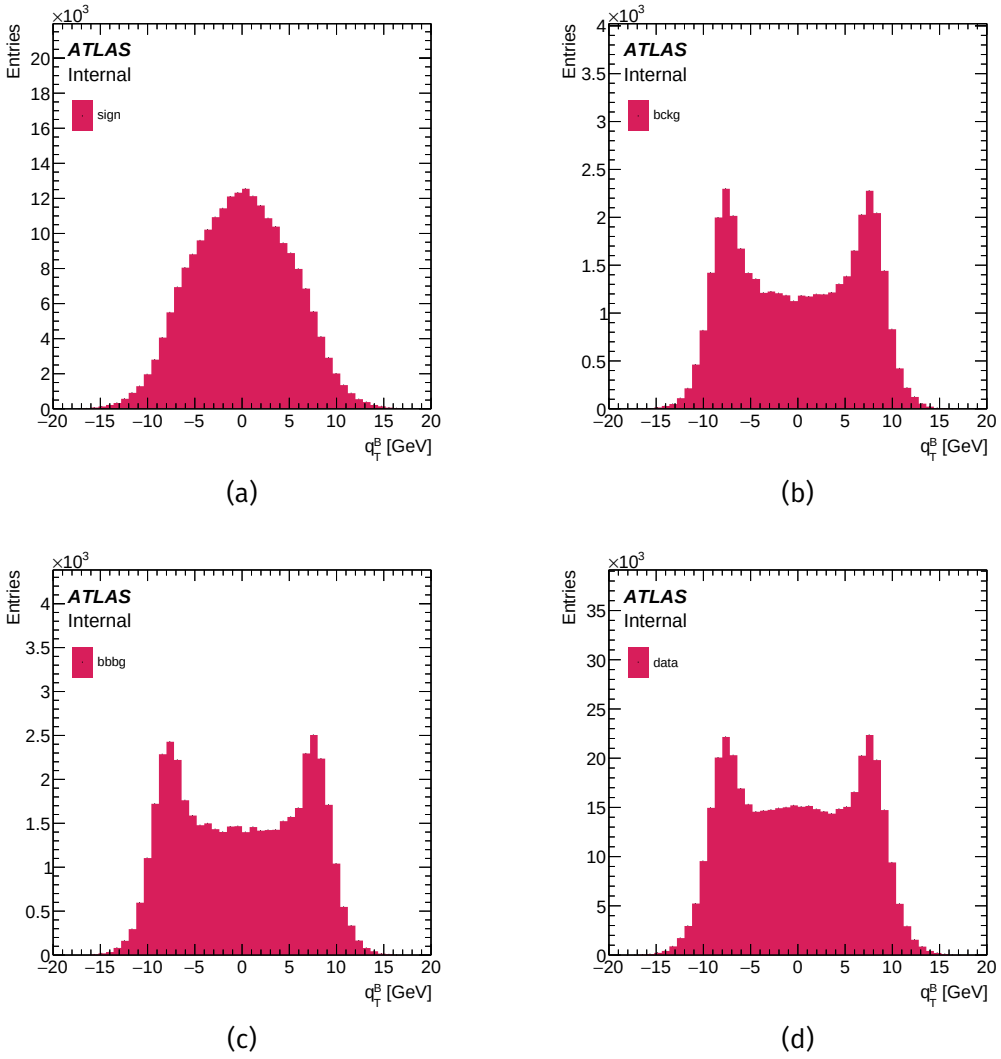


Figure 5.26: Distribution of reconstructed q_T^B, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

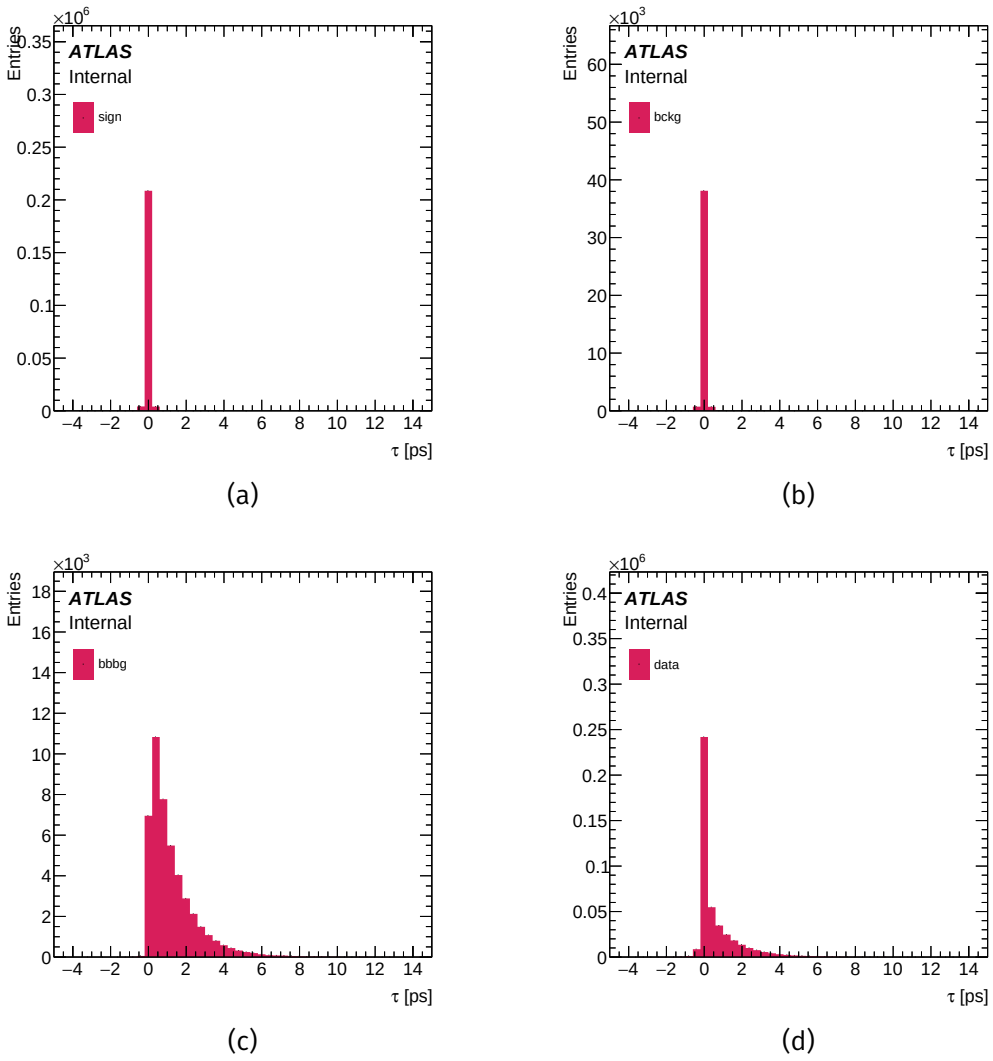


Figure 5.27: Distribution of reconstructed $\tau_{\mu\mu}$, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

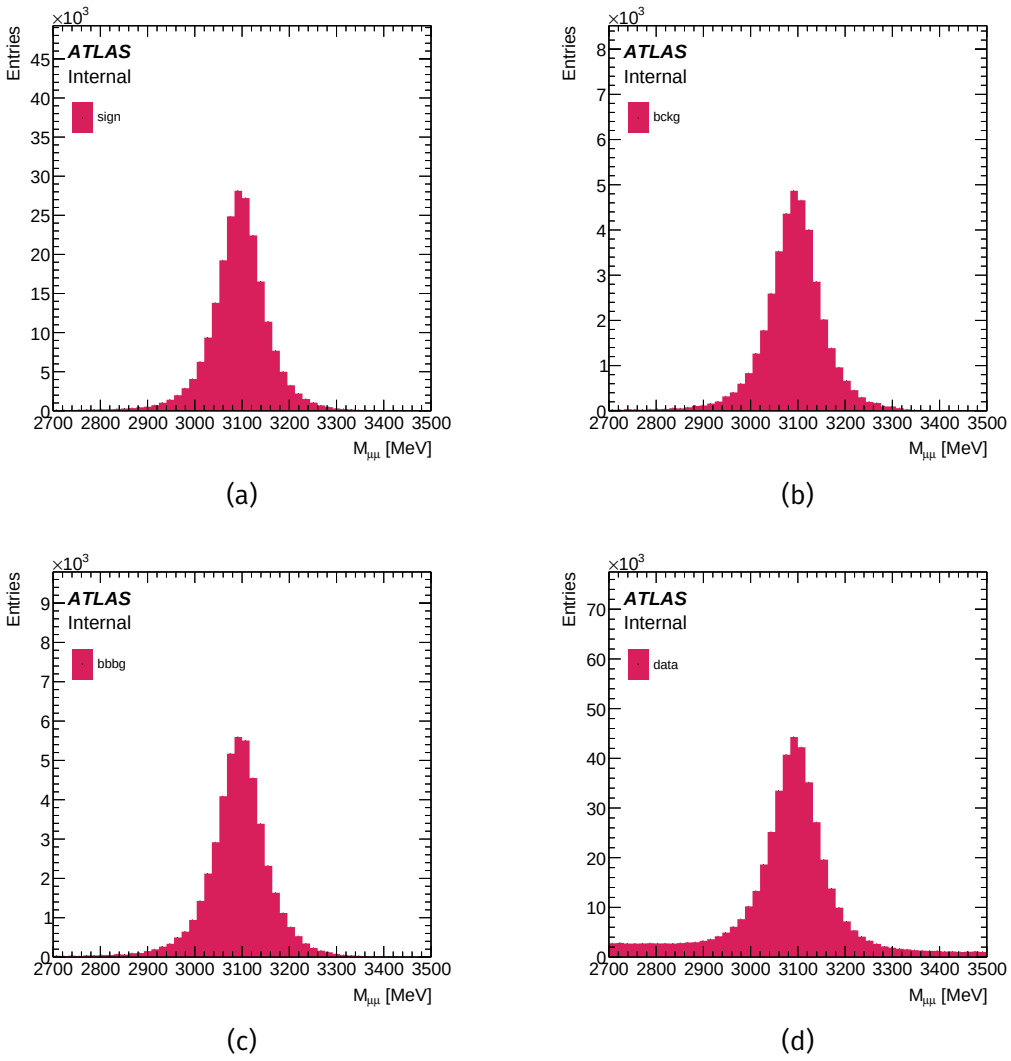


Figure 5.28: Distribution of reconstructed $m_{\mu\mu}$, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

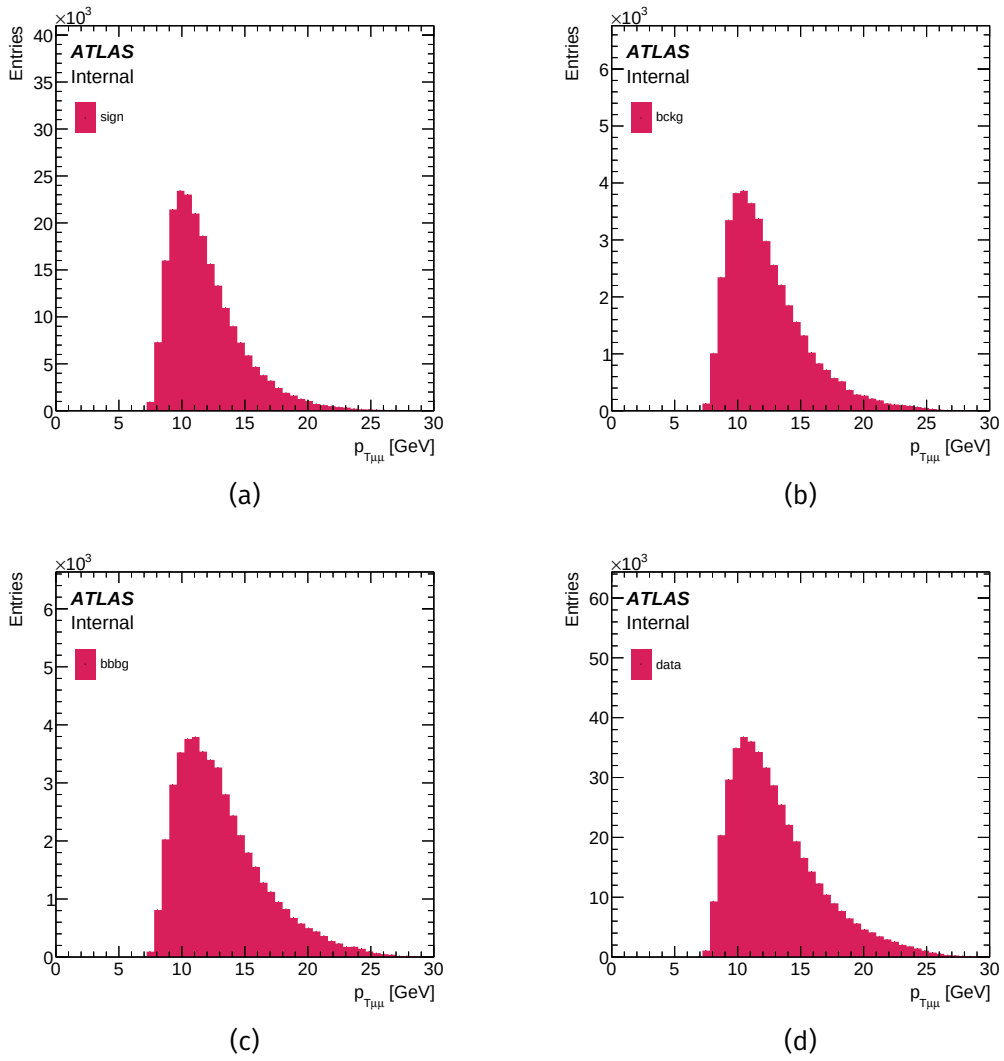


Figure 5.29: Distribution of reconstructed $p_T(\mu\mu)$, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

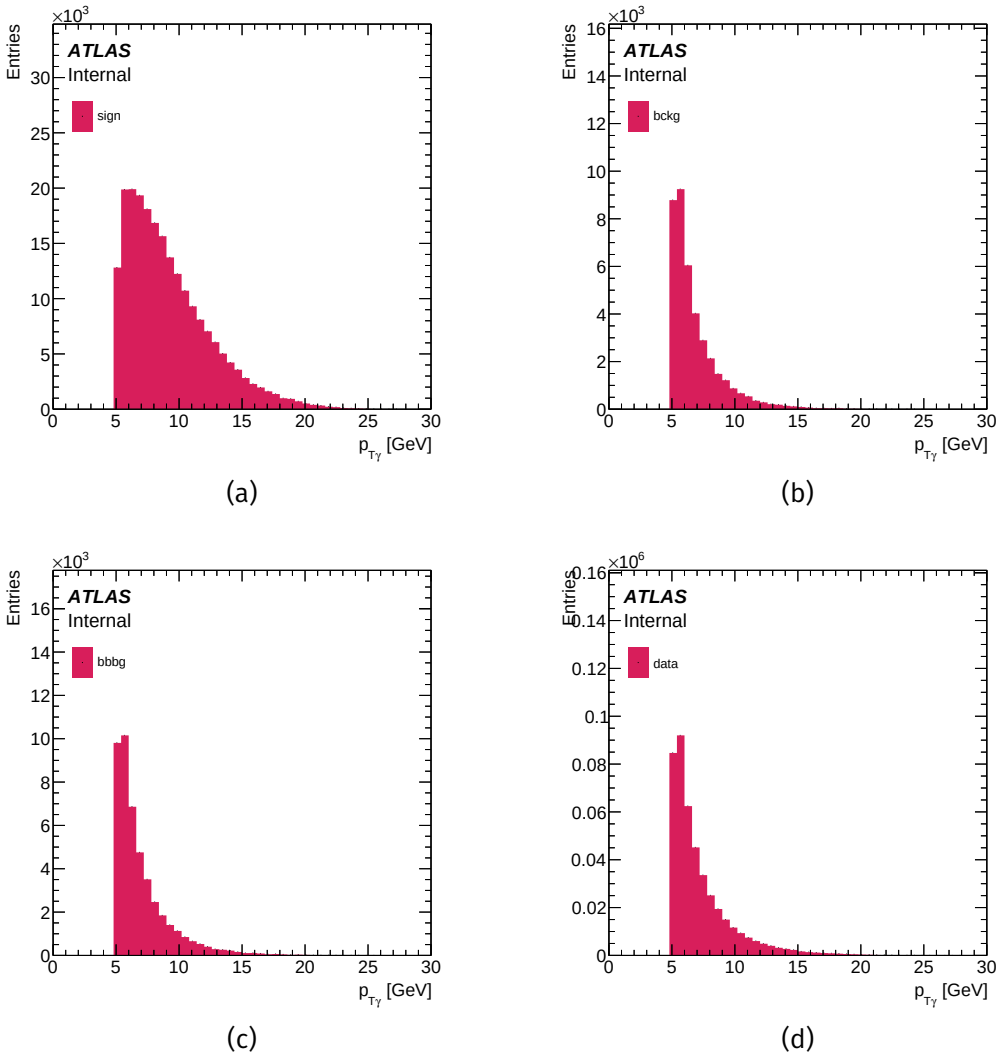


Figure 5.30: Distribution of reconstructed $p_T(\gamma)$, after selection cuts for (a) *sign*, (b) *bckg*, (c) *bbbg* and (d) *data*.

5.3 Background removal

5.3.1 Mass sideband subtraction

Previous studies have made the sideband contribution in the proximity of the J/ψ mass peak well known [83]. The background is modeled well by exponential or linear functions, with the latter used in this analysis. This allows for a simplified procedure for removing the sideband background and the exponential case is used as a systematic variation.

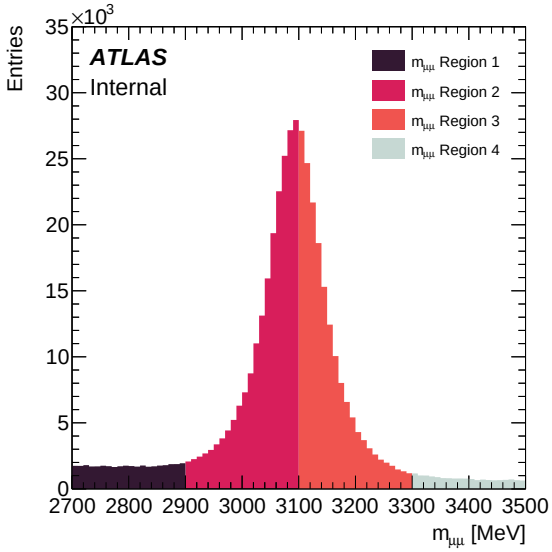


Figure 5.31: Classification of dimuon mass $m_{\mu\mu}$ in the *data* sample, into 4 regions of peak and sideband.

The mass distribution of the dimuon candidates is divided into four equal segments, shown in Figure 5.31. These regions are 200 MeV wide, have edges $2.7 - 2.9 - 3.1 - 3.3 - 3.5$ GeV and are labelled 1 to 4, respectively. Regions are chosen to be approximately equivalent in width to 3σ of the mass resolution, so the regions would cover the entire J/ψ peak. This selection determines that the central regions (2 and 3) have contributions of both signal and the continuum background. The outer regions (1 and 4) contain only background. The combined contribution of continuum background in region 2+3 is equal to that in region 1+4. Hence, the signal yield can be expressed as $(2+3)-(1+4)$. This labelling scheme can be further expanded to become:

$$
\begin{aligned}
(2) + (3) &\rightarrow (7) && s + b \\
(1) + (4) &\rightarrow (8) && b \\
(7) - (8) &\rightarrow (9) && s
\end{aligned}
$$

.

5.3.2 Long lifetime subtraction

From studies of inclusive J/ψ production [83] the pseudo-proper lifetime of J/ψ candidates peaks around 0 ps. The peak is formed from the prompt J/ψ contribution and the quasi-exponential tail. This tail sources from the non-prompt production of the J/ψ from B-hadron decays. The slope is roughly equivalent to the typical B-hadron lifetime.

Fits in the inclusive studies determine the effective lifetime slope to be around $1.2 - 1.3$ ps, corresponding to a half-life τ_0 of 0.86 ± 0.07 ps. Selecting this as a working point splits the non-prompt contribution in two, an equal number of candidates above and below. Similar to the mass sideband subtraction, the region below the half-life working point is labelled 1, and the region above is labelled 2. The removal of non-prompt contributions is provided by $2-1$, leaving a good estimate of the number of prompt candidates.

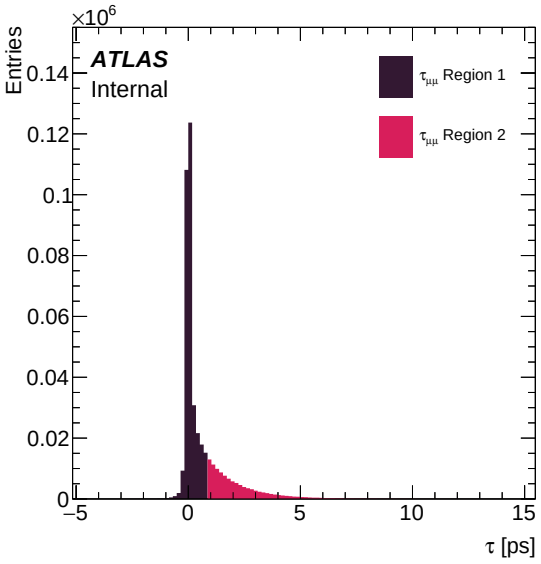


Figure 5.32: Distribution of dimuon pseudo-proper lifetime τ segmented into prompt and non-prompt regions.

The subtraction of both the long lifetime and mass sidebands are performed simultaneously, if the indexing from the two processes is combined (mass-lifetime), the combination becomes:

$$[(21) + (31) - (11) - (41)] - [(22) + (32) - (12) - (42)] \to (99). \tag{5.4}$$

Once this combination is expanded out, the events in regions premultiplied with +1 or -1 are given this weight in the analysis.

The statistical errors require no advanced considerations to propagate; the statistical uncertainty is proportional to the square root of the total number of events in each bin. This is due to weights only being ± 1. As continuum contributions are low, and the non-prompt contributions are not severe, the relative increase in statistical error incurred is not cause for concern. Alternatively to this, the removal of these contributions using 2D maximum-likelihood fits has been attempted in the process of another measurement. This was used for inclusive J/ψ and ψ' analyses, though would be exceedingly complex for this analysis. The likelihood fit process was used during the work following [83]. Comparing with the subtraction procedure, the likelihood procedure distributions were within 1σ uncertainty.

The efficacy of the subtraction procedure can be seen in its effects on the *bbbg* sample. After reweighting the distribution the non-prompt contribution is near-negligible, this is visible in Figure 5.33.

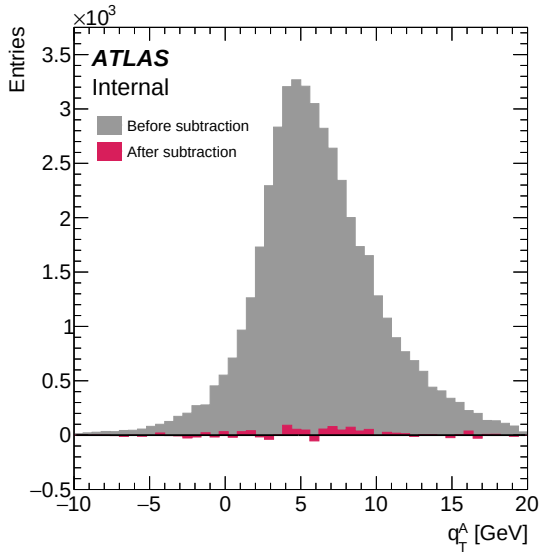


Figure 5.33: Distribution of q_T^A in the *bbbg* sample before (grey) and after (pink) the application of the subtraction process.

The procedure is applied to all samples and the integral of each sample in the

region is given in Table 5.3. In this table, each region is labeled by its mass-lifetime index; the lowest mass region at high lifetime being labelled 12, for example. Non-prompt contribution can be considered removed after this step, and the bbbg sample is not involved in the analysis after this point.

$m_{\mu\mu} - \tau$ **Index**	*data*	*sign*	*bckg*	*bbbg*
00	438187	217241	39611	45920
11	20644	2724	595	341
12	14425	1	0	363
21	163816	114010	20965	12325
22	50065	44	20	11960
31	134767	99516	17811	10433
32	38719	74	15	10288
41	11778	867	204	119
42	3973	3	1	91
71	298583	213526	38776	22758
72	88784	118	35	22248
81	32422	3591	799	460
82	18398	4	1	454
91	266161	209935	37977	22298
92	70386	114	34	21794
99	195775	209821	37943	504

Table 5.3: Number of events in each sample after subtraction operation is performed.

5.4 Boosted Decision Trees

The separation of the single proton scattering (SPS) signal contribution corresponding to the $J/\psi + \gamma$ subprocess, from that of the uncorrelated and double parton scattering (DPS) background contributions is a two-step process. First, a boosted decision tree (BDT) is used, followed by a series of likelihood fits. A boosted decision tree is a multivariate classifier based on several principles [91, 92]. Decision trees make up the core of the technique, they represent a sequential series of conditions applied to an event to classify it. Each condition

is typically a cut on a variable that is held by an event, these are known as discriminating variables. Trees are constructed from nodes and branches and take events as input, they are a common type of 'weak learner'. Each node holds a condition that decides which branch an event travels down towards either another decision node or a leaf node. Events are fed into the root node, and flow down the branches to a leaf node which classifies the event. In the case of a BDT, a large number of decision trees are created. Each time a decision tree is created it influences the construction of the next decision tree. The combined output predictions of all these trees form the output prediction of the BDT. When acting together the trees form a strong learner.

The process of building a BDT using the AdaBoost technique [93] is well established and will only be covered briefly. An input sample of n events is taken and split evenly into a training and testing set. Each event, x, is given an equal weight w so that the total sum of weights is 1. Similarly, each event will have a class k and can belong to an arbitrary number of classes m. Then, a decision tree h is built with the full training set. In this process, each discriminating variable is checked for its effectiveness in separating the events by class using the Gini Index:

$$G = 1 - \sum_{k=0}^{m} (p_k)^2 \tag{5.5}$$

Where G is the Gini Index and p_k is the fraction of events in that class. The discriminators of the lowest Gini index are chosen to form each node, this denotes nodes with low entropy and high separation. When a tree is constructed, its total error is assessed, using predicted classes of events from the tree: κ.

$$\epsilon = \sum_{j=0;\kappa\neq k}^{n} w_j \tag{5.6}$$

This is the sum of the weights of incorrectly classified events. The total error α determines the importance a tree has in the process:

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon} \tag{5.7}$$

Once a tree has been constructed with this method the events used to train it are reweighted. The weight of the j^{th} event after the construction of the i^{th} tree becomes

$$w_j^i = w_j^{i-1} e^{\pm\alpha} \tag{5.8}$$

Correctly classified events are weighted with the declining exponential, reducing its importance in the creation of the next learner. The converse is done for incorrectly classified weights. Weights are then renormalized to sum to 1 and the process of creating another weak learner repeats. When all the weak learners have been created, the resulting classification from the strong learner K, is given as:

$$K(x) = \frac{1}{N} \cdot \sum_{i=0}^{N} \alpha_i h_i(x), \tag{5.9}$$

with N learners. The learning rate of the AdaBoost algorithm β is an exponent that is attached to the say, $\alpha \rightarrow \alpha^{\beta}$. A learning rate is selected to avoid overshoot of a possible minimum, if new learners are attached with large say they may dominate the classifier.

5.5 Likelihood Fits

This analysis uses the `TRExFitter` Framework for model building, fitting, and statistical inference [94]. `TRExFitter` performs binned template profile likelihood fits [95] and bases its test statistics on those outlined in [96]. The basics of likelihood fits and the workings of `TRExFitter` are outlined here for reference.

The concept of likelihood is most intuitively understood as the probability of making an observation given some hypothesis or model. In terms of a model containing some parameter θ and some data x, the likelihood takes the notation $L(\theta|x)$. Likelihoods can indicate the viability of a certain model or their minimisation/maximisation by varying parameters θ can be used for fitting models to data. These techniques should be familiar as maximum-likelihood methods and the derived parameters (denoted $\hat{\theta}$) are often known as maximum likelihood estimates (MLE). Often in particle physics the likelihood ratio is used as a test statistic. In the simplest case, with hypotheses $H_{0,1} : \theta = \theta_{0,1}$, and no unknowns, the likelihood ratio becomes:

$$R = \frac{L(\theta_0|x)}{L(\theta_1|x)} \tag{5.10}$$

Here H_0 and H_1 would be selections of parameters that define the null hypothesis and some model under test, respectively.

In searches for new physics smaller values of R indicate a preference for the

model over the null hypothesis. Some test statistic based on the value of R is used to reject or accept a hypothesis. Typically in particle physics, the hypotheses in use are a “background-only" model H_0 and “signal+background" model H_1. These models will also contain more than one parameter ($\theta \rightarrow \vec{\theta}$) as well as ‘nuisance parameters’ that exist as unknowns and must also be fitted from the data.

5.5.1 The profile likelihood method

Though prevalent, tests such as the likelihood ratio have their limitations [97]. As models may be unable to describe all effects present and acting on the data, additional nuisance parameters are added to systematic uncertainties. As the presence of additional nuisance parameters will reduce sensitivity to those parameters of interest in a model, errorn parameters of interest will increase. In cases with large numbers of nuisance parameters or when the shape of the log-likelihood function is extremely different from Gaussian, accuracy is reduced. Profile-likelihood methods offers some resistance to this. The parameter vector is separable into two categories: $\vec{\theta} = (\vec{\mu}, \vec{\theta})$. The first category are Parameters of Interest (PoI) $\vec{\mu}$, which may represent a cross-section or signal normalisation. The second category contains Nuisance Parameters (NPs) $\vec{\theta}$, such as background normalisations. To create a ‘profiled’ likelihood, nuisance parameters are constrained to values that maximise the likelihood for particular values of μ:

$$\hat{\hat{\theta}} = \arg\max_{\theta}(L(\mu, \theta)) \tag{5.11}$$

With the profiled NPs $\hat{\hat{\theta}}$ a likelihood dependent only on the PoI can be built:

$$L_p(\mu) = L(\mu, \hat{\hat{\theta}}) \tag{5.12}$$

This new profile-likelihood L_p (also often called a constrained likelihood) can be the same way a regular likelihood function can be used for finding intervals and regions of PoI. To evaluate μ, or rather regions of μ space, the profile-likelihood ratio λ is used:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}. \tag{5.13}$$

Here, $\hat{\mu}$ represents the profiled POI μ. For larger values of λ it can be said that μ is increasingly compatible with the data. Similarly applicable is an asymptotic

test statistic $t_\mu = -2 \ln \lambda(\mu)$, where lower values indicate more compatibility. According to Wilks' theorem[98], with sufficiently large samples this test statistic, along with some condition, allows μ to be rejected or accepted for the entire space of θ values.

5.5.2 Systematics, statistics, and the full likelihood

`TRExFitter` can be seen as an extension of another framework `HistFactory` [99], expanded to provide profile likelihood fits. The likelihood fit at its core still relies on a maximum likelihood method, though some modifications are performed. Starting from the case of a single bin with x true data events, the probability of observing ν events is described by the Poisson distribution:

$$P(\nu|x) = \frac{\nu^x e^{-\nu}}{x!} \tag{5.14}$$

When there is only 1 bin overall (as with a counting analysis) this Poisson distribution simply becomes the likelihood. To expand to n bins of a histogram, the likelihood becomes the product of the Poisson distribution of each bin:

$$L(\vec{\nu}|\vec{x}) = \prod_i^n P(\nu_i|x_i). \tag{5.15}$$

With observed data $\vec{x} = (x_i, ..., x_n)$, $\vec{\nu} = (\nu_i, ..., \nu_n)$ can be associated with some prediction, and hence some model, dependent on parameters $\vec{\nu}(\vec{\theta})$. Minimising the negative log likelihood (NLL), defined as $-\ln(L(\vec{\nu}(\vec{\theta})|\vec{x}))$, provides the maximum likelihood estimators. However, this does not account for any systematic effects which must be introduced using prior information from independent experiment. The systematic effects, included as nuisance parameters $\vec{\theta}$, are constrained by a Gaussian distribution in order to provide a model of their uncertainty:

$$G(\theta_j) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{\theta_j^2}{2\sigma_j^2}}, \tag{5.16}$$

where θ_j is one of m nuisance parameters, and σ_j sets the shape of the Gaussian to be equivalent to $\pm 1\sigma$ of the uncertainty. If a θ parameter takes a value ± 1, this corresponds to $\pm 1\sigma$ variation. Now, the likelihood has extended to:

$$L(\vec{\nu}|\vec{x}) = \prod_i^n P(\nu_i(\vec{\theta})|x_i) \cdot \prod_j^m G(\theta_j) \tag{5.17}$$

The final step is to include the statistical uncertainty associated with limited MC statistics. Uncertainties between individual histogram bins can be considered independent, and the relative systematic uncertainty is derived as $\sigma_i = \frac{\Delta x_i}{x_i}$. From this, the effective number of MC events is derived $m_i = 1/\sigma_i^2$ and nuisance parameters γ are included into the final likelihood [94]:

$$L(\vec{\nu}|\vec{x}) = \prod_i^n P(\gamma_i \nu_i(\vec{\theta})|x_i) \cdot P(\gamma m_i|m_i) \cdot \prod_j^m G(\theta_j) \tag{5.18}$$

5.5.3 Goodness of fit and assessing systematics

To provide some idea of goodness of fit a χ^2 measure is used. This is computed with the use of a saturated model [100] which has a likelihood with increased flexibility that can provide an ideal fit. The saturated likelihood does not suffer from the constraints on NP's and presumes the prediction perfectly matches the data. Taking the ratio of the two likelihoods provides a sufficient goodness of fit test as the values follow a χ^2 distribution asymptotically.

Systematics and their nuisance parameters $\vec{\theta}$ are evaluated through their pulls and impact on parameters of interest μ. The impact of a NP θ_j on μ is the change in the PoI when another fit is performed with θ_j set to a specific value. For the 'pre-fit' impact a variation associated with the pre-fit uncertainty is applied to the MLE: $\theta_j = \hat{\theta}_j \pm 1$. On the post-fit side the uncertainty $\pm\Delta\hat{\theta}$ is used to vary the MLE before the fit, this is a change equivalent to the post-fit uncertainty in $\hat{\theta}$. These impacts are often presented in context of the 'pull' $(\hat{\theta} - \theta_0)/\Delta\theta$, the difference of the MLE from the nominal.

Part III

Measurement of Gluon Transverse Momentum Dependent Parton Distribution Functions

Analysis Structure & Strategy

As mentioned, the goal of this analysis is to measure the intrinsic transverse momentum distribution of gluons within the proton. The width of this distribution, equivalent to $\sqrt{\langle k_T^2 \rangle}$, is used for the purpose of parameterising Transverse Momentum Dependent Parton Distribution Functions. A final state of a $J/\psi + \gamma$ is used for this purpose and the analysis uses ATLAS Run 2 data from 2015. The analysis and $\sqrt{\langle k_T^2 \rangle}$ measurement takes the following structure:

1. MC generation, data collection, reconstruction, and application of ATLAS-specific pre-processing (derivation).
2. Event selection and signal efficiency construction.
3. BDT construction and scoring.
4. Mass sideband and long lifetime removal.
5. Likelihood fit of BDT templates.
6. TMD parameterisation and fiducial cross section measurement.

All the analysis techniques utilised herein have been discussed in Part II. Steps 1 and 2 have been covered in chapter 5, in sections 5.1-5.2. Step 4 is detailed in section 5.3. The foundations of steps 3 and 5 are introduced in Sections 5.4 and 5.5, though these will be futher expanded upon in this part of the document. Chapter 6 discusses the construction of the BDT and the scoring. This includes details of how the multivariate analysis framework is configured, the construction of the discriminating variables, the selected hyperparameters, and the analysis of the BDT scoring and efficacy. Chapter 7 details the differential application of the likelihood fits to the constructed BDT templates: the configuration, construction of input distributions, rebinning process, and statistical uncertainties. Chapter 8 is a guide of the systematics analysis, both those systematics processed inside the likelihood fits, and those outside. Chapters 9 and 10 present the aggregated results and concludes the analysis, respectively.

6. BDTs For SPS & DPS Discrimination

The *data* sample may contain a mixture of correlated J/ψ and γ pairs from the subprocess, and uncorrelated pairs from backgrounds. These are the SPS and DPS contributions, respectively, and these two event types are not physically distinguishable on an event-by-event basis. Hence, separation method must take advantage of the statistical differences between the two types of events and exploit correlations between observables.

The differing distributions of variables in the *sign* MC sample and the *bckg* MC are utilized. These samples are used to train a Boosted Decision Tree to classify events as SPS-like or DPS-like. All the information available from both samples is condensed into a single discriminating variable. Hence, this discriminator contains the separating power that may come from an arbitrary number of selection cuts

6.1 BDT Structure and Training process

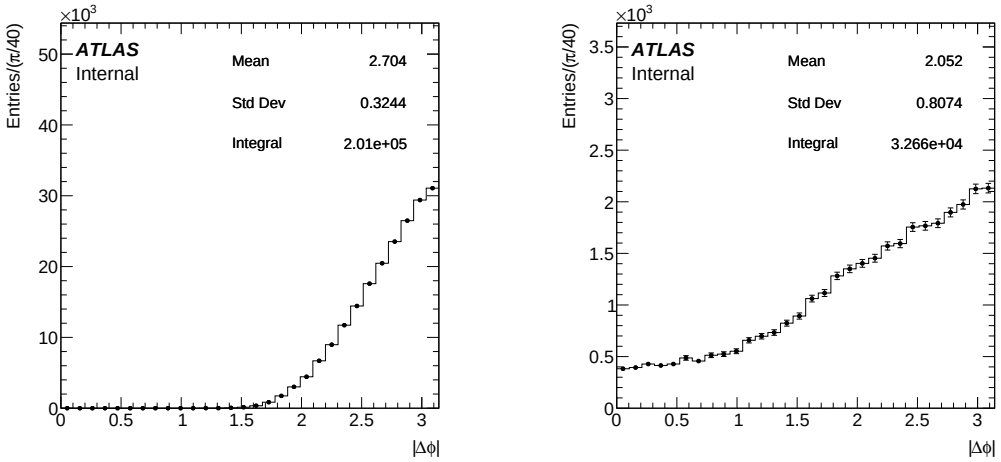
A single set of hyperparameters are selected for this analysis, these are detailed in Table 6.1. The BDT creation process utilises the full *sign* and *bckg* samples, before any background subtraction process is applied, or before the cuts in Section 5.2.2.2 are applied. Events in the input samples are split in half into two subsets: *train* and *test*. Training events are used for the construction of the BDT, testing events are used in its verification. Events in each sample are weighted in the construction process to account for the different size of the two samples. Though common, k-folding cross-validation [102] was not utilised.

Hyperparameter	Value	Explanation
BoostType	AdaBoost	Selected boosting algorithm, common tree boosting algorithm
AdaBoostBeta	0.15	Learning rate of the AdaBoost algorithm
Separation Type	Gini Index	Choice of node splitting metric
NTrees	400	Number of weak learners (decision trees)
MaxDepth	4	Limit of decision tree depth
MinNodeSize	5%	Fraction of total training events that propagate to each node
NCuts	100	Number of points over the range of a variable considered for finding the best point to cut
PruneMethod	No Pruning	No method selected for removing ineffective nodes and branches
Remaining	Default	Remaining relevant defaults are contained in [101]

Table 6.1: Hyperparameters input into the TMVA (Toolkit for Multivariate Data Analysis) framework for BDT construction.

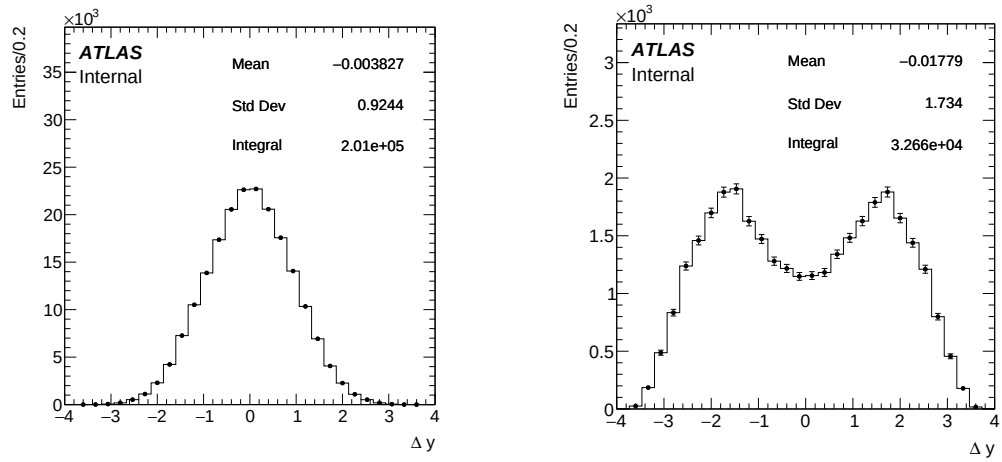
6.2 Discriminating Variables

Previous discussion elucidates that SPS contributions are more likely to follow a back-to-back pattern in $\Delta\phi$. Similarly, these events should have limited spread in Δy. For the DPS and uncorrelated contributions, eponymously, there is no expected correlation between the y and ϕ of the J/ψ and γ. The differences in the $\Delta\phi$ and Δy in the main MC samples in this analysis are visible in Figures 6.1 and 6.2 The acceptance cuts on the muons and photon in this analysis are known to cause distortion to the measured variables, including the $\Delta\phi$ distributions.



(a) SPS contribution (b) DPS contribution

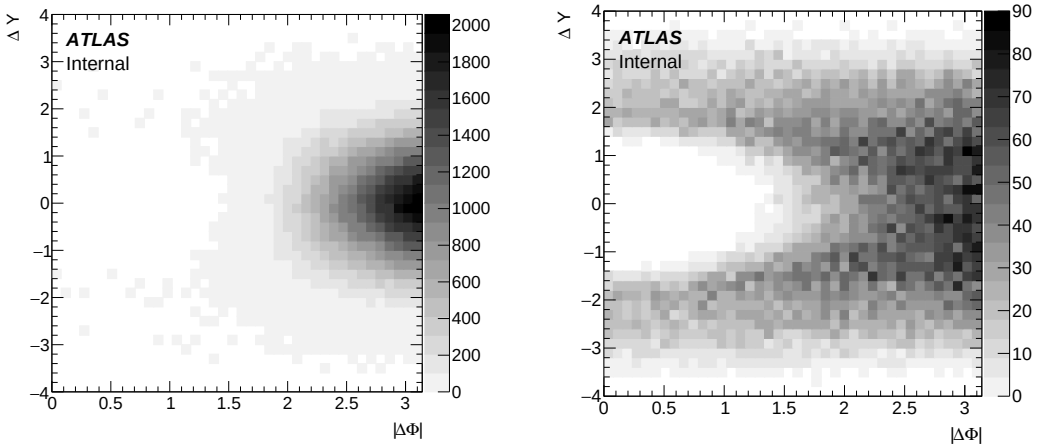
Figure 6.1: Distributions of main discriminating variable $|\Delta\phi|$ for the two MC samples, the sign sample corresponding to the SPS contribution (left) and the bckg sample corresponding to the DPS contribution (right).



(a) SPS contribution (b) DPS contribution

Figure 6.2: Distributions of the main discriminating variable Δy for the two MC samples, the sign sample corresponding to the SPS contribution (left) and the bckg sample corresponding to the DPS contribution (right).

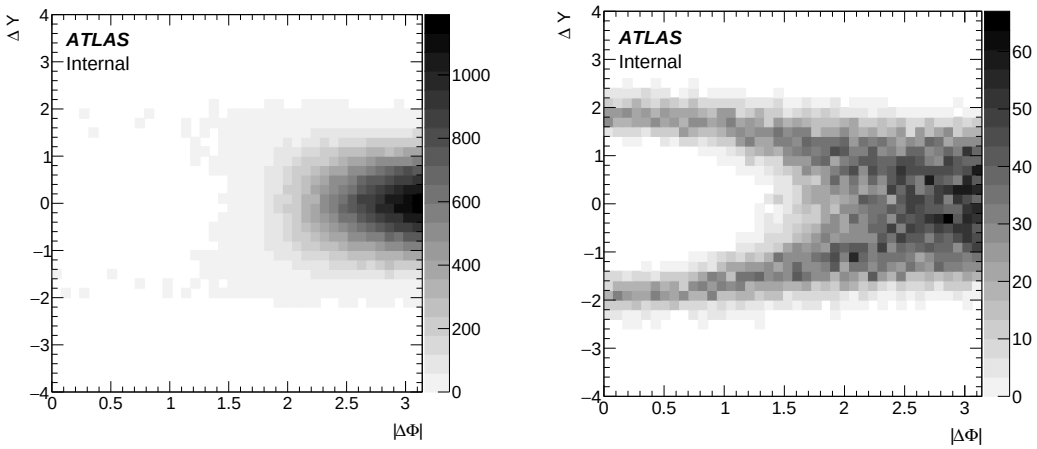
Despite distortion, differences between the two samples are still visible. These differences are more visible when comparing the 2D maps of $\Delta\eta - \Delta Y$, shown in Figure 6.3.



(a) *sign* (b) *bckg*

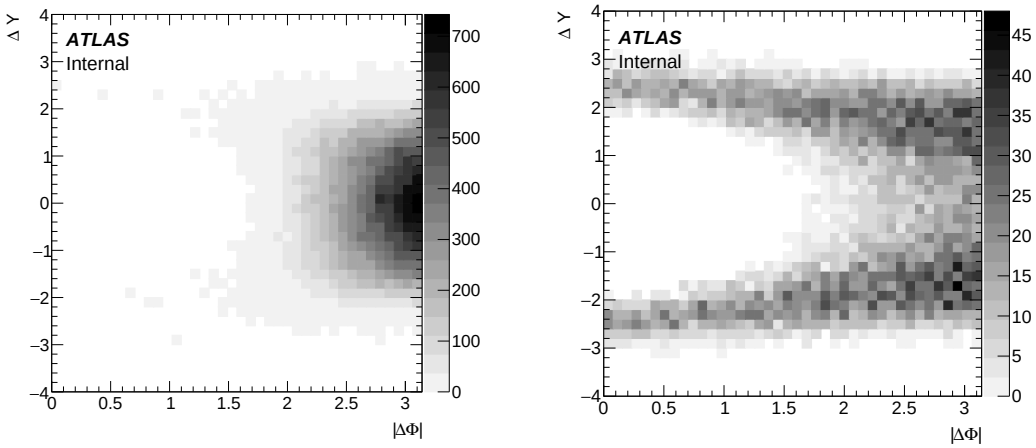
Figure 6.3: 2D distribution of the BDT discriminating variables $\Delta\phi$ and Δy.

These variables $\Delta\phi$ and ΔY are expected to be the main discriminators. The dependence of these two angular variables is correlated with other kinematic variables such as λ and q_T^2. This correlation is shown in Figures 6.4-6.6, where $\Delta\phi - \Delta y$ are shown in progressively increasing mass regions defined as Q3, (defined as the λ range 25-50), Q4 (λ range 50-100) and Q5 (λ range 100-200).



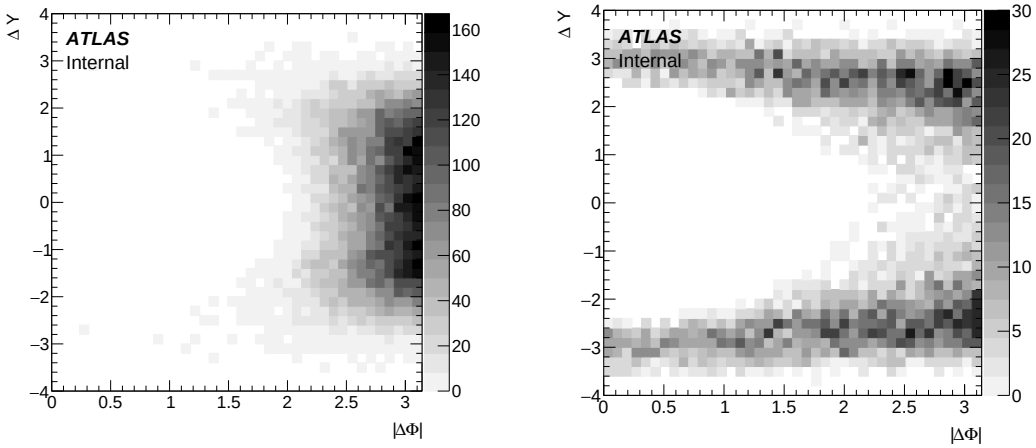
(a) Q3 sign (b) Q3 bckg

Figure 6.4: 2D distributions of $\Delta\phi$ vs ΔY in λ Q3 range for the two MC samples: (left) the *sign* sample corresponding to the SPS contribution, and (right) the *bckg* sample corresponding to the DPS contribution.



(a) Q4 sign (b) Q4 bckg

Figure 6.5: 2D distributions of $\Delta\phi$ vs ΔY in λ Q4 range for the two MC samples: (left) the *sign* sample corresponding to the SPS contribution, and (right) the *bckg* sample corresponding to the DPS contribution.



(a) Q5 sign (b) Q5 bckg

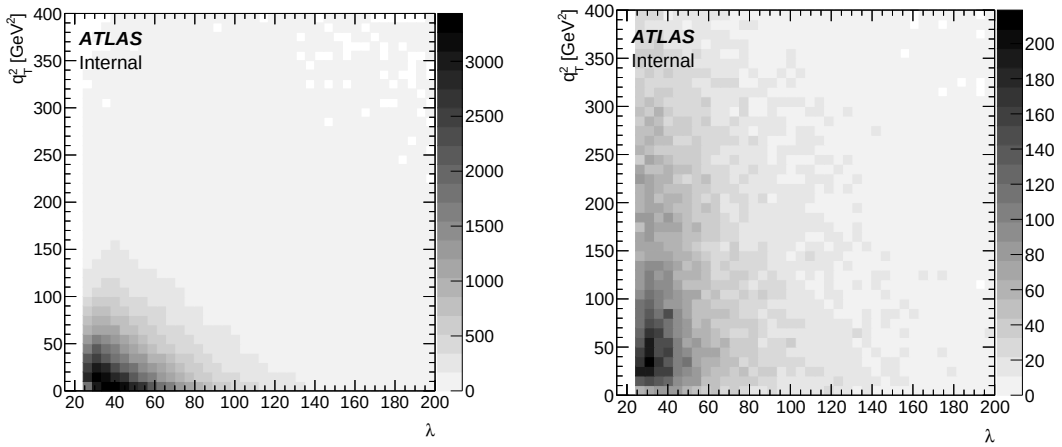
Figure 6.6: 2D distributions of $\Delta\phi$ vs ΔY in λ Q5 range for the two MC samples: (left) the *sign* sample corresponding to the SPS contribution, and (right) the *bckg* sample corresponding to the DPS contribution.

This line of reasoning encourages the creation of additional variables in order to enhance discrimination power. A number of extra discriminators are created, they are outlined in second half of Table 6.2. These are constructed by combining the angular variables with various p_T components of the $J/\psi + \gamma$ system in the lab frame. This approach hopes to build more powerful presentation of information when compared with the two variables alone. Additional discriminating variables are used only in the construction of the BDT.

Variable	**Definition**	**Comment**
$\lvert\Delta\phi\rvert$	Table 5.1	-
Δy	"	-
λ	"	-
q_T^2	"	-
$\lvert\phi\rvert$	-	$J/\psi + \gamma$ system ϕ
$\lvert\cos\theta\rvert$	-	$J/\psi + \gamma$ system $\cos(\theta)$
q_x^{ψ}	$p_T(J/\psi) + \cos(\Delta\phi) \cdot p_T(\gamma)$	x-component of q_T along the J/ψ p_T direction
q_y^{ψ}	$p_T(\gamma) \cdot \sin(\Delta\phi)$	y-component of q_T along J/ψ direction
q_x^{γ}	$p_T(\gamma) + \cos(\Delta\phi) \cdot p_T(J/\psi)$	x-component of q_T along γ p_T direction
q_y^{γ}	$p_T(J/\psi) \cdot \sin(\Delta\phi)$	y-component of q_T along γ p_T direction
q_x^{Σ}	$(p_T(J/\psi) + p_T(\gamma)) \cdot (1 + \cos(\Delta\phi))$	Sum of q_x components
q_y^{Σ}	$(p_T(J/\psi) + p_T(\gamma)) \cdot (\sin(\Delta\phi))$	Sum of q_y components

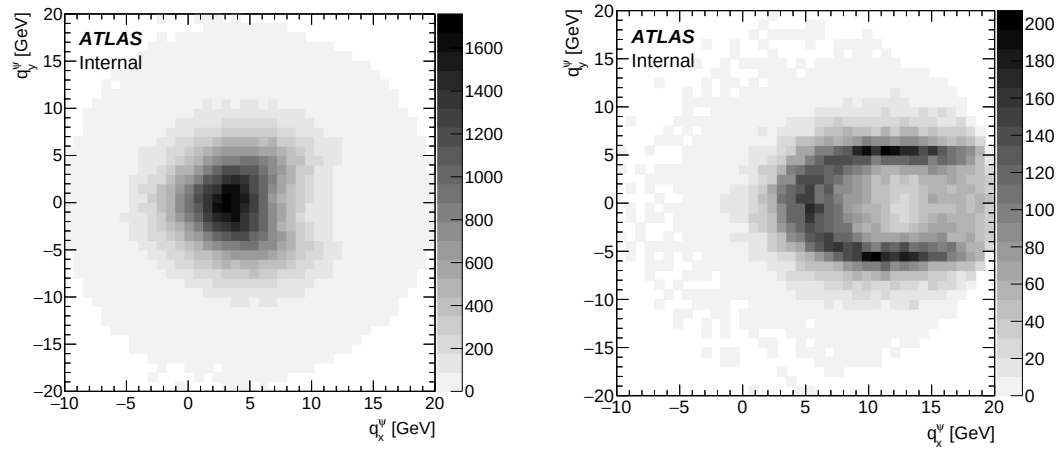
Table 6.2: Variable combinations to create discriminators unique to the BDT process.

A full list of discriminators is given in Table 6.2, the new discriminators are defined in the second half and are shown in maps through Figures 6.7-6.10.



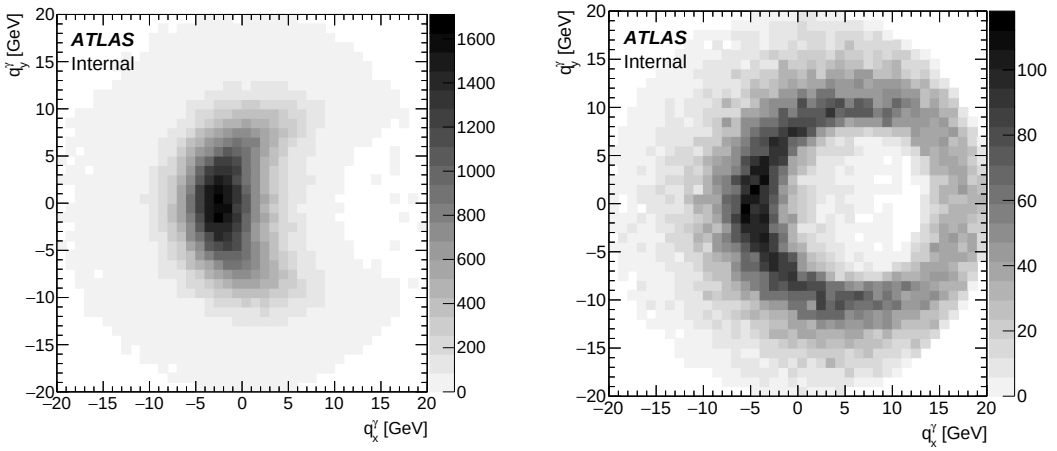
(a) *sign* (b) *bckg*

Figure 6.7: 2D distribution of the BDT discriminating variables λ and q_T^2, the minimum of cut of $\lambda > 25$ is visible on the left.



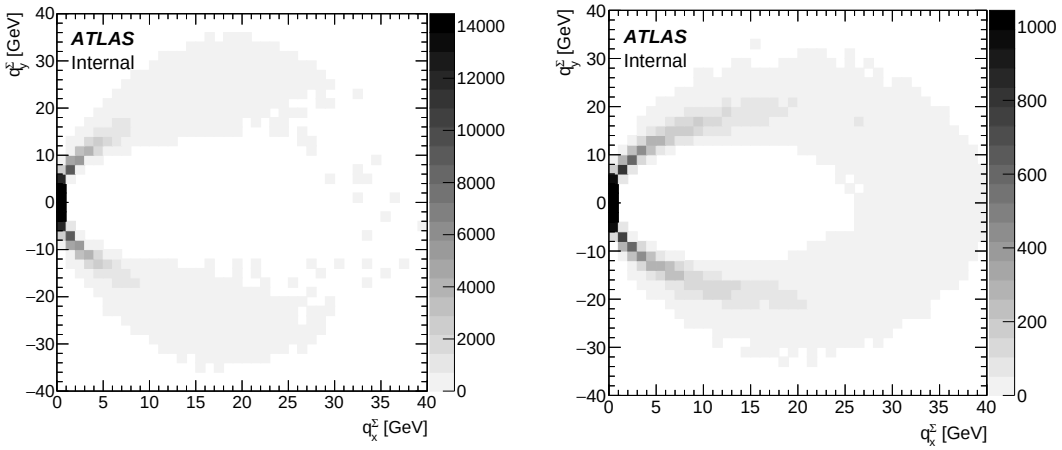
(a) *sign* (b) *bckg*

Figure 6.8: 2D distribution of the BDT discriminating variables q_x^ψ and q_y^ψ.



(a) *sign* (b) *bckg*

Figure 6.9: 2D distribution of the BDT discriminating variables q_x^γ and q_y^γ.



(a) *sign* (b) *bckg*

Figure 6.10: 2D distribution of the BDT discriminating variables q_x^Σ and q_y^Σ.

6.3 Efficacy and Score Distributions

Before the BDT scoring is applied to all samples and used for signal and background separation, some test of the efficacy should be conducted. First a measure of the effect of the derived discriminators should be evaluated. The importance of each discriminator to the BDT process can be ranked by how frequently it is selected as the best choice to split a node, this is shown in Table 6.3. The power of the variable combinations can be seen here, though

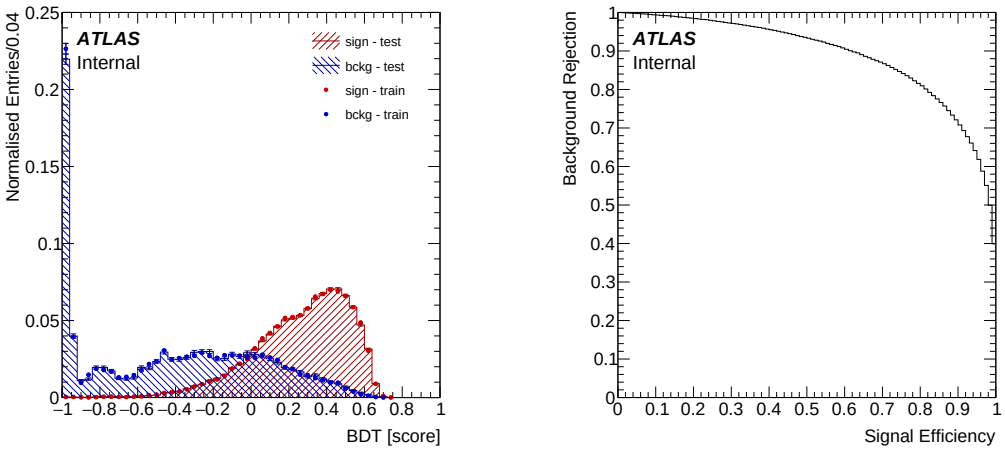
Rank	Variable	Importance
1	$\lvert\Delta\phi\rvert$	0.453
2	q_x^{ψ}	0.223
3	$\lvert\cos\theta\rvert$	0.185
4	q_x^{γ}	0.045
5	q_x^{Σ}	0.042
6	λ	0.022
7	q_T^2	0.008
8	$\lvert\phi\rvert$	0.006
9	q_y^{ψ}	0.006
10	Δy	0.003
11	q_y^{γ}	0.002
12	q_y^{Σ}	0.002

Table 6.3: BDT discriminating variables ranked by importance, the fraction of the total BDT nodes that use a specific discriminating variable to classify at a node.

some are relatively ineffective, namely the q_y components. q_x components demonstrate some good discriminating power, though they do not dominate splits throughout the BDT.

Next, to perform a check that the BDT has consistent behaviour, the overlay of scores of the training and testing samples is examined. The similarity of the two pairs of distributions in Figure 6.11a demonstrates that the BDT has modelled the datasets. For an overall measure of the efficacy of the classifier in separating events, consult Figure 6.11b showing the Receiver Operating Characteristic (ROC) curve. The ROC shows the performance of a classifier across the spectrum of possible cuts used to discriminate between two categories of data. Here, the

curve shows a classifier that offers a good balance of efficiency and background rejection. However, as an overall test the integral of this curve (AUC) can be used to assess the total performance of the classifier in separating events. A perfect classifier would score an AUC of 1. The BDT trained with the *sign* and *bckg* samples has an AUC of 0.890. Tests removing the less important discriminators have been performed. When the BDT is simplified to four main variables, the AUC reduces slightly to 0.883. The full set of discriminators is retained for the increase in discriminating power.



(a) Training and Testing distributions

(b) Rejection and Efficiency

Figure 6.11: (a) Normalised BDT score distributions for the *sign* and *bckg* samples and (b) the efficiency/rejection curve for the BDT.

7. Differential Binned Template Likelihood Fits

Once all events in all samples have been scored by the BDT the mass sideband and non-prompt subtraction procedure is performed and the cut on $|q_T^B|$ is applied. The remaining events, in a high signal/background region, are used to perform the likelihood fit.

7.1 Differential Fits and the Fit Model

Currently, if the likelihood fit were to be performed using the BDT score, the results would amount to an analysis of fiducial cross-section. However, as the aim is to parameterise the TMD f_1^g, multiple fits are performed. 15 histograms of BDT score are created, each representing a 2 GeV wide bin in the q_T^A distribution, across the range of -10 to 20 GeV. A likelihood fit is performed using the BDT template histograms for *sign*, *bckg* and *data* in each q_T^A bin separately. Then, the extracted signal contribution in each bin is used to reconstruct the SPS distribution of q_T^A that exists within the *data* sample. Following this, the resultant distribution can be used to parameterise f_1^g. In this context, the BDT can be thought of less as a classifier and more of a tool to condense all useful information about an event down into a single one-dimensional variable.

The likelihood fits applied here differ from typical new-physics searches and those which are performed with a 'background-only' hypothesis. Here, it is known that there must be some transverse component to the gluon dynamics, though its nature is not explicit. Therefore, each fit of BDT score aims to rescale both the *sign* and *bckg* templates.

The process is equivalent to solving for the expression:

$$x = \mu \cdot sign + \kappa \cdot bckg \tag{7.1}$$

x, as in Section 5.5, represents the *data* template and μ is the parameter of interest (PoI) for rescaling the signal template. κ is the PoI for rescaling the background.

The fitting framework that manages the fits offers a high level of configurability. Considering that the fit is performed independently in multiple bins, the same suite of options is used in each fit. Aside from the selection of nuisance parameters and POI's, a number of additional configuration options have been selected. Gamma NP's are removed when their influence is below a threshold of 1%. This simplified the fit space and improves convergence. Similar constraints on the NP's associated with systematics is performed, those below 0.5% are pruned from the fit. An automated binning procedure is used: TransfoD with settings 2.0:2.0, this is discussed in Section 7.2.

7.2 Template Construction and Binning

Templates of BDT score in each q_T^A bin from the *sign*, *bckg*, and *data* must be formed. As mentioned, a $|q_T^B| < 6$ GeV cut has been applied. A number of additional weights are applied to each distribution:

Subtraction weight. As detailed in Section 5.3, a weight for removing the non-prompt and dimuon continuum contributions from *data* is generated. This weight is applied to events from all three input samples.

Pileup weight. A corrective weight to adjust for differences in pileup from realistic data conditions, derived from ATLAS standard tools. Applied to the events within the MC samples, *sign* and *bckg*.

Photon ID Efficiency Scale Factor. Weight to account for efficiency differences between MC and data conditions according to photon behaviour. Provided by EGamma ATLAS working group. Applied to events from *sign* and *bckg* samples. This is reconstruction related and independent of the `2mu4` trigger.

Muon Trigger Scale Factor. Correction for trigger efficiency differences between MC and data. Dependent on the p_T and rapidity of single muons and muon pairs. Derived from standard ATLAS tools, applied to events from *sign* and *bckg* samples.

Systematic variations of these weights are explored in Chapter 8.

Once the distributions have been formed, the fitting framework provides procedures for rebinning the distributions to provide some optimisation of the starting conditions for a more performant fit. The algorithm utilised for this process is known as Transformation D [103], and is guided by the following:

$$Z = z_s \frac{n_s}{N_s} + z_b \frac{n_b}{N_b}. \tag{7.2}$$

Z defines a measure of the optimisation. The procedure will continually merge bins until $Z < 1$ across the full range. Denominators N_s and N_b are the total count of signal and background events. Input parameters z_s and z_b are bound by the constraint $z_s + z_b = N$ were N is the total number of bins. These input z parameters also control the maximum fraction of signal and background events in a bin. n_s and n_b are the total number of events in a candidate merged bin. The trend of equal z parameters is followed.

7.3 Extracted Yields

Table 7.1 details the extracted yields of *sign* and *bckg* in each q_T^A bin. This fit is performed with the full array of systematics available to the likelihood. However, Table 7.1 gives only the statistical error on each yield, until the systematics can be explored in detail in Chapter 8.

As validation of the full analysis process, a study using toy-model data samples is conducted. This study shows that the extracted signal yields are representative of the signal present, the details are given in Appendix C.

Bin	q_T^A (GeV)	*sign*	*bckg*	*data*
1	-10 – -8	1 ± 1	174 ± 13	175 ± 13
2	-8 – -6	78 ± 20	177 ± 22	255 ± 16
3	-6 – -4	448 ± 34	154 ± 29	602 ± 25
4	-4 – -2	464 ± 51	676 ± 53	1140 ± 34
5	-2 – 0	899 ± 60	1827 ± 67	2726 ± 52
6	0 – 2	1643 ± 85	5336 ± 105	6979 ± 84
7	2 – 4	2518 ± 115	15114 ± 160	17632 ± 133
8	4 – 6	2643 ± 126	19490 ± 181	22133 ± 149
9	6 – 8	2136 ± 112	13806 ± 155	15942 ± 126
10	8 – 10	981 ± 86	8102 ± 121	9083 ± 95
11	10 – 12	465 ± 78	4806 ± 102	5271 ± 73
12	12 – 14	370 ± 58	2653 ± 75	3023 ± 55
13	14 – 16	88 ± 44	1795 ± 60	1883 ± 43
14	16 – 18	32 ± 33	1069 ± 48	1101 ± 33
15	18 – 20	253 ± 35	439 ± 38	692 ± 26

Table 7.1: Post-Fit Yields and absolute statistical error, yields of all samples from `TRExFitter` in all 15 analysis q_T^A bins.

8. Systematics

A multitude of systematic variations and effects have been considered from all stages of the analysis. Systematics evaluations made herein are related to the choices made around fixed parameters, assumptions, reconstruction efficiency, and the uncertainties in MC samples. The tools available for the fitting process allow us to handle a large fraction of the systematic errors within the `TRExFitter` framework. Remaining systematics are included by different methods. Application of reconstruction efficiency correction is discussed in the first subsection. This delineation creates two categories of systematics that are detailed as follows:

Internal. These are systematics that can be included as variations in the `TRExFitter` process. They are given nuisance parameters in the likelihood. Systematics of the analysis process are only included in this category if they meet specific criteria: they can be altered without producing an effect on the *data* sample. This group, by construction, is a collection of the systematics associated with the MC samples. Examples include the systematic errors associated with the muon trigger scale factors. As these systematics are managed almost entirely by the `TRExFitter` framework, only input histograms that outline the $+1\sigma$ variation associated with the systematic must be provided. This is typically called an 'upper' variation. When input into `TRExFitter`, the upper variation is mirrored around the nominal to create a symmetric variation. For each sample, *sign* and *bckg*, and each systematic, an upper variation is created.

External. The other set of systematics consists of those that have broader effects. These affect the *sign*, *bckg*, and additionally the *data* sample. This category is named External. Each systematic is a group of variations

of a single analysis step or assumption performed before the likelihood fitting process. The full fitting process is then repeated for each of these variations and the effects on the extracted signal yield are conservatively expected to be distributed uniformly. In a group the difference between the maximal variation and minimum variation of the yields over $\sqrt{3}$ is taken as the absolute error. This absolute error derives from the uncertainty over a uniform probability distribution over some interval. The error calculated from these groups are added in quadrature with the final error from the fits with the internal systematics included. Steps are taken to avoid double counting of systematics and second order effects. This requires performing the likelihood fit step without the presence of any Internal systematics present in the likelihood.

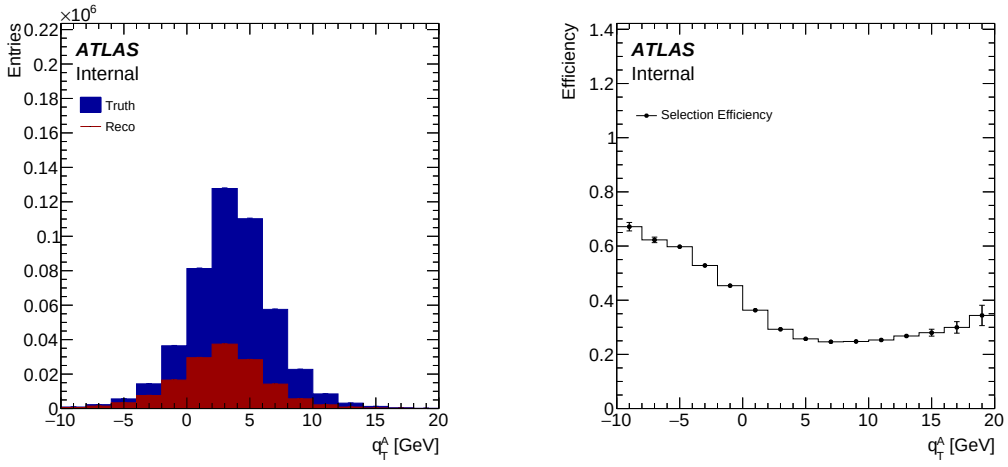
Cut	Description	Truth Signal
0	No cut	8399000
1	Trigger	8399000
2	γ and μ's present	8399000
3	$p_{\mathrm{T}}(\mu^{\pm}) > 4$ GeV	1021150
4	$15 \leq \lambda \leq 200$	998182
5	$0 \leq \cos^2\theta_{cs}$	998182
6	$q_T \leq 20$GeV	996746
7	$0 \leq \lvert\phi_{CS}\rvert < \pi$ rad.	996746
8	2.7 GeV $\leq m_{\mu\mu} \leq$ 3.5 GeV	996746
9	$-5\ ps \leq \tau_{\mu\mu} \leq 15$ ps	996746
10	$-10 < q_T^A < 20$	996746
11	$\lvert q_T^B\rvert < 15$	996746
12	γ and μ, $\lvert\eta\rvert < 2.5$ and $\lvert y\rvert < 2.5$	866652
13	$p_{\mathrm{T}}(\gamma) > 5$ GeV	689988

Table 8.1: Cut flow of *sign* events through all basic selection cuts applied to the truth-level variables in the signal Monte Carlo sample.

8.1 Efficiency

Any extracted signal contribution from the *data* sample must be adjusted for the efficiency of selection cuts detailed in Section 5.2.2. Truth MC information in the *sign* sample is used to provide a determination of this efficiency. Cuts detailed in Table 5.2 are applied to the equivalent truth variables, the resulting truth level cutflow is given in Table 8.1. These cuts are utilised for the extraction of the TMD, though this is reduced to just cuts 1, 2, 3, 12, and 13 for an assessment of the fiducial cross-section.

Efficiency is calculated using the ratio of the truth-level and reconstruction-level distributions of q_T^A. As events in each bin of the truth and reconstruction histograms are not independent, the error is propagated in a binomial manner. The truth and reconstruction distributions of q_T^A is shown in Figure 8.1a and the derived selection efficiency is shown in Figure 8.1b.



(a) (b)

Figure 8.1: Distributions of (a) truth and reconstructed q_T^A *sign* and (b) the calculated efficiency.

8.2 Event Selection Systematics

The systematics surrounding event selection and background removal are collected into a single external group. Some parameters chosen in the background

removal process have a known uncertainty, such as the half-life of the non-prompt contributions. There are four variations: one from the width of the sideband mass windows detailed in Section 8.2.1, two from the non-prompt subtraction working point in Section 8.2.2, and a final variation from the sideband contribution modelling in Section 8.2.3.

8.2.1 Mass Sideband Removal Systematics

The mass sideband windows, 200 MeV wide, are initially given the following bin edges: $2.7-2.9-3.1-3.3-3.5$ GeV. Mass windows delineating signal+background and background-only regions are chosen to be roughly 3σ of the J/ψ mass peak. A variation to reduce the width of the mass windows is used, and the boundaries are changed to $2.74 - 2.92 - 3.1 - 3.28 - 3.46$ GeV. Reducing the width of the boundaries is one half of the variation; however, an opposite variation that increases the bounds cannot be conducted. Processing provided by the `HLT_2mu4_bJpsimumu_noL2` trigger limits the dimuon mass range from $2.6 - 3.5$ GeV. Below the 2.7 GeV analysis bound the trigger distorts the mass distribution and the analysis window already takes the upper bound of the available invariant mass. The signal yields from the likelihood process are shown in the mass column of Table 8.2.

8.2.2 Non-Prompt Removal Systematics

The efficacy of the non-prompt subtraction process has previously been demonstrated in Figure 5.33. For the non-prompt subtraction process, the working point is the half-life of the B-hadron contributions, $\tau_0 = 0.86$ ps. This half-life was derived from a number of trial fits of inclusive J/ψ production [83]. The inclusive J/ψ production range covered a 1.24 ± 0.07 ps effective lifetime range, covering all bins of the non-prompt J/ψ production. Two variations were made in this step, the cutoff value was altered to $\tau_1 = 0.81$ ps and $\tau_2 = 0.91$ ps. As with the mass sideband windows, the full likelihood fit is performed again with these changes in place. The results are collated in columns 'tau1' and 'tau2' of Table 8.2.

8.2.3 Mass Sideband Modelling Systematics

Contributions from the background $\mu\mu$ continuum is assumed to be linear in shape during the subtraction process. As the continuum shape can reasonably be fit with an exponential, this is explored as a variation. Linear sideband modelling produces weights on the events in mass windows that are ± 1. If the background is not linear, the ratio of the contributions of continuum in sidebands and under the peak will not be 1. Instead, the ratio will follow the exponential curve. Figure 8.2 shows a fit of the J/ψ mass window with an exponential continuum and a Voigt distribution for the J/ψ peak.

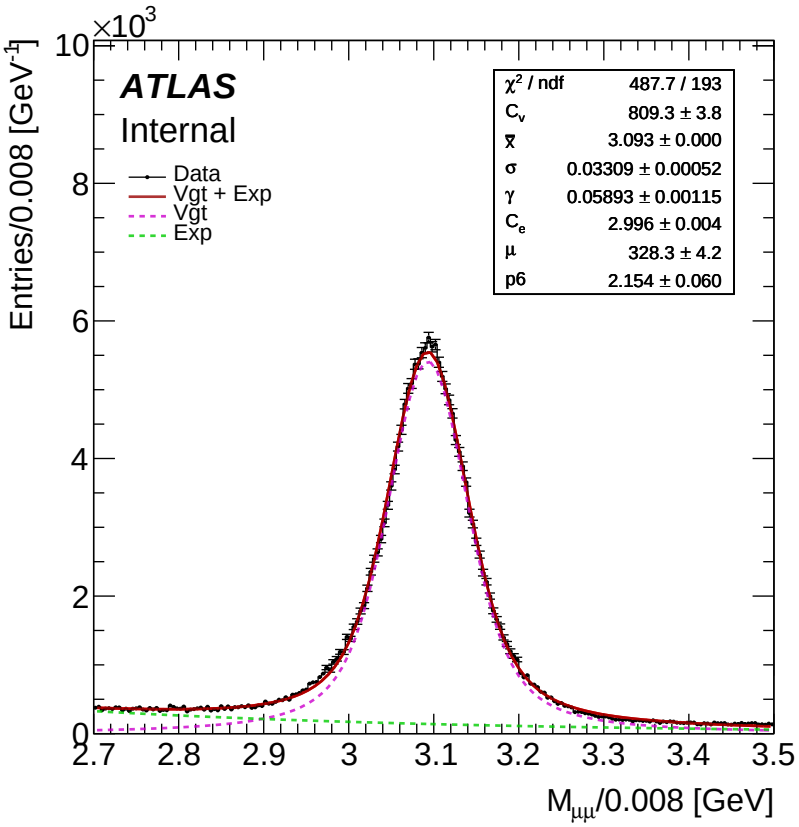
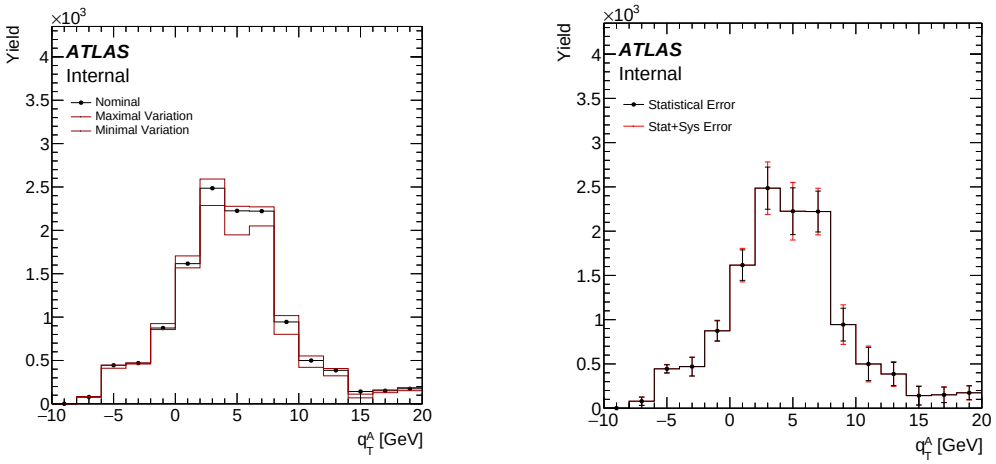


Figure 8.2: Plot of a combined Voigt distribution and declining exponential to match the distribution of the $\mu\mu$ mass in the data sample over the full q_T^A range.



(a) (b)

Figure 8.3: (a) Maximal and minimal yields associated with the event selection systematics and relevant nominal distributions across all q_T^A bins in the analysis and (b) the results with the error bars which show statistical and quadratically combined statistical plus event selection systematic errors.

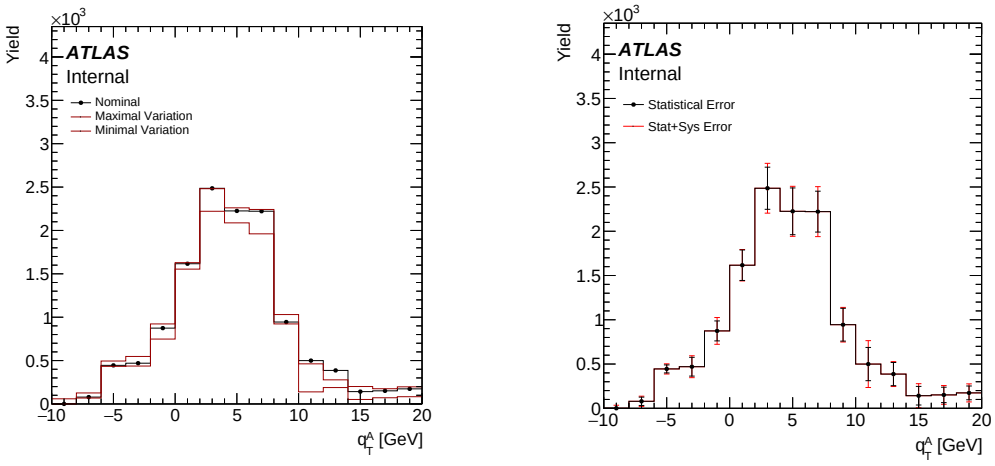
Fitting of the $\mu\mu$ distribution for the correction weight is performed in each q_T^A bin. The ratio of the integrals in the sidebands can be used to deduce the corrected weights. The results of this process are collected in 'yexp' Table 8.2. Additionally, all variations and the errors associated with this external group are shown in Figure 8.3.

8.3 BDT Systematics

The second set of variations associated with an external group are those of the BDT variables. Variations of the input discriminators in the tree construction process are explored in two forms. The first is the addition or removal of discriminators that may be highly correlated with other variables. The second type of variation is the alteration of variables that enter into the tree construction in and out of their absolute value form. When these discriminators are provided to the AdaBoost algorithm in absolute value form, discriminatory power left in that asymmetry could have been folded out.

Index	q_T^A GeV	Yield	mass	tau1	tau2	yexp	Error
1	-10 – -8	0.56	-0.01	-0.01	-0.01	-0.01	0.01
2	-8 – -6	79.02	0.36	-3.08	5.75	-0.07	5.10
3	-6 – -4	444.78	-34.27	-15.27	-0.64	1.83	20.84
4	-4 – -2	469.76	-8.69	-8.90	1.74	-10.34	6.98
5	-2 – 0	874.11	27.74	-17.60	51.16	19.73	39.70
6	0 – 2	1616.29	-5.81	-48.58	89.24	-5.61	79.57
7	2 – 4	2485.81	-58.08	-199.16	105.87	-44.93	176.11
8	4 – 6	2225.16	-89.25	-277.24	52.23	-103.44	190.22
9	6 – 8	2221.57	-35.77	-170.60	49.28	-81.87	126.95
10	8 – 10	944.30	-6.29	-142.51	73.17	-39.82	124.52
11	10 – 12	499.58	-78.67	-61.69	52.61	-30.13	75.79
12	12 – 14	386.50	-63.23	-54.91	20.93	-32.69	48.59
13	14 – 16	142.20	-72.70	-30.06	-65.33	-57.40	24.62
14	16 – 18	151.27	-16.10	-4.25	9.58	-22.55	18.55
15	18 – 20	174.00	-22.02	-17.80	13.26	-10.79	20.37

Table 8.2: Yields of the external systematic baseline and the change in yield for each subtraction systematic in each q_T^A bin, and a total error from the combination of all four systematics.



(a) (b)

Figure 8.4: (a) Maximal and minimal yields associated with the event selection systematics and relevant nominal distributions across all q_T^A bins in the analysis and (b) the results with the error bars which show statistical and quadratically combined statistical plus BDT systematic errors.

The angular variables, not necessarily the discriminators, are $\Delta\phi$, Δy, ϕ, $\cos\theta$. As discriminators these are $|\Delta\phi|$, Δy, $|\phi|$, and $|\cos\theta|$ The large number of variation included at this stage does not necessarily introduce an increased statistical error. As the error is assumed to be distributed uniformly, variations attempted that do not extend range of variations do not contribute. The input discriminating variables for the BDT were previously given in Table 6.2, and the set of variations is given in Tables 8.3. Resultant differences in the extracted yield between the nominal and the variations are given in Tables 8.4. The maximal and minimal variations and the errors associated with this external group are shown in Figure 8.4.

Variation name	Discriminator set
Nominal	$\|\Delta\phi\|, \Delta y, \lambda, q_T^2, \|\phi\|, \|\cos\theta\|, q_x^\psi, q_y^\psi, q_x^\gamma, q_y^\gamma, q_x^\Sigma, q_y^\Sigma$
aaf	$\|\Delta\phi\|, \Delta y, \lambda, q_T^2, \|\phi\|, \cos\theta, q_x^\psi, q_y^\psi, q_x^\gamma, q_y^\gamma, q_x^\Sigma, q_y^\Sigma$
afa	$\|\Delta\phi\|, \Delta y, \lambda, q_T^2, \phi, \|\cos\theta\|, q_x^\psi, q_y^\psi, q_x^\gamma, q_y^\gamma, q_x^\Sigma, q_y^\Sigma$
faa	$\Delta\phi, \Delta y, \lambda, q_T^2, \|\phi\|, \|\cos\theta\|, q_x^\psi, q_y^\psi, q_x^\gamma, q_y^\gamma, q_x^\Sigma, q_y^\Sigma$
noqt2	$\|\Delta\phi\|, \Delta y, \lambda, \|\phi\|, \|\cos\theta\|, q_x^\psi, q_y^\psi, q_x^\gamma, q_y^\gamma, q_x^\Sigma, q_y^\Sigma$
aaDPDY	$\|\Delta\phi\|, \|\Delta y\|, \lambda, q_T^2, \|\phi\|, \|\cos\theta\|, q_x^\psi, q_y^\psi, q_x^\gamma, q_y^\gamma, q_x^\Sigma, q_y^\Sigma$
faDPDY	$\Delta\phi, \|\Delta y\|, \lambda, q_T^2, \|\phi\|, \|\cos\theta\|, q_x^\psi, q_y^\psi, q_x^\gamma, q_y^\gamma, q_x^\Sigma, q_y^\Sigma$

Table 8.3: Discriminator set variations used to train the BDT.

8.4 Scale Factor and Reconstruction Systematics

MC simulations are estimations of true physical processes. In this analysis these are used to reproduce the $J/\psi + \gamma$ subprocess, inclusive J/ψ process, and J/ψ production from non-prompt B-hadrons. By nature these will have a limited fidelity so differences between data and MC simulations are corrected with reweighting. The weights used in this process are commonly known as scale factors (SF) and in the ATLAS Collaboration these are produced by a multitude of Combined Performance (CP) groups, each for a specific purpose. There are multiple SF's that are utilised in this analysis and systematic uncertainty associated with each is explored herein. Sections 8.4.1, 8.4.2, 8.4.3 and 8.4.4

Index	q_T^A GeV	Yield	aaf	afa	faa	noqt2	aaDPDY	faDPDY	Error
1	-10 – -8	0.56	52.21	-0.01	58.73	52.94	-0.01	7.03	33.91
2	-8 – -6	79.02	12.29	27.89	47.87	-10.84	6.82	14.39	33.90
3	-6 – -4	444.78	-11.42	9.52	-9.48	10.84	47.38	50.16	35.56
4	-4 – -2	469.76	72.01	-32.76	-9.37	77.73	12.35	0.73	63.79
5	-2 – 0	874.11	-71.94	-101.64	47.28	-126.16	48.32	15.95	100.73
6	0 – 2	1616.29	-44.25	-62.07	-33.36	-2.69	12.55	-5.77	43.08
7	2 – 4	2485.81	-264.75	-70.73	-56.10	-4.64	-67.15	-37.56	150.17
8	4 – 6	2225.16	-119.71	3.70	-76.13	-11.64	35.55	-137.73	100.04
9	6 – 8	2221.57	-260.65	-201.53	-58.84	-51.86	-69.26	20.38	162.25
10	8 – 10	944.30	45.47	56.50	15.85	86.22	-21.78	57.76	62.36
11	10 – 12	499.58	-360.39	-74.47	-236.72	-37.29	-217.90	-200.81	186.54
12	12 – 14	386.50	-126.36	-151.49	-108.63	-195.13	-161.09	-197.51	51.31
13	14 – 16	142.20	-54.75	58.78	-90.43	-58.64	-72.45	-24.99	86.15
14	16 – 18	151.27	-47.15	1.26	2.72	-3.45	24.20	-79.90	60.11
15	18 – 20	174.00	-90.48	-30.36	23.46	-34.89	9.36	-54.23	65.78

Table 8.4: Yields of the external systematic baseline and the change in yield for each BDT systematic in each q_T^A bin, and a total error from the combination of all BDT systematics.

cover the muon reconstruction, pileup reweighting, muon trigger, and photon reconstruction, respectively.

8.4.1 Muon Reconstruction

SF's associated with the muon reconstruction process were explored in a previous J/ψ cross-section measurement in [83]. This was done by examing the scale factor maps provided by the muon combined-performance group. The SF's here were found to be small, well below a 1% error. Further study was not continued in this direction, and a 1% variation was conservatively selected as an internal variation in `TRExFitter`.

8.4.2 Pileup Reweighting Systematics

The data sample utilised in this analysis consisted of ATLAS Run 2 2015 data which, advantageously, exhibited a relatively low level of pileup. Trigger studies with respect to pileup were conducted in [83]. Due to the minimal effects

pileup corrections were small, and as a consequence further systematic study was not considered necessary after the pileup correction was applied. This analysis follows the same conclusion, no systematic for pileup is inserted into the likelihood.

8.4.3 Muon Trigger Scale Factors

An SF weighting must be made for the muon trigger. Differences between MC and Data for the muon trigger are dependent on multiple variables. There are four separate determining angular and kinematic values, the p_T of each muon, rapidity η, charge, and angular distance between two muon tracks ΔR. SF maps for the `HLT_2mu4_bJpsimumu_noL2` trigger are provided by the ATLAS BPhysics and Light States Trigger subgroup.

These scale factor maps have an associated statistical and systematic uncertainty. The process for deriving the systematic uncertainty is prescribed by the ATLAS BLS Trigger Group. Along with the provided SF maps, 100 variations of the SF are provided, based on a toy model. So, 100 BDT score distributions are created, each with a SF weight from the toy model. The standard deviation of the score in each bin is used to calculate the magnitude of the systematic uncertainty. For demonstratory purposes, the mean muon trigger SF are shown for each of the variables it depends on, and q_T^A, in Figure 8.5. This is added in quadrature with the statistical error associated with the SF, and provided as a variation.

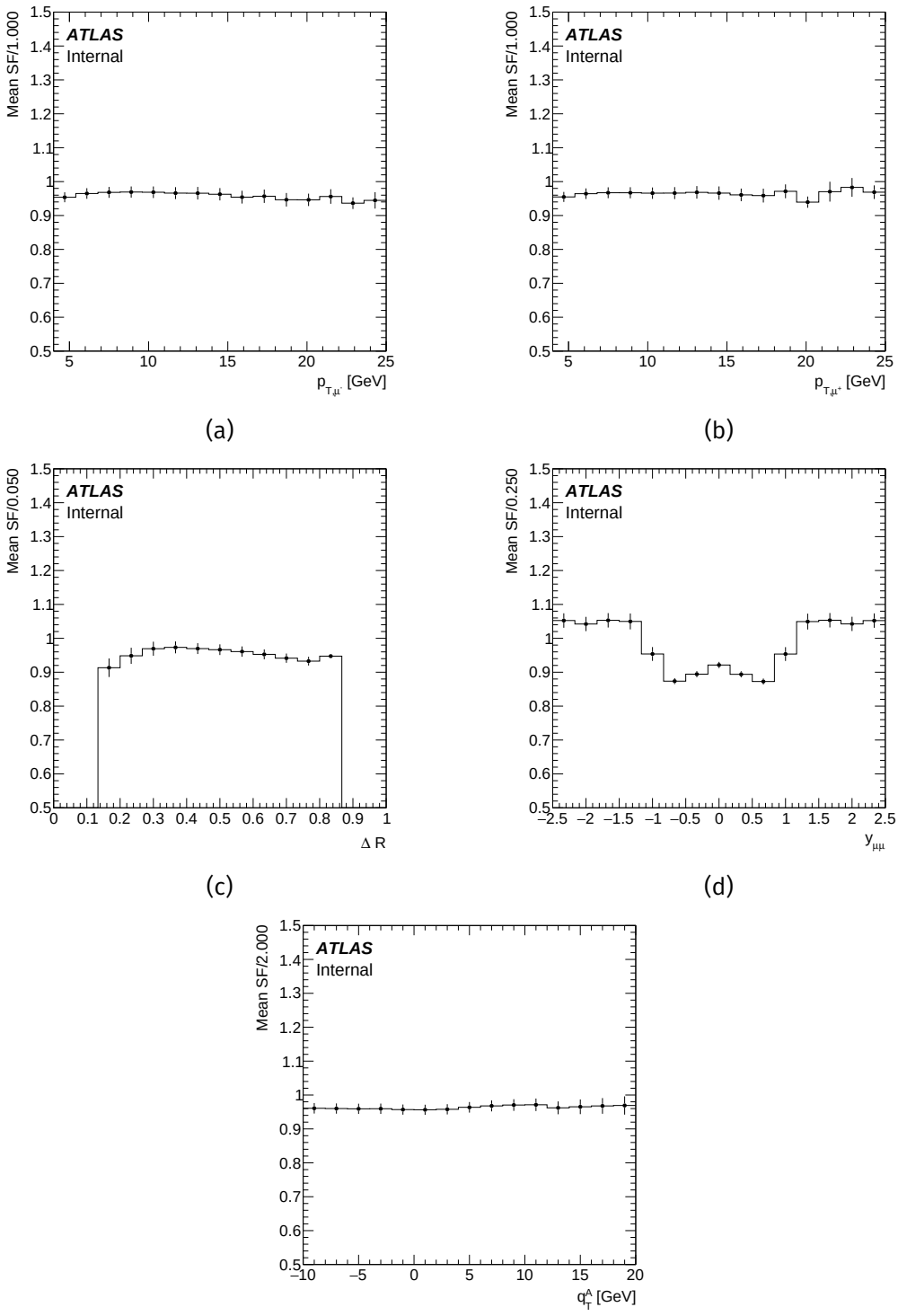
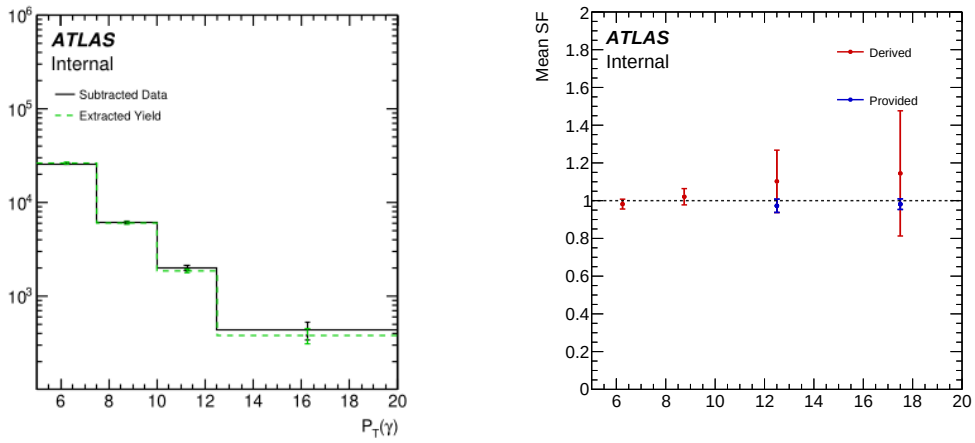


Figure 8.5: Mean muon trigger scale factor values with respect to (a) $p_{T,\mu}^{+}$, (b) $p_{T,\mu}^{-}$, (c) ΔR, (d) y, and (e) q_T^A.

8.4.4 Photon ID Scale Factors and Expansion

Much like the muon trigger SF's, the ATLAS EGamma Group provides SF maps for photon reconstruction. These maps, based on $p_T(\gamma)$ and $\eta(\gamma)$, provide SF's to again be applied to the BDT score distribution. This is applied to only the BDT response for two reasons. Firstly, reapplication to a single variable before the BDT step introduces a possible compounding of uncertainty, not allowing the BDT systematics to be examined in isolation. Secondly, the only variable that directly enters into the final systematics process is the BDT response, hence it must be used as the medium for the correction. The EGamma group provides an uncertainty for these values that is used to construct a variation for `TRExFitter`. The SF value combined with its uncertainty is used to create a 1σ 'upper' variation that is mirrored around the nominal value.

SF values for $p_T(\gamma) < 10$ GeV are not provided by the EGamma group, and approximately $\sim 65\%$ of photons in this analysis fall below this threshold. To provide SF's in this region a `TRExFitter` validation region is used. The framework allows fits to be performed in one variable, and the results to be projected out to other variables, allowing the fit in BDT score to reconstruct the $p_T(\gamma)$ distribution to be recreated from the data. A pilot fitting procedure must be performed to extract this distribution, so all other internal systematics are removed from the likelihood to avoid double-counting and second-order effects. The fitted $p_T(\gamma)$



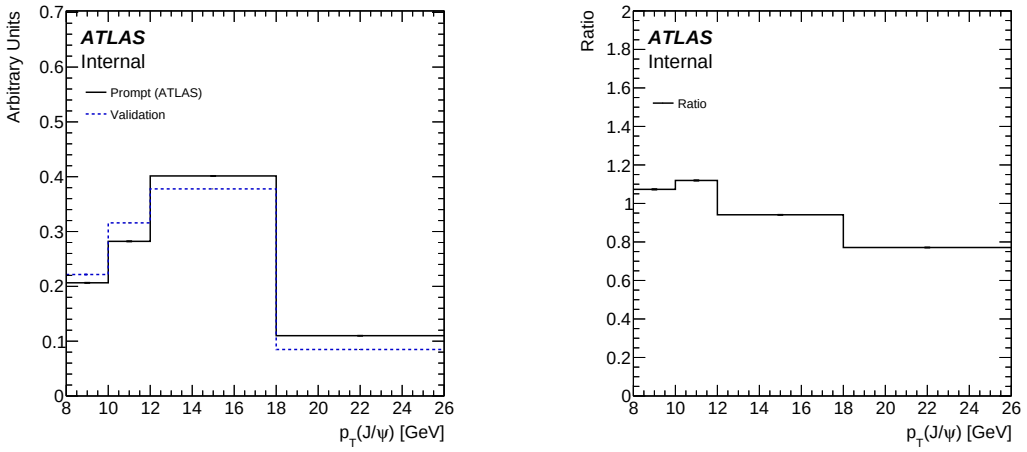
(a) (b)

Figure 8.6: Closure tests for *data*. (a) $p_T(\gamma)$, and (b) provided and derived SF's.

distribution is summed over the 15 q_T^A analysis bins and is then compared with that of the *data* sample distribution, shown in Figure 8.6a. Taking the ratio of these two distributions provides a correction factor, Figure 8.6b, encapsulating the difference between MC and data. This correction factor is used as a scale factor for regions below 10 GeV. The error here is also used to supplement the SF's and create the upper variation which is fed into the final fitting procedure.

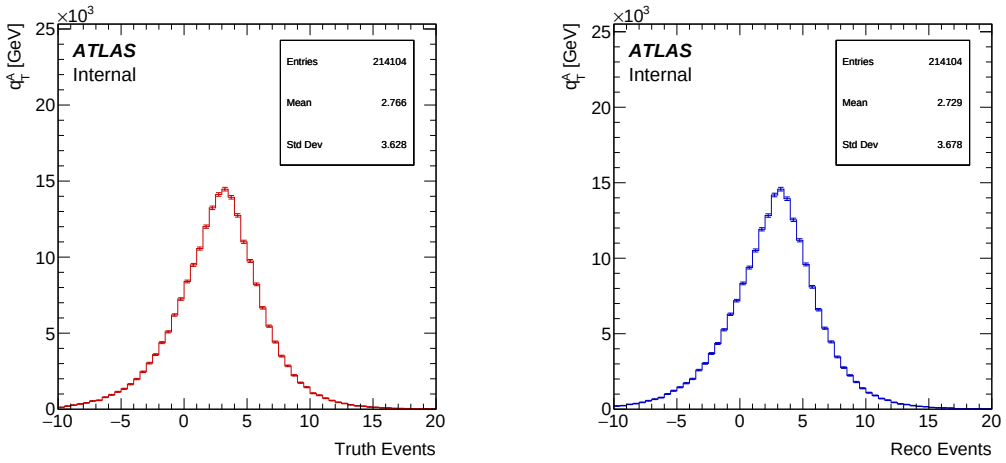
8.5 J/ψ p_T Validation

To provide some quantifiable uncertainty for the modelling of the J/ψ distribution in the uncorrelated MC sample, *bckg*, a data driven test is performed. Using the same pilot fit as in Section 8.4.4, the distribution of $p_T(J/\psi)$ is evaluated. In other words, differences between *bckg* and the prompt-inclusive $p_T(J/\psi)$ distribution of data collected at ATLAS should be quantified in the likelihood of the final fit. A previous analysis [83] measured the prompt inclusive spectrum and this is compared with the projected $p_T(J/\psi)$ distribution summed over all 15 q_T^A bins from the validation procedure in Figure 8.7. Taking the ratio between the two distributions in Figure 8.7a gives a factor that can be included in the likelihood of the final fit.



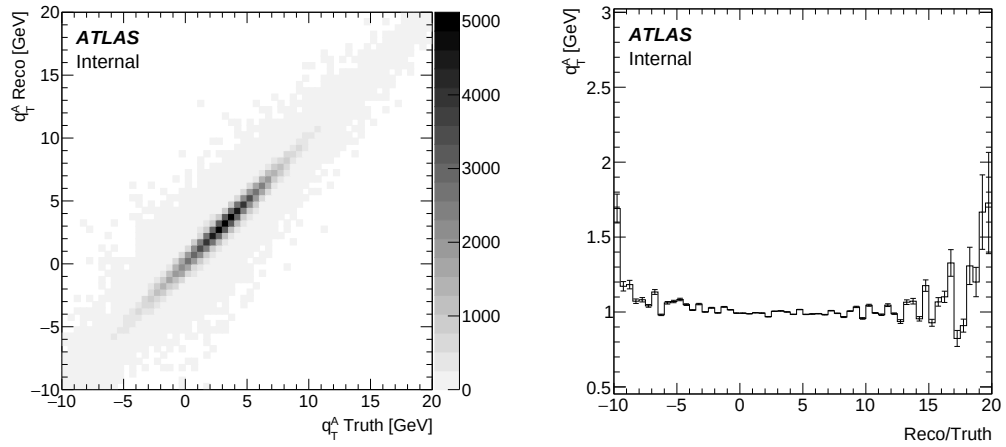
(a) (b)

Figure 8.7: Closure tests for *bckg*. (a) dimuon p_T, and (b) ratio of validation to prompt.



(a) (b)

Figure 8.8: (a) Truth q_T^A events passing selection and reconstruction, (b) reconstructed q_T^A distribution of the same events



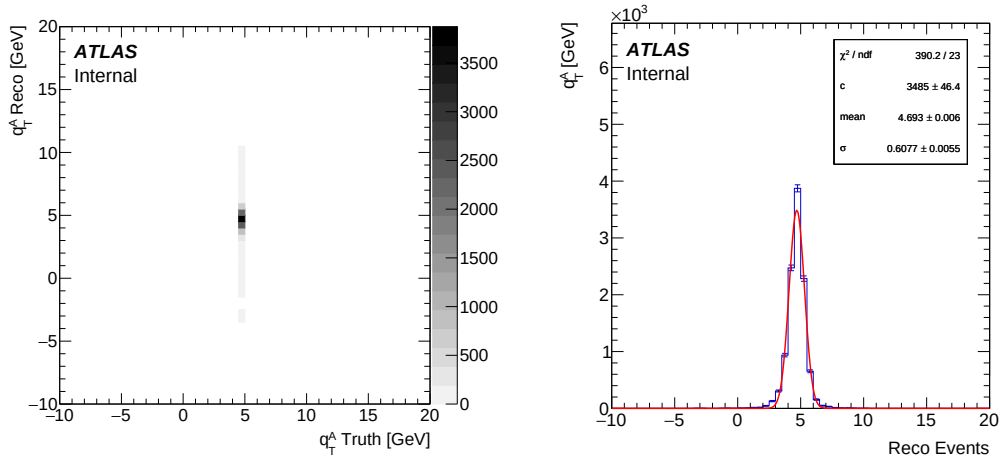
(a) (b)

Figure 8.9: (a) Derived migration matrix, and (b) ratio of projected migration matrix distributions.

8.6 Bin Migration and Resolution Study

As a supplementary verification of the reconstruction process, the effects of bin migration and resolutions are examined. The study herein is performed

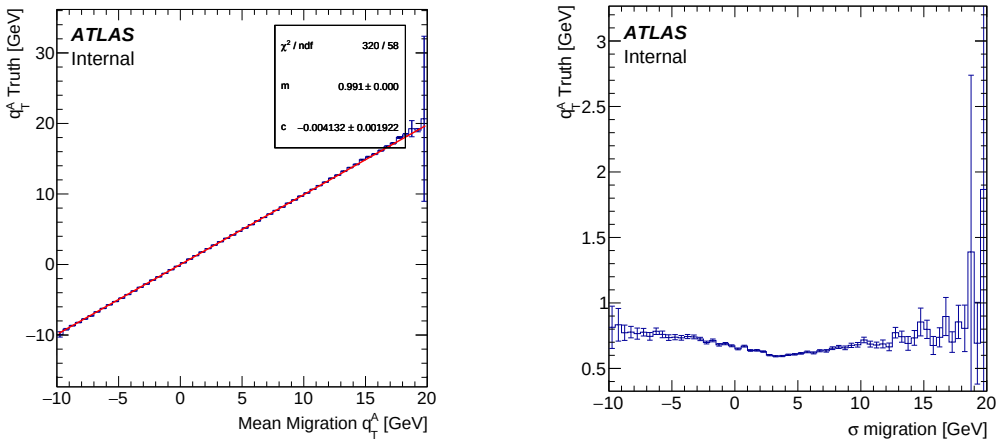
using the *sign* sample. Effects of a bin migration are mitigated by the selection efficiency correction process though these changes should be understood. As q_T^A is the main medium for parameterising the TMD, the migration of this distribution is explored. A migration matrix is constructed from the 2D histogram the truth-q_T^A and reco-q_T^A, these are shown in Figures 8.8-8.9. The ratio of the projections on the truth and reconstruction axes emphasises the migration effect. The projection is limited to 0.5 GeV windows of truth-q_T^A in Figure 8.10 allowing the migration effects to be examined in minutiae across the full range. Expecting the resolution effects to be normally distributed and fitting the



(a) (b)

Figure 8.10: (a) Migration of q_T^A at a single point in the truth distribution and (b) the same reconstructed q_T^A projected from a single point in truth q_T^A.

projection accordingly, the evolution of the parameters can be seen in Figure 8.11. Any non-linear effects of the migration would be visible in some bias of the fit parameters across the q_T^A range. When limited statistics do not interfere with fit quality, the width (σ) of the projected distributions vary in the range of $0.6 - 0.8$ GeV. With the nominal width of the q_T^A distribution at $\sigma(q_T^A) \simeq 3.68$ GeV, the effect is approximately $\sqrt{3.68^2 + 0.8^2} = 3.76$ GeV. This amounts to a 2.3% variation. The mean of the projected variation evolves 1:1 with the centre of the truth window and no bias is present in the range. Resolution effects here are corrected during the efficiency stage and no further systematic study is deemed necessary.



(a) (b)

Figure 8.11: Plots of (a) mean of fitted gaussian migration distribution and (b) resolution of migration.

8.7 Aggregation of internal systematics.

The systematics which enter into the likelihood fit, complete with the yields and statistical error are presented in Table 8.5. Similarly, the pulls of all the internal systematics are shown in Figures 8.12-8.15. In these Figures, the green shaded area represents the space where the pulls are between 1 and -1, where as the yellow represents the areas where pulls are greater in magnitude than 1, but less than 2. These are rule-of-thumb guidelines for goodness of fit setup.

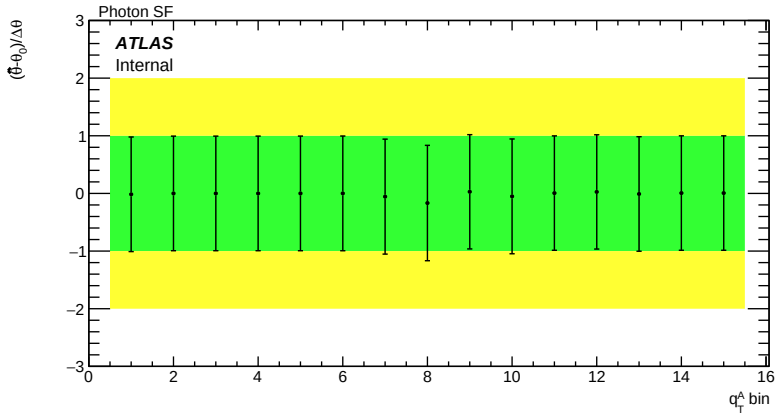


Figure 8.12: Pulls of the internal systematics in the likelihood fits across all q_T^A bins for Photon ID SF.

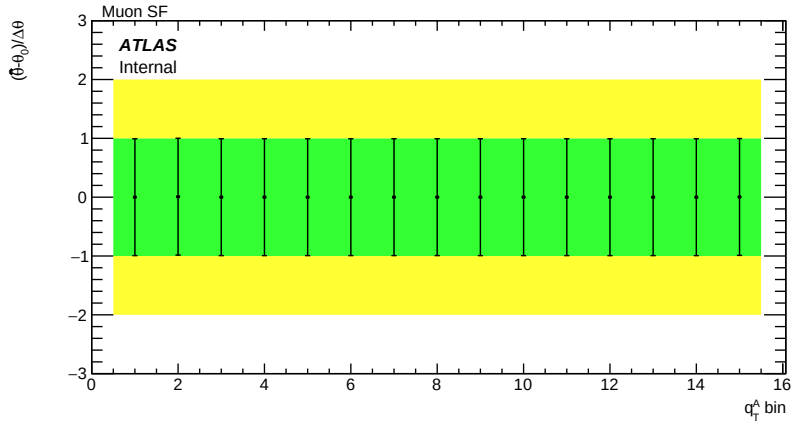


Figure 8.13: Pulls of the internal systematics in the likelihood fits across all q_T^A bins for Muon Trigger SF.

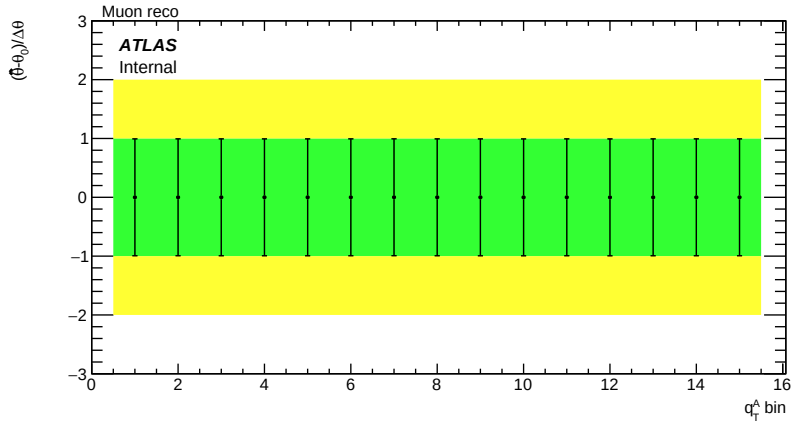


Figure 8.14: Pulls of the internal systematics in the likelihood fits across all q_T^A bins for Muon reconstruction systematics.

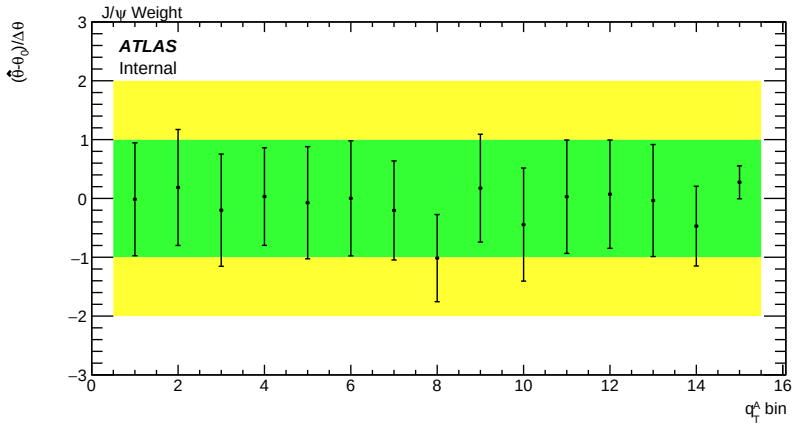


Figure 8.15: Pulls of the internal systematics in the likelihood fits across all q_T^A bins for $p_T(J/\psi)$ closure systematic.

Index	q_T^A GeV	Yield	Statistical Error	Internal Systematic Error	Combined Error
1	-10 – -8	1	1	1	1
2	-8 – -6	78	20	44	48
3	-6 – -4	448	34	36	49
4	-4 – -2	464	51	118	128
5	-2 – 0	900	60	117	131
6	0 – 2	1642	85	245	259
7	2 – 4	2518	115	303	324
8	4 – 6	2643	126	430	448
9	6 – 8	2136	112	288	308
10	8 – 10	981	86	218	234
11	10 – 12	465	78	265	276
12	12 – 14	370	58	124	137
13	14 – 16	88	44	102	111
14	16 – 18	32	33	111	115
15	18 – 20	253	35	116	116

Table 8.5: Yields, errors on extracted signal with validation systematics included, and total internal systematic errors in q_T^A.

9. Interpretation of Results

With the signal contribution in each q_T^A bin extracted, the differential cross-section with respect to q_T^A is reconstructed. From this distribution, after applying efficiency and luminosity corrections, the TMD can be parameterised. Corrections are detailed in equation 9.1:

$$B(J/\psi \to \mu^+\mu^-)\frac{d\sigma_{\text{eff}}(pp \to J/\psi + \gamma + X)}{dq_T^A} = \frac{N_{\text{fit}}}{L \cdot \epsilon \cdot \Delta q_T^A} \tag{9.1}$$

Here L is the luminosity of the *data* sample, computed to be $2.6\ \text{fb}^{-1}$ with an uncertainty of 1.13% [104]. Δq_T^A is the width of the q_T^A bins, 2 GeV, and ϵ represents the selection efficiency in this bin. The extracted signal yield is given as N_{fit}.

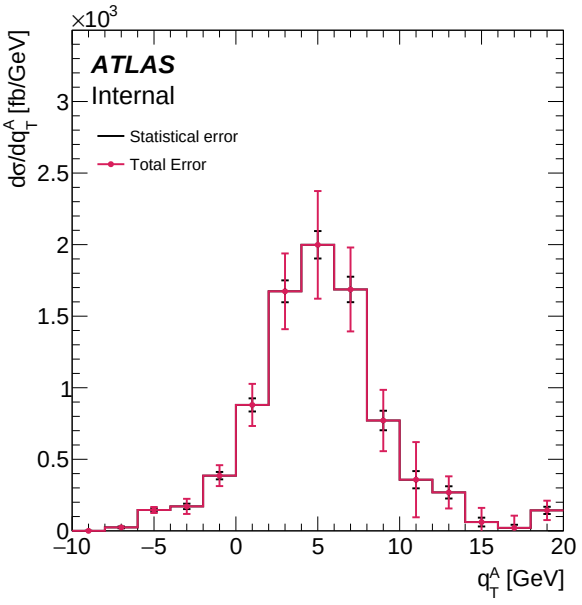


Figure 9.1: Extracted TMD yields as function of q_T^A, with q_T^B cut present.

Applying these corrections and combining the uncertainty of the yields from the

profile likelihood fits with the uncertainty from the external group systematics, the final q_T^A distribution is recovered. The yields and cross-sections are shown in Table 9.1 and Figure 9.1.

q_T^A	Yield	Statistical Error	Internal Systematic Error	External Systematic Error	Efficiency (%)	Total error	Cross-section fb/GeV
1	1	1	1	34	67 ± 16	34	0 ± 10
2	78	20	44	34	62 ± 10	59	24 ± 19
3	448	34	36	41	60 ± 7	64	146 ± 21
4	464	51	118	64	53 ± 4	143	171 ± 53
5	900	60	117	108	45 ± 3	170	386 ± 73
6	1642	85	245	90	36 ± 2	275	880 ± 147
7	2518	115	303	231	29 ± 1	398	1674 ± 265
8	2643	126	430	215	26 ± 1	497	1999 ± 376
9	2136	112	288	206	25 ± 2	371	1687 ± 293
10	981	86	218	139	25 ± 3	272	771 ± 214
11	465	78	265	201	25 ± 5	342	357 ± 263
12	370	58	124	71	27 ± 8	154	269 ± 112
13	88	44	102	90	28 ± 13	143	61 ± 99
14	32	33	111	63	30 ± 21	131	21 ± 85
15	253	35	116	116	34 ± 38	116	143 ± 66

Table 9.1: Fully corrected yields, efficiencies, and errors for the cross-section in each q_T^A bin. Note efficiency values here are rounded before calculations are performed.

Studies conducted in literature commonly model the TMD f_1^g as a Gaussian probability distribution multiplying the original collinear PDF $G(x)$. Measuring the cross-section is equivalent to measuring the convolution of two f_1^g. Equation 3.31 is given again for reference:

$$\mathcal{C}\left[f_1^g f_1^g\right] = \frac{G(x_1)G(x_2)}{2\langle k_T^2\rangle\sqrt{2\pi}} \exp\Big(-\frac{\bar{q_T}^2}{2(2\langle k_T^2\rangle)}\Big). \tag{3.31}$$

Any resultant q_T^A distribution that is reconstructed from the likelihood fit should logically be fitted to a Gaussian. However, this would not account for the hard initial state radiation (ISR) contributions that will still be present in the distribution. To resolve this, with the ISR assumed to be distributed around the

tails of q_T^A, the sum of two Gaussians is used to extract $\sqrt{\langle k_T^2 \rangle}$:

$$g(x) = c_1 \cdot e^{\frac{(m-x)^2}{2\cdot\sigma_1}} + c_2 \cdot e^{\frac{(m-x)^2}{2\cdot\sigma_2}}. \tag{9.2}$$

The two Gaussians here are constrained to have a common centre, and cover different ranges in σ. Of the two sets of values that are extracted only the width of the narrower Gaussian will be used to parameterise the TMD. Here, the wider Gaussian is used to collect the ISR effect, though reduced statistics in the periphery regions of q_T^A may reduce how accurately the ISR effects can be assessed. The results of the two-Gaussian fit is shown in Figure 9.2. Using σ_1,

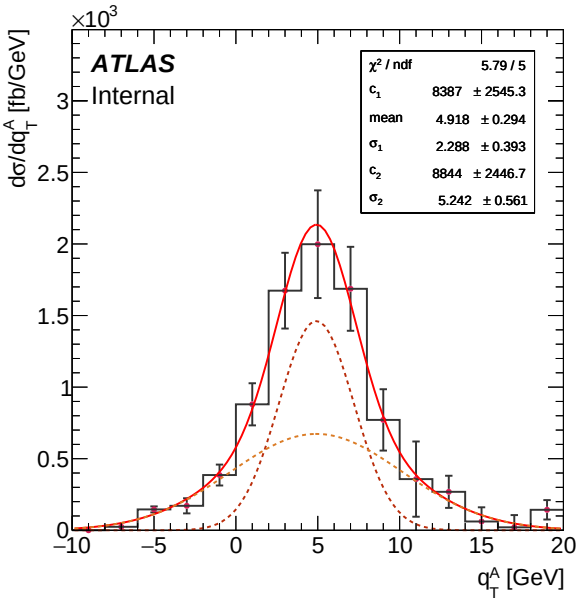


Figure 9.2: Fits of the q_T^A distribution to parameterise the TMD, with q_T^B cut present. The solid red curve represents the double Gaussian, the two dashed curves are the individual Gaussian components.

which will represent the convolution of two f_1^g's, one per proton, the width of the intrinsic gluon distribution is:

$$\sqrt{\langle k_T^2 \rangle} = \frac{\sigma_1}{\sqrt{2}} = 1.61 \pm 0.26 \text{ GeV}. \tag{9.3}$$

Additionally, the fiducial cross-section of the $J/\psi + \gamma$ SPS subprocess can be extracted. Efficiency corrections representing the fiducial cuts are applied to the same central values for raw event yields obtained in the previous subsection are used to determine the fiducial cross-section. This correction here will differ from

that used in the TMD measurement as a consequence of the change $|q_T^B| < 6$ GeV cut, determined to have an efficiency of 76.67 ± 0.15%. The integral of Equation 9.1, with $\epsilon \rightarrow \epsilon_{\text{fid}}$ gives a fiducial cross-section of 22.8 ± 1.9 pb. The fiducial cuts applied are equivalent to $y < 2.4$ and $p_T > 4$ GeV on the muons, and $\eta < 2.5$ and $p_T > 5$ GeV on the photon.

10. Summary

To conclude, inclusive associated production of a J/ψ meson and a photon was used to measure the transverse-momentum-dependent parton distribution function (TMD) of unpolarised gluons in an unpolarised proton, $f_1^g(k_T)$.

This measurement used data collected by the ATLAS experiment at the centre-of-momentum energy of 13 TeV during Run 2 of the LHC, in 2015. Events are collected with a dimuon trigger with threshold $p_T(\mu) > 4$ GeV and a series of selection cuts to isolate a high signal-to-background region. The *data* sample corresponds to an effective integrated luminosity of 2.6 fb^{-1}.

Following the convention in literature that the transverse component of $f_1^g(k_T)$ is normally distributed, the mean transverse momentum of the gluons was found to be $\langle k_T^2 \rangle^{\frac{1}{2}} = 1.61 \pm 0.28$ GeV.

Supplementarily, the cross-section of the single-parton scattering process ($p + p \rightarrow J/\psi + \gamma + X$), corresponding to the fiducial region of this analysis, is measured These fiducial cuts correspond to $p_T(\mu) > 4$ GeV, $p_T(\gamma) > 5$ GeV and $|\eta| < 2.4$. This cross-section is measured to be 22.8 ± 1.9 pb.

References

[1] M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory*. Advanced book classics. Avalon Publishing, 1995. ISBN: 9780201503975. URL: https://books.google.co.uk/books?id=_H-oPv1raioC.

[2] W. N. Cottingham and D. A. Greenwood. *An Introduction to the Standard Model of Particle Physics*. 2nd ed. Cambridge University Press, 2007.

[3] Franz Mandl and Graham Shaw. *Quantum Field Theory*. 1985.

[4] Anthony Zee. *Group Theory in a Nutshell for Physicists*. USA: Princeton University Press, Mar. 2016. ISBN: 978-1-4008-8118-5, 978-0-691-16269-0.

[5] Ian Aitchison and Anthony Hey. *Gauge theories in particle physics, 40th Anniversary Edition: A practical introduction, volume 1: From relativistic quantum mechanics to QED, Fifth Edition*. CRC Press, 2024.

[6] Ian Aitchison and Anthony Hey. *Gauge theories in particle physics, 40th Anniversary Edition: A practical introduction, volume 2: Non-Abelian Gauge Theories: QCD and the electroweak theroy, Fifth Edition*. CRC Press, 2024.

[7] Michael Creutz. *Quarks, Gluons and Lattices*. Oxford University Press, 1983. ISBN: 978-1-009-29039-5, 978-1-009-29038-8, 978-1-009-29037-1, 978-0-521-31535-7. DOI: 10.1017/9781009290395.

[8] G. C. Donald et al. "Precision tests of the J/ψ from full lattice QCD: Mass, leptonic width, and radiative decay rate to η". In: *Physical Review D* 86.9 (Nov. 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.86.094501. URL: http://dx.doi.org/10.1103/PhysRevD.86.094501.

[9] Constantia Alexandrou et al. “Light-Cone Parton Distribution Functions from Lattice QCD”. In: *Phys. Rev. Lett.* 121 (11 Sept. 2018), p. 112001. DOI: 10.1103/PhysRevLett.121.112001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.112001.

[10] John C. Collins, Davison E. Soper, and George Sterman. *Factorization of Hard Processes in QCD*. 2004. arXiv: hep-ph/0409313 [hep-ph]. URL: https://arxiv.org/abs/hep-ph/0409313.

[11] F. Halzen and Alan D. Martin. *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*. 1984. Chap. 10. ISBN: 978-0-471-88741-6.

[12] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. Chap. 8. ISBN: 978-1-107-03426-6. DOI: 10.1017/CBO9781139525367.

[13] Tung-Mow Yan and Sidney D. Drell. “The parton model and its applications”. In: *International Journal of Modern Physics A* 29.30 (Dec. 2014), p. 1430071. ISSN: 1793-656X. DOI: 10.1142/s0217751x14300713. URL: http://dx.doi.org/10.1142/S0217751X14300713.

[14] Wolfgang Kurt Hermann Panofsky. “Low q electrodynamics, elastic and inelastic electron (and muon) scattering”. In: *14th International Conference on High Energy Physics*. 1968, pp. 23–42.

[15] E. D. Bloom et al. “High-Energy Inelastic $e-p$ Scattering at 6° and 10°”. In: *Phys. Rev. Lett.* 23 (16 Oct. 1969), pp. 930–934. DOI: 10.1103/PhysRevLett.23.930. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.930.

[16] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. New York, NY: Wiley, 2008. URL: https://cds.cern.ch/record/111880.

[17] Sidney D. Drell and Tung-Mow Yan. “Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies”. In: *Phys. Rev. Lett.* 25 (5 Aug. 1970), pp. 316–320. DOI: 10.1103/PhysRevLett.25.316. URL: https://link.aps.org/doi/10.1103/PhysRevLett.25.316.

[18] V. M. Braun. *Lecture notes on Quantum Chromodynamics*. Feb. 2023. URL: https://homepages.uni-regensburg.de/~scj58118/.

[19] Davison E. Soper. “Parton distribution functions”. In: *Nuclear Physics B - Proceedings Supplements* 53.1–3 (Feb. 1997), pp. 69–80. ISSN: 0920-5632. DOI: 10.1016/s0920-5632(96)00600-7. URL: http://dx.doi.org/10.1016/S0920-5632(96)00600-7.

[20] Renaud Boussarie et al. *TMD Handbook*. 2023. arXiv: 2304.03302 [hep-ph]. URL: https://arxiv.org/abs/2304.03302.

[21] R. Angeles-Martinez et al. “Transverse Momentum Dependent (TMD) parton distribution functions: status and prospects”. In: *Acta Phys. Polon. B* 46.12 (2015), pp. 2501–2534. DOI: 10.5506/APhysPolB.46.2501. arXiv: 1507.05267 [hep-ph].

[22] A. Airapetian et al. “Observation of the Naive-T-Odd Sivers Effect in Deep-Inelastic Scattering”. In: *Phys. Rev. Lett.* 103 (15 Oct. 2009), p. 152002. DOI: 10.1103/PhysRevLett.103.152002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.103.152002.

[23] Dennis Sivers. “Single-spin production asymmetries from the hard scattering of pointlike constituents”. In: *Phys. Rev. D* 41 (1 Jan. 1990), pp. 83–90. DOI: 10.1103/PhysRevD.41.83. URL: https://link.aps.org/doi/10.1103/PhysRevD.41.83.

[24] Cristian Pisano et al. *A first determination of the unpolarized quark TMDs from a global analysis*. 2018. arXiv: 1801.08443 [hep-ph]. URL: https://arxiv.org/abs/1801.08443.

[25] John Collins. *Foundations of Perturbative QCD*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2011.

[26] Wilco J. den Dunnen et al. “Accessing the Transverse Dynamics and Polarization of Gluons inside the Proton at the LHC”. In: *Physical Review Letters* 112.21 (May 2014). Comment: 5 pages, 3 figures; final version to appear in Physical Review Letters, p. 212001. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.112.212001. arXiv: 1401.7611 [hep-ex, physics:hep-ph]. (Visited on 08/23/2023).

[27] Daniël Boer et al. “Linearly Polarized Gluons and the Higgs Transverse Momentum Distribution”. In: *Physical Review Letters* 108.3 (Jan. 2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.108.032002. URL: http://dx.doi.org/10.1103/PhysRevLett.108.032002.

[28] P. J. Mulders and J. Rodrigues. “Transverse momentum dependence in gluon distribution and fragmentation functions”. In: *Phys. Rev. D* 63 (9 Apr. 2001), p. 094021. DOI: 10.1103/PhysRevD.63.094021. URL: https://link.aps.org/doi/10.1103/PhysRevD.63.094021.

[29] P. Schweitzer, T. Teckentrup, and A. Metz. “Intrinsic transverse parton momenta in deeply inelastic reactions”. In: *Physical Review D* 81.9 (May 2010). ISSN: 1550-2368. DOI: 10.1103/physrevd.81.094019. URL: http://dx.doi.org/10.1103/PhysRevD.81.094019.

[30] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024.

[31] Lyndon R Evans and Philip Bryant. “LHC Machine”. In: *JINST* 3 (2008). This report is an abridged version of the LHC Design Report (CERN-2004-003), S08001. DOI: 10.1088/1748-0221/3/08/S08001. URL: https://cds.cern.ch/record/1129806.

[32] D. Lincoln. *The Quantum Frontier: The Large Hadron Collider*. The Quantum Frontier. Johns Hopkins University Press, 2009. ISBN: 9780801896446. URL: https://books.google.co.uk/books?id=M2m-Knuf_rAC.

[33] Karlheinz Schindl. “The injector chain for the LHC”. In: (1999). URL: https://cds.cern.ch/record/397574.

[34] Maurizio Vretenar et al. *Linac4 design report*. Vol. 6. CERN Yellow Reports: Monographs. Geneva: CERN, 2020. DOI: 10.23731/CYRM-2020-006. URL: https://cds.cern.ch/record/2736208.

[35] Vincenzo Forte. “Performance of the CERN PSB at 160 MeV with H^- charge exchange injection”. Presented 03 Jun 2016. U. Blaise Pascal, Clermont-Ferrand, 2016. URL: https://cds.cern.ch/record/2194937.

[36] Michael Benedikt et al. “The PS complex produces the nominal LHC beam”. In: (2000). URL: https://cds.cern.ch/record/452081.

[37] Hannes Bartosik and Giovanni Rumolo. *Performance of the LHC injector chain after the upgrade and potential development*. Tech. rep. Contribution to Snowmass 2021. 2022. arXiv: 2203.09202. URL: https://cds.cern.ch/record/2847538.

[38] Ewa Lopienska. *The CERN Accelerator Complex, Layout in 2022. Complexe Des Accélérateurs Du CERN En Janvier 2022*. General Photo. 2022. (Visited on 08/23/2023).

[39] A Verweij et al. "Retraining of the 1232 Main Dipole Magnets in the LHC". In: *IEEE Trans. Appl. Supercond.* 26.4 (2016), p. 4000705. DOI: 10.1109/TASC.2016.2514598. URL: https://cds.cern.ch/record/2262862.

[40] L Rossi. "The LHC Superconducting Magnets". In: (2003). URL: https://cds.cern.ch/record/630341.

[41] B.J. Holzer. "Lattice Design in High-energy Particle Accelerators". In: (2014). 40 pages, contribution to the CAS - CERN Accelerator School: Advanced Accelerator Physics Course, Trondheim, Norway, 18-29 Aug 2013. arXiv admin note: substantial text overlap with arXiv:1303.6514, pp. 61–100. DOI: 10.5170/CERN-2014-009.61. arXiv: 1601.04913. URL: https://cds.cern.ch/record/1982419.

[42] G L Bayatian et al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical design report. CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757.

[43] G L Bayatian et al. "CMS Physics: Technical Design Report Volume 2: Physics Performance". In: *J. Phys. G* 34 (2007). revised version submitted on 2006-09-22 17:44:47, pp. 995–1579. DOI: 10.1088/0954-3899/34/6/S01. URL: https://cds.cern.ch/record/942733.

[44] A Augusto Alves et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08005. DOI: 10.1088/1748-0221/3/08/S08005. URL: https://cds.cern.ch/record/1129809.

[45] K Aamodt et al. "The ALICE experiment at the CERN LHC. A Large Ion Collider Experiment". In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08002. DOI: 10.1088/1748-0221/3/08/S08002. URL: https://cds.cern.ch/record/1129812.

[46] R Alemany-Fernandez et al. “Operation and Configuration of the LHC in Run 1”. In: (2013). URL: https://cds.cern.ch/record/1631030.

[47] Jorg Wenninger. “Operation and Configuration of the LHC in Run 2”. In: (2019). URL: https://cds.cern.ch/record/2668326.

[48] G Aad et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08003. DOI: 10.1088/1748-0221/3/08/S08003. URL: https://cds.cern.ch/record/1129811.

[49] Sascha Mehlhase. *ATLAS detector slice (and particle visualisations)*. 2021. URL: https://cds.cern.ch/record/2770815.

[50] *ATLAS inner detector: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331063.

[51] S Haywood et al. *ATLAS inner detector: Technical Design Report, 2*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331064.

[52] *ATLAS central solenoid: Technical Design Report*. Technical design report. ATLAS. Electronic version not available. Geneva: CERN, 1997. DOI: 10.17181/CERN.ZZVJ.2JYE. URL: https://cds.cern.ch/record/331067.

[53] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. DOI: 10.17181/CERN.FWRW.FOOQ. URL: https://cds.cern.ch/record/331061.

[54] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. DOI: 10.17181/CERN.JRBJ.7028. URL: https://cds.cern.ch/record/331062.

[55] *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068.

[56] G. Aad et al. “The ATLAS Inner Detector commissioning and calibration. The ATLAS Inner Detector commissioning and calibration”. In: *Eur. Phys. J. C* 70 (2010). Comments: 34 pages, 25 figures, pp. 787–821. DOI: 10.1140/epjc/s10052-010-1366-7. arXiv: 1004.5293. URL: https://cds.cern.ch/record/1262789.

[57] Heinz Pernegger. *The Pixel Detector of the ATLAS Experiment for LHC Run-2*. Tech. rep. 06. Geneva: CERN, 2015. DOI: 10.1088/1748-0221/10/06/C06012. URL: https://cds.cern.ch/record/1985432.

[58] Marco Battaglia. "Performance of the ATLAS Pixel Detector at the Start of LHC Run 3". In: *PoS* (2023), p. 006. DOI: 10.22323/1.420.0006. URL: https://cds.cern.ch/record/2870192.

[59] A Ahmad et al. *The Silicon Microstrip Sensors of the ATLAS SemiConductor Tracker*. Tech. rep. 1. Geneva: CERN, 2007. DOI: 10.1016/j.nima.2007.04.157. URL: https://cds.cern.ch/record/1019885.

[60] A Abdesselam and T Akimoto. *The Barrel Modules of the ATLAS SemiConductor Tracker*. Tech. rep. Geneva: CERN, 2006. DOI: 10.1016/j.nima.2006.08.036. URL: https://cds.cern.ch/record/974073.

[61] A. Abdesselam et al. "The ATLAS semiconductor tracker end-cap module". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 575.3 (2007), pp. 353–389. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2007.02.019. URL: https://www.sciencedirect.com/science/article/pii/S0168900207003270.

[62] Georges Aad et al. "Operation and performance of the ATLAS semiconductor tracker. Operation and performance of the ATLAS semiconductor tracker". In: *JINST* 9 (2014). Comments: 60 pages plus author list (80 pages total), 33 figures, 10 tables, submitted to JINST. All figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/IDET-2013-01/ v2: replaced unreadable labels in Fig. 2, P08009. DOI: 10.1088/1748-0221/9/08/P08009. arXiv: 1404.7473. URL: https://cds.cern.ch/record/1698966.

[63] Bartosz Mindur. *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*. Tech. rep. Geneva: CERN, 2017. DOI: 10.1016/j.nima.2016.04.026. URL: https://cds.cern.ch/record/2139567.

[64] Joao Pequenao. "Computer generated image of the ATLAS inner detector". 2008. URL: https://cds.cern.ch/record/1095926.

[65] Henric Wilkens and (on behalf ofthe ATLAS LArg Collaboration). "The ATLAS Liquid Argon calorimeter: An overview". In: *Journal of Physics: Conference Series* 160.1 (Apr. 2009), p. 012043. DOI: 10.1088/1742-6596/160/1/012043. URL: https://dx.doi.org/10.1088/1742-6596/160/1/012043.

[66] ATLAS LAr Collaboration and Margherita Spalla. *ATLAS LAr Calorimeter Performance in LHC Run 2*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2707741.

[67] B Aubert et al. "Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter". In: *Nucl. Instrum. Methods Phys. Res., A* 558 (2006), pp. 388–418. DOI: 10.1016/j.nima.2005.11.212. URL: https://cds.cern.ch/record/883909.

[68] Joao Pequenao. "Computer Generated image of the ATLAS calorimeter". 2008. URL: https://cds.cern.ch/record/1095927.

[69] Ana Maria Henriques Correia. *The ATLAS Tile Calorimeter*. Tech. rep. Geneva: CERN, 2015. DOI: 10.1109/ANIMMA.2015.7465554. URL: https://cds.cern.ch/record/2004868.

[70] Louise Heelan and ATLAS Collaboration. *Performance of the ATLAS Tile Calorimeter*. Tech. rep. 1. Geneva: CERN, 2015. DOI: 10.1088/1742-6596/623/1/012014. URL: https://cds.cern.ch/record/2005212.

[71] Georges Aad et al. *Operation and performance of the ATLAS tile calorimeter in LHC Run 2*. Tech. rep. 77 pages in total, author list starting page 60, 36 figures, 4 tables, submitted to EPJC. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TCAL-2021-01/. Geneva: CERN, 2024. arXiv: 2401.16034. URL: https://cds.cern.ch/record/2887689.

[72] T Barillari. "The ATLAS liquid argon hadronic end-cap calorimeter: construction and selected beam test results". In: *Nucl. Phys. B, Proc. Suppl.* 150 (2006). Comments: 4 pages, 2 figures,IPRD04 conference, pp. 102–105. DOI: 10.1016/j.nuclphysbps.2004.10.087. URL: https://cds.cern.ch/record/768750.

[73] *Performance of the ATLAS Hadronic Endcap Calorimeter in beam tests: Selected results*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-LARG-PUB-2022-001. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2811731.

[74] D M Gingrich et al. *Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter*. Tech. rep. Geneva: CERN, 2007. DOI: 10.1088/1748-0221/2/05/P05005. URL: https://cds.cern.ch/record/1026994.

[75] *Performance of the muon spectrometer alignment in 2017 and 2018 data*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-MUON-PUB-2021-002. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2753329.

[76] J Wotschack. *ATLAS Muon Chamber Construction Parameters for CSC, MDT, and RPC chambers*. Tech. rep. Back-up document for the ATLAS Detector Paper. Geneva: CERN, 2008. URL: https://cds.cern.ch/record/1099400.

[77] Georges Aad et al. "Performance of the ATLAS muon triggers in Run 2". In: *JINST* 15 (2020). 60 pages in total, author list starting page 44, 34 figures, 5 tables, submitted to JINST. All figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TRIG-2018-01, P09015. DOI: 10.1088/1748-0221/15/09/p09015. arXiv: 2004.13447. URL: https://cds.cern.ch/record/2716326.

[78] Joao Pequenao. "Computer generated image of the ATLAS Muons subsystem". 2008. URL: https://cds.cern.ch/record/1095929.

[79] Georges Aad et al. "Operation of the ATLAS trigger system in Run 2". In: *JINST* 15.10 (2020). 60 pages in total, author list starting page 44, 19 figures, 3 tables, submitted to JINST, All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TRIG-2019-04/, P10004. DOI: 10.1088/1748-0221/15/10/P10004. arXiv: 2007.12539. URL: https://cds.cern.ch/record/2725146.

[80] Morad Aaboud et al. “Performance of the ATLAS Trigger System in 2015. Performance of the ATLAS Trigger System in 2015”. In: *Eur. Phys. J. C* 77.5 (2017). 77 pages in total, author list starting page 61, 50 figures, 1 table. Published in Eur. Phys. J. C. All figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TRIG-2016-01/, p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661. URL: https://cds.cern.ch/record/2235584.

[81] *Trigger Menu in 2016*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2017-001. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2242069.

[82] *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2016-001. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2136007.

[83] Vato Kartvelishvili et al. *Measurement of the production cross-section of J/ψ and ψ(2S) mesons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2806663.

[84] Georges Aad et al. “Muon reconstruction and identification efficiency in ATLAS using the full Run 2 *pp* collision data set at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J., C* 81 (2021). 64 pages in total, author list starting page 42, auxiliary material starting at page 59, 34 figures, 3 tables. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/MUON-2018-03/, p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578. URL: https://cds.cern.ch/record/2746302.

[85] Morad Aaboud et al. “Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016”. In: *Eur. Phys. J. C* 79.3 (2019). 55 pages in total, author list starting page 39, 18 figures, 2 tables, published in EPJC. All figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/PERF-2017-02, p. 205. DOI: 10.1140/epjc/

s10052-019-6650-6. arXiv: 1810.05087. URL: https://cds.cern.ch/record/2643391.

[86] S Albrand et al. *ATLAS Dataset Nomenclature*. Tech. rep. The scope of the document covers:- (1) Monte-Carlo datasets (2) Real Data Datasets. a. Primary b. Super datasets (including relational event collections) (3) User datasets (4) Group datasets (5) Conditions datasets (6) Application Internal datasets. Geneva: CERN, 2007. URL: https://cds.cern.ch/record/1070318.

[87] Jonathan Pumplin et al. "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis". In: *Journal of High Energy Physics* 2002.07 (July 2002), pp. 012–012. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2002/07/012. URL: http://dx.doi.org/10.1088/1126-6708/2002/07/012.

[88] E Barberio and Zbigniew Was. "PHOTOS: a universal Monte Carlo for QED radiative corrections; version 2.0". In: *Comput. Phys. Commun.* 79 (1994), pp. 291–308. DOI: 10.1016/0010-4655(94)90074-4. URL: https://cds.cern.ch/record/254343.

[89] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[90] S. Navas et al. *Review of particle physics*. Vol. 110. 3. 2024. Chap. $c\bar{c}$ Mesons, p. 030001. DOI: 10.1103/PhysRevD.110.030001.

[91] Jerome H. Friedman. "Stochastic gradient boosting". In: *Computational Statistics & Data Analysis* 38.4 (2002). Nonlinear Methods and Data Mining, pp. 367–378. ISSN: 0167-9473. DOI: https://doi.org/10.1016/S0167-9473(01)00065-2. URL: https://www.sciencedirect.com/science/article/pii/S0167947301000652.

[92] Yann Coadou. "Boosted Decision Trees". In: *Artificial Intelligence for High Energy Physics*. WORLD SCIENTIFIC, Feb. 2022, pp. 9–58. ISBN: 9789811234033.

DOI: 10.1142/9789811234033_0002. URL: http://dx.doi.org/10.1142/9789811234033_0002.

[93] Yoav Freund and Robert E Schapire. "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting". In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139. ISSN: 0022-0000. DOI: https://doi.org/10.1006/jcss.1997.1504. URL: https://www.sciencedirect.com/science/article/pii/S002200009791504X.

[94] Michele Pinamonti et al. *Documentation for TRExFitter*. URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs/.

[95] Kyle Cranmer. "Practical Statistics for the LHC". In: (2015). Comments: presented at the 2011 European School of High-Energy Physics, Cheile Gradistei, Romania, 7-20 September 2011 I expect to release updated versions of this document in the future, pp. 267–308. DOI: 10.5170/CERN-2014-003.267. arXiv: 1503.07622. URL: https://cds.cern.ch/record/2004587.

[96] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[97] S. Navas et al. *Review of particle physics*. Vol. 110. 3. 2024. Chap. 40, pp. 694–710. DOI: 10.1103/PhysRevD.110.030001.

[98] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. ISSN: 00034851, 21688990. URL: http://www.jstor.org/stable/2957648 (visited on 02/15/2025).

[99] Kyle Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., 2012. DOI: 10.17181/CERN-OPEN-2012-016. URL: https://cds.cern.ch/record/1456844.

[100] Robert D. Cousins. "Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms". In: 2013. URL: https://api.semanticscholar.org/CorpusID:5936965.

[101] Andreas Hocker et al. *TMVA - Toolkit for Multivariate Data Analysis with ROOT: Users guide. TMVA - Toolkit for Multivariate Data Analysis*. Tech. rep. TMVA-v4 Users Guide: 135 pages, 19 figures, numerous code examples and references. Geneva: CERN, 2007. URL: https://cds.cern.ch/record/1019880.

[102] M. Stone. "Cross-Validatory Choice and Assessment of Statistical Predictions". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 36.2 (Dec. 2018), pp. 111–133. ISSN: 0035-9246. DOI: 10.1111/j.2517-6161.1974.tb00994.x. eprint: https://academic.oup.com/jrsssb/article-pdf/36/2/111/49096683/jrsssb_36_2_111.pdf. URL: https://doi.org/10.1111/j.2517-6161.1974.tb00994.x.

[103] Thomas Philippe Calvet. "Search for the production of a Higgs boson in association with top quarks and decaying into a b-quark pair and b-jet identification with the ATLAS experiment at LHC". Presented 08 Nov 2017. Aix-Marseille U., 2017. URL: https://cds.cern.ch/record/2296985.

[104] G. Aad et al. "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: *The European Physical Journal C* 83.10 (Oct. 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11747-w. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11747-w.

A. Definition of Collins-Soper Frame

The Collins-Soper frame provides a basis for measuring angular asymmetries in cross-sections. In the lab frame, the angular distribution of J/ψ's and γ's do not have physical meaning for the study of TMDs. This is due to there being no relation between the lab-frame angles of either particle and the transverse momentum component of the gluons in the initiating protons. The Collins-Soper frame introduces angles that include the context of the $J/\psi + \gamma$ system, and hence include information about TMDs.

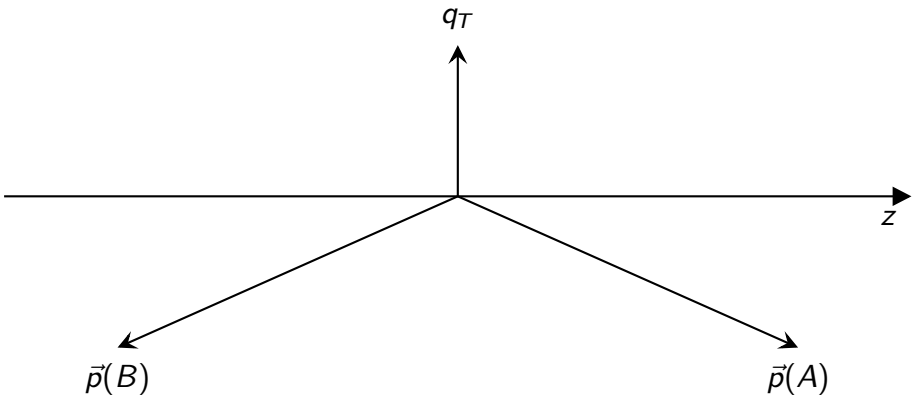


Figure A.1: Diagram of Collins-Soper frame.

Figure A.1 details the Collins-Soper frame. $\vec{p}(A)$ and $\vec{p}(B)$ are the momenta of the incoming partons in the J/ψ rest frame. The z-axis bisects the angle between $\vec{p}(A)$ and $-\vec{p}(B)$. q_T and the y-axis are normal to the plane formed by these two vectors. The x-axis completes a right-handed coordinate system. θ_{cs} and ϕ_{cs} are

defined as:

$$\cos\theta_{cs} = \frac{2}{Q\sqrt{Q^2 + p_T(J/\psi+\gamma)}}(P^+(J/\psi)P^-(\gamma) - P^-(J/\psi)P^+(\gamma)) \tag{A.1}$$

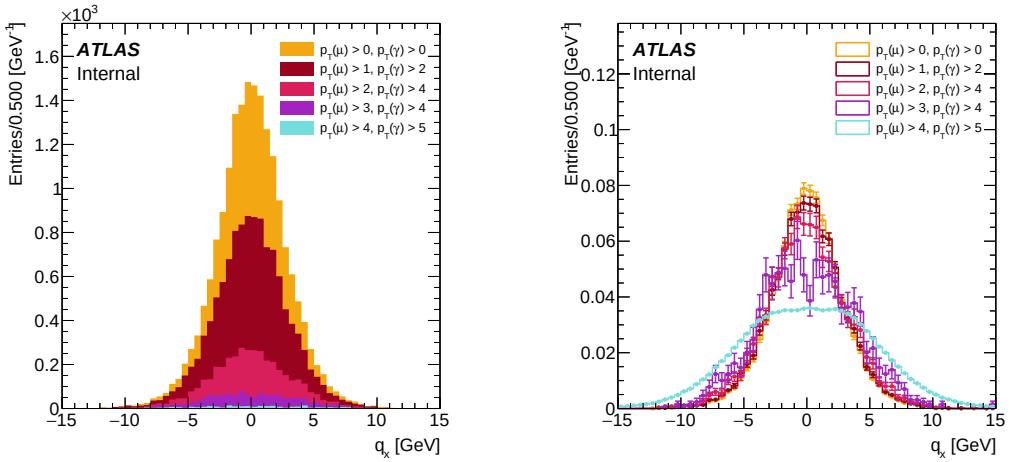
$$\phi_{cs} = \tan^{-1}\left\{\frac{Q^2 + p_T(J/\psi+\gamma)^2}{Q}\frac{\bar{\Delta}_T \cdot \hat{q}_T}{\bar{\Delta}_T \cdot \hat{p}_T}\right\}. \tag{A.2}$$

Here, $P^{\pm} = \frac{1}{\sqrt{2}}(E \pm p_z)$, with p_z as the longitudinal momentum and E as the energy of a particle. $\Delta^j = p^j_{z,q} - p^j_{z,2}$, $\hat{p}_T$ is the unit vector of the $J/\psi+\gamma$ transverse momentum, and $\hat{q}_T$ the unit vector for the y-axis, again.

B. Acceptance Effects on q_T^A and q_T^B

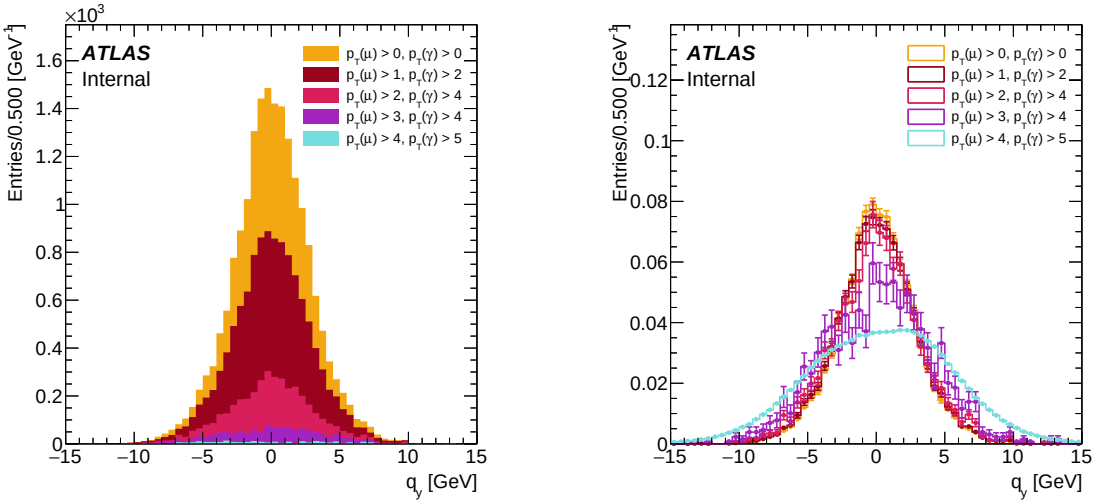
$p_T(\mu)$	$p_T(\gamma)$	**Events**	$\bar{q}_T^A$	$\bar{q}_T^B$	$\bar{q}_x$	$\bar{q}_y$	$\sigma(q_T^A)$	$\sigma(q_T^B)$	$\sigma(q_x)$	$\sigma(q_y)$	**Notes**
0	0	18810	0.06	0.01	0.02	0.05	2.66	2.65	2.82	2.82	Photon count, η, and y, cuts applied
1	2	11879	0.21	0.03	-0.01	0.03	2.73	2.81	2.94	2.94	-
2	4	4021	0.62	0.02	-0.05	0.01	2.60	3.37	3.23	3.24	-
3	4	1295	2.05	0.03	0.02	0.09	2.66	3.87	3.73	3.92	-
4	5	257	3.08	0.42	0.14	0.10	3.05	4.77	4.49	5.03	-
4	5	689988	3.43	0.01	0.05	0.29	3.33	4.85	5.03	4.98	Sample with increased generator level cuts used

Table B.1: Total events, mean, and standard deviation of the q_T^A, q_T^B, q_x, and q_y distributions through progressive simulated acceptance cuts.



(a) (b)

Figure B.1: Distributions of (a) q_x through progressive acceptance cuts without normalisation, showing statistics reduction and (b) with normalisation, showing effects on shape.



(a) (b)

Figure B.2: Distributions of (a) q_y through progressive acceptance cuts without normalisation, showing statistics reduction and (b) with normalisation, showing effects on shape.

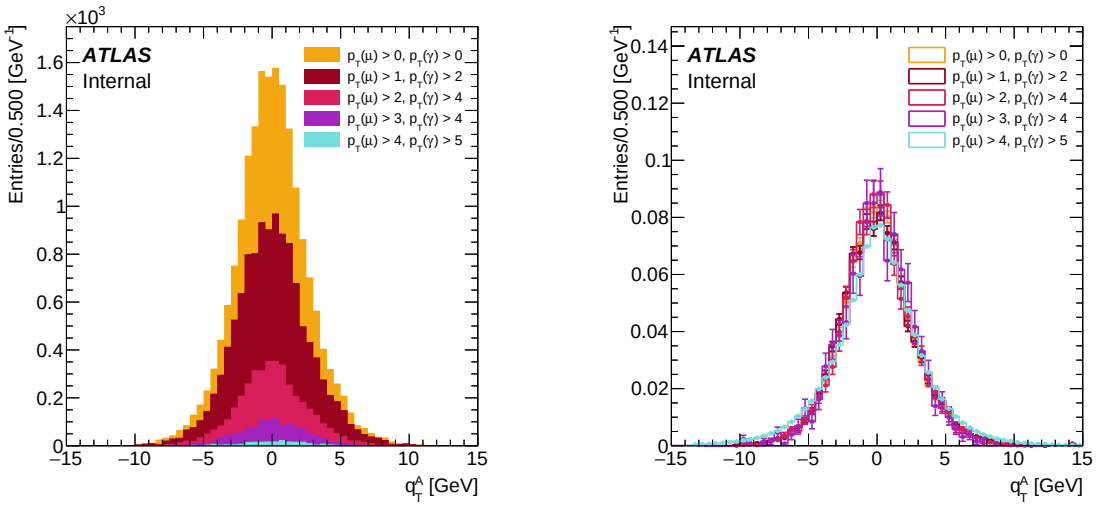


Figure B.3: Distributions of (a) q_T^A through progressive acceptance cuts without normalisation, showing statistics reduction and (b) with normalisation, showing effects on shape.

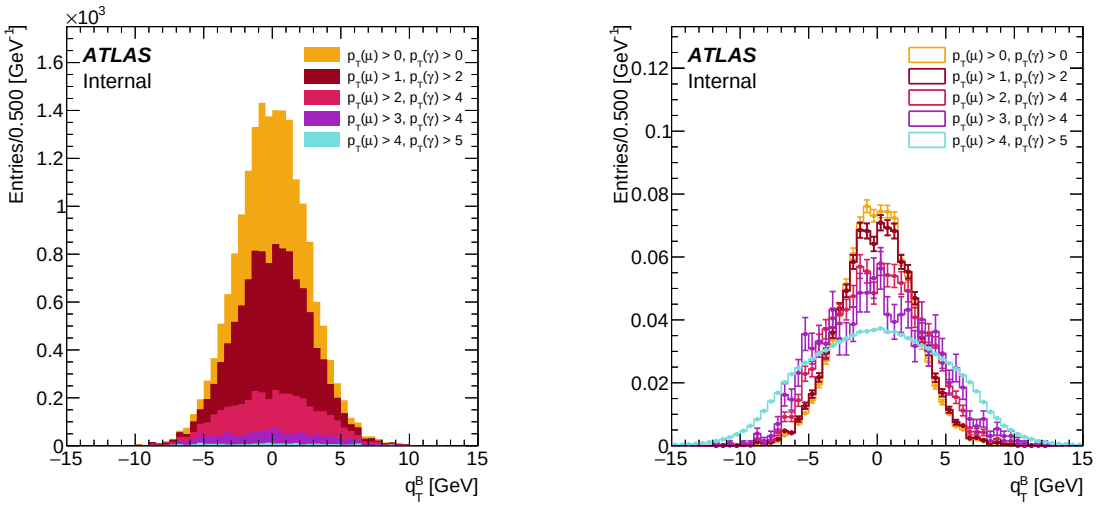


Figure B.4: Distributions of (a) q_T^B through progressive acceptance cuts without normalisation, showing statistics reduction and (b) with normalisation, showing effects on shape.

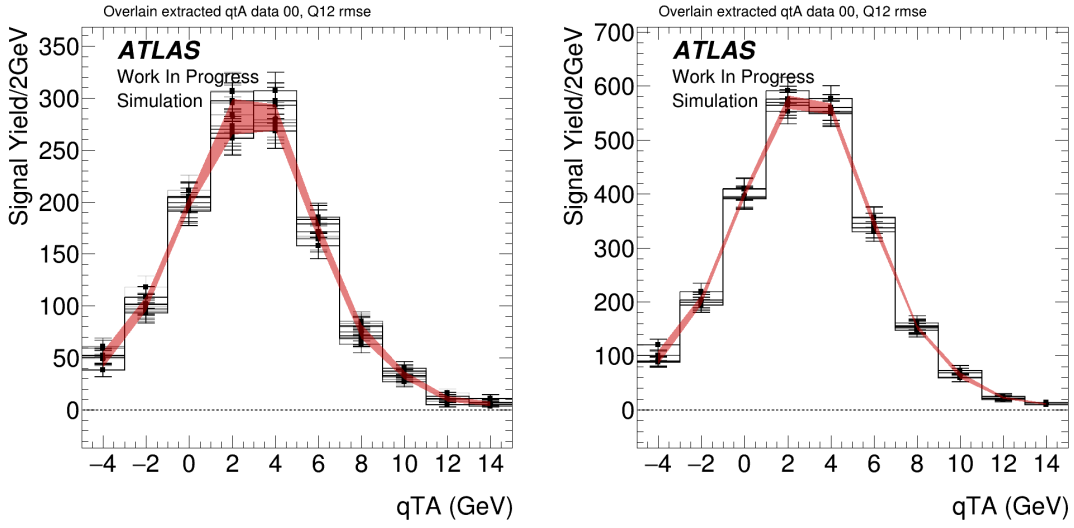
C. Validation of Analysis Procedure

As a proof of concept, the analysis strategy was tested using toy samples. Since this study was conducted, the framework managing the minimisation procedure was changed over to `TRExFitter` instead of `HistFactory` directly, though the rest of the chain is identical. The analysis chain is tested to ensure that the signal yield that is extracted, is linearly dependent on the signal yield present, and that there are no deleterious effects. To do this, create a number of toy data samples, built from a mix of events of MC signal and MC background. These toy 'data' samples are built directly from the *sign* and *bckg* samples.

First, the events of *sign* and *bckg* samples are randomly split into three categories. The first category uses the two subsets of events, one from *sign* and one from *bckg*, as training samples for building the BDT. The second category of events are used as *sign* and *bckg* for the fitting procedure. Lastly, the third category of events are used to build toy data samples, by mixing the *sign* and *bckg* events into a single sample.

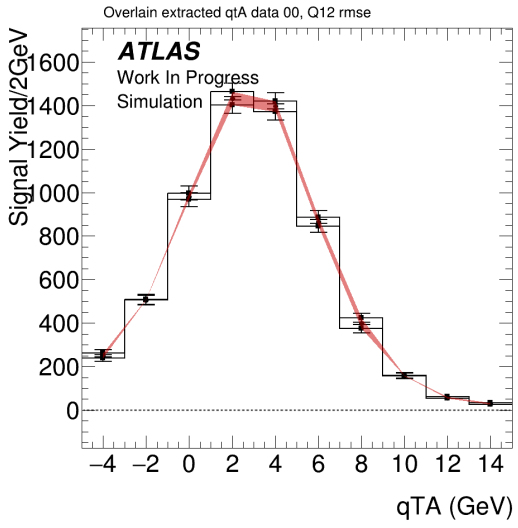
A number of variations of this toy data sample are created in order to test the linearity of the analysis strategy, this is done by introducing different ratios of the *sign* and *bckg* events. Use the same 13000 *sign* events throughout the process, and create a baseline sample of an equal amount of *sign* and *bckg* events, call this sample *s50*. Next, distribute these *sign* events into multiple files with different signal to background ratios, creating samples named *s5*, *s10*, *s25*. Taking the example of *s5*, there 10 samples, each with $1/10^{\text{th}}$ of the *sign* events that are allotted to *s50* and the same 13000 *bckg* events. Recombining the extracted yields from all the *s5* (or other toy samples) after the fitting procedure,

should yield the same results as the *s50* sample. Any divergence from this should reveal potential non-linearities or biases are present in the analysis strategy.

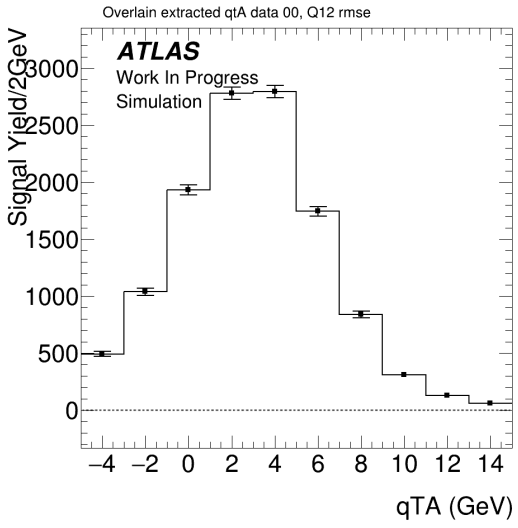


(a) *s5*

(b) *s10*



(c) *s25*

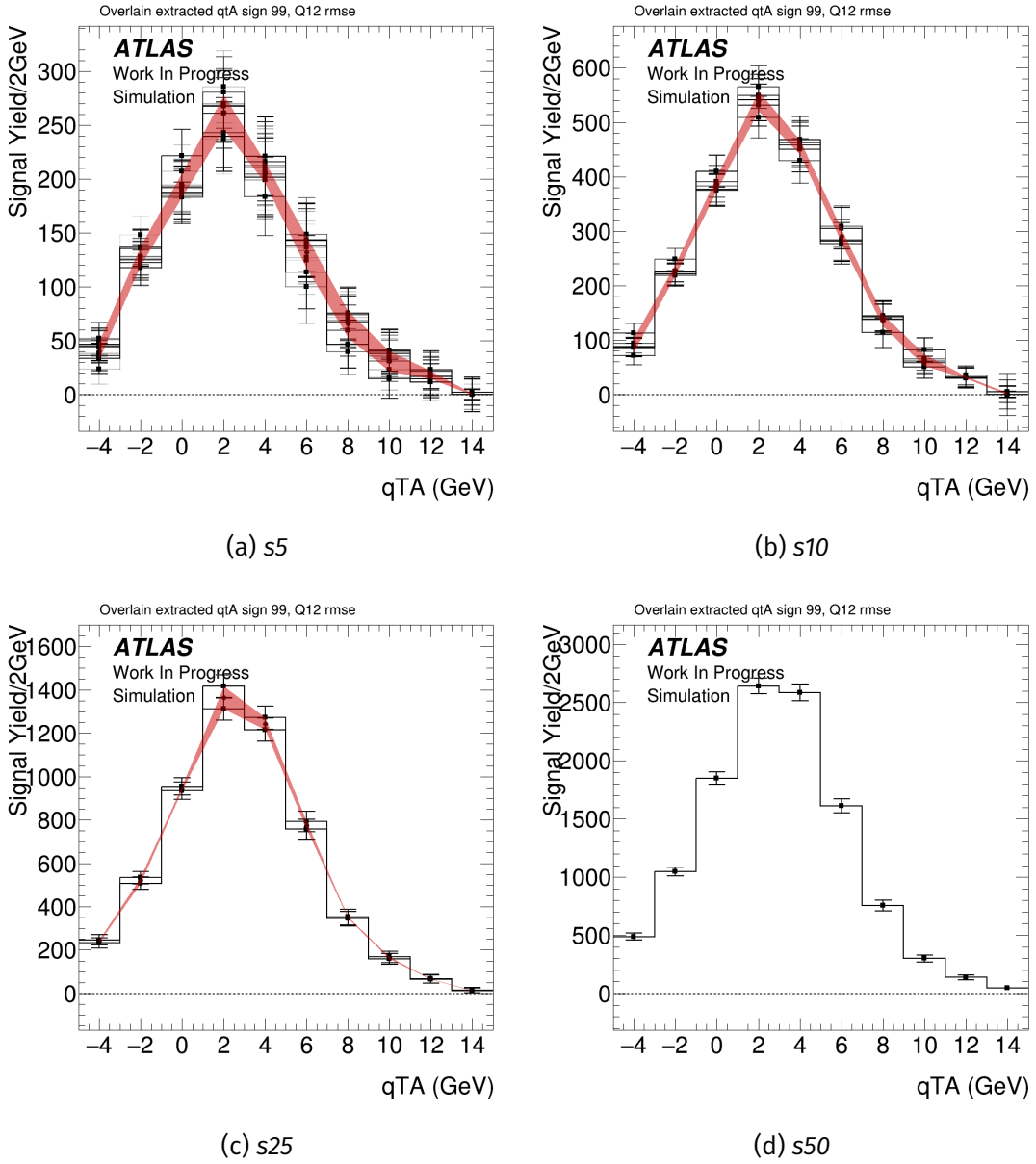


(d) *s50*

Figure C.1: Overlain q_T^A distributions for the signal component of constructed toy data samples for (a) s5, (b) s10, (c) s25, and (d) s50

Figure C.1 shows overlain q_T^A peaks for the signal component of the toy samples, the error in red is the RMS of entries in each q_T^A bin. After performing the analysis procedure, BDT training, mass sideband and lifetime subtraction, and signal

extraction, the resulting q_T^A distributions are shown in Figure C.2.



(a) *s5*

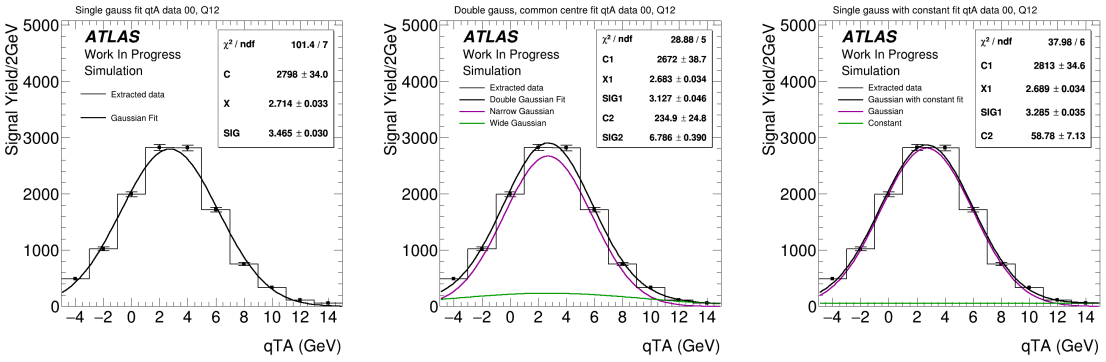
(b) *s10*

(c) *s25*

(d) *s50*

Figure C.2: Overlain q_T^A distributions for the signal extracted from the for (a) s5, (b) s10, (c) s25, and (d) s50

Examining the distribution of the signal components in the toy data, there is no visible distortion of the spread between *s5, s10*, etc. First, check the recombined s05 samples:



(a)

Figure C.3: Left fit, regular Gaussian, central fit, two Gaussians with a common centre, right fit, single Gaussian with a constant.

Trivially, when these fits are applied to all *s10, s25* and *s50*, the results are the same. The widths of the Gaussians are 3.465 ± 0.030 GeV, 3.127 ± 0.046 GeV, and 3.285 ± 0.035 GeV, for the regular, double and single+constant Gaussian fits, respectively.

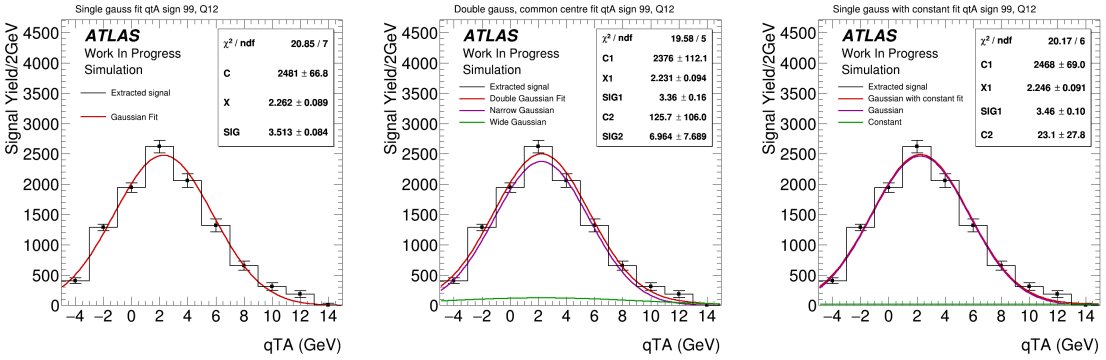


Figure C.4: Left fit, regular Gaussian, central fit, two Gaussians with a common centre, right fit, single Gaussian with a constant for the recombined s05 samples.

The three Gaussians are then fit to the recombined extracted signal from the *s05* samples after the analysis procedure is conducted. It can be seen that the q_T^A distribution shifts slightly leftwards, though when comparing with the previous fittings, the most of the sigmas remain close to their comparable values. Means of fitted Gaussians are now grouped around 2.2GeV, and the widths are 3.513 ± 0.084 GeV, 3.36 ± 0.16 GeV, and 3.46 ± 0.10 GeV.

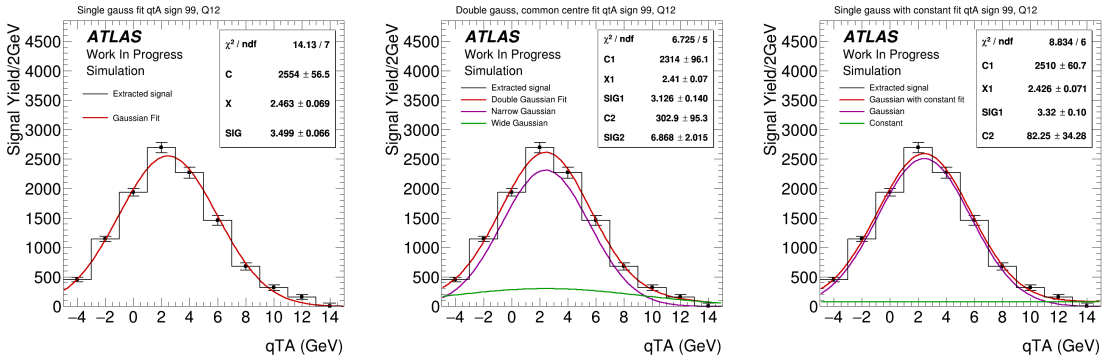


Figure C.5: Left fit, regular Gaussian, central fit, two Gaussians with a common centre, right fit, single Gaussian with a constant for the recombined s10 samples.

The same shift effect is present, though the leftward shift has reduced from 0.5 GeV to 0.3 GeV. From the definition of q_T^A, this shifts corresponds for a reduction of J/ψ momentum, or an increase in photon momentum, or both. The widths remain within the same ranges as before.

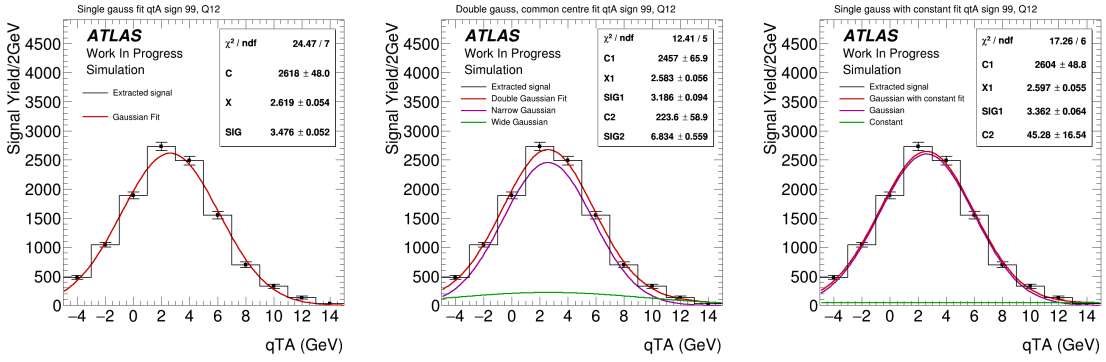


Figure C.6: Left fit, regular Gaussian, central fit, two Gaussians with a common centre, right fit, single Gaussian with a constant for the recombined s25 samples.

For s25 in Figure C.6, the shift has reduced to ~ 0.1 GeV, and becomes negligible in s50, Figure C.7.

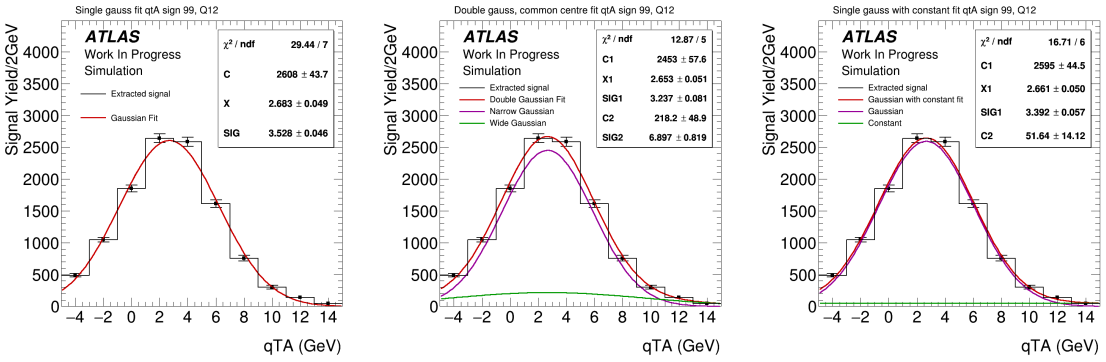


Figure C.7: Left fit, regular Gaussian, central fit, two Gaussians with a common centre, right fit, single Gaussian with a constant for the s50 sample.

For clarity, when the evolution of parameters from these fits with respect to the s value is viewed in Figure C.8, a change in the mean is clearly visible. From the observations in the main text, the amount of signal in the analysis data sample is on the order of approximately $10 - 25\%$, corresponding to a small deviation of about 0.3 GeV.

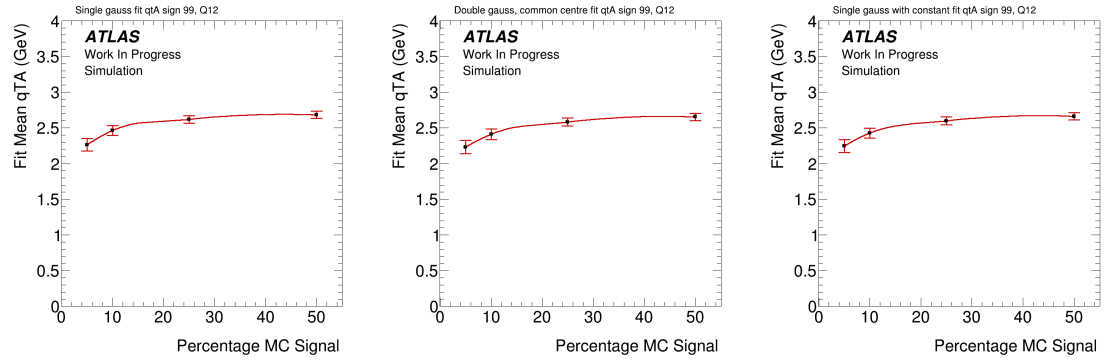


Figure C.8: Central values of the Gaussians in the three fits from left to right across all recombined samples.

Figure C.9 shows the evolution of the Gaussian widths. The TMD and ISR Gaussians in the double Gaussian fit are shown in red and blue, respectively, in the central plot. It can be observed that the widths of Gaussians is invariant with respect to the s values.

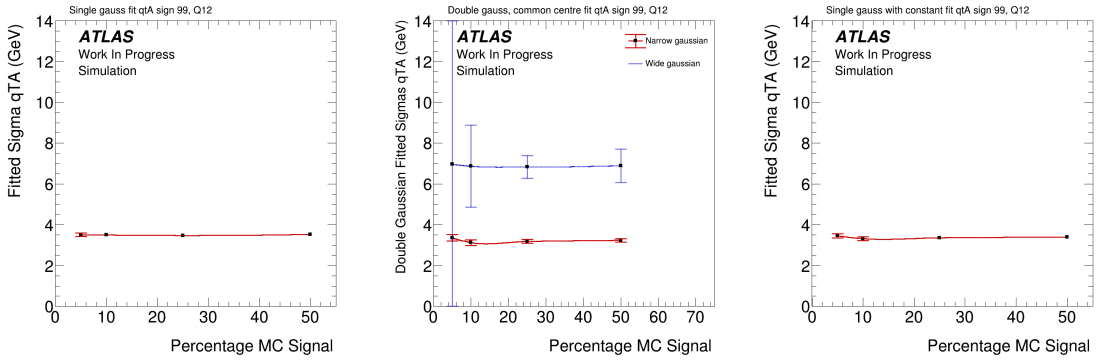


Figure C.9: Sigma values of the Gaussians in the three fits from left to right across all recombined samples.

This process of validation confirms the analysis procedure is sound.

D. Tables of TRExFitter Yields

	Analysis region	Validation Region 1	Validation Region 2
sign	0.55 ± 0.04	0.308 ± 0.007	0.544 ± 0.013
bckg	174 ± 13	81 ± 21	170 ± 40
Total	175 ± 13	81 ± 21	170 ± 40
Data 2015	175	78	176

Table D.1: `TRExFitter` output fit tables for q_T^A bin 1.

	Analysis region	Validation Region 1	Validation Region 2
sign	80 ± 50	60 ± 40	80 ± 50
bckg	180 ± 50	180 ± 60	170 ± 60
Total	255 ± 16	240 ± 50	250 ± 50
Data 2015	255	194	251

Table D.2: `TRExFitter` output fit tables for q_T^A bin 2.

	Analysis region	Validation Region 1	Validation Region 2
sign	450 ± 50	420 ± 50	430 ± 50
bckg	150 ± 50	150 ± 60	150 ± 50
Total	602 ± 25	570 ± 50	580 ± 50
Data 2015	602	554	586

Table D.3: `TRExFitter` output fit tables for q_T^A bin 3.

	Analysis region	Validation Region 1	Validation Region 2
sign	460 ± 130	460 ± 120	450 ± 110
bckg	680 ± 130	670 ± 190	630 ± 180
Total	1140 ± 34	1130 ± 130	1080 ± 120
Data 2015	1140	1108	1093

Table D.4: `TRExFitter` output fit tables for q_T^A bin 4.

	Analysis region	Validation Region 1	Validation Region 2
sign	900 ± 130	890 ± 120	870 ± 110
bckg	1830 ± 130	1820 ± 280	1770 ± 270
Total	2730 ± 50	2710 ± 260	2640 ± 260
Data 2015	2726	2719	2627

Table D.5: `TRExFitter` output fit tables for q_T^A bin 5.

	Analysis region	Validation Region 1	Validation Region 2
sign	1640 ± 260	1640 ± 200	1620 ± 200
bckg	5340 ± 270	5300 ± 800	5100 ± 800
Total	6980 ± 80	7000 ± 800	6700 ± 700
Data 2015	6979	6975	6703

Table D.6: `TRExFitter` output fit tables for q_T^A bin 6.

	Analysis region	Validation Region 1	Validation Region 2
sign	2520 ± 320	2500 ± 280	2490 ± 280
bckg	15110 ± 330	14900 ± 1400	14700 ± 1400
Total	17630 ± 130	17400 ± 1300	17200 ± 1300
Data 2015	17632	17631	17414

Table D.7: `TRExFitter` output fit tables for q_T^A bin 7.

	Analysis region	Validation Region 1	Validation Region 2
sign	2600 ± 400	2600 ± 400	2600 ± 400
bckg	19500 ± 500	18200 ± 1500	18200 ± 1500
Total	22130 ± 150	20800 ± 1200	20800 ± 1200
Data 2015	22133	22133	22133

Table D.8: `TRExFitter` output fit tables for q_T^A bin 8.

	Analysis region	Validation Region 1	Validation Region 2
sign	2140 ± 310	2100 ± 400	2100 ± 400
bckg	13810 ± 330	13700 ± 600	13700 ± 600
Total	15940 ± 130	15800 ± 700	15800 ± 700
Data 2015	15942	15942	15942

Table D.9: `TRExFitter` output fit tables for q_T^A bin 9.

	Analysis region	Validation Region 1	Validation Region 2
sign	980 ± 230	1030 ± 320	1030 ± 320
bckg	8100 ± 240	8400 ± 600	8400 ± 600
Total	9080 ± 100	9400 ± 800	9400 ± 800
Data 2015	9083	9083	9083

Table D.10: `TRExFitter` output fit tables for q_T^A bin 10.

	Analysis region	Validation Region 1	Validation Region 2
sign	470 ± 280	460 ± 330	460 ± 330
bckg	4810 ± 280	4800 ± 400	4800 ± 400
Total	5270 ± 70	5300 ± 600	5200 ± 600
Data 2015	5271	5271	5272

Table D.11: `TRExFitter` output fit tables for q_T^A bin 11.

	Analysis region	Validation Region 1	Validation Region 2
sign	370 ± 140	360 ± 170	350 ± 160
bckg	2650 ± 150	2600 ± 600	2600 ± 600
Total	3020 ± 60	3000 ± 700	3000 ± 700
Data 2015	3023	3023	3002

Table D.12: `TRExFitter` output fit tables for q_T^A bin 12.

	Analysis region	Validation Region 1	Validation Region 2
sign	90 ± 110	90 ± 110	80 ± 100
bckg	1790 ± 140	1800 ± 600	1800 ± 600
Total	1880 ± 50	1900 ± 600	1900 ± 500
Data 2015	1883	1883	1840

Table D.13: `TRExFitter` output fit tables for q_T^A bin 13.

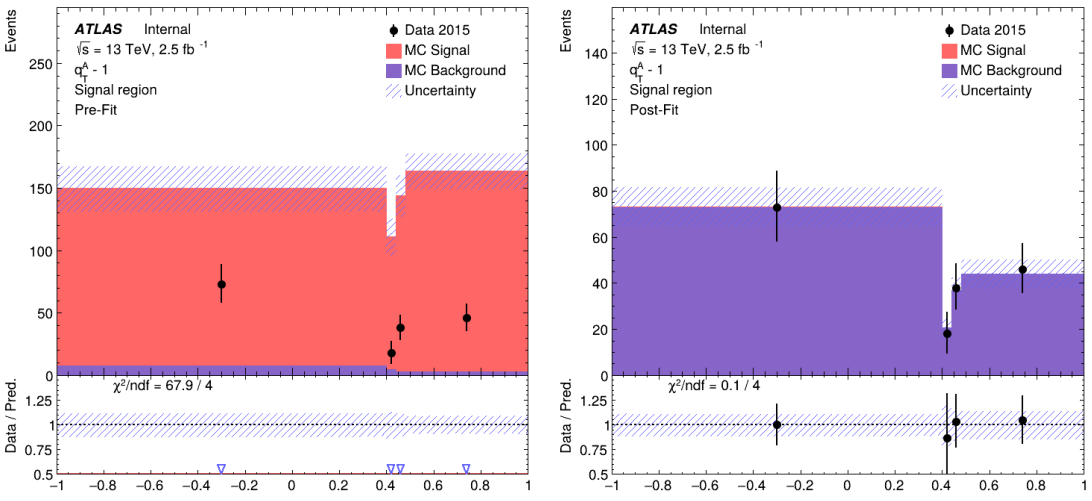
	Analysis region	Validation Region 1	Validation Region 2
sign	30 ± 120	30 ± 120	30 ± 100
bckg	1070 ± 210	1200 ± 400	1100 ± 400
Total	1100 ± 100	1240 ± 340	1180 ± 340
Data 2015	1101	1101	1020

Table D.14: `TRExFitter` output fit tables for q_T^A bin 14.

	Analysis region	Validation Region 1	Validation Region 2
sign	250 ± 120	270 ± 150	150 ± 80
bckg	440 ± 120	420 ± 130	370 ± 110
Total	692 ± 26	700 ± 70	520 ± 60
Data 2015	692	696	527

Table D.15: `TRExFitter` output fit tables for q_T^A bin 15.

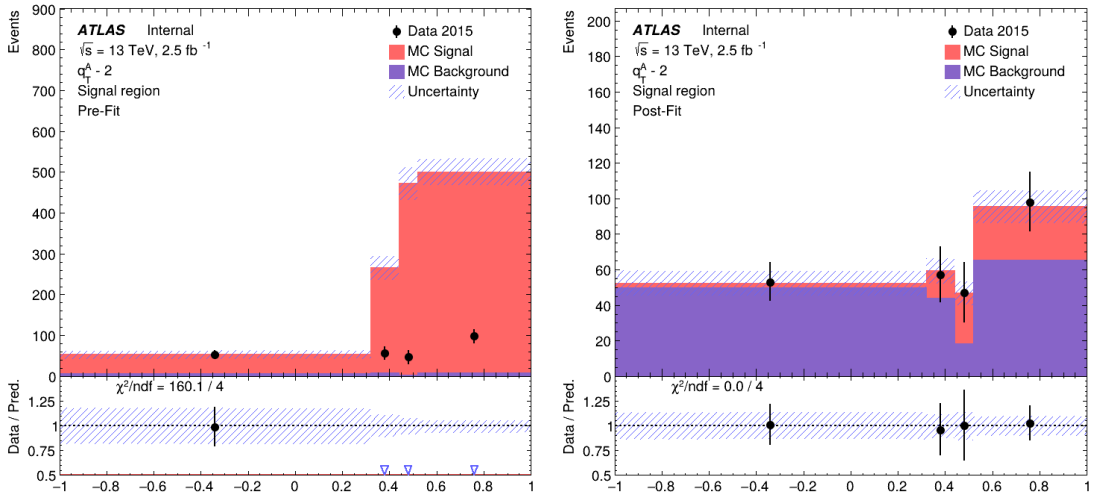
E. Plots of Differential Fits in BDT Score



(a) Pre-fit

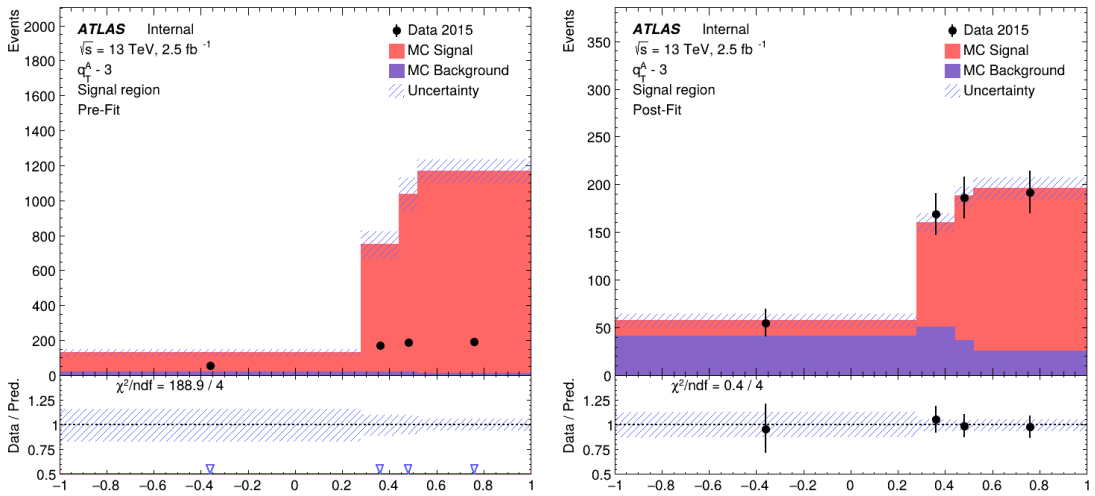
(b) Post-fit

Figure E.1: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 1.



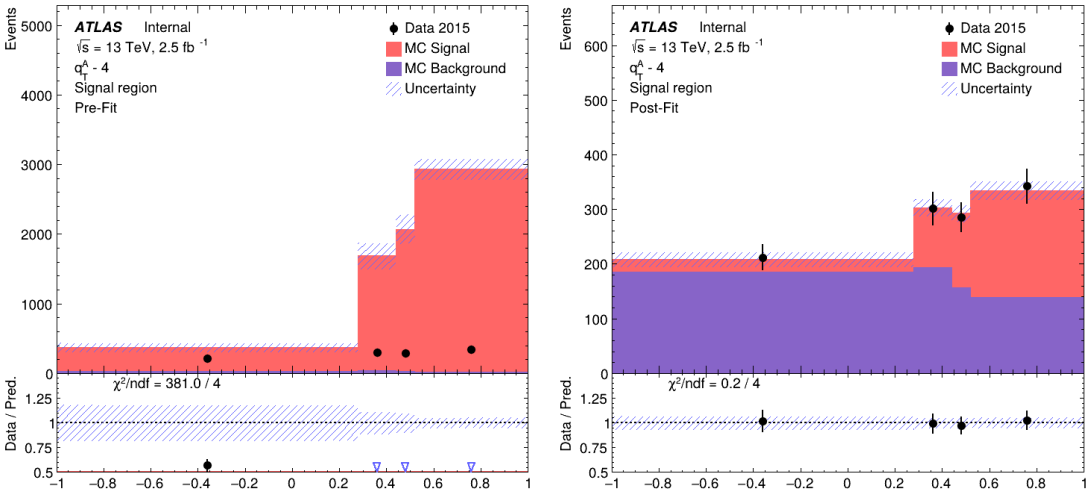
(a) Pre-fit (b) Post-fit

Figure E.2: TRExFitter fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 2.



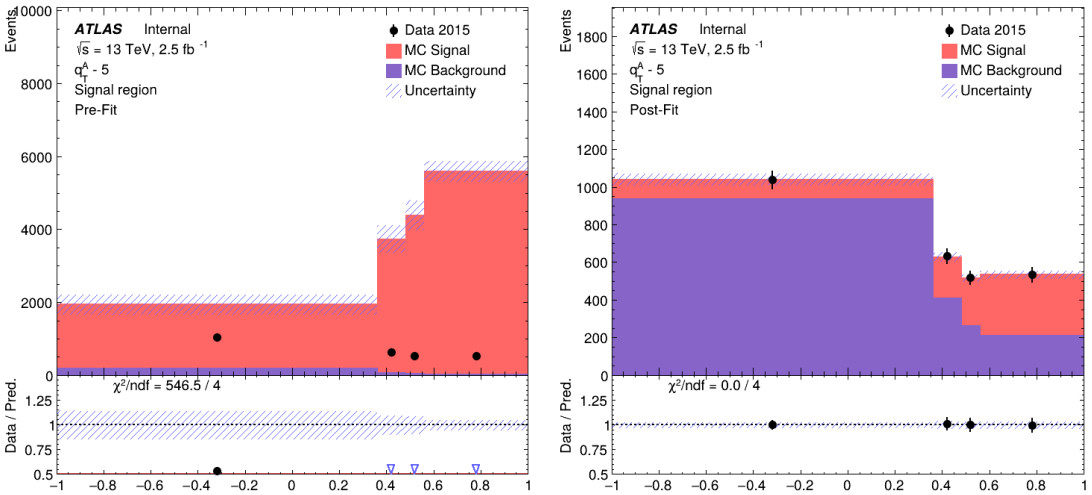
(a) Pre-fit (b) Post-fit

Figure E.3: TRExFitter fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 3.



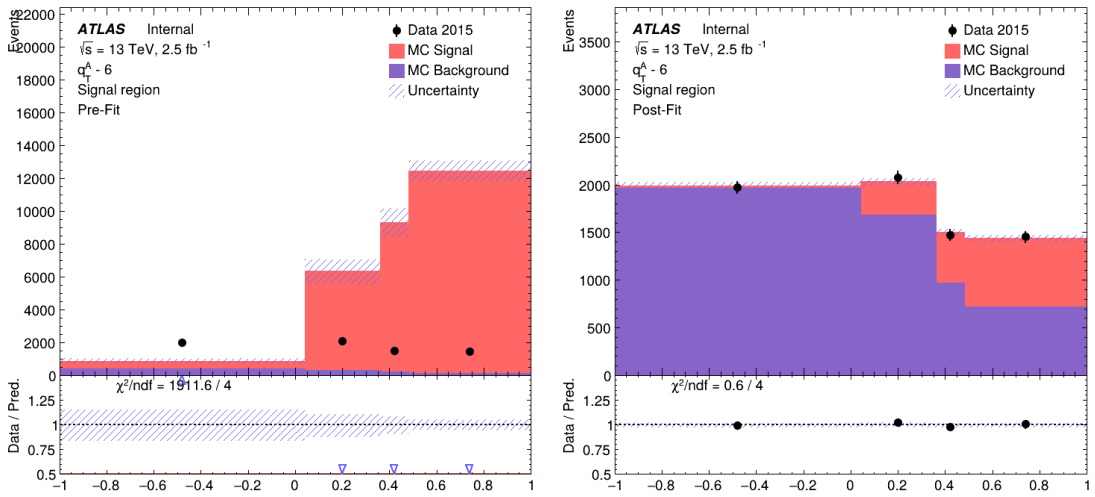
(a) Pre-fit (b) Post-fit

Figure E.4: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 4.



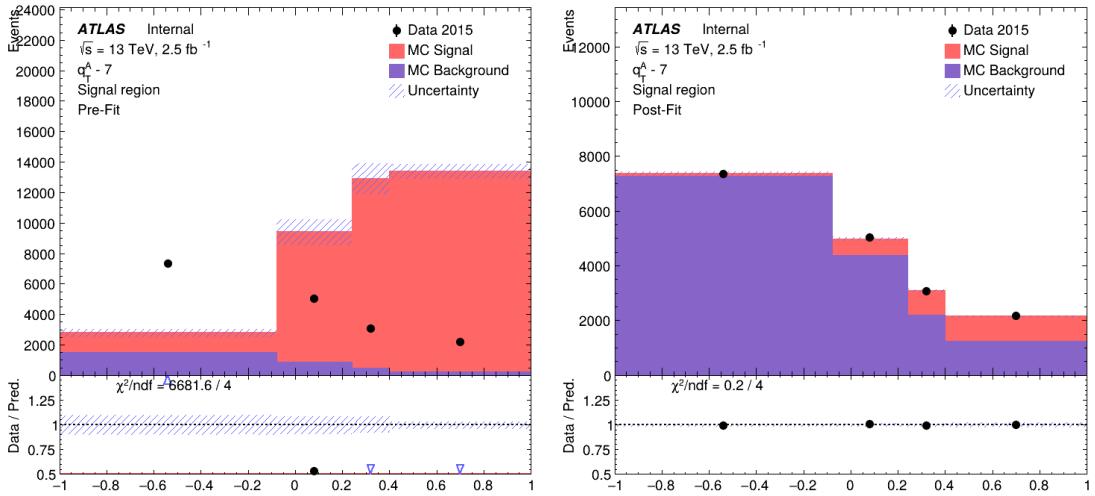
(a) Pre-fit (b) Post-fit

Figure E.5: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 5.



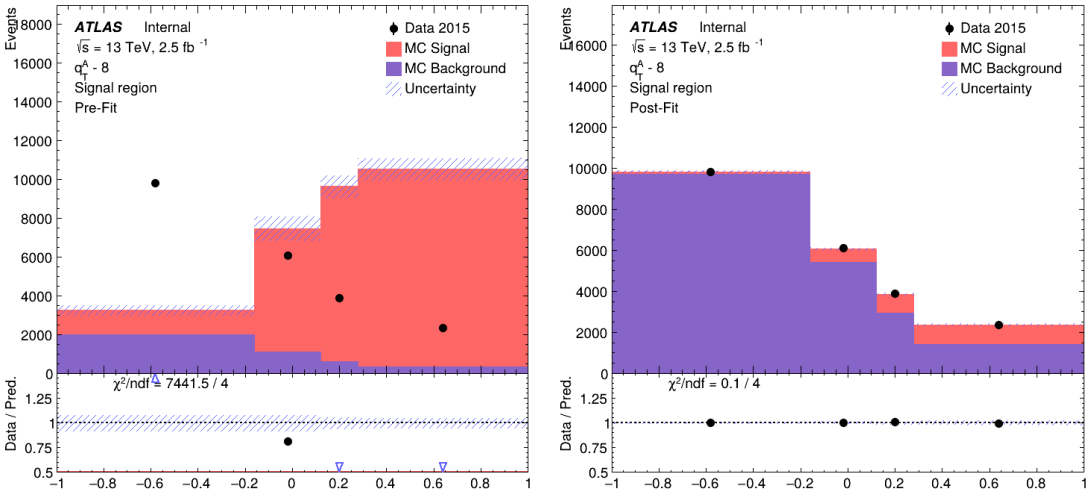
(a) Pre-fit (b) Post-fit

Figure E.6: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 6.



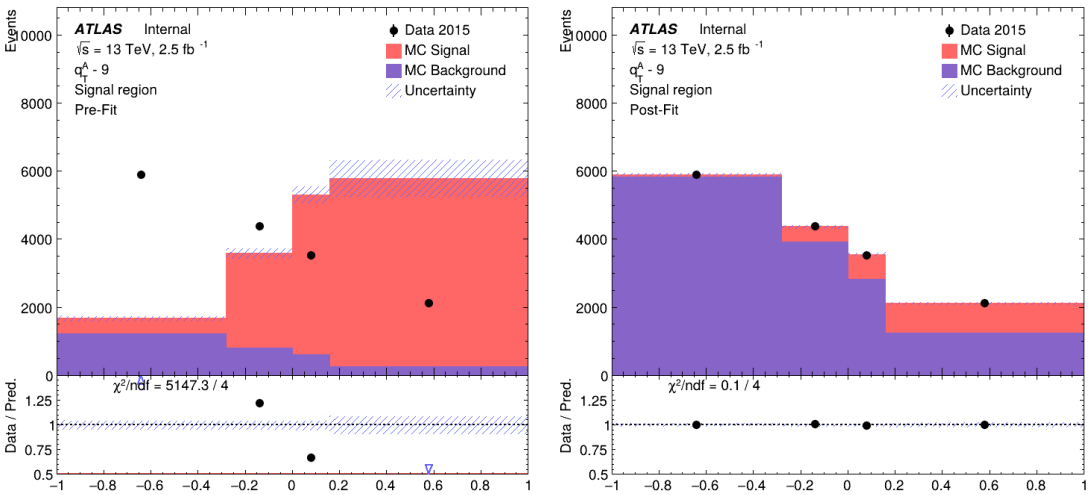
(a) Pre-fit (b) Post-fit

Figure E.7: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 7.



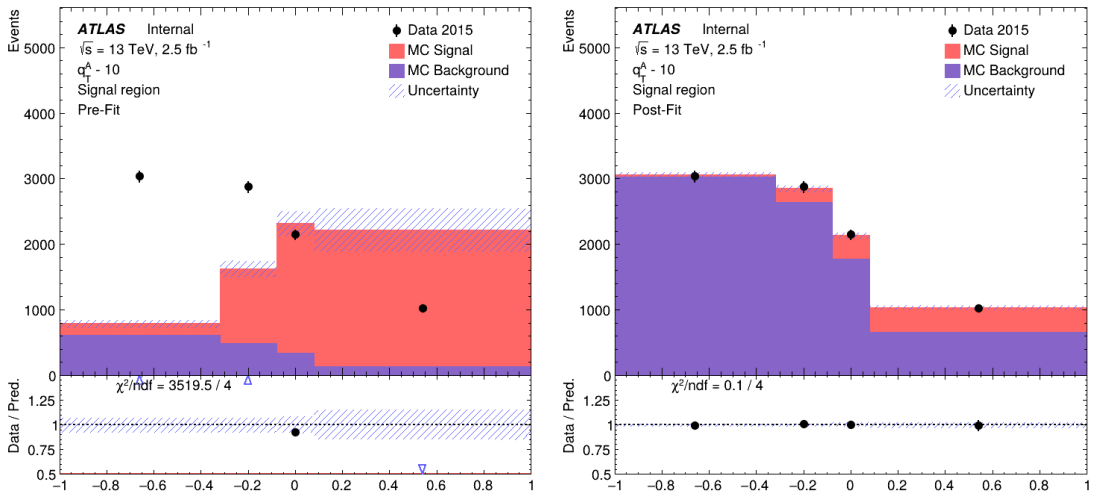
(a) Pre-fit (b) Post-fit

Figure E.8: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 8.



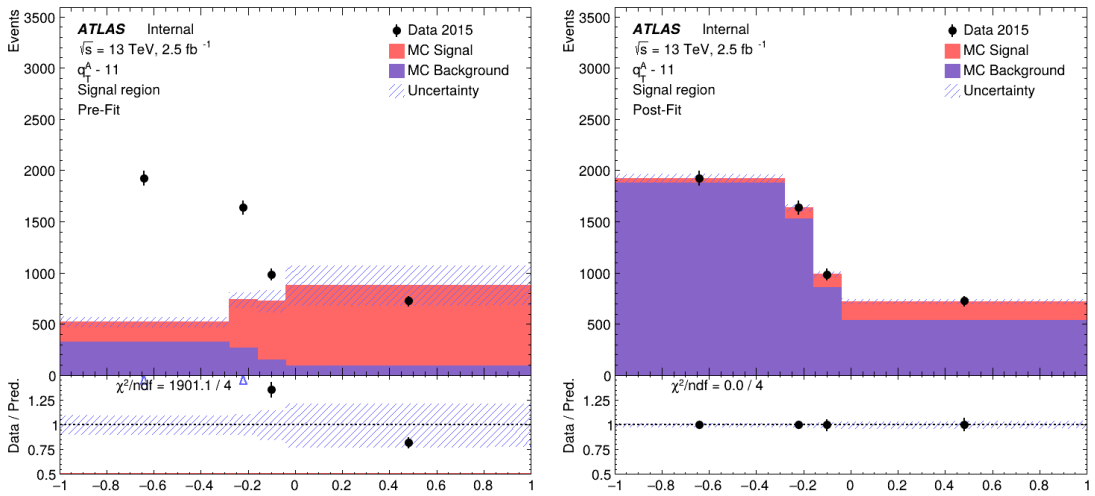
(a) Pre-fit (b) Post-fit

Figure E.9: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 9.



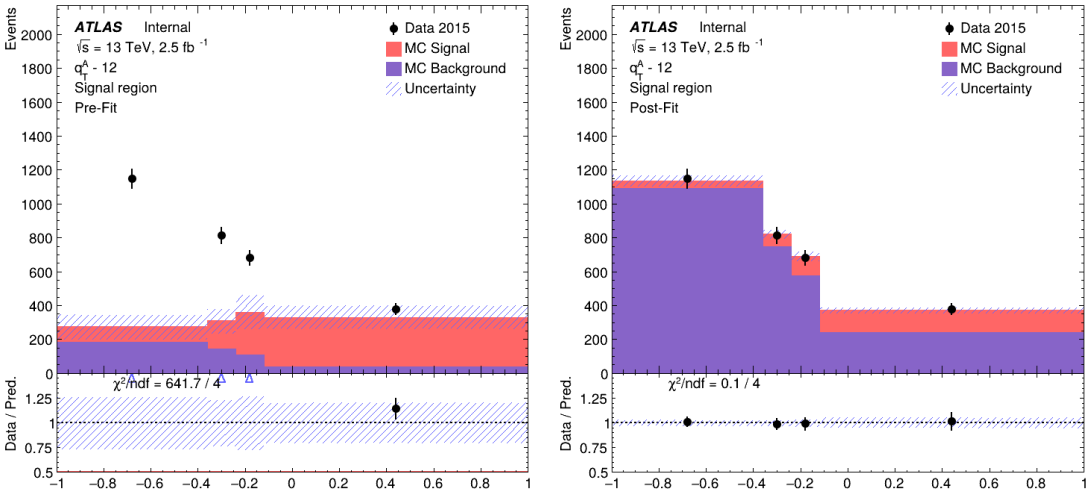
(a) Pre-fit (b) Post-fit

Figure E.10: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 10.



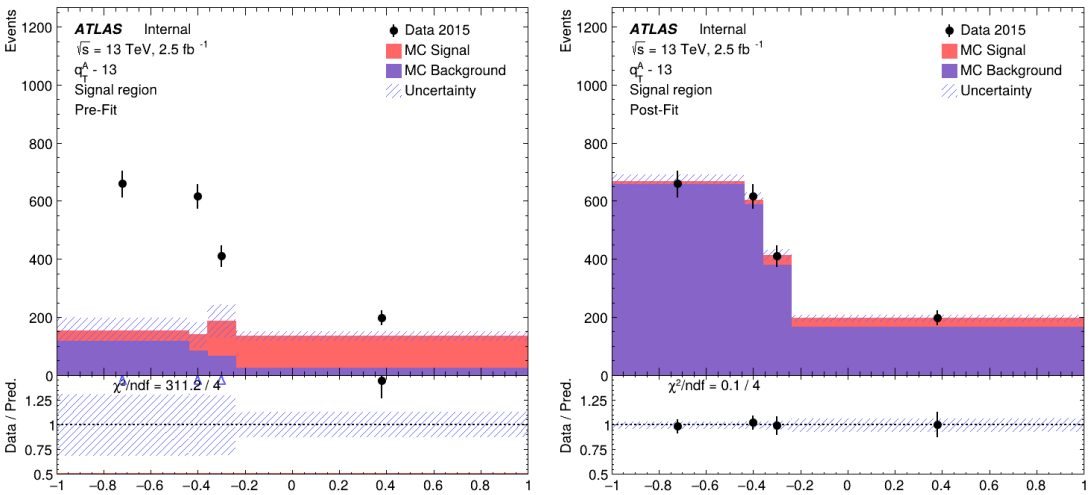
(a) Pre-fit (b) Post-fit

Figure E.11: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 11.



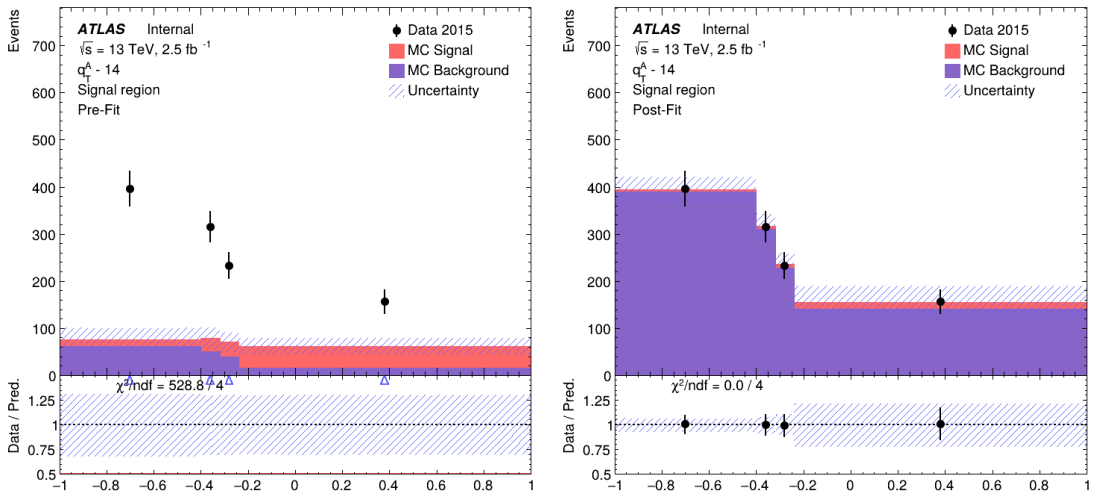
(a) Pre-fit (b) Post-fit

Figure E.12: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 12.



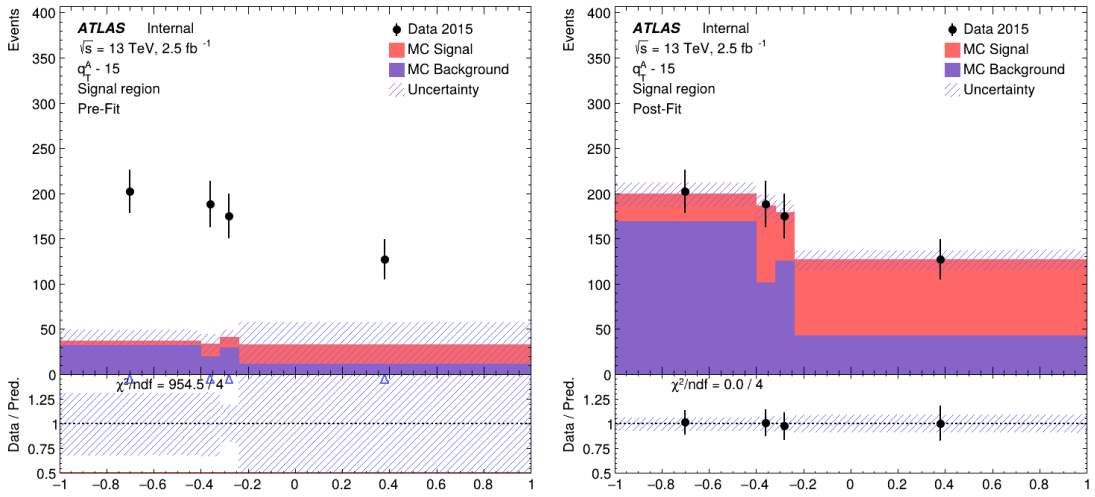
(a) Pre-fit (b) Post-fit

Figure E.13: `TRExFitter` fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 13.



(a) Pre-fit (b) Post-fit

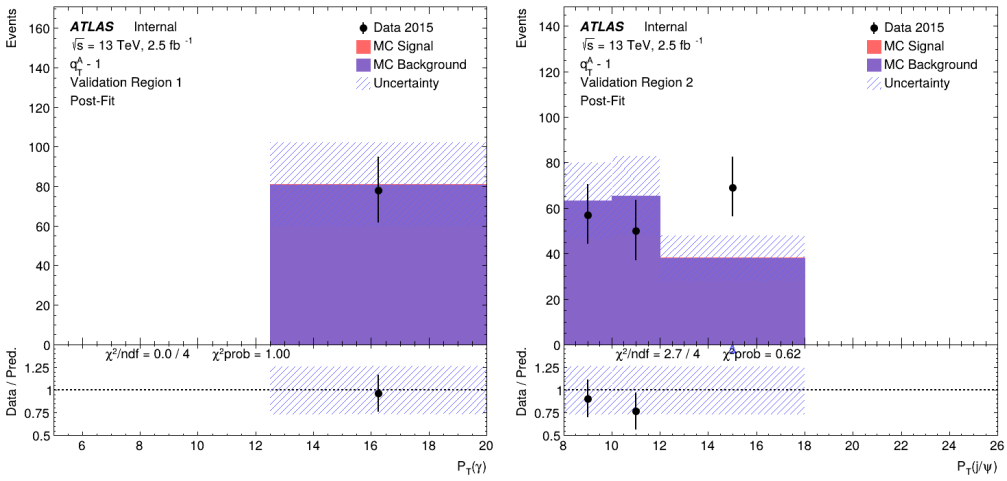
Figure E.14: TRExFitter fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 14.



(a) Pre-fit (b) Post-fit

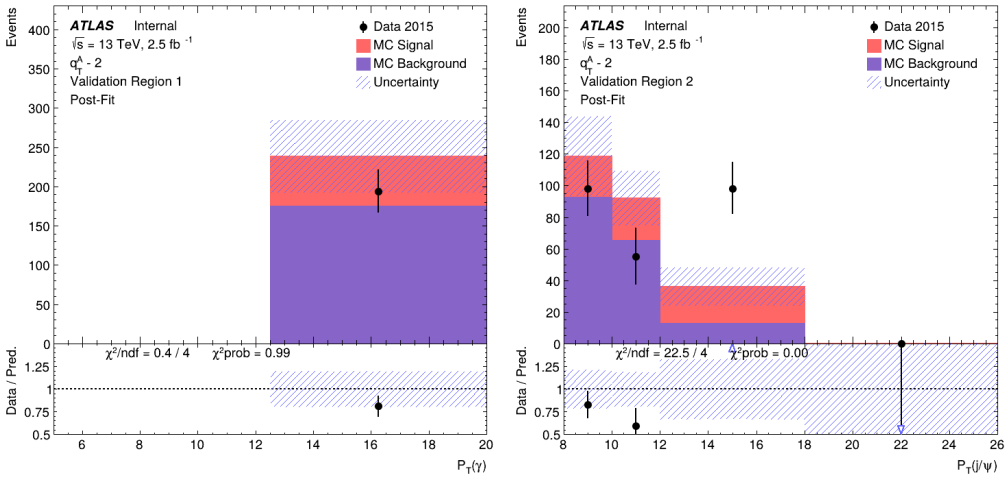
Figure E.15: TRExFitter fit plots for (a) pre-fit and (b) post-fit, in q_T^A bin 15.

F. Plots of Differential Fits in $p_T(J/\psi)$ and $p_T(\gamma)$



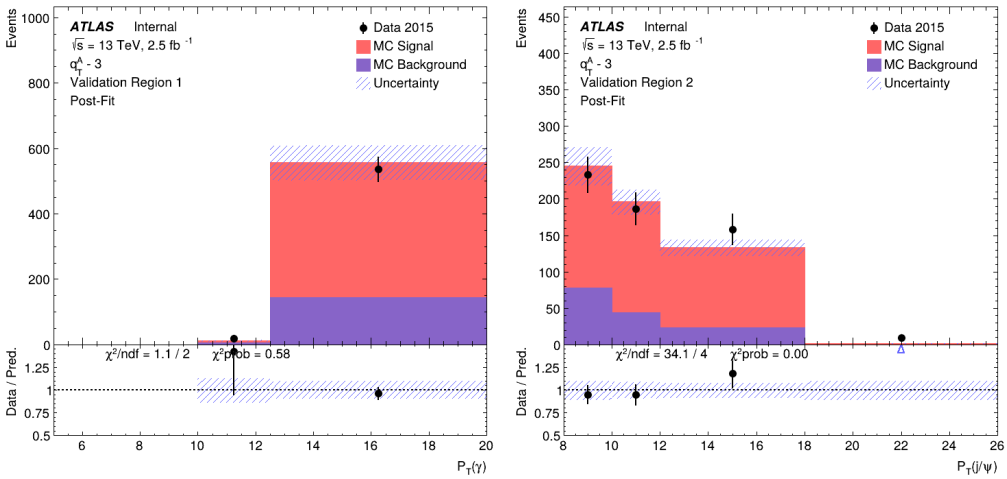
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.1: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 1.



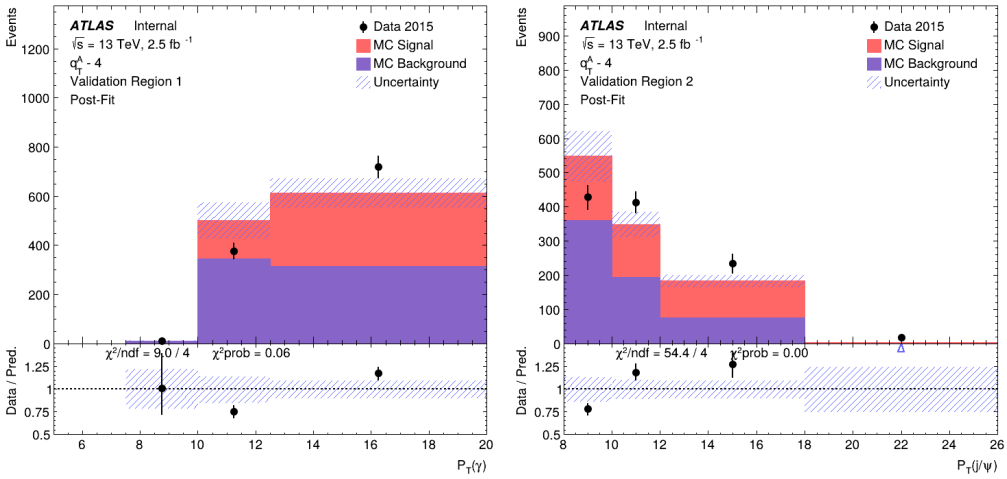
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.2: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 2.



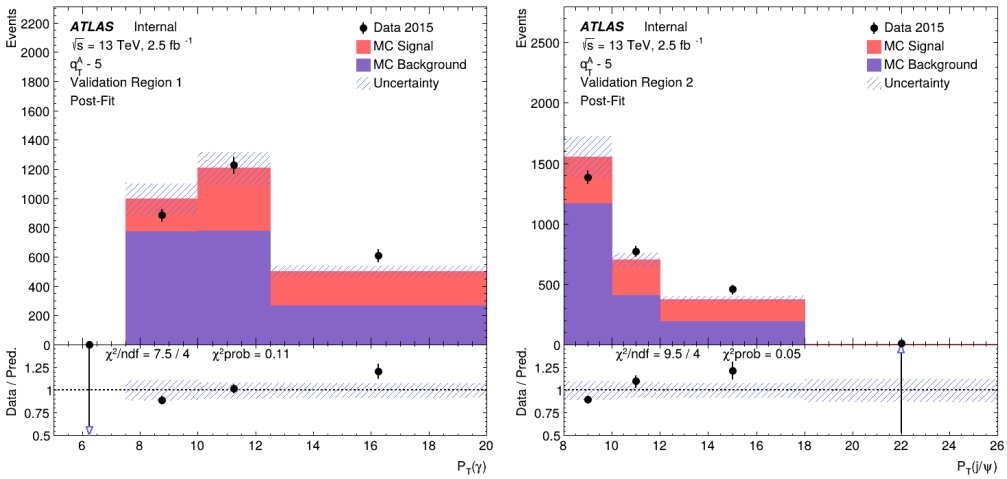
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.3: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 3.



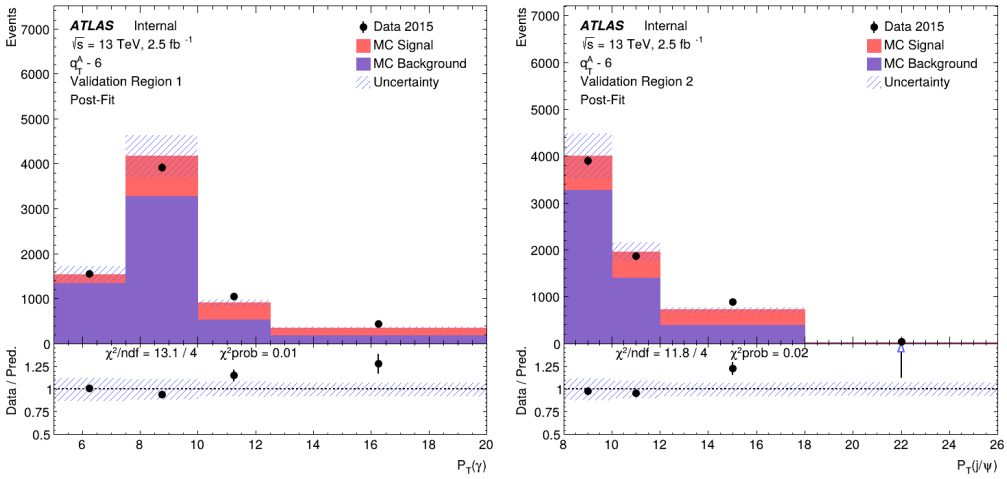
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.4: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 4.



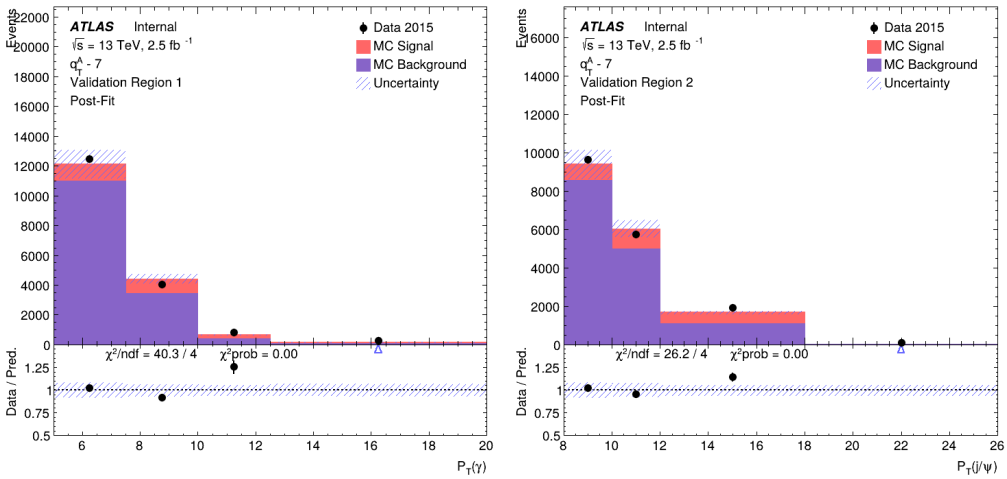
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.5: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 5.



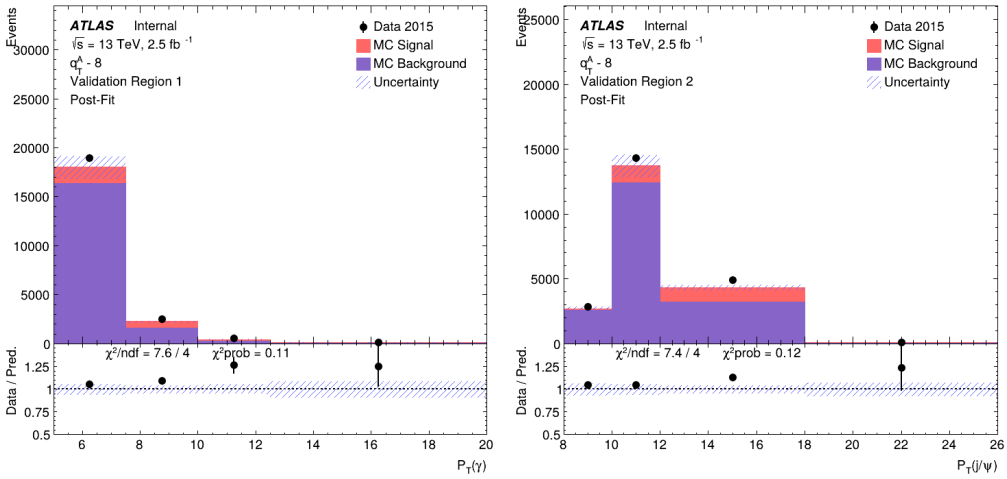
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.6: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 6.



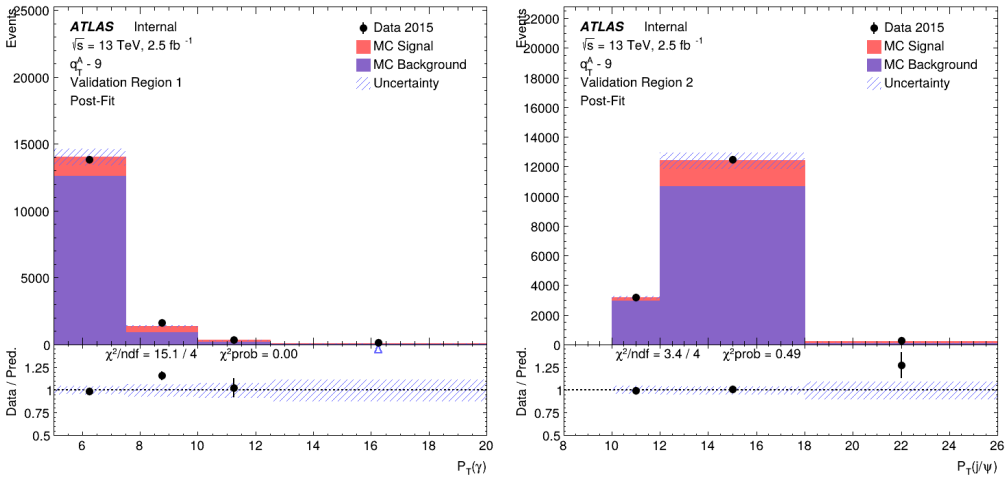
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.7: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 7.



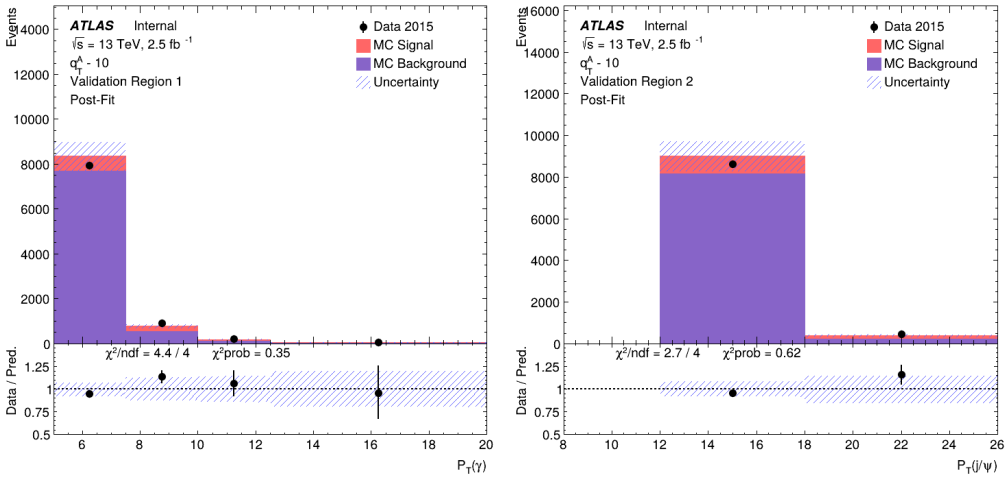
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.8: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 8.



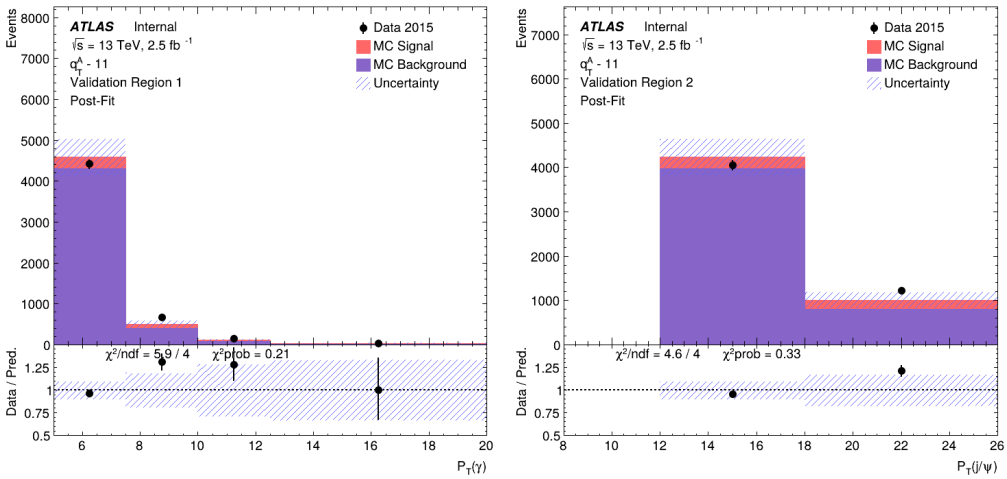
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.9: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 9.



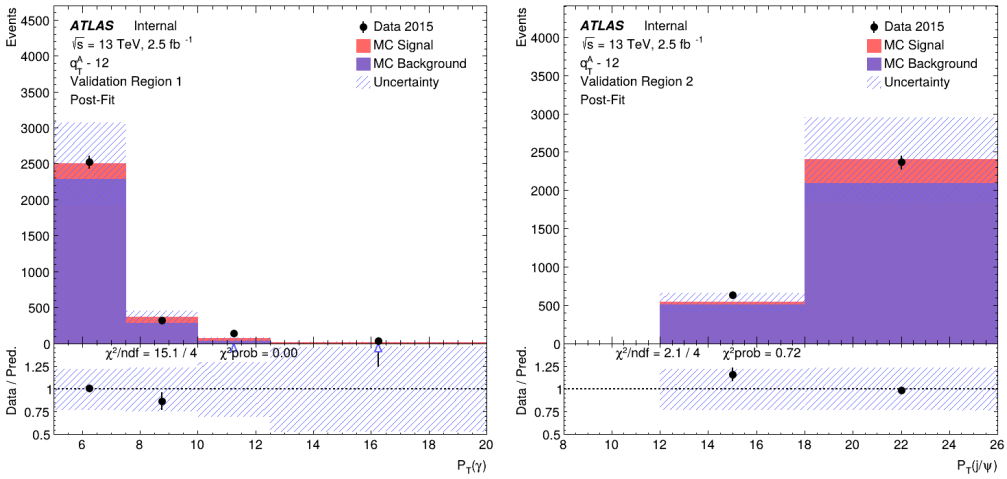
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.10: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 10.



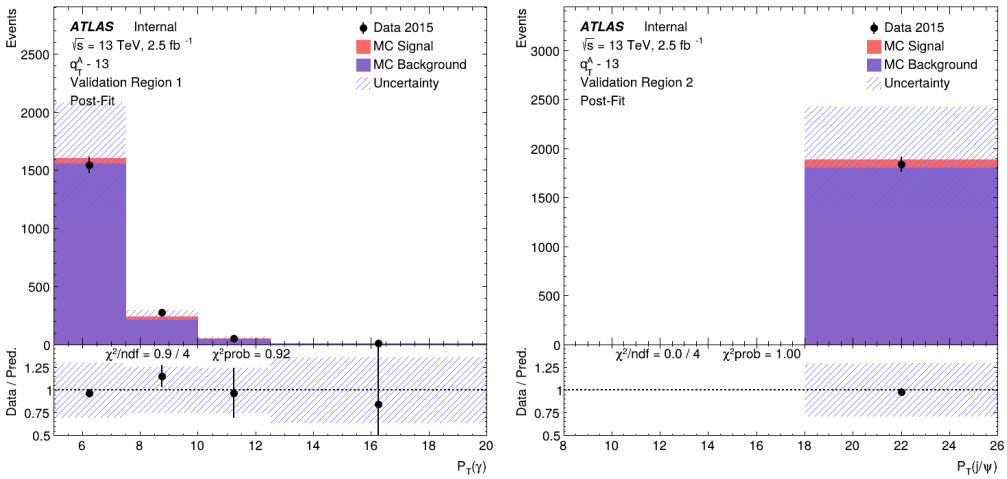
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.11: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 11.



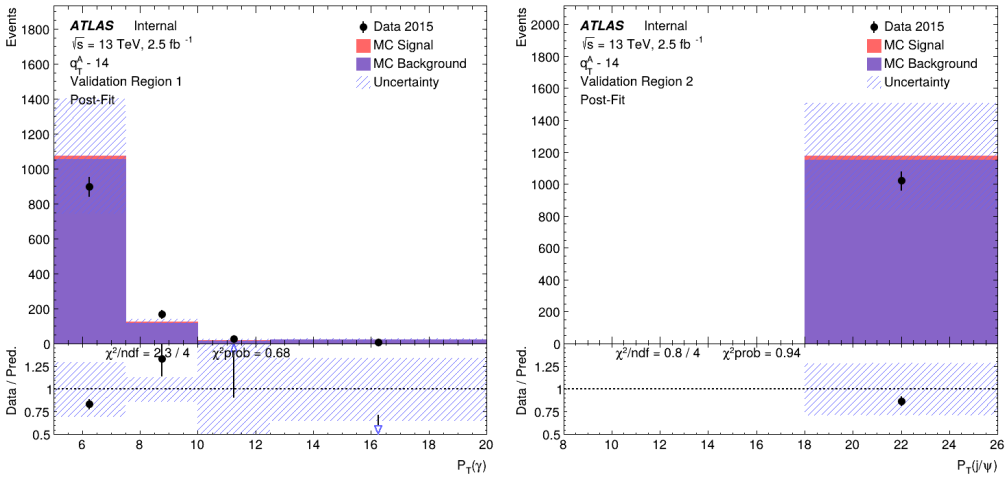
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.12: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 12.



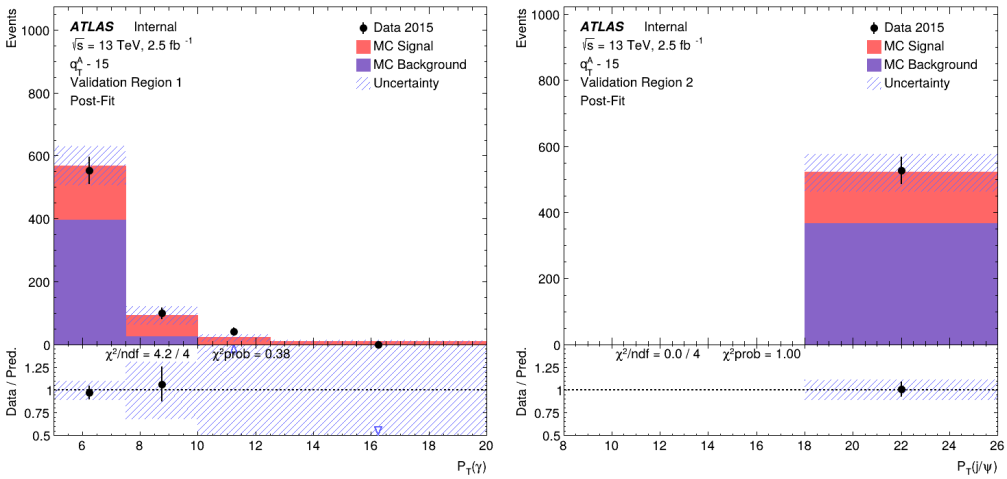
(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.13: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 13.



(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

Figure F.14: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 14.

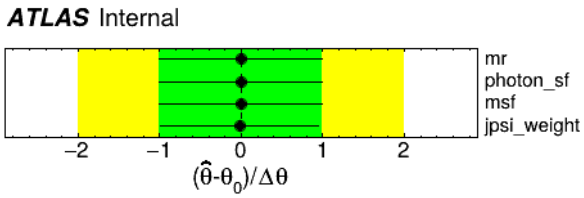


(a) $p_{T,\gamma}$ (b) $p_{T,\mu\mu}$

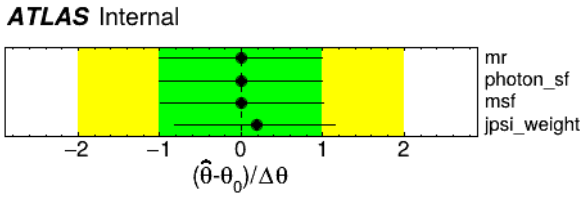
Figure F.15: Plots output from `TRExFitter` framework detailing the validation regions, post-fit (a) $p_{T,\gamma}$ and (b) $p_{T,\mu\mu}$ in q_T^A bin 15.

G. Plots of Nuisance Parameters Associated with Differential Fits

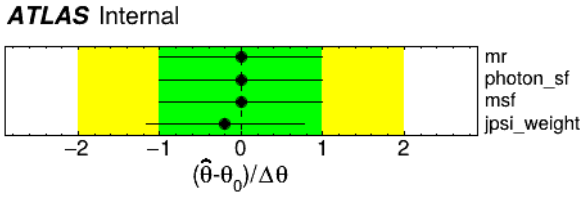
G.1 Nuisance Parameter Pulls



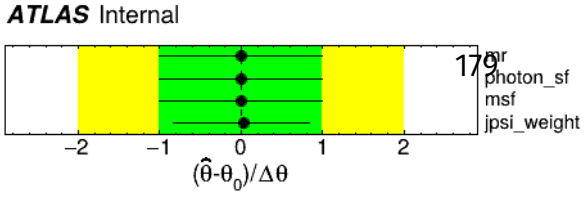
(a) $q_T^A - 1$



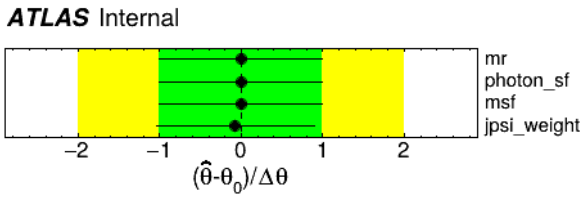
(b) $q_T^A - 2$



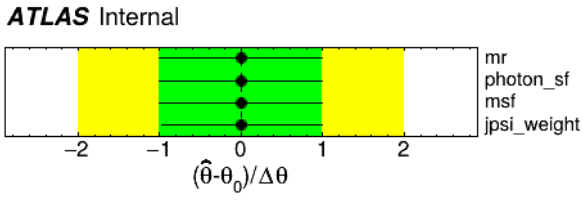
(c) $q_T^A - 3$



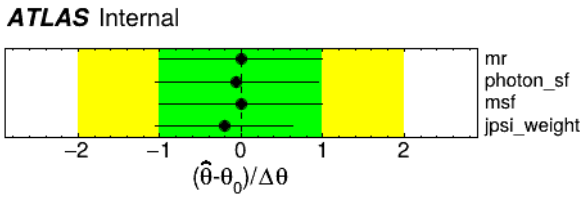
(d) $q_T^A - 4$



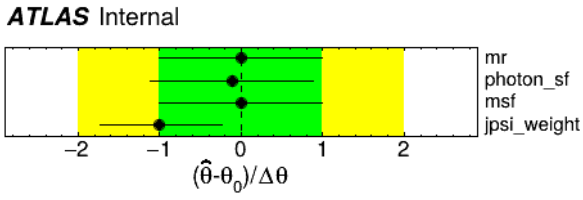
(e) $q_T^A - 5$



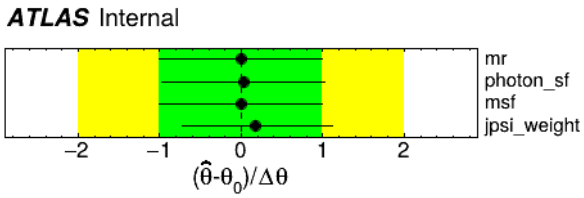
(f) $q_T^A - 6$



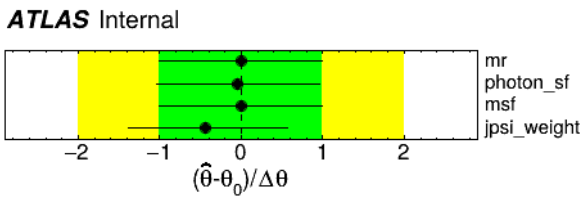
(g) $q_T^A - 7$



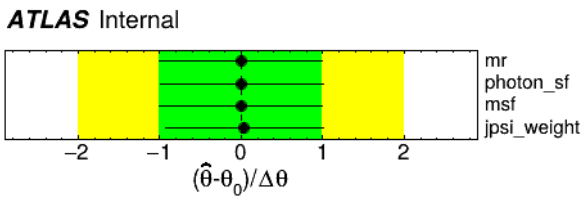
(h) $q_T^A - 8$



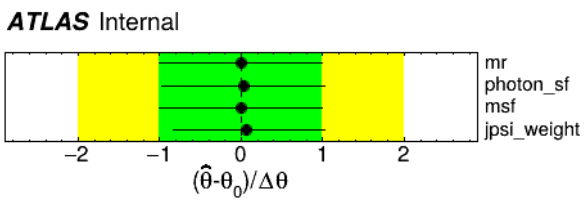
(i) $q_T^A - 9$



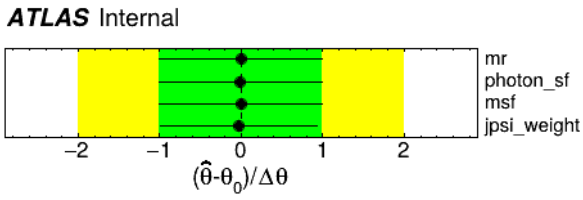
(j) $q_T^A - 10$



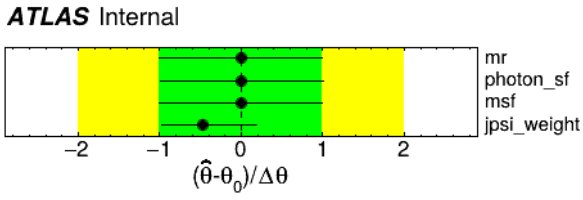
(k) $q_T^A - 11$



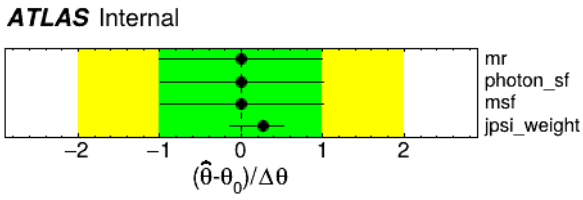
(l) $q_T^A - 12$



(m) $q_T^A - 13$



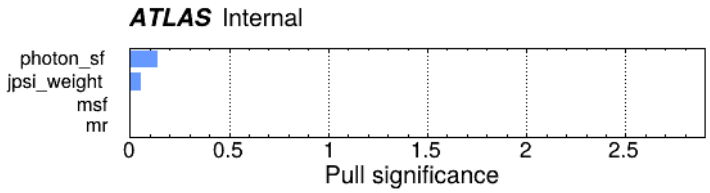
(n) $q_T^A - 14$



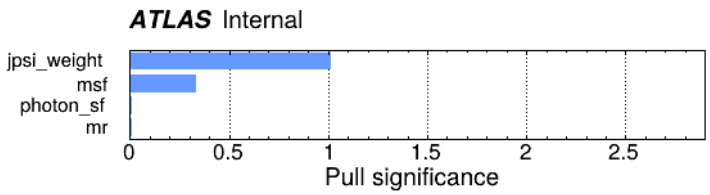
(o) $q_T^A - 15$

Figure G.1: Nuisance Parameter plots for all q_T^A bins

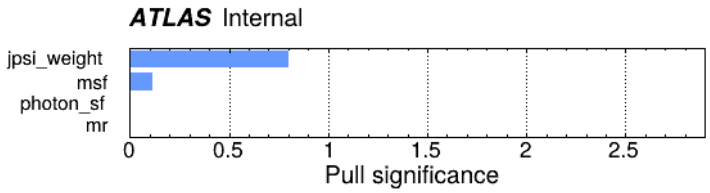
G.2 Nuisance Parameter Significances



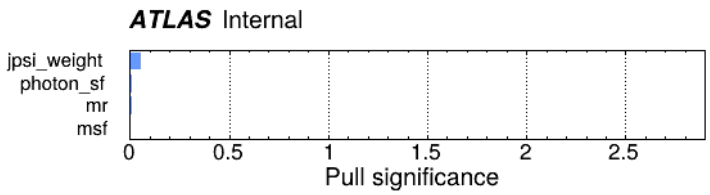
(a) $q_T^A - 1$



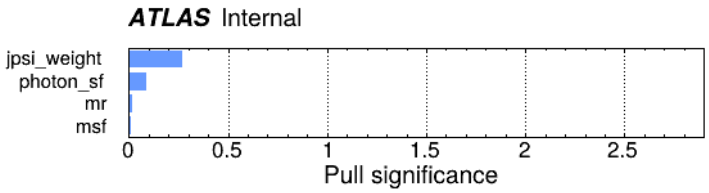
(b) $q_T^A - 2$



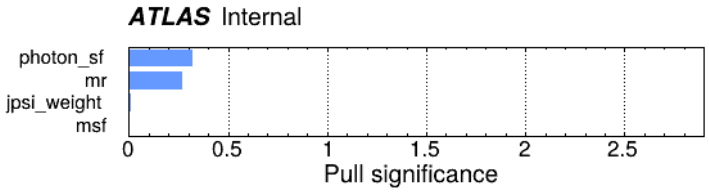
(c) $q_T^A - 3$



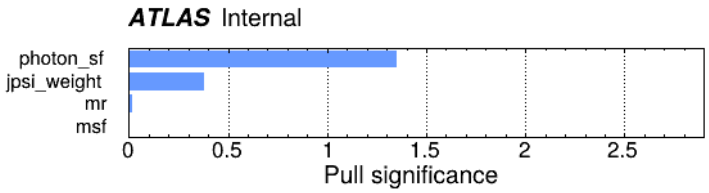
(d) $q_T^A - 4$



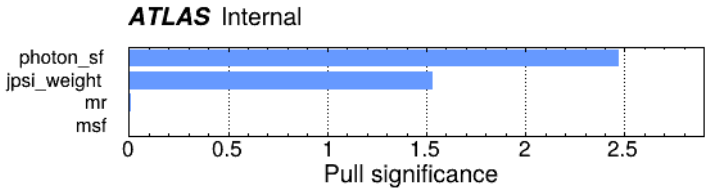
(e) $q_T^A - 5$



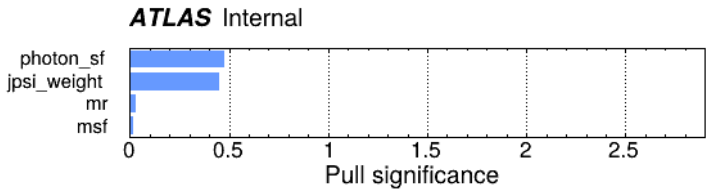
(f) $q_T^A - 6$



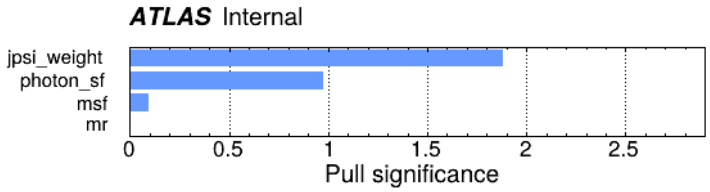
(g) $q_T^A - 7$



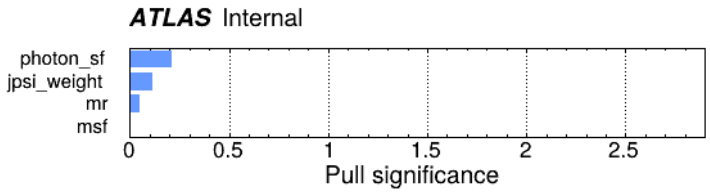
(h) $q_T^A - 8$



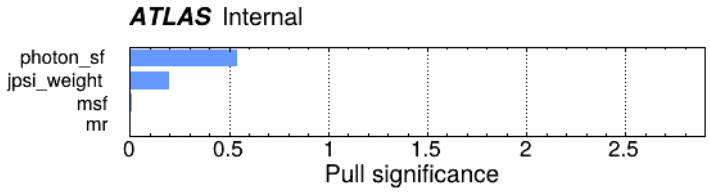
(i) $q_T^A - 9$



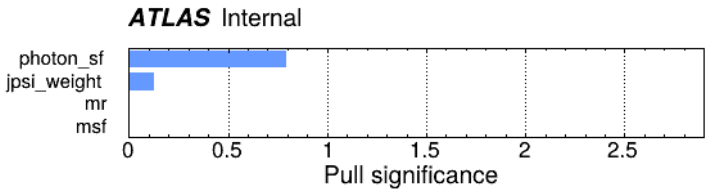
(j) $q_T^A - 10$



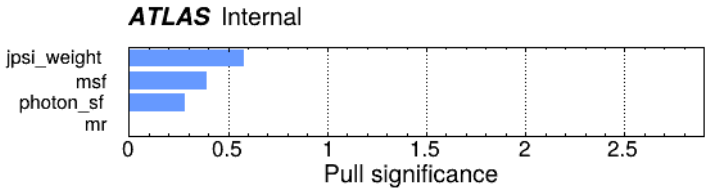
(k) $q_T^A - 11$



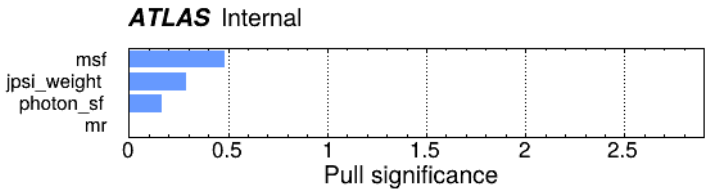
(l) $q_T^A - 12$



(m) $q_T^A - 13$



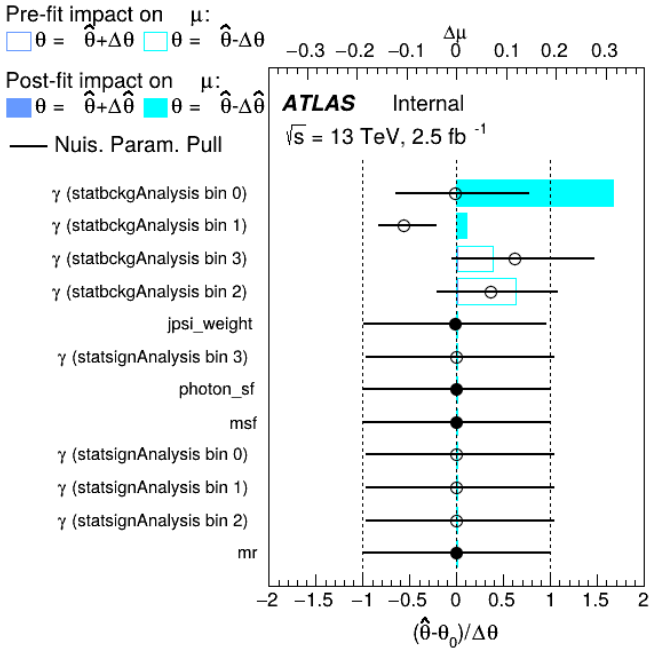
(n) $q_T^A - 14$



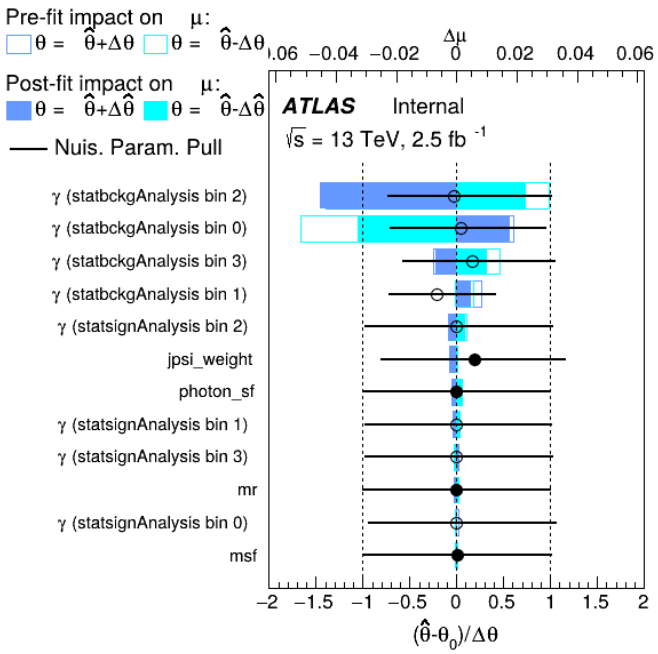
(o) $q_T^A - 15$

Figure G.2: Nuisance Parameter Pull Significances for all q_T^A bins

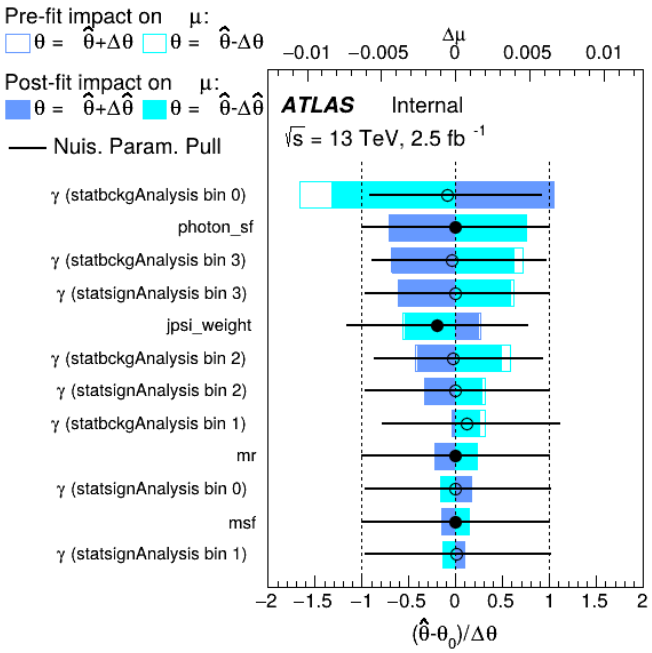
G.3 Nuisance Parameter Ranking Plots



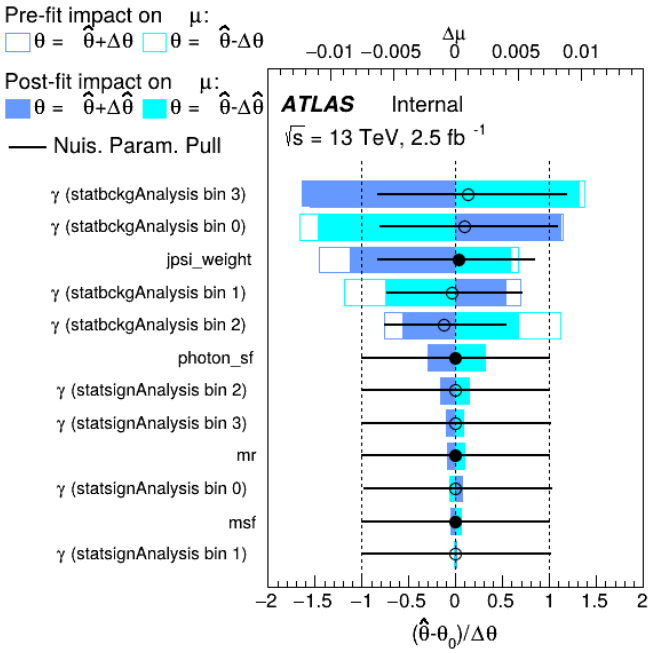
(a) $q_T^A - 1$



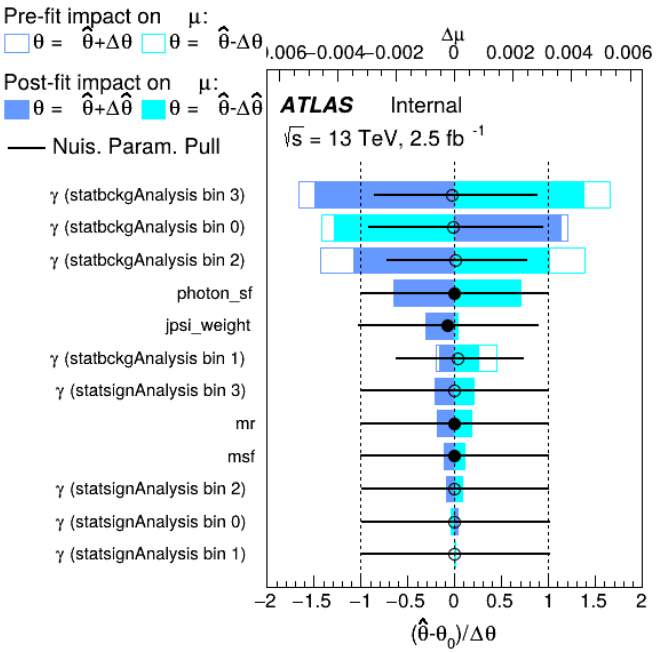
(b) $q_T^A - 2$


Pre-fit impact on μ:
θ = θ̂+Δθ
θ = θ̂-Δθ
Post-fit impact on μ:
θ = θ̂+Δθ̂
θ = θ̂-Δθ̂
Nuis. Param. Pull
ATLAS Internal
√s = 13 TeV, 2.5 fb -1
Δμ
−0.01 −0.005 0 0.005 0.01
γ (statbckgAnalysis bin 0)
photon_sf
γ (statbckgAnalysis bin 3)
γ (statsignAnalysis bin 3)
jpsi_weight
γ (statbckgAnalysis bin 2)
γ (statsignAnalysis bin 2)
γ (statbckgAnalysis bin 1)
mr
γ (statsignAnalysis bin 0)
msf
γ (statsignAnalysis bin 1)
−2 −1.5 −1 −0.5 0 0.5 1 1.5 2
(θ̂-θ0)/Δθ

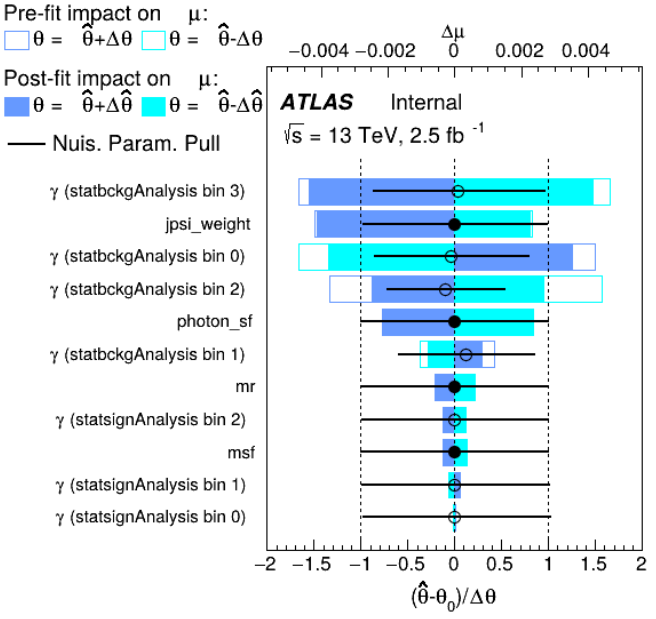
(c) $q_T^A - 3$


Pre-fit impact on μ:
θ = θ̂+Δθ
θ = θ̂-Δθ
Post-fit impact on μ:
θ = θ̂+Δθ̂
θ = θ̂-Δθ̂
Nuis. Param. Pull
ATLAS Internal
√s = 13 TeV, 2.5 fb -1
Δμ
−0.01 −0.005 0 0.005 0.01
γ (statbckgAnalysis bin 3)
γ (statbckgAnalysis bin 0)
jpsi_weight
γ (statbckgAnalysis bin 1)
γ (statbckgAnalysis bin 2)
photon_sf
γ (statsignAnalysis bin 2)
γ (statsignAnalysis bin 3)
mr
γ (statsignAnalysis bin 0)
msf
γ (statsignAnalysis bin 1)
−2 −1.5 −1 −0.5 0 0.5 1 1.5 2
(θ̂-θ0)/Δθ

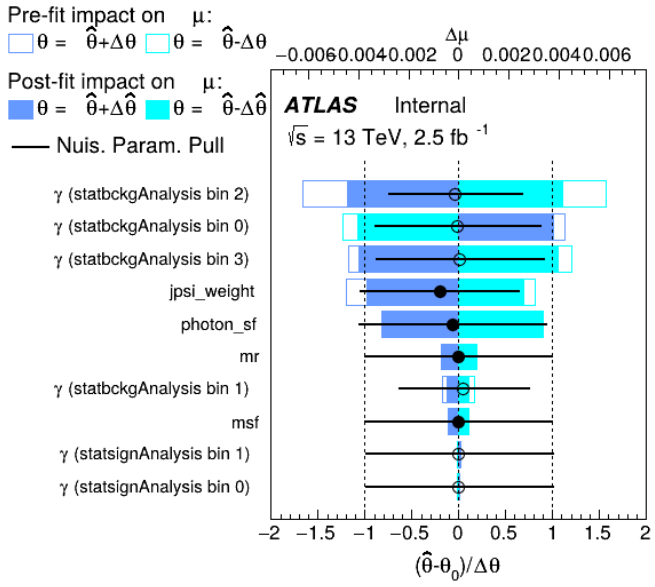
(d) $q_T^A - 4$


Pre-fit impact on μ:
θ = θ̂+Δθ
θ = θ̂-Δθ
Post-fit impact on μ:
θ = θ̂+Δθ̂
θ = θ̂-Δθ̂
Nuis. Param. Pull
ATLAS Internal
√s = 13 TeV, 2.5 fb -1
Δμ
γ (statbckgAnalysis bin 3)
γ (statbckgAnalysis bin 0)
γ (statbckgAnalysis bin 2)
photon_sf
jpsi_weight
γ (statbckgAnalysis bin 1)
γ (statsignAnalysis bin 3)
mr
msf
γ (statsignAnalysis bin 2)
γ (statsignAnalysis bin 0)
γ (statsignAnalysis bin 1)
(θ̂-θ0)/Δθ

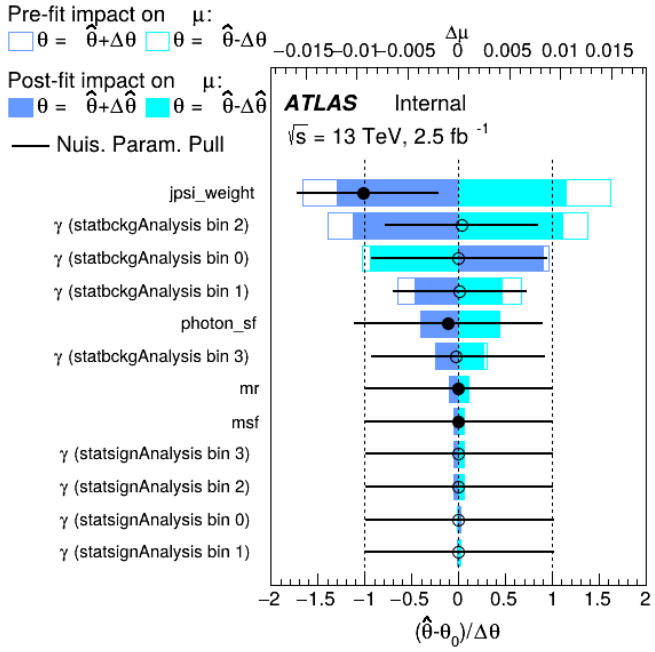
(e) $q_T^A - 5$


Pre-fit impact on μ:
θ = θ̂+Δθ
θ = θ̂-Δθ
Post-fit impact on μ:
θ = θ̂+Δθ̂
θ = θ̂-Δθ̂
Nuis. Param. Pull
ATLAS Internal
√s = 13 TeV, 2.5 fb -1
Δμ
γ (statbckgAnalysis bin 3)
jpsi_weight
γ (statbckgAnalysis bin 0)
γ (statbckgAnalysis bin 2)
photon_sf
γ (statbckgAnalysis bin 1)
mr
γ (statsignAnalysis bin 2)
msf
γ (statsignAnalysis bin 1)
γ (statsignAnalysis bin 0)
(θ̂-θ0)/Δθ

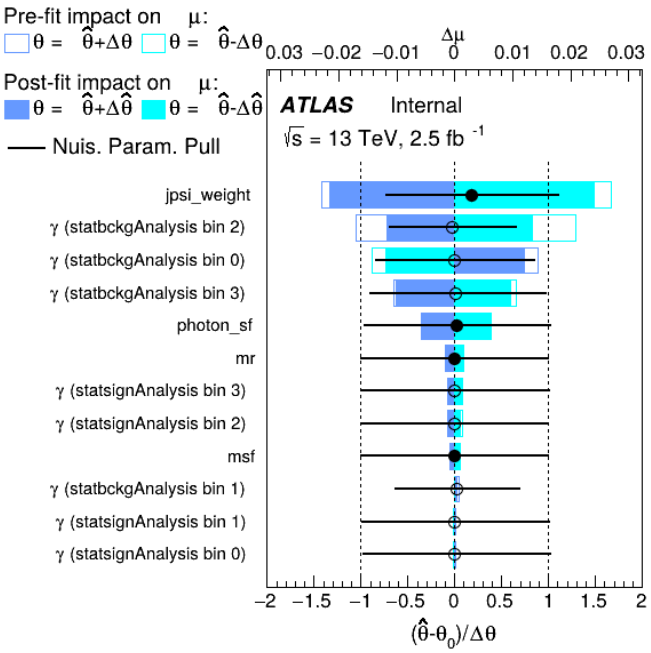
(f) $q_T^A - 6$


Pre-fit impact on μ:
θ = θ̂+Δθ θ = θ̂-Δθ
Post-fit impact on μ:
θ = θ̂+Δθ̂ θ = θ̂-Δθ̂
Nuis. Param. Pull
Δμ
−0.006 −0.004 −0.002 0 0.002 0.004 0.006
ATLAS Internal
√s = 13 TeV, 2.5 fb -1
γ (statbckgAnalysis bin 2)
γ (statbckgAnalysis bin 0)
γ (statbckgAnalysis bin 3)
jpsi_weight
photon_sf
mr
γ (statbckgAnalysis bin 1)
msf
γ (statsignAnalysis bin 1)
γ (statsignAnalysis bin 0)
−2 −1.5 −1 −0.5 0 0.5 1 1.5 2
(θ̂-θ0)/Δθ

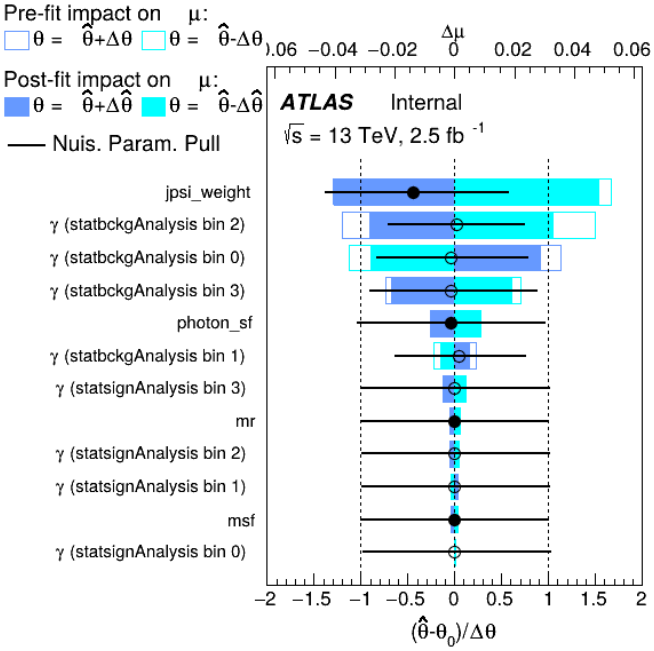
(g) $q_T^A - 7$


Pre-fit impact on μ:
θ = θ̂+Δθ θ = θ̂-Δθ
Post-fit impact on μ:
θ = θ̂+Δθ̂ θ = θ̂-Δθ̂
Nuis. Param. Pull
Δμ
−0.015 −0.01 −0.005 0 0.005 0.01 0.015
ATLAS Internal
√s = 13 TeV, 2.5 fb -1
jpsi_weight
γ (statbckgAnalysis bin 2)
γ (statbckgAnalysis bin 0)
γ (statbckgAnalysis bin 1)
photon_sf
γ (statbckgAnalysis bin 3)
mr
msf
γ (statsignAnalysis bin 3)
γ (statsignAnalysis bin 2)
γ (statsignAnalysis bin 0)
γ (statsignAnalysis bin 1)
−2 −1.5 −1 −0.5 0 0.5 1 1.5 2
(θ̂-θ0)/Δθ

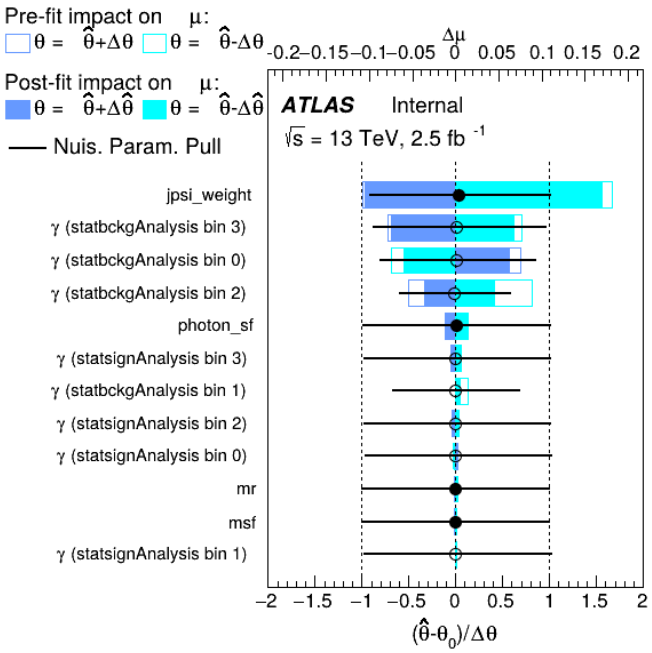
(h) $q_T^A - 8$



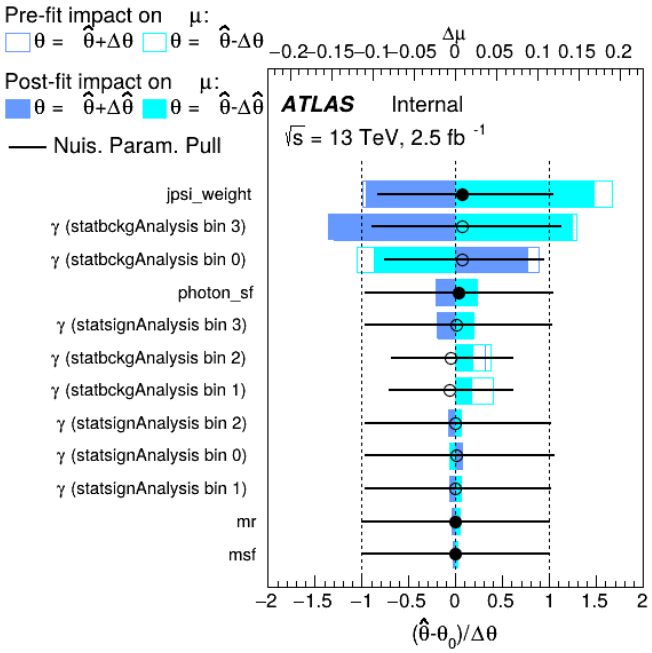
(i) $q_T^A - 9$



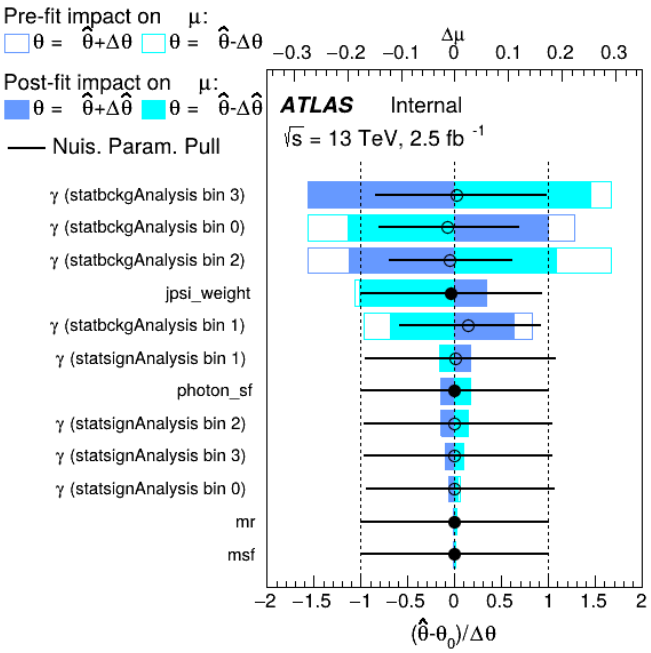
(j) $q_T^A - 10$



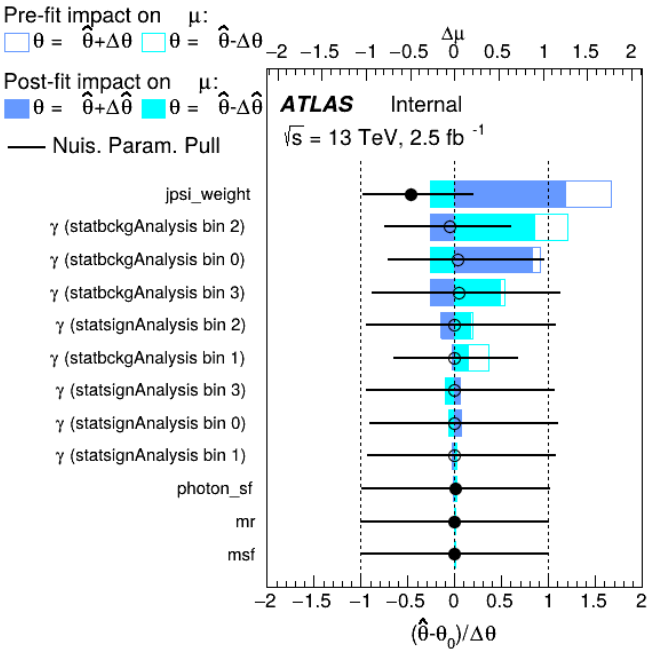
(k) $q_T^A - 11$



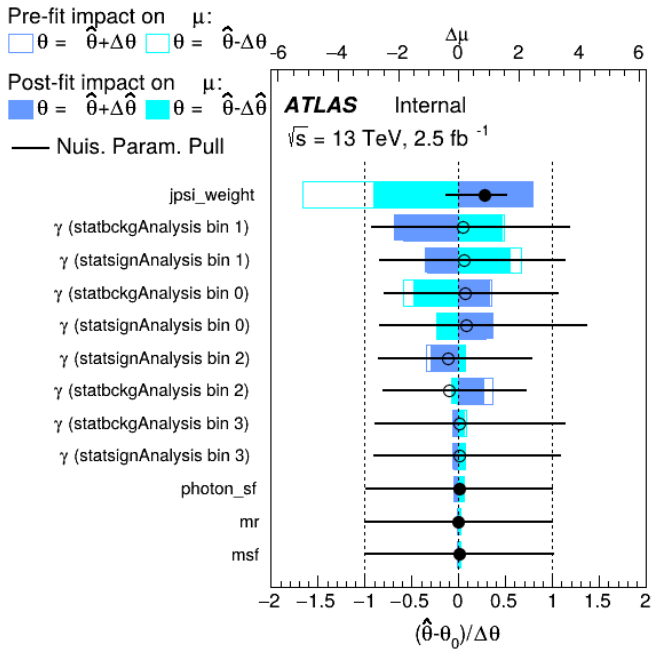
(l) $q_T^A - 12$



(m) $q_T^A - 13$



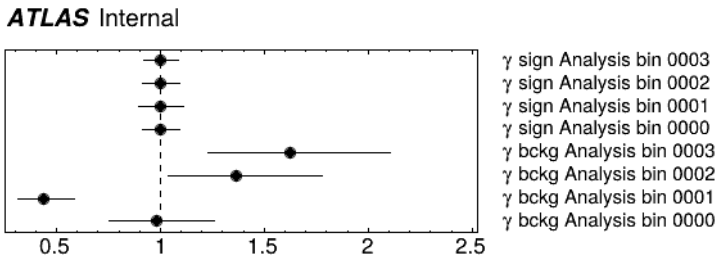
(n) $q_T^A - 14$



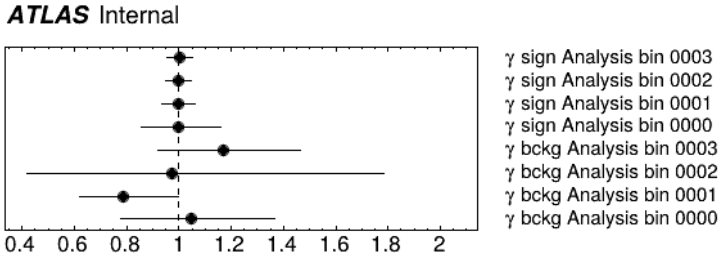
(o) $q_T^A - 15$

Figure G.3: Nuisance Parameter Rankings for all q_T^A bins.

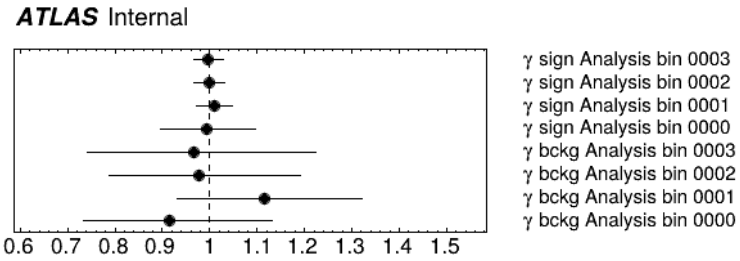
G.4 Gammas



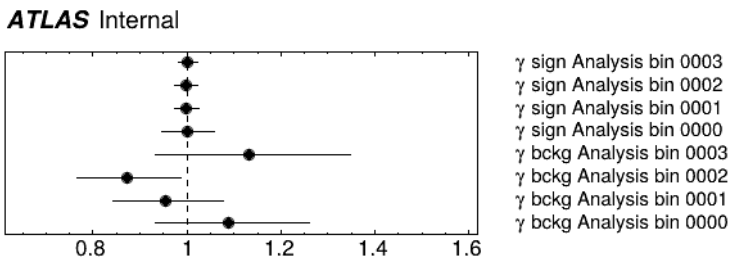
(a) $q_T^A - 1$



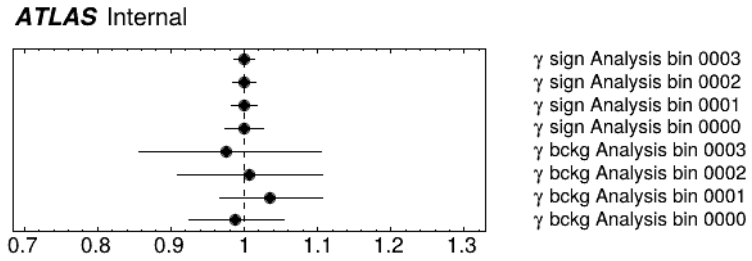
(b) $q_T^A - 2$



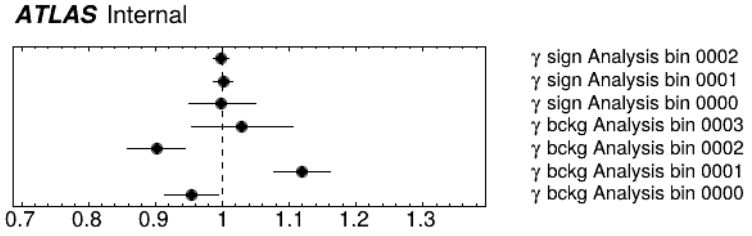
(c) $q_T^A - 3$



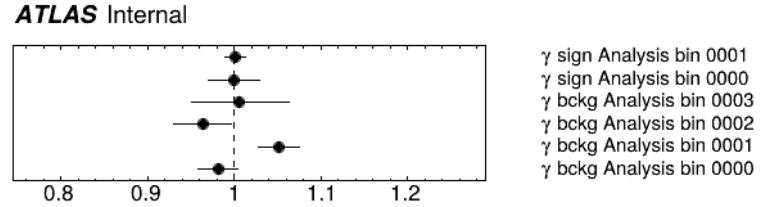
(d) $q_T^A - 4$



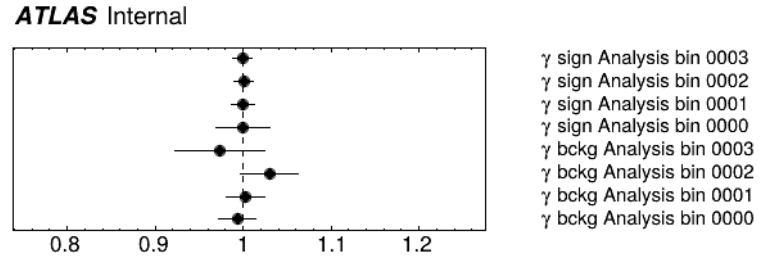
(e) $q_T^A - 5$



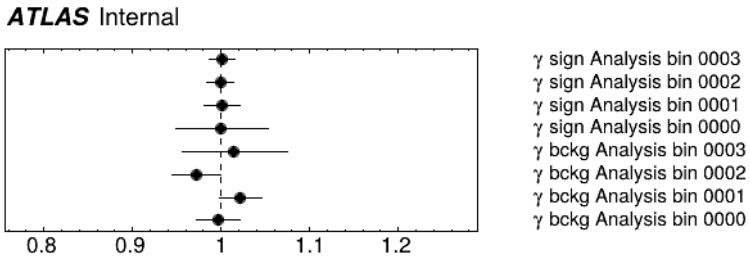
(f) $q_T^A - 6$



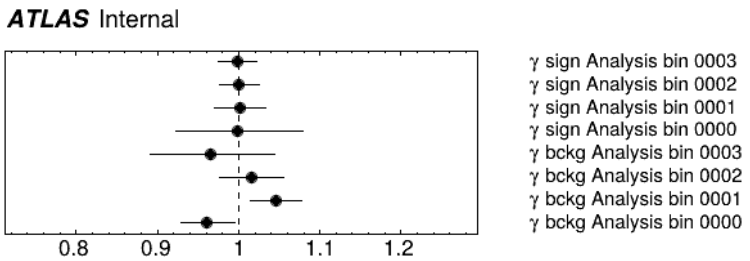
(g) $q_T^A - 7$



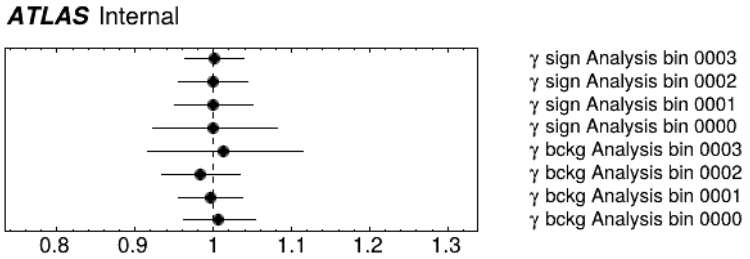
(h) $q_T^A - 8$



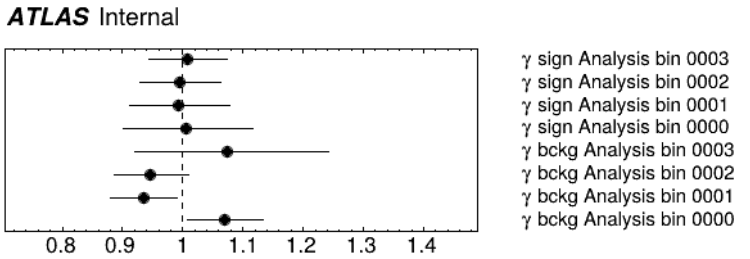
(i) $q_T^A - 9$



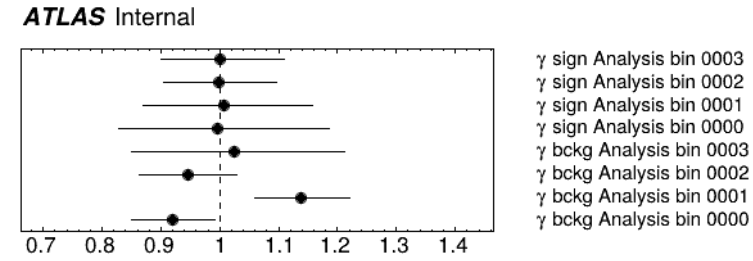
(j) $q_T^A - 10$



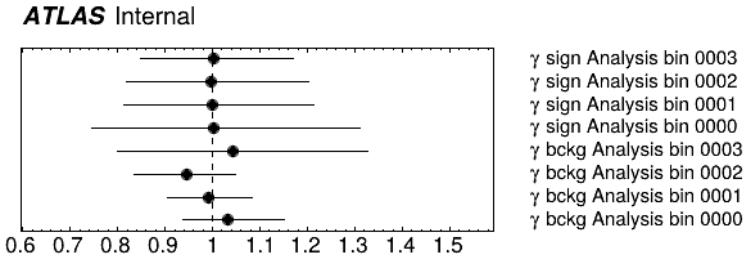
(k) $q_T^A - 11$



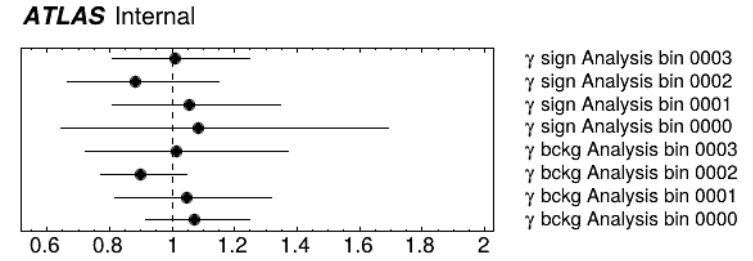
(l) $q_T^A - 12$



(m) $q_T^A - 13$



(n) $q_T^A - 14$



(o) $q_T^A - 15$

Figure G.4: Gamma pull plots for all q_T^A bins