Strange particle production in $t\bar{t}$ final states in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC

Facultad de Ciencias
Departamento de Física Teórica
Universidad Autónoma de Madrid

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor in Physics by

Sergio Calvente López

Supervisor: Fernando Barreiro Alonso

7 de Octubre del 2019

Producción de partículas extrañas en estados finales $t\bar{t}$ en colisiones protón protón a $\sqrt{s} = 7$ TeV con el detector ATLAS en el LHC

Facultad de Ciencias
Departamento de Física Teórica
Universidad Autónoma de Madrid

Memoria presentada para la obtención del título de Doctor en Ciencias Físicas por

Sergio Calvente López

Director: Fernando Barreiro Alonso

7 de Octubre del 2019

A mi familia.
Por ser siempre el mejor apoyo posible.

"Science and industry, knowledge and application, discovery and practical realization leading to new discoveries, cunning of brain and of hand, toil of mind and muscle – all work together. Each discovery, each advance, each increase in the sum of human riches, owes its being to the physical and mental work of the past and the present. By what right can anyone whatever appropiate the least morsel and say – This is mine, not yours?"[1]

Piotr Kropotkin, *The Conquest of Bread.*

"We are at the very beginning of time for the human race. It is not unreasonable that we grapple with problems. But there are tens of thousands of years in the future. Our responsibility is to do what we can, learn what we can, improve the solutions, and pass them on."[2]

Richard P. Feynman

[1] "Ciencia e industria, saber y aplicación, descubrimiento y realización práctica que conduce a nuevas invenciones, trabajo o cerebral y trabajo manual, idea y labor de los brazos, todo se enlaza. Cada descubrimiento, cada progreso, cada aumento de la riqueza de la humanidad, tiene su origen en el conjunto del trabajo manual y cerebral, pasado y presente. Entonces, con qué derecho alguien puede apropiarse la menor partícula de ese inmenso todo y decir: 'Esto es mío y no vuestro'"

[2] "Estamos al principio de los tiempos para la humanidad. No es poco razonable que lidiemos con problemas. Pero quedan decenas de miles de años por delante. Es responsabilidad nuestra hacer lo que podamos, aprender lo que podamos, mejorar las soluciones y transmitirlas a las generaciones siguientes."

Acknowledgements - Agradecimientos

This thesis would have not been possible without the help of many people. Firstly, I would like to thank my supervisor, and I hope that after all these years also friend, Fernando Barreiro. He made this thesis possible with his enormous help, advice and hard work.

I would like to thank all the people in the UAM ATLAS team: Juan Terrón, Claudia Glasman, José del Peso, and Almudena Montiel for their help at both professional and personal levels. And the other PhD students with whom I have shared many moments of work and leisure: Manu, Dani, Miguel, Héctor, Josu, Ana and last but not least, Javier LLorente, who has been like my second supervisor, and has helped me through the completion of this thesis.

I would like to mention all the Top and Standard Model conveners I met during these years: Kevin Finelli, Deepak Kar, Ulla Blumenschein, Matthias Schott, Jan Kretzschmar, James W. Monk, Rafal Staszewski, Frederic Deliot, Reinhard Schwienhorst, Elizaveta Shabalina, Wolfgang Wagner, Veronique Boisvert, James W. Howarth, Dimitris Varouchas, Michele Faucci Giannelli and Tom Neep. All of them contributed in many ways to shape the final form of the analysis presented in this thesis. The chairman of the Editorial Board, Thorsten Kuhl, deserves my wholehearted thanks.

I would not like to forget the ATLAS LAr team, in particular Luis Hervás, who was so helpful during my stays at CERN, Pavol Strizenec who supervised my authorship qualification tasks, and the members of the LADIeS team, specially Benjamin Trocme, Emma Sian Kuwertz and Clement Camincher.

I thank CERN for the very successful operation of the LHC, as well as all the members of the ATLAS collaboration responsible for the efficient performance of the detector.

Además de todos estos agradecimientos, quiero recordar ahora a todas las personas que no estaban directamente relacionadas con mi trabajo de tesis, pero han sido importantes de una manera u otra a lo largo de estos años. Quiero empezar por las personas con las que he compartido mis estancias en el CERN, y que son responsables de que hayan sido tan buena experiencia. Por esos viernes de cervezas en la cafetería después de una dura semana. Esos cafés fuera en cuanto salía un poco el sol. Las barbacoas y baños en el lago, las escapadas a la montaña o un simple paseo por Ginebra. Gracias a los que siempre necesitaban un *coffee break*: Mer, Roger, Ceci, Andrea, Javi Jiménez, Ana, Josu, Neus, Javi Montejo, Carlos (también por introducirme en el mundo de Bujinkan por lo que siempre te estaré agradecido), Pili, Santi, Emma, Jordi, Mar, Sabrina, Jose Luis, Daniel, María, Bea, y mucha más gente que seguro que me dejo, pero también estuvo ahí. También a mis compañeros de piso de *Champs-Fréchets* Javi, Jorge y Maria.

Como no mencionar aquí a mis compañeras que amenizaron tantas comidas en la UAM (aunque *una beca no dé casi ni para comer*), y fuera de ella, viajes incluidos. Tanto a las antiguas conocidas del grado, que no se si se quedaron tantos años en la universidad por amor a la Física o por el cariño que nos hemos cogido. Porque nos sigamos viendo muchos años más, y ya van 10: Iris, Laura S., Laura R., María, Aitor, Alex, Elena, David, Danky, Gus, Juan Fran, Javi Cuerda, y la gente de la carrera que no se quedó pero dejó su huella.

Y como ya he dicho que las condiciones para las doctorandas no siempre son las mejores o incluso dignas, quiero agradecer a toda la gente que durante este tiempo ha sacado tiempo que no tenía para luchar por lo que es justo. A la gente de la sección sindical de CGT Enseñaza de la UAM, Adrián, Jacobo, Ángel, Rocío, María, Yolanda, Noelia, Lorena, Mario, Miguel, Amaia, Nagore, Ricardo, José, Marimar, y toda la gente que se haya pasado en algún momento aunque no se haya quedado. Porque además de compañeras de lucha, hemos sido casi una familia y un importante apoyo moral. Espero que sigáis con este trabajo combativo y que mantengamos los vínculos de cuidados y afectivos.

Y dentro de esta lucha no puedo olvidarme de todas las doctorandas de la Asamblea Dignidad Predoctoral UAM, que no conseguimos todos los objetivos, pero sí una buena victoria parcial con el nuevo *EPIF*, que ha mejorado las condiciones de mucha gente. Queda todo por hacer, pero ya es el turno de las siguientes. Muchas gracias a las

Irenes, Itziar, Inés, Ana, Aida, Sheila, Mónica, Dani, Carlos, Edu, Trunkel, Marta, Leo, Olga, Pablo, Lucía, Gonzalo, y toda la gente que apoyó o se pasó por las asambleas.

No quiero olvidarme de toda la gente con la que he compartido, comparto y compartiré (espero) otras luchas (dentro y fuera de la universidad), y mucho más. A Sol, Álvaro, Adri, Arturo, Lucía, Duran, Raúl, Andrés, Miguel, Diego, Sthela, Molly, Irene, los Javis, y mucha más gente que me dejo pero sintiéndolo mucho no tengo espacio infinito en estas líneas.

Hablando de luchas, quiero agradecer a mi maestro, David, y a todos mis compañeros del dojo Churyu: Camilo, Jorge, Ana, Isma, Juanjo, Fran, Ale, Ramón y Carlos. Y al resto de gente de Bujinkan, con la que he coincidido en entrenamientos y Taikais, y de la que he aprendido tanto. Gracias por compartir vuestro tiempo, para aprender juntos y crecer como personas. Ese tiempo me ha sido muy necesario para mantener el equilibrio y la calma en momentos tan estresantes como puede ser a veces una tesis.

Para ir acabando, volvemos al principio, a los que llevan ahí desde siempre, porque pueden ser esa *gentuza* pero se les quiere mucho: Alicia, Sergi, Edu, Irene, Robin, Iván, Berny, Alex y Laura. A toda la gente que también están presentes aunque nos veamos menos a menudo: Juanjo, Adri, Jorge, Yalda, Cris, Ire, Luis, Antuán, Navid, Javi, y a todo el que se me haya pasado en estas líneas pero hemos pasado cosas importantes juntos. Perdonadme el olvido, pero es fruto de los tensos momentos en los que escribo esto.

Por último, a mi familia, en especial mi padre, mi madre, mi hermana, que me han criado y por todo el amor que me han dado. A mis abuelas, que ya no están para verlo, pero que no sería quién soy sin lo que aprendí de ellas. Mis tíos y tías, mis primos y primas, con los que hemos pasado tantas celebraciones, vacaciones, fines de semana en el camping, la parcela, o en cualquier sitio con tal de juntarse. Y te dejo para el final, pero no por ello eres menos importante, más bien al revés, Ale. Llegaste hace menos, pero has sido un apoyo imprescindible en esta última etapa, espero que sigas aquí mucho tiempo y que continúen las aventuras.

Madrid, Octubre 2019

Introducción

La construcción del LHC en el CERN (Ginebra) tiene como finalidad, de una parte la búsqueda del bosón de Higgs, de otra permitir comprobaciones experimentales del modelo estándar (SM) a energías y con precisión nunca hasta ahora alcanzadas y finalmente llevar a cabo búsquedas más allá del modelo estándar (BSM). Estas últimas han permitido empujar los límites inferiores para las masas de partículas predichas en la mayoría de los modelos BSM por encima del TeV. Dos grandes detectores multipropósito, ATLAS y CMS, han sido construidos para llevar a cabo este ambicioso programa.

El descubrimiento del bosón de Higgs supone sin duda alguna, uno de los hitos más importantes en la historia de la Física de Partículas Elementales. Otro de los campos donde la precisión alcanzada no tiene parangón es el estudio de los estados finales con quarks top producidos a pares o en solitario.

Una de las posibles contribuciones que puede aportar la experimentación en el LHC en el campo de la Física del quark top es la de medir directamente los elementos de la matriz de Cabibbo-Kobayashi-Maskawa que involucran a este quark, en particular V_{ts}. Ali et al (ABL, Phys. Lett. B 693 (2010) 44-51) han publicado un estudio en el que proponen un método para la medida de este elemento de matriz. Según estos autores sería posible distinguir los estados finales $t\bar{t} \rightarrow WbWs$ del 'fondo dominante' $t\bar{t} \rightarrow WbWb$ estudiando la producción de partículas extrañas neutras en los chorros observados en los estados hadrónicos $t\bar{t}$. En efecto las partículas extrañas neutras serían 'leading' en los chorros asociados a quarks s y serían subleading en los chorros iniciados por quarks b que se desintegran secuencialmente según $b \rightarrow c \rightarrow s$. Este estudio propone el uso de un análisis multivariado en el que junto a la producción de partículas extrañas se utilizan otras variables con poder discriminatorio entre quarks b y quarks s como la existencia de posibles vértices secundarios o la de 'soft leptons'. Los análisis multivariados dependen naturalmente de las simulaciones Monte Carlo utilizados para el proceso de 'training'.

El estudio al que hacemos referencia es cualitativamente correcto, cuantitativamente sin embargo, al ser un estudio a nivel generador, no es del todo realista. El objeto de la presente tesis es doble.

- De una parte y como continuación a la tesis del Dr. Javier Llorente, que midió los 'jet shapes' para chorros etiquetados como b y los ligeros y observó que los primeros son más anchos que los segundos, hemos estudiado a nivel generador la mejora que aporta la introducción de los 'jets shapes' en el MVA propuesto por ABL. Estos estudios se presentaron en la ICHEP 2014 en Valencia,Nucl. Part. Phys. Proc. 273-275 (2016) 2761-2763. El capítulo 4 de la presente tesis se ocupa de estos resultados.

- De otra parte hemos estudiado, utilizando los datos de ATLAS a 7 TeV del 2011, la producción de partículas extrañas en los estados hadrónicos finales $t\bar{t}$ donde ambos tops se desintegran leptónicamente. Este estudio acaba de ser aprobado por la colaboración y publicado en EPJC, Eur. Phys. J. C 79 (2019) 1017. Encontramos que las distribuciones cinemáticas de las partículas extrañas como su momento transverso, rapidez, o energía relativa a la energía del chorro vienen bien descritas por modelos Monte Carlo al uso como POWHEG+PYTHIA6, POWHEG+PYTHIA8, MC@NLO+HERWIG, POWHEG+HERWIG7, aMC@NLO+HERWIG7 o SHERPA cuando se pueden adscribir a chorros estas partículas extrañas, independientemente de que estos chorros estén o no etiquetados como b's. Estos resultados refuerzan la propuesta de ABL de utilizar un MVA para medir V_{ts}. Más aún encontramos que la producción de partículas extrañas no asociadas a chorros no está bien descrita por los generadores al uso que la infraestiman en un 30 %. Estos datos por tanto servirán para afinar los parámetros que describen la fragmentación de las interacciones multipartónicas. El capítulo 5 de la presente tesis contiene una descripción completa de estas investigaciones.

Por completitud, y sin ánimo de originalidad, los tres primeros capítulos se ocupan de resumir los aspectos más relevantes del modelo estándar, capítulo 1, del detector ATLAS y el LHC, capítulo 2, y de la reconstrucción de objetos (electrones, muones, 'jets', b-jets, energía faltante), capítulo 3.

Introduction

The LHC has opened up a new era in particle physics in as much it paved the way for:

- the discovery of the Higgs boson, the missing piece in the so called standard model (SM)
- pushing the lower limits for particles predicted in many beyond the standard model (BSM) scenarios to values well above the TeV scale
- allowing precision measurements at unprecedented energies

This ambitious scientific program has been carried out with the construction of two big multipurpose detectors: ATLAS and CMS.
The discovery of the Higgs boson is, no doubt, one of the most important milestones in the history of particle physics. Another field where one can reach unprecedented precision is in the study of the top quark.
With the advent of unprecedented statistics for top quark pair production, one can think of directly measuring the Cabibbo-Kobayashi-Maskawa matrix elements where the top quark is involved, in particular V_{ts}. Ali et al (ABL, Phys. Lett. B 693 (2010) 44-51) made the proposal to determine this matrix element by discerning the signal $t\bar{t} \to WbWs$ final state from the dominant background process $t\bar{t} \to WbWb$. This could be done by looking at neutral strange particle production which turns out to be leading in s-quark jets while non-leading in b-quark jets as b-quarks decay sequentially $b \to c \to s$.
This proposal makes use of MVA techniques, with the possible existence of jet secondary vertices and soft leptons as additional variables with discriminant power. Therefore this proposal is dependent on the good agreement between data and Monte Carlo generators used in the training phase.
However the proposal we refer to is qualitatively right, but quantitatively is somewhat overoptimistic as it is based on generator level data. The aim of the present thesis is twofold:

- on the one hand, as a continuation of the thesis presented in 2015 by Dr. Javier Llorente, who measured b- and light-jet shapes in $t\bar{t}$ events, we have investigated the improvements in the performance of the MVA discussed above when adding the jet shapes as a discriminant variable. This was presented at ICHEP 2014 in Valencia, Nucl. Part. Phys. Proc. 273-275 (2016) 2761-2763. This is the basis for Chapter 4 in this thesis.

- on the other hand, we made use of the ATLAS data at 7 TeV taken in 2011, to study neutral strange particle production in the dileptonic $t\bar{t}$ final states. This study has been recently approved for publication by the ATLAS collaboration and published by EPJC, Eur. Phys. J. C 79 (2019) 1017. It is found that neutral strange particle production embedded in jets, b-tagged or not, is well described by current MC models like the following generators: Powheg+Pythia6, Powheg+Pythia8, MC@NLO+Herwig, Powheg+Herwig7, aMC@NLO+Herwig7 or Sherpa. This lends support to the idea put forward by ABL that one could attempt to determine V_{ts} with a TMVA. Furthermore, neutral strange particles produced outside jets, the dominant contribution, in understimated by current MC generators by 30 %. Thus, these data have the potential to help tune models for MPI. This is summarised in chapter 5.

For the sake of completeness and without any pretension to originality, the first three chapters are devoted to present the most relevant aspects of the SM, chapter 1, the ATLAS detector and the LHC, chapter 2, and the objects (leptons, jets, missing energy) reconstruction, chapter 3.

Contents

List of Figures

List of Tables

Chapter 1

Theoretical framework

A brief introduction to the Standard Model (SM) of Particle Physics is presented in this chapter.

1.1 The Standard Model: A brief introduction

The Standard Model of Particle Physics [1, 2] is a quantum field theory [3] summarising our present understanding of the interactions among the building blocks of matter, excluding gravity. The dynamics of the SM is dictated by the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The first part, $SU(3)_C$, properly describes the strong interaction, while the second $SU(2)_L \times U(1)_Y$ describes the electroweak sector of the theory.

The elementary particles in the SM are divided into two main subgroups, fermions and bosons. Fermions are the elementary constituents of matter, and have a fractional spin. Bosons, with integer spin, are considered to be the force carriers. Figure 1.1, which supersedes Mendeleev table, illustrates the elementary particles and their interactions.

Fermions are classified into quarks and leptons. Quarks are subject to the strong and electroweak interaction in as much they have fractional electric charge Q, colour charge C, as well as weak hypercharge Y and no lepton number L. They are the constituents of hadrons, in particular nucleons. Leptons have integer electric charge and lepton number as well as weak hypercharge, but no colour, and therefore they are subject to the electroweak force.

The $SU(3)_C$ symmetry is exact and therefore the gauge bosons associated to the strong interaction, gluons, are massless. In contrast the $SU(2)_L \times U(1)_Y$ is broken and the electroweak bosons $W^{\pm}$ and Z have to be massive. In the Standard Model they acquire mass according to the Higgs mechanism. And so do the fermions.

This symmetry breaking is introduced with the help of a new field Φ, which trans-

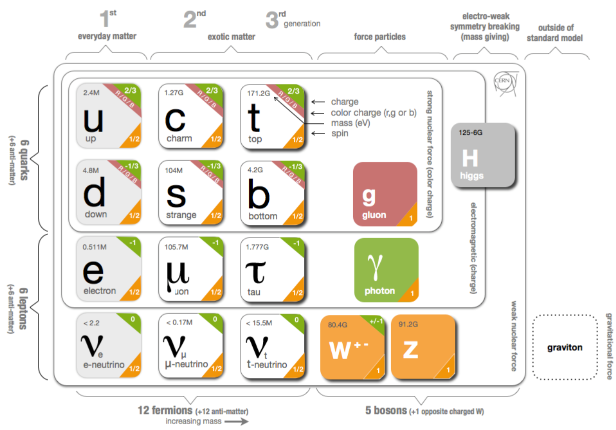


Figure 1.1: The fundamental constituents and their interactions in the Standard Model.

forms as an $SU(2)$ doublet and has a non-vanishing vacuum expectation value $v = \langle 0|\Phi|0\rangle \neq 0$.

1.1.1 Gauge symmetry and dynamics: QED as an example

In order to discuss the relationship between symmetry and dynamics, let us consider the lagrangian for a free fermion field:

$$\mathcal{L}_\Psi = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi \tag{1.1}$$

This lagrangian is obviously invariant under global phase transformations. However it is not invariant under local phase transformaions of the form:

$$\Psi(x) \mapsto \Psi'(x) = e^{i\omega(x)}\Psi(x) \tag{1.2}$$

If we want to preserve the invariance of the Lagrangian we must seek a modified derivative D_μ which transforms covariantly. One introduces a gauge field A_μ such that defining D_μ as:

$$D_\mu = \partial_\mu - ieA_\mu \tag{1.3}$$

forces:

$$A_\mu \mapsto A_\mu + \frac{1}{e}\partial_\mu \omega(x) \tag{1.4}$$

The Lagrangian:

$$\mathcal{L}_\Psi = \bar{\Psi}(i\gamma^\mu D_\mu - m)\Psi \tag{1.5}$$

is now invariant under local phase transformations as long as the gauge field A_μ remains massless. If one identifies this gauge field with the photon field one shoud add to the lagranfian its kinetic energy term. This kinetic energy term should be also invariant under local phase transformations: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Thus we are led to the following lagrangian for QED:

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\Psi}(i\gamma^\mu D_\mu - m)\Psi \tag{1.6}$$

The term in $F^{\mu\nu}F_{\mu\nu}$ in Eq. 1.6 encodes the dynamics of the photon field. However, for non-commutative groups such as $SU(N)$, this term has more fundamental physical consequences. The coupling of the gauge bosons with themselves, as we will see in the next section, is encoded in the new terms arising from the non-trivial structure constants of the group.

1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the quantum field theory describing the strong interactions [4] and it is based on the gauge group $SU(3)_C$. Unlike QED, whose Lagrangian has been described before, the Lie algebra $\mathfrak{su}(3)$ associated with the gauge group is non-commutative. The fundamental representation of this algebra is given by the Gell-Mann matrices:

$$\lambda^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda^4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda^5 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda^6 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\lambda^7 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda^8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{1.7}$$

They satisfy the following commutation relations:

$$[\lambda^a, \lambda^b] = i f_{abc} \lambda^c \tag{1.8}$$

where f_{abc} are the so-called structure constants of the group $SU(3)$. In terms of these matrices, the gauge transformation for a fermion field Ψ_j can therefore be written as

$$\Psi_j(x) \mapsto \Psi'_j(x) = \sum_k \left(e^{i\Gamma(x)}\right)_{jk} \Psi_k(x) \tag{1.9}$$

where $\Gamma(x) \equiv G_a(x)\lambda^a$ is an element of $\mathfrak{su}(3)$ and the indices j, k cover all quark flavours. The Lagrangian of the theory, which is left invariant by the transformations in Eq. 1.9, can be written as

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a + \sum_{j,k} \bar{\Psi}_j \left(i\gamma^\mu D_\mu - m\right)_{jk} \Psi_k \tag{1.10}$$

where the indices j, k cover all quark flavours. The covariant derivative is defined as:

$$D_\mu = \partial_\mu - i g_s \sum_a G^a_\mu(x)\lambda_a \tag{1.11}$$

and the field strength tensor is given by:

$$F_{\mu\nu} = \frac{i}{g_s}[D_\mu, D_\nu] = \partial_\mu G_\nu - \partial_\nu G_\mu + g_s f_{abc} G^b_\mu G^c_\nu \lambda^a \tag{1.12}$$

Here, the vector fields $G^a_\mu(x)$ represent the gluons, the carriers of the strong force, and have $n = 8$ degrees of freedom. The first term in 1.10 contains the self-couplings of the gluon fields arising from the non-Abelian nature of the theory. The strong coupling constant α_s is the fundamental parameter of the theory, giving the interaction strength in the vertices of the gluon-gluon and quark-gluon couplings. This constant can be written in terms of the coupling g_s in equation 1.11 as

$$\alpha_s = \frac{g_s^2}{4\pi} \tag{1.13}$$

1.2.1 Running coupling constant

The value of α_s is not constant, but rather depends on the energy scale Q involved in the interaction process. The dependence of the strong coupling constant with the scale is very different to that of the fine structure constant. The strong coupling constant grows as the scale becomes smaller i.e. at large distances, this explains confinement, and decreases as the scale grows i.e.at short distances, asymptotic freedom.

Let us consider a dimensionless observable R which depends on the energy scale Q. If R is dimensionless, the dependence on Q can only be achieved via the ratio Q^2/μ^2, where μ is an arbitrary parameter introduced by the renormalisation procedure. Because the QCD lagrangian 1.10 does not make explicit reference on the parameter μ, neither should any physical observable such as R. Thus, the independence of $R(Q^2/\mu^2, \alpha_s)$ on μ can be expressed as

$$\mu^2 \frac{d}{d\mu^2} R(Q^2/\mu^2, \alpha_s) \equiv \left[\mu^2 \frac{\partial}{\partial \mu^2} + \mu^2 \frac{\partial \alpha_s}{\partial \mu^2} \frac{\partial}{\partial \alpha_s}\right] R = 0 \tag{1.14}$$

We can transform this equation into a more compact form by introducing the notations

$$t = \log\left(\frac{Q^2}{\mu^2}\right); \quad \beta(\alpha_s) = \mu^2 \frac{\partial \alpha_s}{\partial \mu^2} \tag{1.15}$$

With these definitions, Eq. 1.14 can be rewritten as

$$\left[-\frac{\partial}{\partial t} + \beta(\alpha_s)\frac{\partial}{\partial \alpha_s}\right] R(e^t, \alpha_s) = 0 \tag{1.16}$$

The running of the strong coupling constant α_s is thus determined by the renormalisation group equation

$$\beta(\alpha_s) = Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \frac{\partial \alpha_s}{\partial \log Q^2} \tag{1.17}$$

Here, the function $\beta(\alpha_s)$ can be expanded in perturbation theory as a power series in α_s

$$\beta(\alpha_s) = -\alpha_s^2(\beta_0 + \beta_1\alpha_s + \beta_2\alpha_s^2 + \mathcal{O}(\alpha_s^3)) \tag{1.18}$$

The coefficients β_i in Eq. 1.18 are extracted from the higher order corrections to the vertices of the theory. Fig. 1.2 shows the one-loop diagrams contributing to the leading-order (LO) β function

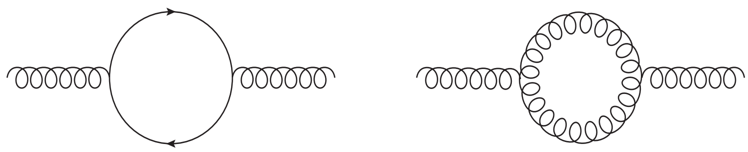


Figure 1.2: Graphs contributing to the β function in the one-loop approximation.

The result for the coefficients up to two loops is given by [5]

$$\beta_0 = \frac{1}{4\pi}\left[11 - \frac{2}{3}n_f\right]; \quad \beta_1 = \frac{1}{(4\pi)^2}\left[102 - \frac{38}{3}n_f\right] \tag{1.19}$$

In the one-loop approximation, i.e. keeping only the first term in the power expansion 1.18, the differential equation 1.17 can be integrated out, leading to the solution

$$\alpha_s(Q^2) = \frac{\alpha_s(Q_0^2)}{1 + \beta_0\alpha_s(Q_0^2)\log\left(\frac{Q^2}{Q_0^2}\right)} \tag{1.20}$$

which relates the values of α_s at two scales scale Q and Q_0. Let us define the QCD scale Λ_{QCD} as

$$\Lambda_{\text{QCD}}^2 = Q_0^2 \exp\left[-\frac{1}{\beta_0\alpha_s(Q_0^2)}\right] \tag{1.21}$$

Then, equation 1.20 can be rewritten in terms of Λ_{QCD} as

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \log\left(\frac{Q^2}{\Lambda^2_{\text{QCD}}}\right)} \tag{1.22}$$

For completeness, the two-loop expression for $\alpha_s(Q)$ in terms of Λ_{QCD} is also quoted [5, 6]

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \log z}\left[1 - \frac{\beta_1}{\beta_0^2}\frac{\log(\log z)}{\log z}\right]; \quad z = \frac{Q^2}{\Lambda^2_{\text{QCD}}} \tag{1.23}$$

The running of α_s has been tested by many experiments in e^+e^-, ep and hadron-hadron colliders as illustrated in Figure 1.3.

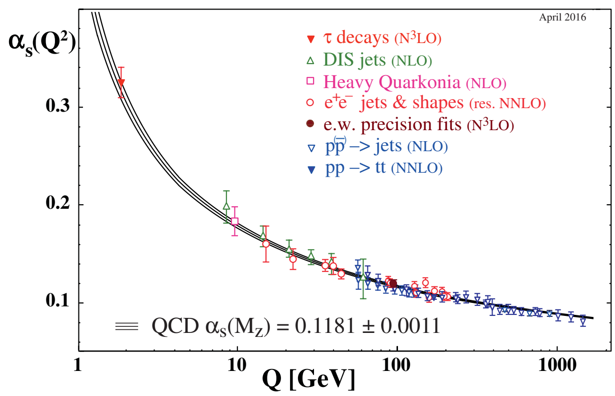


Figure 1.3: The running of the strong coupling α_s as a function of the scale Q. The shaded band represents the theoretical calculation, while the data points represent several measurements used as input for the 2016 α_s world average [6]

1.2.2 Proton structure: Deep inelastic scattering

The classical way to photographing a charge distribution is by scattering an electron beam off it. Thus, the internal structure of the proton has been studied in detail in deep inelastic scattering (NC DIS), [7], see Figure 1.4.

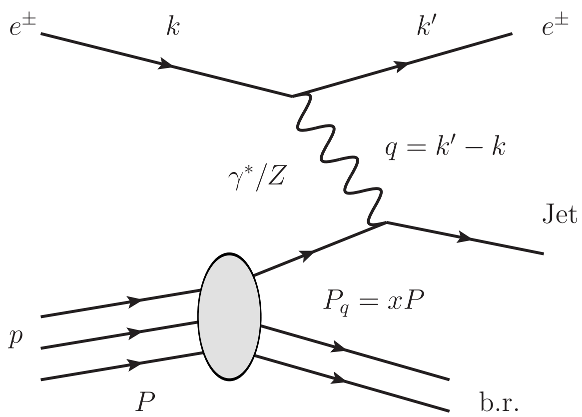


Figure 1.4: Electron-proton deep inelastic scattering

The process occurs via the exchange of a virtual photon γ^* or a Z boson with momentum q. This neutral current is absorbed by a parton carrying a fraction x of the proton momentum. The struck parton develops into a jet whose properties are measured by the detector. The scattered electron can also be detected, measuring its final energy and momentum, thus allowing us to determine the scale $Q^2 \equiv -q^2$ relevant for the process. The observation in the early 60s of precocius scaling by the SLAC-MIT experiment brought the first evidence that protons are made of spin 1/2 fermions [8].

The DIS cross-section can be expressed in terms of a set of three Lorentz-invariant observables

$$Q^2 = -q^2 = (k' - k)^2; \quad x = \frac{Q^2}{2(P \cdot q)}; \quad y = \frac{Q^2}{sx} \tag{1.24}$$

where s is the squared center-of-mass energy of the electron-proton system. In terms of these 3 variables, the double-differential NC DIS cross section takes the form

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left[xy^2F_1 + (1-y)F_2 + \left(y - \frac{y^2}{2}\right)F_3\right] \tag{1.25}$$

where $F_i(x, Q^2)$ are referred to as the structure functions of the proton. These are related to the probability that a quark interacts with the neutral current boson and are given by

$$F_1(x, Q^2) = \frac{1}{2}\sum_i \left[f_{q_i}(x, Q^2) + f_{\bar{q}_i}(x, Q^2)\right] C_{q_i}(Q^2) \tag{1.26}$$

$$F_2(x, Q^2) = x\sum_i \left[f_{q_i}(x, Q^2) + f_{\bar{q}_i}(x, Q^2)\right] C_{q_i}(Q^2) \tag{1.27}$$

$$F_3(x, Q^2) = \sum_i \left[f_{q_i}(x, Q^2) - f_{\bar{q}_i}(x, Q^2)\right] D_{q_i}(Q^2) \tag{1.28}$$

The coefficients C_{q_i}, D_{q_i} are functions of the weak angle θ_W, the weak isospin T_3 and the propagator $P(Q^2)$ [4]. The functions $f_i(x, Q^2)$ are called parton distribution functions (PDFs) and can be interpreted as the probability of finding different parton flavours, quarks or gluons, with a fraction x of the proton momentum at a given scale Q^2.

1.2.3 The DGLAP evolution equations

QCD does not predict the parton distribution functions analytical form but rather their evolution with the energy scale Q^2. The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [9] describe this scale evolution of the PDFs for both quarks and gluons as a convolution $f \otimes P$ of the PDFs at the present scale configuration with the DGLAP kernels $P_{ab}(z)$. They can be written as

$$\frac{\partial f_{q_i}(x, Q^2)}{\partial \log Q^2} = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[f_{q_i}(\xi, Q^2) P_{qq}\left(\frac{x}{\xi}\right) + f_g(\xi, Q^2) P_{qg}\left(\frac{x}{\xi}\right) \right] \quad (1.29)$$

$$\frac{\partial f_g(x, Q^2)}{\partial \log Q^2} = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[\sum_i f_{q_i}(\xi, Q^2) P_{gq}\left(\frac{x}{\xi}\right) + f_g(\xi, Q^2) P_{gg}\left(\frac{x}{\xi}\right) \right] \quad (1.30)$$

Equation 1.29 describes the scale evolution of the i-th active flavour quark density inside the proton, while 1.30 describes the running of the gluon PDF. The splitting functions $P_{ab}(z)$ parametrise the probability of a parton b emitting a parton a with momentum fraction z via the processes $q \to qg$ (gluon radiation), $g \to q\bar{q}$ (pair production) or $g \to gg$ (gluon splitting). The LO diagrams for these processes are illustrated in Figure 1.5.

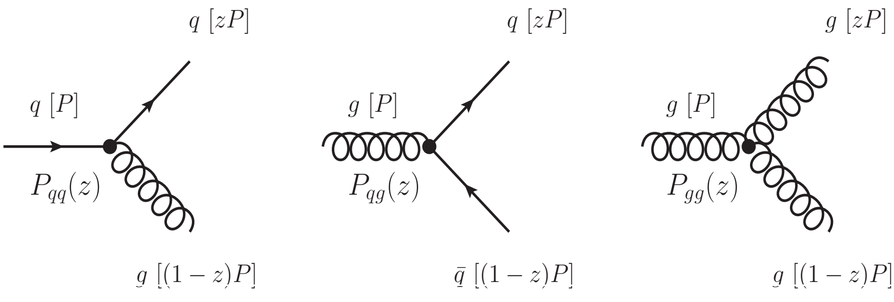


Figure 1.5: Leading-order diagrams contributing to the parton splitting functions.

The parton splitting functions can be expressed as a power series in the strong coupling constant α_s

$$P_{ij}(z) = P_{ij}^{(0)}(z) + \left(\frac{\alpha_s}{2\pi}\right) P_{ij}^{(1)}(z) + \mathcal{O}(\alpha_s^2) \quad (1.31)$$

The leading order contributions were calculated in Ref. [9], and are given by

$$P_{qq}^{(0)}(z) = C_F\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right] \tag{1.32}$$

$$P_{gq}^{(0)}(z) = C_F\left[\frac{1+(1-z)^2}{z}\right] \tag{1.33}$$

$$P_{qg}^{(0)}(z) = T_R\left[z^2+(1-z)^2\right] \tag{1.34}$$

$$P_{gg}^{(0)}(z) = 2C_A\left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right] + \delta(1-z)\frac{11C_A - 4N_f T_R}{6} \tag{1.35}$$

where factors of the form $(1-z)^{-1}$ have been regularised by reinterpreting them as distributions defined in the following way

$$\int_0^1 \frac{f(z)dz}{(1-z)_+} \equiv \int_0^1 \frac{f(z)-f(1)}{1-z}dz = \int_0^1 \log(1-z)\frac{df}{dz}dz \tag{1.36}$$

1.2.4 PDFs and the factorisation theorem

In order to extract the PDFS defined in the previous subsection, fits are performed to the DIS data at a given scale Q_0^2, which in the case of HERAPDF discussed here is taken to be $Q_0^2 = 1.9$ GeV^2. This ensures that the starting scale is below the charm mass threshold. An educated guess for the parametrical form of the PDFs at the fixed scale Q_0^2 is chosen to be [10]

$$xf(x) = Ax^B(1-x)^C(1+Ex^2) \tag{1.37}$$

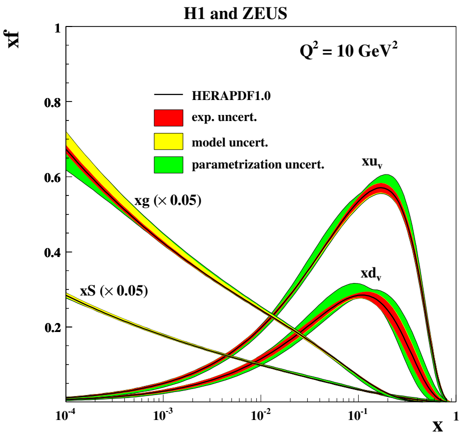
The initial choice of these parameters for the gluon $xf_g(x)$ and valence quarks $xf_{q_v}(x)$ yields

$$xf_g(x) = A_g x^{B_g}(1-x)^{C_g} \tag{1.38}$$

$$xf_{u_v}(x) = A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}(1+E_{u_v}x^2) \tag{1.39}$$

$$xf_{d_v}(x) = A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}} \tag{1.40}$$

Once the PDFs are determined at the scale Q_0^2, their form at an arbitrary scale $Q^2 > Q_0^2$ is derived by solving the DGLAP equations at next-to-leading order (NLO) or next-to-next-to-leading order (NNLO). Figure 1.6 (left) shows the HERAPDF1.0 fits at $Q^2 = 10$ GeV^2, together with the experimental, model and parametrisation uncertainties. In the right part of Figure 1.6, the PDFs are evolved at $Q^2 = 10000$ GeV^2, i.e. at the energy scale relevant for hadron colliders as Tevatron or the LHC, using the DGLAP equations up to NNLO [$\mathcal{O}(\alpha_s^6)$].



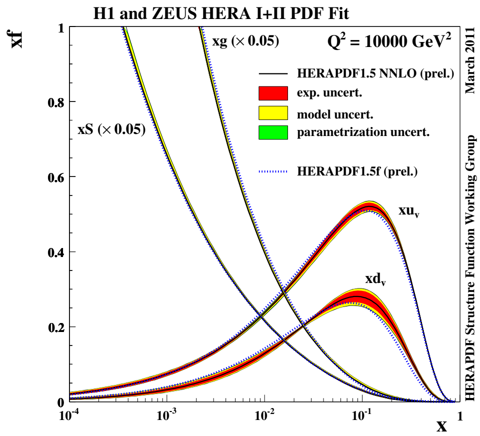


Figure 1.6: The HERAPDF parton distribution functions $xf(x, Q^2)$ as a function of x for valence quarks, gluons and sea quarks obtained for different values of the interaction scale Q^2. The gluon and sea PDFs have been scaled down by a factor of 20 [10]

The parton distribution functions are relevant in the calculation of hadron-hadron cross sections $\sigma(pp \to X)$ from the partonic process $\hat{\sigma}(ij \to Y)$. This can be so done because of the factorisation theorem which states that the perturbative and non-perturbative contributions to the physical cross sections in hadron-hadron collisions can be separated [11].

Thus, the differential cross-section for a given process in a hadron-hadron collision can be calculated as the convolution of the partonic cross section $\hat{\sigma}_{ij}$ with the parton distribution functions, after summing over all possible initial-state configurations:

$$\frac{d\sigma}{dX} = \sum_{i,j} \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \frac{d\hat{\sigma}_{ij}}{dX} \left(x_1, x_2, \alpha_s(\mu_R^2), \frac{Q^2}{\mu_R^2} \right) \quad (1.41)$$

where X is any physical observable. The parameters μ_R and μ_F are the renormalisation and the factorisation scales.

1.3 Electroweak interactions

1.3.1 Weak charged currents

Weak interactions are responsible for nuclear beta decay, for $\pi^+ \to l^+ \nu_l$ with $l^+ = e^+, \mu^+$ or for the sequential decay $\mu^+ \to e^+ \nu_e \bar{\nu_\mu}$. Fermi (1932) postulated that nuclear beta decay could be described as a contact interaction between two vector weak currents responsible for the metamorphose $n \to p$ and $\nu_l \to l$. However, this ansatz was able to explain the lepton energy spectrum but not their angular distributions, as shown by the famous Wu Co^{60} experiment in 1956. In fact, this experiment did show that parity was violated maximally. Feynman and Gell-Mann and independently Sudershan and Marshak found that the correct description could be recovered by giving the weak current a V-A structure:

$$J_\mu^\pm = \bar{\chi_L}\gamma_\mu\tau_\pm\chi_L \tag{1.42}$$

where $\bar{\chi_L} = (\bar{\nu_l}\ l^+)_L$, the subscript L is used to denote left-handed spinors and record the V-A nature of the charged currents, and τ refers to the Pauli matrices. Weak interaction amplitudes were thought to be of the form:

$$\mathcal{M} = \frac{4G_F}{\sqrt{2}} J_\mu J_\mu^\dagger \tag{1.43}$$

where G_F is Fermi's constant.
This ansatz however violates unitarity. Already in the 60s people thought that the Fermi modified model for weak interactions could be the low energy limit of a theory where charged current interactions were mediated by a massive gauge boson $W^\pm$, such that:

$$\mathcal{M} = \frac{g}{\sqrt{2}} J_\mu \frac{1}{M_W^2 - q^2} \frac{g}{\sqrt{2}} J_\mu^\dagger \tag{1.44}$$

where g is a dimensionless weak coupling, q is the momentum carried by the weak boson, and M_W denotes its mass.
The following relation should hold:

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} \tag{1.45}$$

So in the late 50s leptons and quarks participated in weak interactions through charged V-A currents constructed from the following pairs of fermion states: (ν_e e), (ν_μ μ) and (u d). In 1963, Cabibbo in order to explain the decay $K^+ \to \mu^+\nu_\mu$, while preserving the universality of the Fermi coupling, postulated that the u quark couples to the following admixture:

$$d' = d\ \ cos\theta_C + s\ \ sin\theta_C \tag{1.46}$$

This introduces an arbitrary parameter θ_C, the quark mixing or Cabibbo angle, which turns out to be approximately 13°.

The true relevance of the Cabibbo angle was understood by Glashow, Iliopoulos and Maiani who in order to further explain the experimentally observed suppression of strangeness changing neutral currents postulated the existence of a new quark, charm (c), such that the following modification of the charged weak current is postulated (GIM mechanism):

$$J_\mu^+ = (\bar{u} \quad \bar{c})\gamma_\mu \tau_- \begin{pmatrix} cos\theta_C & sin\theta_C \\ -sin\theta_C & cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{1.47}$$

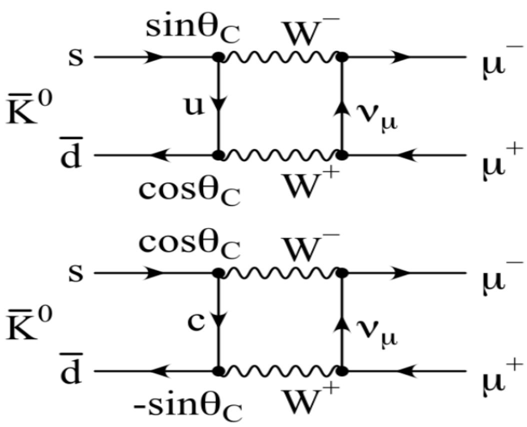


Figure 1.7: Feynman diagrams of $K^0 \to \mu^+\mu^-$.

The discovery of the third generation of leptons (τ lepton at SPEAR) and quarks (bottom (b) quark at Fermilab) let to Kobayashi and Maskawa to the well known generalisation to a 3x3 quark mixing matrix, the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix V_{CKM} [12, 13]:

$$V_{CKM} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.48}$$

The subsequent measurement at SLD and LEP of the Z boson width, which determined the number of families to be three, made further generalisations of the CKM matrix to higher dimensions not necessary. It can be seen that a 3×3 unitary matrix is given in terms of three real parameters and a phase. In the Wolfenstein parametrisation [14]

$$V_{CKM} \simeq \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda(1 + iA^2\lambda^4\eta) & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2(1 + i\lambda^2\eta) & 1 \end{pmatrix} \tag{1.49}$$

where A, λ, ρ and η are the Wolfenstein parameters. They are determined from experimental data on weak decays as it will be discussed later. The most up to date values are: $\lambda = 0.22453 \pm 0.00044$, $A = 0.836 \pm 0.015$, $\rho = 0.122^{+0.018}_{-0.017}$ and $\eta = 0.355^{+0.012}_{-0.011}$.
Since V_{CKM} is unitary, the complex vectors asssociated with its rows or columns must be orthonormal, this can be expressed pictorially with the well known unitarity triangles in the complex plane. The most interesting of the unitarity relations is $V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0$ which is represented in Figure 1.9. This is in excellent agreement with the SM. Thus, though the SM does not explain the origin of $\mathcal{CP}$ violation, it explains the pattern of $\mathcal{CP}$ violation observed in the strange and B meson systems.

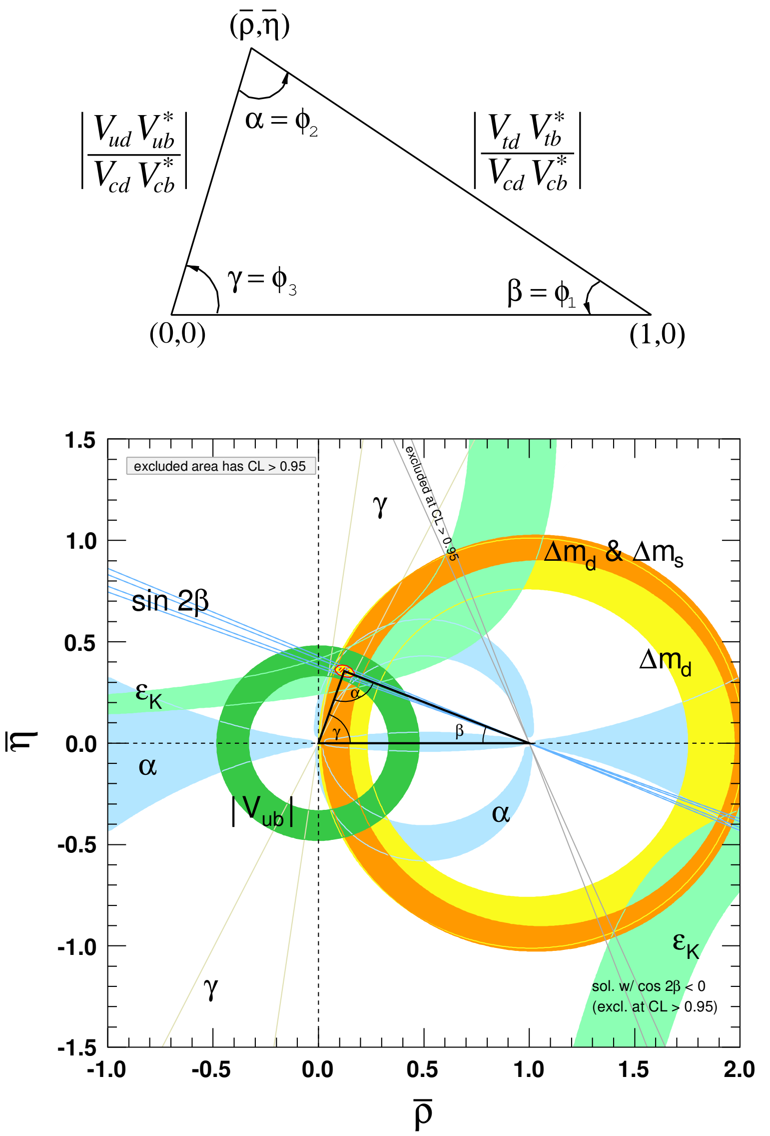


Figure 1.8: Theoretical (top) and experimental (bottom) representations of one of the unitary triangle relationship from the V_{CKM} matrix [6].

1.3.2 Electroweak unification

The weak charged currents discussed above suggest a possible $SU(2)_L$ structure in which the third component of this triplet should be of the form:

$$J^3_\mu = \bar{\chi_L}\gamma_\mu\tau_3\chi_L \tag{1.50}$$

This idea was very bold at the time because no neutral weak currents of the type $\nu_\mu e \to \nu_\mu e$ or $\nu_\mu p \to \nu_\mu p$. However, the discovery by Gargamelle in 1963 of this type of processes gave further support to these ideas. Experimentally, it is found that the weak neutral current has a small right-hand component so that it is not possible to ascribe it to the J^3_μ in the equation above. S. Glashow's idea was to build a mixture of J^3_μ and j^{em} (electromagnetic current) in such a way that the Gell-Mann and Nishijima rule is preserved i.e.:

$$j^{em} = J^3_\mu + \frac{1}{2} j^Y_\mu \tag{1.51}$$

where j^Y_μ is a new hypercharge (Y) current.
S. Weinbeg and A. Salam extended this idea and proposed that the electroweak interaction Lagrangian be written as:

$$\mathcal{L}_{int} = -ig(J^i)^\mu W^i_\mu - i\frac{g'}{2}(j^Y)^\mu B_\mu \tag{1.52}$$

The fields $W^\pm_\mu = \sqrt{\frac{1}{2}}(W^1_\mu \mp iW^2_\mu)$ should be associated with the charged weak bosons and the two neutral fields W^3_μ and B_μ should mix as:

$$\begin{pmatrix} A_\mu \\ \\ Z_\mu \end{pmatrix} = \begin{pmatrix} cos\theta_W & sin\theta_W \\ \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ \\ W^3_\mu \end{pmatrix} \tag{1.53}$$

with θ_W such that $J^{NC}_\mu = J^3_\mu - sin^2\theta_W j^{em}_\mu$. The vector and axial-vector coupling constants of a fermion pair f to the Z^0 should be:

$$c^f_V = T^3_f - 2sin^2\theta_W Q_f \tag{1.54}$$

$$c^f_A = T^3_f \tag{1.55}$$

with T^3_f being the third component of fermion f weak isospin. In order that electrodynamics is properly incorporated into this electroweak interacion the following relation has to hold $gsin\theta_W = g'cos\theta_W = e$. Fermi's constant G_F and g should satisfy the following relation:

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M^2_W} \tag{1.56}$$

while the relative coupling of the neutral weak currents with respect to the charged ones is given by the so called ρ parameter:

$$\rho = \frac{M^2_W}{M^2_Z cos^2\theta_W} \quad ; \quad \rho = 1 \textit{ in GWS to the lowest order.} \tag{1.57}$$

Since the $W^\pm$ and Z bosons should be massive, the $SU(2)_L \otimes U(1)_Y$ symmetry must be broken. In the GWS model this symmetry breaking is enforced via the introduction of a Higgs potential. In 1973 M. Veltman and G. t'Hooft proved that the GWS theory is renormalizable.

1.3.3 Spontaneous symmetry breaking: The Higgs mechanism

The way to generate masses by spontaneous symmetry breaking can be very easily illustrated with the following example. Let us consider a scalar field in the potential (V)

$$V = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{1.58}$$

with λ positive. The lagrangian is given by

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - (\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4) \tag{1.59}$$

Clearly this lagrangian is invariant under $\phi \mapsto -\phi$. The case μ^2 positive is uninteresting, it describes a self interacting scalar particle with mass μ. The ground state correspons to $\phi = 0$ such that the vacuum is also invariant under the symmetry operation discussed above.
The more interesing case occurs for μ^2 negative. The potential has two minima:

$$\phi = \pm v \tag{1.60}$$

with

$$v = \sqrt{-\mu^2/\lambda} \tag{1.61}$$

One says that the symmetry has been spontaneously broken, as any of these two minima are not invariant under the considered symmetry. Perturbative calculations can be carried out via expansions around one of these two minima. If one writes

$$\phi = v + \eta(x) \tag{1.62}$$

the lagrangian takes the expression

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\eta)^2 - \lambda v^2\eta^2 - \lambda v\eta^3 - \frac{1}{4}\lambda\eta^4 \tag{1.63}$$

The field η has a correct mass term:

$$m_\eta = \sqrt{2\lambda v^2} \tag{1.64}$$

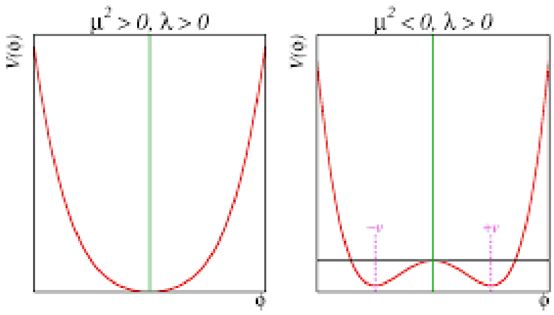


Figure 1.9: Representation of the scalar potential for different values of the parameters μ and λ.

For the electroweak lagrangian the most economic choice to spontaneously break the $SU(2)_L \otimes U(1)_Y$ symmetry is to introduce an isospin doublet with $Y = -1$:

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.65}$$

with the familiar Higgs potential discussed previously. The choice for the vacuum expectation value is:

$$\phi_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.66}$$

Expanding $\phi(x)$ around this vacuum:

$$\phi_(x) = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.67}$$

and sparing the reader the algebra which can be found in any textbook, one finds:

$$M_W = \frac{1}{2} g v \tag{1.68}$$

$$M_H = 2\lambda v^2 \tag{1.69}$$

with $v = 246\ GeV$. The photon remains massless. The mass of the fermions are not predicted. The coupling of the Higgs boson to fermions is found to be proportional to their masses:

$$H_{ff} = -i\frac{m_f}{v} \tag{1.70}$$

1.4 Monte Carlo predictions

In order to compare the experimental data with the theoretical predictions, Monte Carlo (MC) generators are used to simulate events for a given physics process. This simulation incorporates the calculation of the parton level cross section $\hat{\sigma}$ at a fixed order in perturbation theory, with higher order corrections taken into account with parton cascades. Non-perturbative effects are considered with the help of hadronisation models. Multiple parton interactions must also be implemented. This is illustrated in Figure 1.10

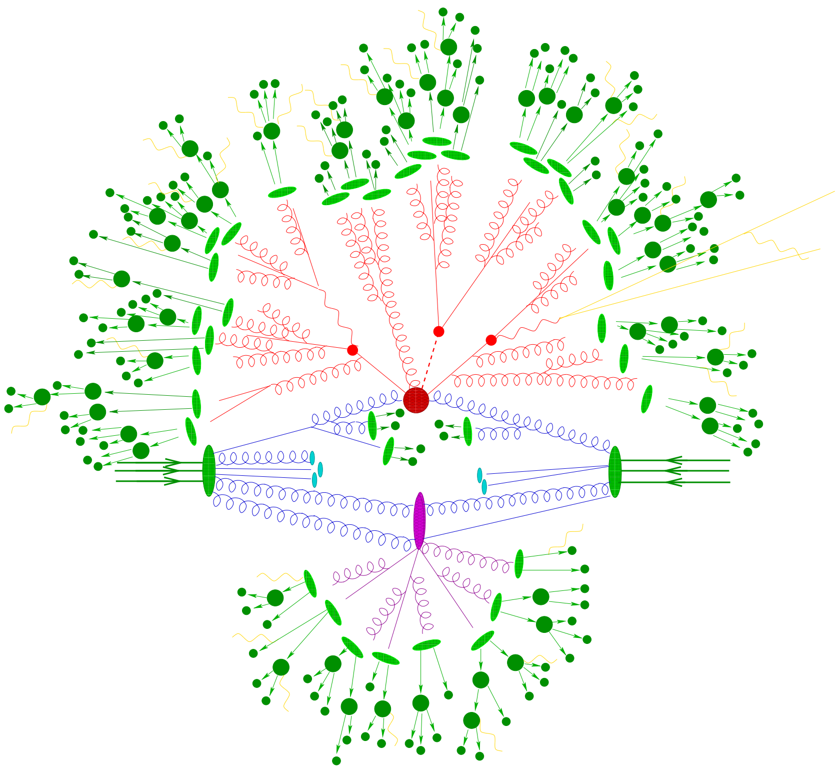


Figure 1.10: Sketch of a proton-proton collision, including the initial state partons (blue), the underlying event (violet), the hard scattering process and subsequent parton shower (red), and the final state hadrons (green). Figure taken from [15]

1.4.1 Parton showers

The parton shower algorithm takes into account higher order corrections to the matrix element (ME) calculation in fixed order perturbation theory. It is driven by the parton splitting functions 1.32 to 1.35 and implemented via Sudakov factors $\Delta(t_2, t_1)$ [16]. This quantity represents the non-splitting probability of a

parton between two scales t_1 and t_2.
Consider the following simplified example, where the scale t is identified as the time axis: If the splitting probability at a time t is given by a function $f(t)$ and one requires that the splitting can only happen at the instant t if it did not happen for $t' < t$, whose probability is $\mathcal{N}(t')$, then the splitting probability $\mathcal{P}(t)$ satisfies

$$\mathcal{P}(t) = \frac{d}{dt}\left(1 - \mathcal{N}(t)\right) = f(t)\mathcal{N}(t) \tag{1.71}$$

Therefore, if the process starts at $t = 0$ with $\mathcal{N}(0) = 1$, the solution of 1.71 is

$$\mathcal{P}(t) = f(t) \exp\left[-\int_0^t f(s)ds\right] \tag{1.72}$$

And therefore, the Sudakov factor $\Delta(t_2, t_1)$ with $t_1 \leq t_2$ is

$$\Delta(t_2, t_1) = 1 - \int_{t_1}^{t_2} \mathcal{P}(s)ds = 1 + \exp\left(-\int_0^t f(s)ds\right)\Bigg|_{t=t_1}^{t_2} \leq 1 \tag{1.73}$$

The parton showers can be angle or p_{T}-ordered, depending on the chosen scale t. In the first case, the angle of emission with respect to the incoming parton is decreased in each step, while in the second, the parton transverse momentum is decreased. The shower stops upon reaching a certain cut-off scale at which the strong coupling constant becomes large. For the initial-state radiation, the scale variation is reversed (angles and momenta are increased), while for final-state radiation, the scales are forced to decrease in each step.

1.4.2 Hadronisation

The hadronisation is the process by which final state partons turn into final state hadrons. This process is based on the parton-hadron duality hypothesis [17], which establishes that, as hadronisation is a long-distance process involving only small momentum transfers, the flow of momentum and quantum numbers at the hadron level tends to follow the flow at the parton level. Thus for example the flavour of a quark initiating a jet should be found in a hadron near the jet axis. The angular distribution of a jet should be determined by the spin of the parent parton.

The hadronisation process is simulated using two broad models, namely the Lund string model [18] and the cluster model [19], which are schematically described in Figure 1.11 and Figure 1.12

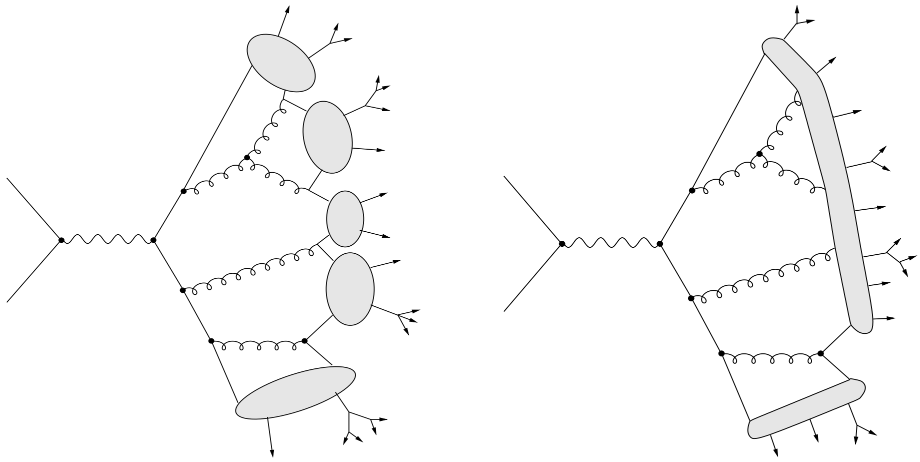


Figure 1.11: Schemes for the cluster (left) and the string (right) hadronisation models [20]

- **The Lund string model:** The string model is most easily described for the process $e^+e^- \to q\bar{q}$. As the quark and antiquark fly in opposite directions, they lose energy to the colour field given by a stringlike configuration between them. The string has a constant energy per unit length, as given by a linear confining potential consistent with quarkonium spectroscopy. As the quarks separate in phase space, this confinement potential increases the energy of the string up to the mass threshold of a new $q\bar{q}$ pair. At this point, the string is broken and the new $q\bar{q}$ pair gives rise to the formation of hadrons. Gluon bremstrahlung is understood to simply produce a kink as now two strings will be formed.

- **The cluster model:** The cluster model is based on the preconfinement property of the shower, by which neighbouring colour connected partons have an asymptotic mass distribution that falls steeply at high masses and is aymptotically independent of Q^2 and universal. The method starts with a non-perturbative splitting $g \to q\bar{q}$, and follows with the association of $q\bar{q}$ pairs into colour singlet combinations, which are assumed to form clusters. These clusters decay into pairs of hadrons following an isotropic pattern.

1.4.2.1 Strangeness suppression in the hadronisation process

It is well known since the 60s that strange particles are produced in pairs in pp or πp collisions with cros sections which are suppressed an order of magnitude relative to those for non strange particles.
In the Lund fragmentation scheme, see the left hand side of Figure 1.12, this is achieved with the help of a free parameter usually denoted by $\gamma_s = u/s$ which

determines the probability to produce an $s\bar{s}$ pair in the string breakup. This parameter, together with those determining the ratio of tensor and vector mesons (baryons) to scalar ones (mesons), or σ_q giving the jet p_T width, are fitted to data. They show a residual dependence on c.m. energy and type of process, e^+e^- vs e^-p vs pp.
In the cluster fragmentation, which is the basis of Herwig as illustrated in the right hand side of Figure 1.12, it is in the last two stages namely cluster fission and decay that the flavour of the fragmentation products is chosen. In the original Herwig model the probability to pick up an $s\bar{s}$ is the same to that of a $u\bar{u}$ or $d\bar{d}$. The strangeness suppression is achieved by phase space considerations i.e. the s quark is heavier than the u and d. In modern tunes, one can introduce two different parameters governing the fission and decay steps.

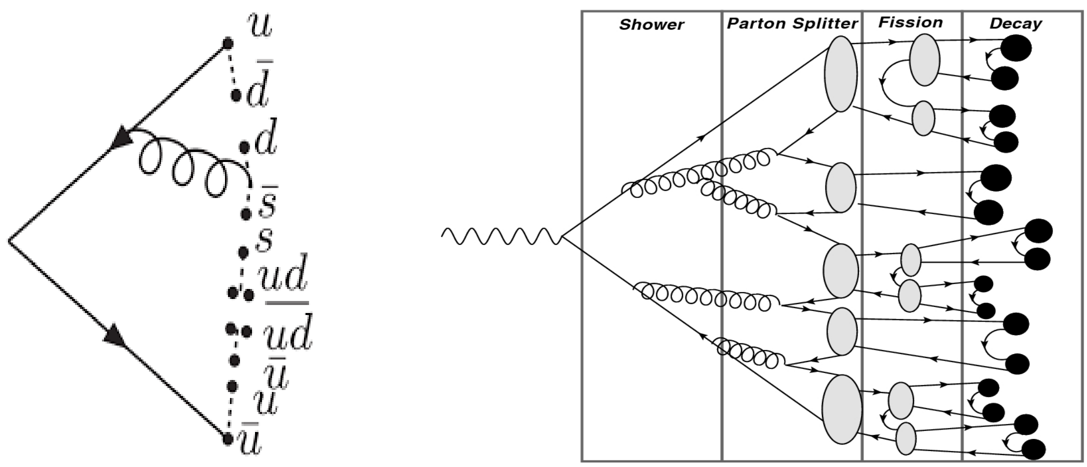


Figure 1.12: Schemes for strangeness production in string (left) and the cluster (right) hadronisation models.

1.4.3 Underlying Event

The underlying event (UE) is a result of the interaction between the remnants of the incoming hadrons. It contributes to the final state hadronic activity by interacting with the colour structure of the hard scattering outcoming partons, as well as creating new hadrons from the hadronisation of its own outcoming partons (see Figure 1.10). The parameters involved in the modelling of the UE have to be tuned using experimental data such as the jet shapes [21] or other sensitive observables [22]

1.4.4 Monte Carlo Programs

The general features described above are used by several MC generator programs to describe the experimental data. Here, a description is given for the most commonly used and, in particular, the ones used in the analyses carried out in this thesis.

1.4.4.1 PYTHIA

PYTHIA [23] is a LO event generator based on ME calculations for $2 \rightarrow 1$ and $2 \rightarrow 2$ processes. They are matched to p_{T}-ordered parton showers accounting for the initial and final state radiation. For the initial (final) state radiation, the parton showers are space-like (time-like). The hadronisation part PYTHIA relies on the Lund string model described in the previous section.
The UE modelling is based on different tunes [24] depending on the values of the QCD parameters such as Λ_{QCD} used for this purpose, and they differ from version 6 to version 8, written in FORTRAN and C++, respectively.

1.4.4.2 HERWIG

The HERWIG program [25] differs from PYTHIA in two important aspects. The parton showers are angular-ordered. This means that, for a branching $a \rightarrow bc$ the relevant scale is given by $Q^2 = 2E_a^2(1 - \cos\theta_{bc})$, where E_a is the energy of the particle a and θ_{bc} is the angle between the branching products b and c. The so-defined scale Q^2 increases with the evolution of the shower, thus the angle between the products of sucessive branchings decreases as the shower evolves.
The hadronisation in HERWIG is implemented using the cluster model, while the underlying event is modelled using minimum-bias interactions. It also implements an interface to JIMMY [26], which uses multiple parton interactions to model the underlying event effects.

1.4.4.3 MC@NLO

The MC@NLO program [27] matches exact NLO calculations to a parton shower. Particular care is given to avoiding double counting, as the parton shower calculation already implement some of the NLO corrections.
Events generated using MC@NLO are associated with a weight which might be $+1$ or -1. The negative weighting accounts for the cancellation of interfering diagrams. The smoothness of the distributions obtained using this scheme is guaranteed since the number of events with negative weights is reasonably small. MC@NLO is commonly matched to the HERWIG parton showers interfaced with JIMMY.

1.4.4.4 POWHEG

In the POWHEG formalism [28], ME calculations for $2 \to 2$ and $2 \to 3$ processes are implemented. The renormalisation and factorisation scales μ_R and μ_F are set to be equal to the transverse momentum of the hard partons ($p_{\mathrm{T}}^{\mathrm{Born}}$). The cut on the minimum transverse momentum of the generated hard partons, $p_{\mathrm{T}}^{\mathrm{Born}}$, may affect the value of the cross section due to the low-p_{T} divergence of the $2 \to 2$ cross section. POWHEG is matched either to PYTHIA or HERWIG for the parton shower step of the event generation. In case of using HERWIG, the UE modelling is tipically performed by the JIMMY program.

1.4.4.5 ALPGEN

The ALPGEN generator [29] is based on a multi-leg approach, where $2 \to n$ multiparton processes are considered. The combination of the different parton multiplicities is performed by weighting each sample with their corresponding cross sections. Then, the generated partonic final states are matched to parton shower algorithms as implemented in PYTHIA or HERWIG.
The addition of the parton shower on top of the n-parton final state can lead to an overlap of the event with the $(n + 1)$-parton sample if the shower produces an additional jet. To avoid this, ALPGEN uses MLM matching to identify the overlaps.

1.5 Jets and jet algorithms

The collimated showers of particles formed by the fragmentation of partons produced in hard scattering processes are known as jets. These jets are the experimental counterpart to the hard scattering partons. Thus, the reconstruction of the kinematic properties of partons is possible using jet algorithms. Their performance is normally dependent on a clustering parameter R, which can be regarded as the jet radius in the η–φ[1] space.
In general terms, a jet algorithm needs to fulfill two conditions namely collinear safety and infrared safety. This is a way to state that the jet topology is not going to be changed by collinear or soft radiation. These two requirements have been successfuly implemented into two different types of jet reconstruction algorithms. The cone algorithms, such as SISCone [30] are based on the maximisation of the energy density inside cones of a given radius R, followed by a merging-splitting step where overlapping configurations of two well-defined jets are removed. This thesis uses another approach based on sequential recombination algorithms.

[1]Pseudorapidity is defined as $\eta = -\log\left(\tan\frac{\theta}{2}\right)$ with θ the polar angle.

The sequential recombination algorithms are based on the iterative merging of pairs of constituents (i, j) in a same object attending to both their transverse momentum k_t and their spatial distance $\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2}$. This is done using a metric d_{ij} defined as

$$d_{ij} = \min\left(k_{t,i}^{2p}, k_{t,j}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2} \tag{1.74}$$

$$d_{i,B} = k_{t,i}^{2p} \tag{1.75}$$

The algorithm works as follows: for a pair of input particles (i, j), the value of d_{ij} is calculated and compared to the value of $d_{i,B}$. If $d_{ij} < d_{i,B}$, the inputs i and j are merged into a single input following a given recombination method. If this condition is not fulfilled, i is identified as a stable jet. Different values of the parameter p lead to different algorithms, such as k_t [31, 32] with $p = 1$, anti-k_t [?] with $p = -1$ or the Cambridge-Aachen [33] algorithm for $p = 0$. Figure 1.13 shows the behaviour of different jet algorithms. All these methods have been implemented in the FastJet program [34].

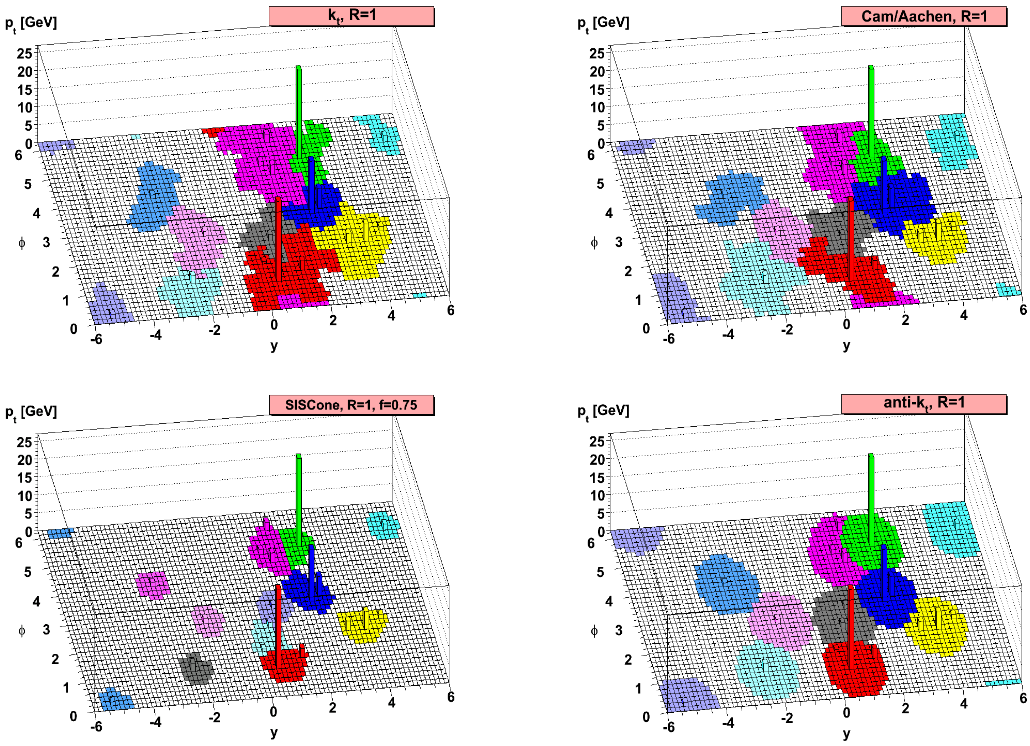


Figure 1.13: Sketch of the transverse profile of the jet areas attending to different algorithms. Figure taken from [35].

Chapter 2

The LHC and the ATLAS experiment

This chapter summarizes some of the most relevant aspects of both, the CERN accelerator complex, in particular the LHC, and the ATLAS detector.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC), see Ref. [36], spanning a 27 km circumference, is the world's largest particle accelerator. It is built underground, between 50 and 175 metres deep, in the franco-swiss border near Geneva. It is designed to accelerate two proton beams up to a nominal centre-of-mass energy $\sqrt{s} = 14$ TeV. The analyses presented in this thesis, uses data taken in 2011 when due to some technical difficuties the c.m. energy was only 7 TeV. Both beams are made to collide inside the 4 main experimental halls housing four detectors ATLAS, ALICE, CMS and LHCb, distributed along the accelerator circumference, as shown in Fig. 2.1.

The proton beams are extracted from Hydrogen gas and their energy ramped up with the help of the CERN accelerator complex from the 50 MeV obtained with the linear accelerators (LINAC), to 1.4 GeV in the Proton Booster, to 25 GeV in the Proton Synchrotron (PS) and 450 GeV in the SPS. When the beams reach the LHC, they fly around in two separate rings inside vacuum chambers passing through 1232 superconducting dipole magnets distributed around the circumference. These dipole modules produce a vertical magnetic field which bends the trajectory of the protons via the Lorentz force. At 7 TeV, the magnitude of the magnetic field is around 8.4 Tesla with a current of 11700 A. To avoid excessive resistive losses due to Joule effect, a large cryogenics system provides the liquid helium which is then injected into the dipoles to keep the magnets cold. A

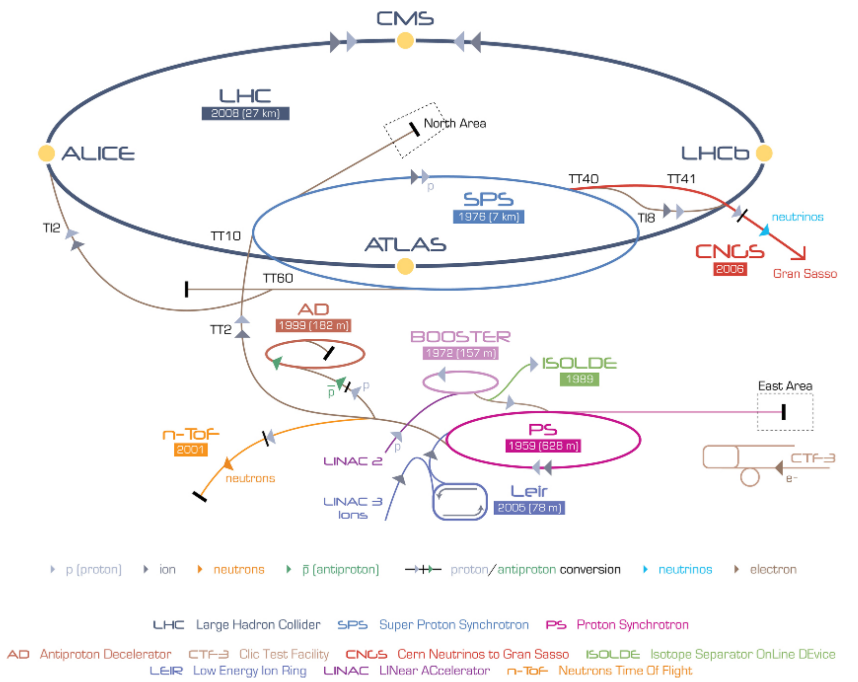


Figure 2.1: The distribution of all CERN accelerators and experiments, including the LHC experiments ATLAS, ALICE, CMS and LHCb along the LHC ring.

standard dipole module is depicted in Fig.2.2

2.1.1 Physics event rate. Luminosity

The motivation for the construction of the LHC is first and foremost to find the Higgs boson, the missing piece needed to complete the SM, and then more generally to explore physical processes at the TeV scale, where new physics beyond the SM, such as Supersymmetry (SUSY) could be found.
Since the beginning of its operation the LHC run in two periods. Run 1 started in 2010 at 7 TeV, continued in 2011 and in 2012-2013 when the energy was pushed up to 8 TeV. After a technical shutdown operations resumed in 2105 at 13 TeV for Run 2.
The figure of merit characterising the performance of a collider is the so called instantaneous luminosity defined as:

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\Sigma_x\Sigma_y} \tag{2.1}$$

The parameters Σ_x and Σ_y define the beam transverse profiles assuming a 2-dimensional gaussian distribution.
For a given physics process, characterised by its cross-section σ, the following

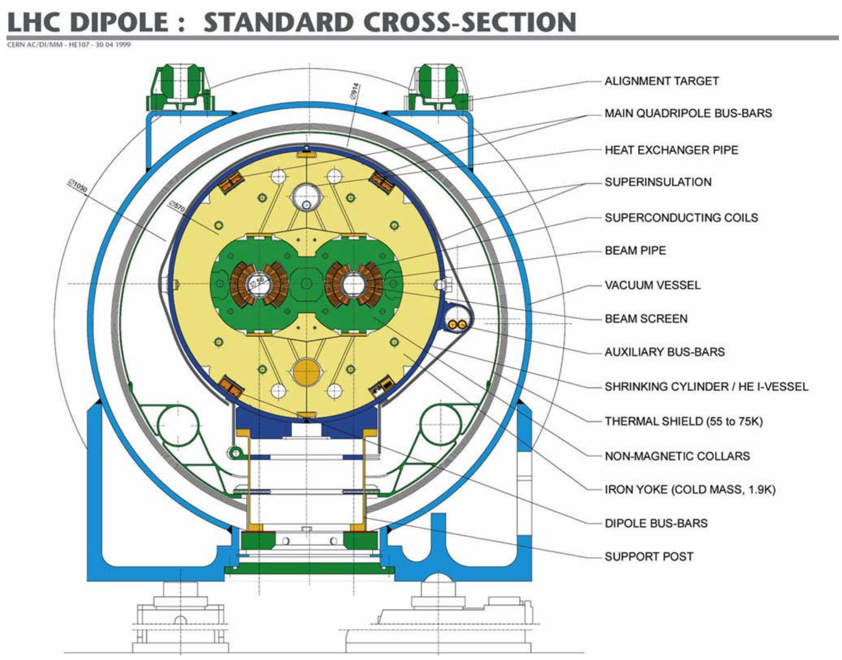


Figure 2.2: The cross section of a dipole module at the LHC

relation holds for the event rate [6]:

$$R = \frac{dN_{\text{ev}}}{dt} = \sigma \mathcal{L} \Rightarrow N_{\text{ev}}(t_1, t_2) = \sigma \int_{t_1}^{t_2} \mathcal{L} dt = \sigma L \tag{2.2}$$

The integrated luminosity L is a key quantity in cross-section measurements and therefore needs to be precisely determined by both the LHC (delivered luminosity) and each experimental detector (recorded luminosity) when the data taking is ready. Fig. 2.3 shows the integrated luminosity delivered to ATLAS as a function of the month in the years 2010, 2011 and 2012, i.e. Run 1.

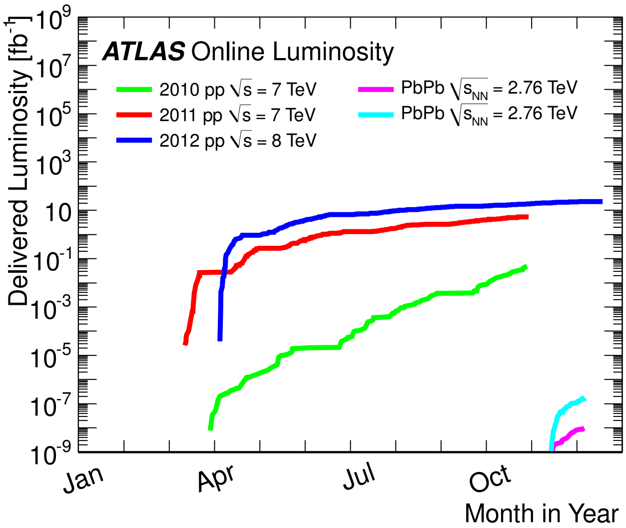


Figure 2.3: Cumulative luminosity versus day delivered to ATLAS during stable beams and for *pp* and Pb-Pb collisions. This is shown for 2010 (green for *pp*, magenta for Pb-Pb), 2011 (red for *pp*, turquoise for Pb-Pb) and 2012 (blue) running. The online luminosity is shown.

For completeness the luminosity delivered to ATLAS including Run 2 is shown in Figure 2.4. It illustrates the magnificent performance of the CERN accelerator group.

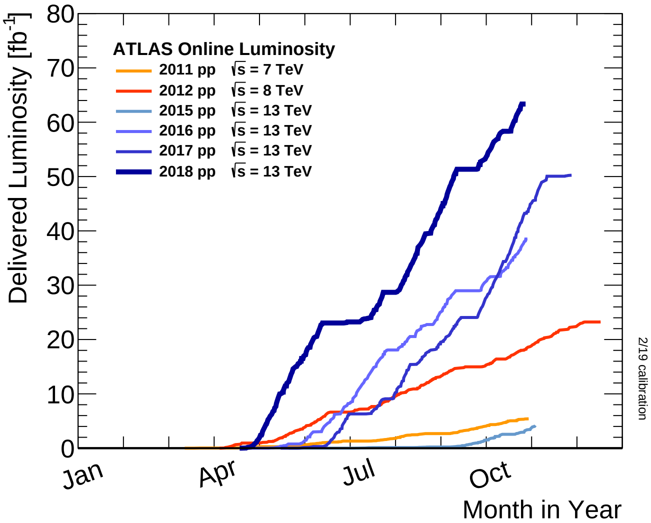


Figure 2.4: Cumulative luminosity versus day delivered to ATLAS during stable beams and for *pp* collisions only. This is shown from 2011 to 2018.

2.1.2 Bunch crossing and pile-up

The proton beams have a bunch train structure with each bunch containing roughly 10^{11} protons. The rate of inelastic events produced in a collider is given by $R_{\text{inel}} = L \times \sigma_{\text{inel}}$, while the average bunch crossing rate is given by the product of the number of bunches per beam and the revolution frequency, $R_{\text{cross}} = N_{\text{bunch}} \times f_{\text{LHC}}$. Thus, one can define the number of interactions per bunch crossing μ as

$$\mu = \frac{R_{\text{inel}}}{R_{\text{cross}}} = \frac{L \times \sigma_{\text{inel}}}{N_{\text{bunch}} \times f_{\text{LHC}}} \tag{2.3}$$

One can distinguish two types of pile-up collisions: in-time (out-of-time) depending on whether the pile-up signal corresponds to the same (a different) bunch crossing as that of the hard scattering signal.

In the data taken in 2010, the amount of out-of-time pile-up interactions is negligible, as the bunch spacing was smaller than in 2011, when it was raised to the nominal value of 25 ns. The average number of interactions per bunch crossing can be up to $\langle\mu\rangle = 25$ for the data presented here. This is smaller than in the 2012 data taking period and certainly much smaller than in Run 2.

2.2 The ATLAS experiment

The ATLAS detector citedetector is a multi-purpose particle detector installed at the LHC Point 1, near the CERN Meyrin site. The design of the ATLAS detector was guided by these general considerations:

- large pseudorapidity and full azimuthal acceptance
- good pattern recognition for charged particles allowing secondary vertices reconstruction for long lived particles
- excellent electromagnetic calorimetry for detection of electrons and photons
- good containment of hadronic showers for good measurement of missing energy needed for many BSM searches
- good muon identification
- fast and radiation hard readout electronics in all components
- excellent trigger capabilities

The defining choices for ATLAS are LAr fine granularity calorimetry and toroidal magnets for muon bending and muon charge determination.

From the particle interaction point outwards, it contains an Inner Detector (ID), designed to precisely measure the momentum of charged particle tracks, an electromagnetic calorimeter for measuring the energy deposition of electrons and photons, a hadronic calorimeter, used for the measurement of strong-interacting particles and a muon spectrometer for measuring the energy and momentum of muons, as shown in Figure 2.5.

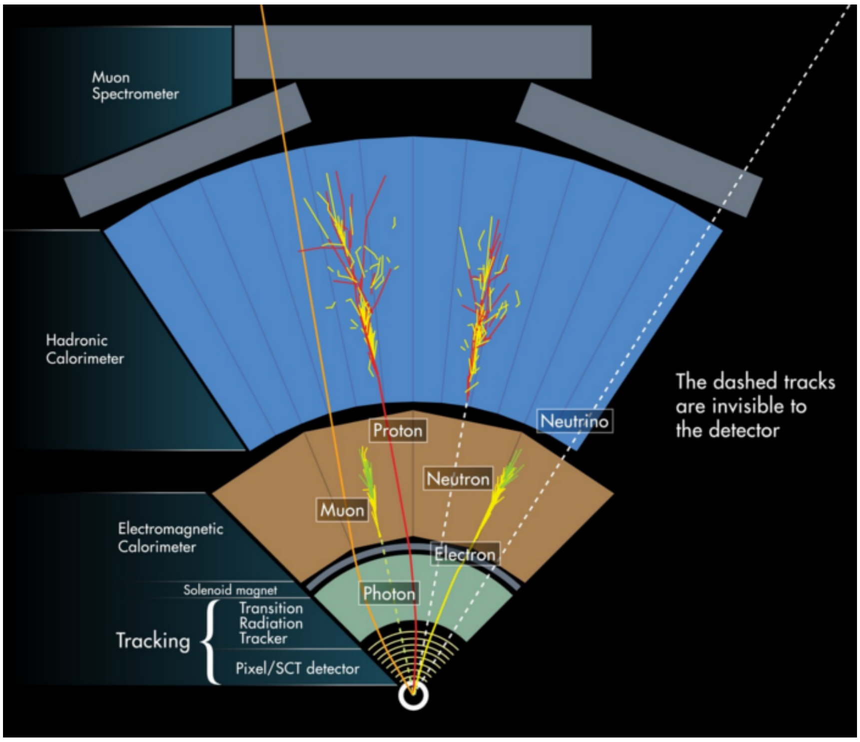


Figure 2.5: Simulation of different particle types crossing the ATLAS detector.

A schematic representation of the detector can be seen in Fig. 2.6.

2.2.1 The Inner Detector

The ID [37] is designed to measure the momentum of charged particles produced in the collisions, as well as for primary and secondary vertex reconstruction. Its η–φ coverage includes the full azimuthal range $-\pi \leq \varphi \leq \pi$ and the pseudorapidity

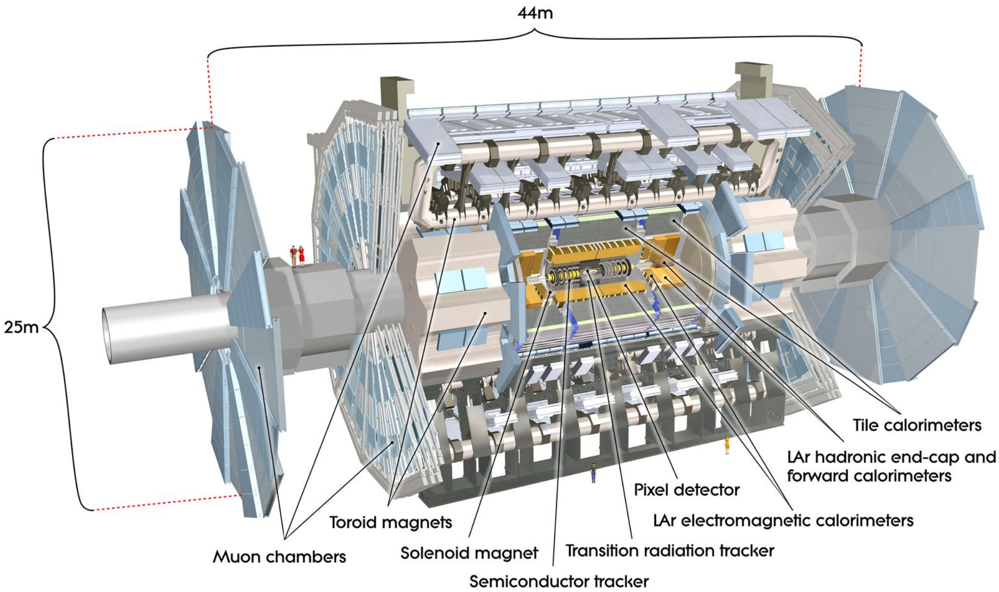


Figure 2.6: General view of the ATLAS detector.

region $|\eta| < 2.5$.[1] It is able to track particles with transverse momentum above 500 MeV, except for some minimum bias analyses where the threshold is 100 MeV . The ID is embedded in an axial magnetic field (2 T) which bends the trajectory of charged particles and allows to measure the charge. It consists of the three following subsystems, starting from the nearest to the IP is the Pixel detector, then the Semiconductor Tracker (SCT) and finally the Transition Radiation Tracker (TRT), as shown in Figure 2.7.

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, φ) are used in the transverse plane, φ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

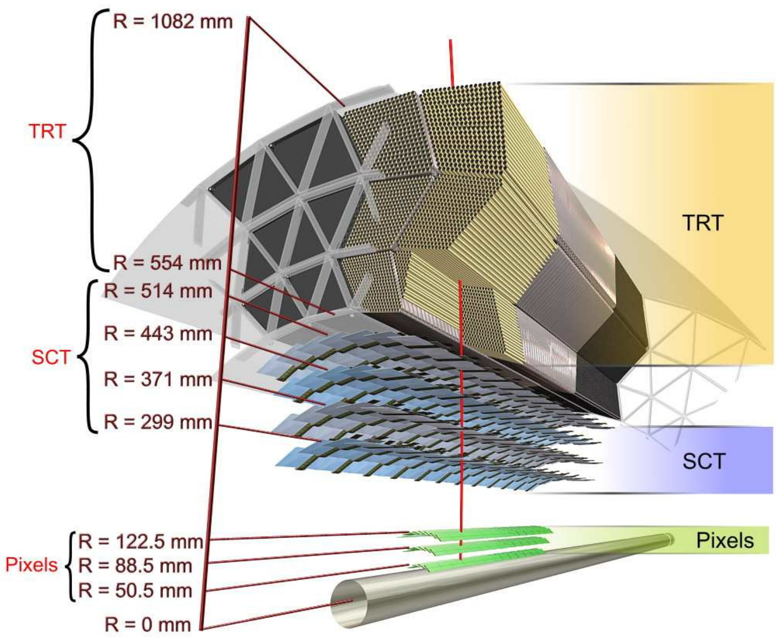


Figure 2.7: Transverse view of the ATLAS Inner Detector subsystems being traversed by a 10 GeV track with $\eta = 0.3$. Each track crosses 3 pixel layers, 4 SCT layers and approximately 35 axial straws in the TRT.

The overall transverse momentum resolution of the ID is:

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.05\% \times p_\mathrm{T} \oplus 1\% \tag{2.4}$$

The main parameters of the subsystems composing the Inner Detector are given in Table 2.1.

Subsystem	Position	Area (m^2)	Resolution σ (μm)	Channels ($\times 10^6$)	η coverage
Pixels	1 barrel layer	0.2	$R\varphi = 12$, $z = 66$	16	± 2.5
	2 barrel layers	1.4	$R\varphi = 12$, $z = 66$	81	± 1.7
	4 end-cap disks on each side	0.7	$R\varphi = 12$, $R = 77$	43	1.7-2.5
Silicon strips	4 barrel layers	34.4	$R\varphi = 16$, $z = 580$	3.2	± 1.4
	9 end-cap wheels on each side	26.7	$R\varphi = 16$, $R = 580$	3.0	1.4-2.5
TRT	Barrel straws		170 (per straw)	0.1	± 0.7
	End-cap straws 36 straws per track		170 (per straw)	0.32	0.7-2.5

Table 2.1: The main parameters of the inner detector subsystems. The resolutions quoted are typical values (the actual resolution depends on $|\eta|$).

2.2.1.1 The Pixel Detector

The main advantage of the pixel detector is its high granularity close to the interaction point. This drives the performance of the ID for finding short-lived particles such as b-quarks and τ-leptons.
The system has a highly modular design, with three barrels containing approximately 1500 identical barrel modules and four disks on each side containing 1000 disk modules. The support structure is of only one type in the barrel and another one in the disks.
The pixel modules for the disks and the barrel are very similar. Each barrel module contains 61440 pixel elements, with a readout system consisting in 16 chips. Each of these modules with a size of 62.4 mm $\times$ 22.4 mm.

2.2.1.2 The Semiconductor Tracker

The SCT contributes to the determination of the track momentum by providing four independent measurements per track in the intermediate radial range. This is also important for the reconstruction of the vertex position.
The barrel SCT contains four layers of silicon microstrip detectors, which provide precision measurements of the track coordinates. Each of them has a surface of 6.36×6.40 cm^2 and contain 768 readout strips parallel to the beam direction, providing measurements of the radial and azimuthal coordinates.

2.2.1.3 The Transition Radiation Tracker

The TRT consists of gaseous straw detectors, measuring the track radial and azimuthal coordinates. It covers the region up to $|\eta| = 2.0$ with a typical resolution of about 130 μm. Each straw is 4 mm in diameter with a maximum length of 150 cm.
The barrel contains 144 cm long straws, parallel to the beam direction with their wires divided into two halves approximately at $\eta = 0$. In the end-cap region, the 37 cm long straws are arranged radially in wheels. The number of readout channels is approximately 351000.

2.2.2 The Calorimeter System

The ATLAS calorimeter consists of two subdetectors: A Liquid Argon (LAr) electromagnetic calorimeter and a hadronic calorimeter. The main parameters are given in Table 2.2. The calorimeters cover the pseudorapidity region $|\eta| < 4.9$ and provide a good containment of electromagnetic and hadronic showers.
The EM calorimeter in the $|\eta|$ region covered by the ID presents a fine granularity, well suited for precision measurements of electrons and photons. The rest of the calorimeter has a coarser granularity able fullfil the physics requrements for jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. Fig. 2.8 shows a view of the ATLAS calorimeter.

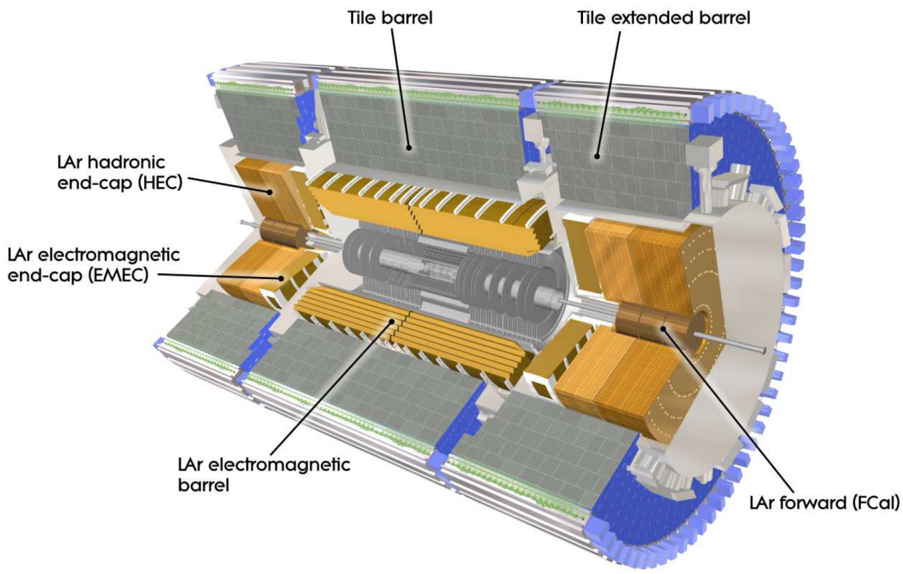


Figure 2.8: Cut-away view of the ATLAS calorimeter system.

2.2.2.1 The LAr electromagnetic calorimeter

The ATLAS electromagnetic calorimeter is divided into a barrel part, with $|\eta| < 1.475$, and two end-caps with $1.375 < |\eta| < 3.2$. The barrel calorimeter consists of two identical half-barrels, separated by a small gap (4 mm) at $z = 0$. Each end-cap is divided into two coaxial wheels: an outer wheel covering the range $1.375 < |\eta| < 2.5$ and an inner one covering $2.5 < |\eta| < 3.2$.

The layers of active material in the calorimeter depend on the pseudorapidity region (see Table 2.2). The granularity $\Delta\eta \times \Delta\varphi$ also depends on the η range. The central region of the barrel, where maximum precision is required, is finely segmented. A sketch of a barrel module is shown in Fig. 2.9, where the granularity is specified for cells in the central region.

A presampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter in the region $|\eta| < 1.8$. The presampler consists of an active LAr layer of 1.1 cm (0.5 cm) thick in the barrel (end-cap).

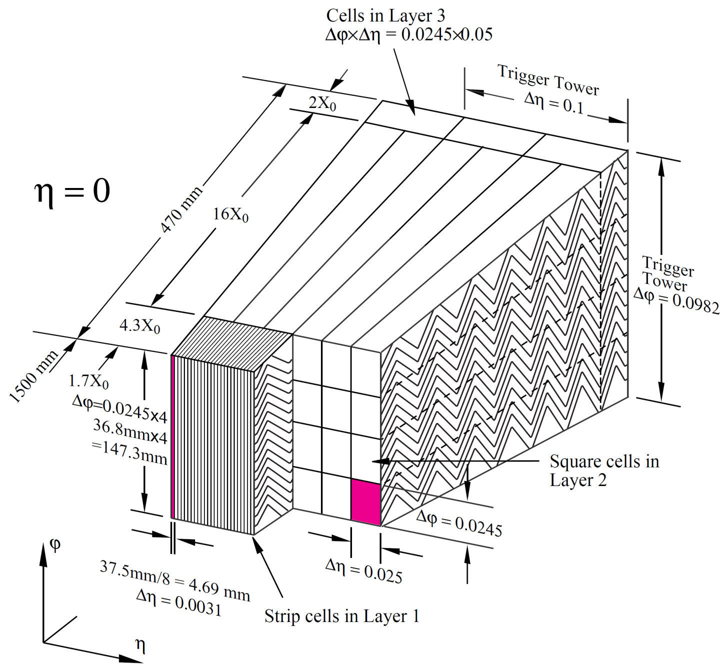


Figure 2.9: Sketch of a barrel module. The granularity in η and φ of the cells in each of the three layers and of the trigger towers is also shown.

The EM calorimeter is a lead-LAr detector with accordion geometry, see Figure 2.11, thus providing a complete azimuthal symmetry without cracks. The lead thickness in the absorber plates is chosen so that the energy resolution meets the requirement:

$$\frac{\sigma_E}{E} = \frac{0.1\ \mathrm{GeV}^{\frac{1}{2}}}{\sqrt{E}} \oplus 0.7\% \tag{2.5}$$

The first term is determined by the lead thickness, the second term, which becomes dominant at high energies, is determined by the resolution in the lead thickness shown in Figure 2.10.

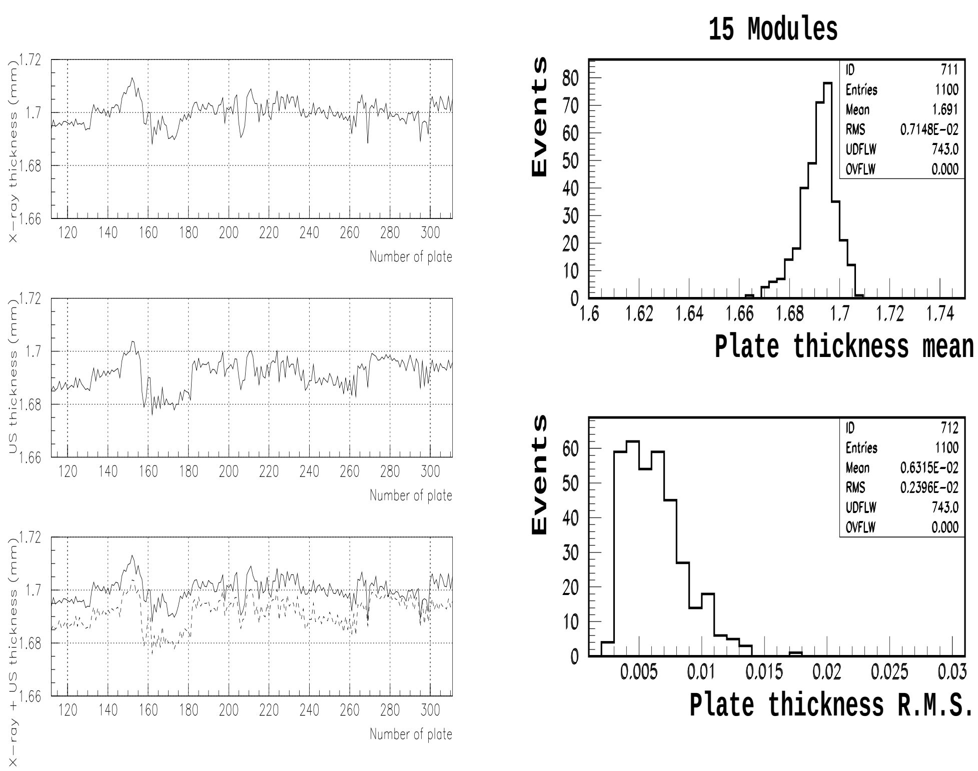


Figure 2.10: X-ray and ultrasound measurements of the lead thickness plates for the LAr endcap (left). Average lead thickness and its RMS (right).

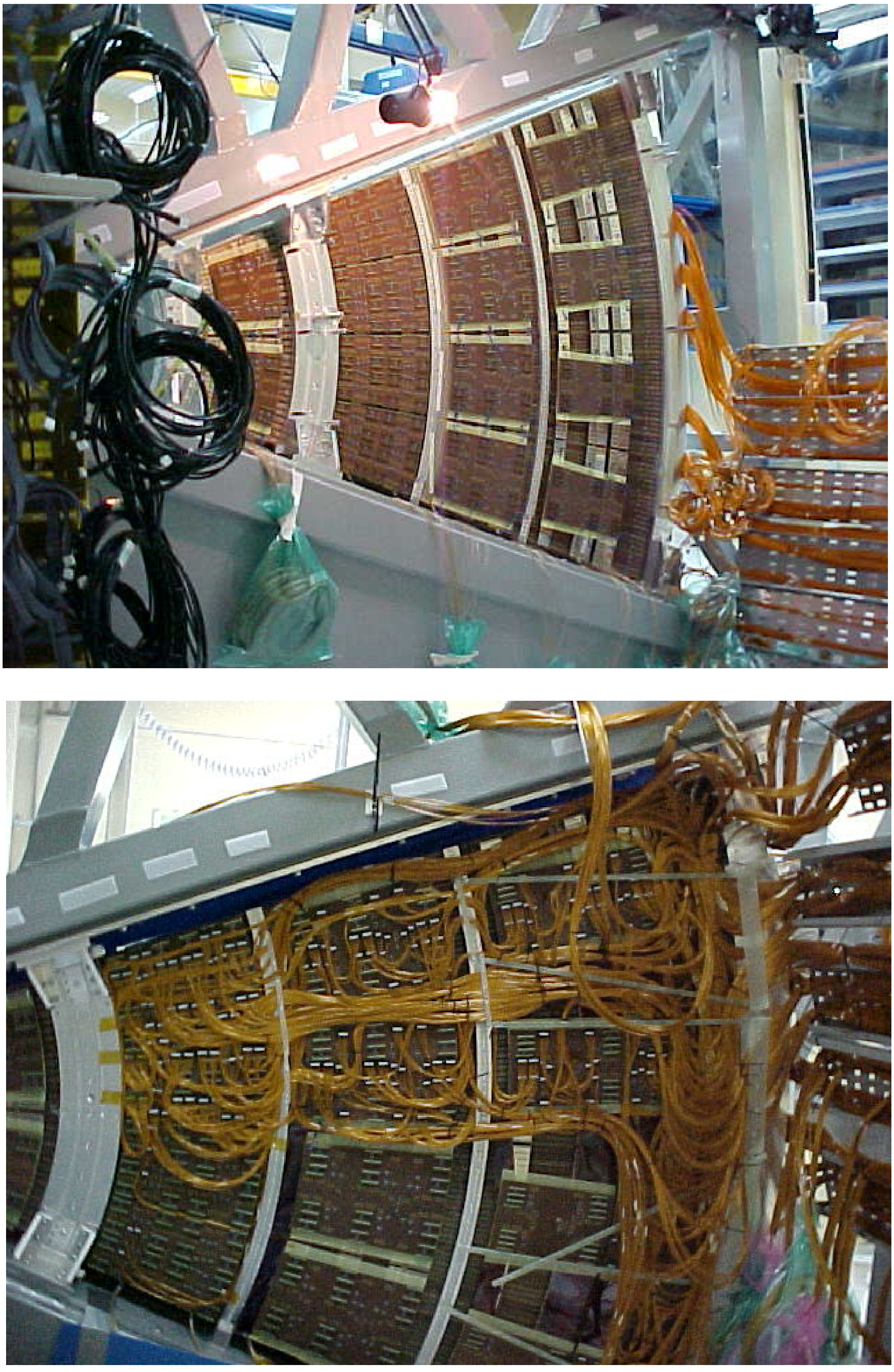


Figure 2.11: Picture of a LAr EC octant with the cold electronics mounted (top) and in the process of cabling (bottom).

2.2.2.2 The hadronic calorimeters

The hadronic ATLAS calorimeter consists of the following subsystems: the tile calorimeter, the LAr hadronic end-cap calorimeter (HEC) and the LAr forward calorimeter (FCal).

	Barrel		**End-cap**	
EM calorimeter				
Number of layers and $\|\eta\|$ coverage				
Presampler	1	$\|\eta\| < 1.52$	1	$1.5 < \|\eta\| < 1.8$
Calorimeter	3	$\|\eta\| < 1.35$	2	$1.375 < \|\eta\| < 1.5$
	2	$1.35 < \|\eta\| < 1.475$	3	$1.5 < \|\eta\| < 2.5$
			2	$2.5 < \|\eta\| < 3.2$
Granularity $\Delta\eta \times \Delta\varphi$ versus $\|\eta\|$				
Presampler	0.025×0.1	$\|\eta\| < 1.52$	0.025×0.1	$1.5 < \|\eta\| < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\|\eta\| < 1.40$	0.050×0.1	$1.375 < \|\eta\| < 1.425$
	0.025×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.1	$1.425 < \|\eta\| < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \|\eta\| < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \|\eta\| < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \|\eta\| < 2.4$
			0.025×0.1	$2.4 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\|\eta\| < 1.40$	0.050×0.025	$1.375 < \|\eta\| < 1.425$
	0.075×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.025	$1.425 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 3rd layer	0.050×0.025	$\|\eta\| < 1.35$	0.050×0.025	$1.5 < \|\eta\| < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$\|\eta\|$ coverage			$1.5\|\eta\| < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\varphi$			0.1×0.1	$1.5 < \|\eta\| < 2.5$
			0.2×0.2	$2.5 < \|\eta\| < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$\|\eta\|$ coverage			$3.1 < \|\eta\| < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			FCal1: 3.0×2.6	$3.15 < \|\eta\| < 4.30$
			FCal1: $\sim$ four times finer	$3.10 < \|\eta\| < 3.15$
				$4.30 < \|\eta\| < 4.83$
			FCal2: 3.3×4.2	$3.24 < \|\eta\| < 4.50$
			FCal2: $\sim$ four times finer	$3.20 < \|\eta\| < 3.24$
				$4.50 < \|\eta\| < 4.81$
			FCal3: 5.4×4.7	$3.32 < \|\eta\| < 4.60$
			FCal3: $\sim$ four times finer	$3.29 < \|\eta\| < 3.32$
				$4.60 < \|\eta\| < 4.75$
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$\|\eta\|$ coverage	$\|\eta\| < 1.0$		$0.8 < \|\eta\| < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\varphi$	0.1×0.1		0.1×0.1	
Last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

Table 2.2: The main parameters of the calorimeter subsystems.

- **Tile calorimeter:** The tile calorimeter barrel covers the region $|\eta| < 1.0$ and its two extended barrels the range $0.8 < |\eta| < 1.7$. It is a sampling calorimeter with steel absorber plates and scintillating tiles as active material. The barrel and extended barrels are divided into 64 modules extended azimuthally. The inner (outer) radius is 2.28 m (4.25 m). It is segmented in depth in three layers.

- **LAr hadronic end-cap:** The HEC is located directly behind the end-cap electromagnetic, both sharing the same LAr cryostats. The HEC overlaps with the forward calorimeter to reduce the drop in the material density at the transition region around $|\eta| = 3.1$, extending up to $|\eta| = 3.2$. Each wheel is built from 32 identical wedge-shaped modules. Each wheel is divided into two segments in depth for a total of four layers per end-cap.

- **LAr forward calorimeter:** The FCal, integrated in the LAr cryostat, is approximately 10 interaction lengths deep, and consists of three modules in each end-cap: the first, made of copper, is optimised for electromagnetic measurements, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions.

The energy resolution of the hadronic calorimeters is given by

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \tag{2.6}$$

For the barrel and end-cap hadronic calorimeters, the values of a and b are $a = 0.5\ \mathrm{GeV}^{\frac{1}{2}}$ and $b = 3\%$, while the forward calorimeter has a poorer resolution with $a = 1\ \mathrm{GeV}^{\frac{1}{2}}$ and $b = 10\%$.
The structure of a cryostat, showing the relative positions of the three forward calorimeters is depicted in Fig. 2.12. The outer radius of the cylindrical cryostat vessel is 2.25 m and the length of the cryostat is 3.17 m. The feed-throughs and front-end crates containing the readout electronics are also shown.

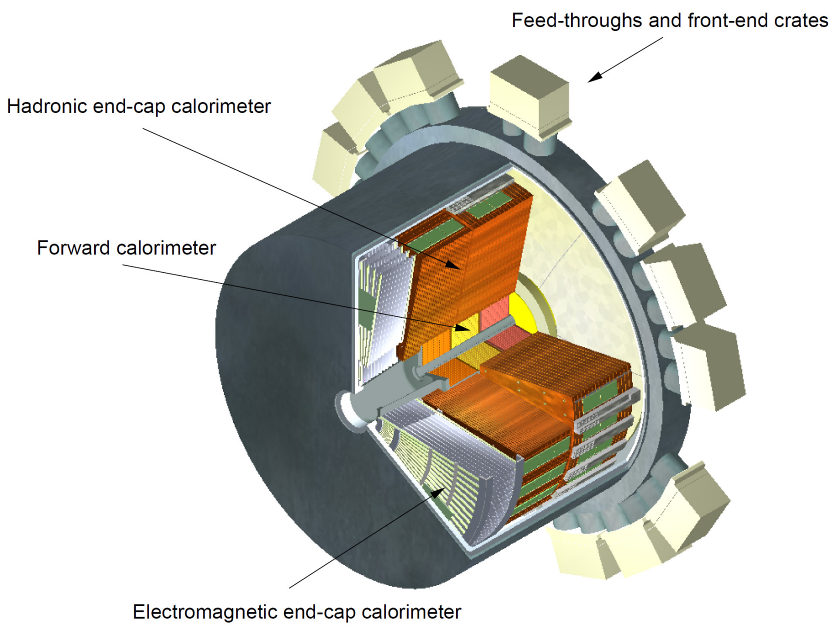


Figure 2.12: Cut-away view of an end-cap cryostat, showing the positions of the three end-cap calorimeters.

2.2.3 The Muon Spectrometer

The ATLAS muon system exploits the deflection of muon tracks in a magnetic field. The bending power is characterised by the line integral $\int \vec{B} \cdot \vec{dl}$. In the range $|\eta| < 1.4$, the bending is provided by the large barrel toroid with a bending power between 1.5 to 5.5 Tm.. For $1.6 < |\eta| < 2.7$, the muon trajectories are deflected by two smaller end-cap magnets inserted at both ends of the barrel toroid and providing a bending power between 1 and 7.5 Tm. In the so called transition region $1.4 < |\eta| < 1.6$, the deflection is provided by a combination of barrel and end-cap fields providing a somewhat lower bending power.
In the barrel region, tracks are reconstructed in chambers arranged in three cylindrical layers around the beam axis. In the transition and end-cap regions, the chambers are installed in planes perpendicular to the beam, also in three layers. This is illustrated in Fig. 2.13.

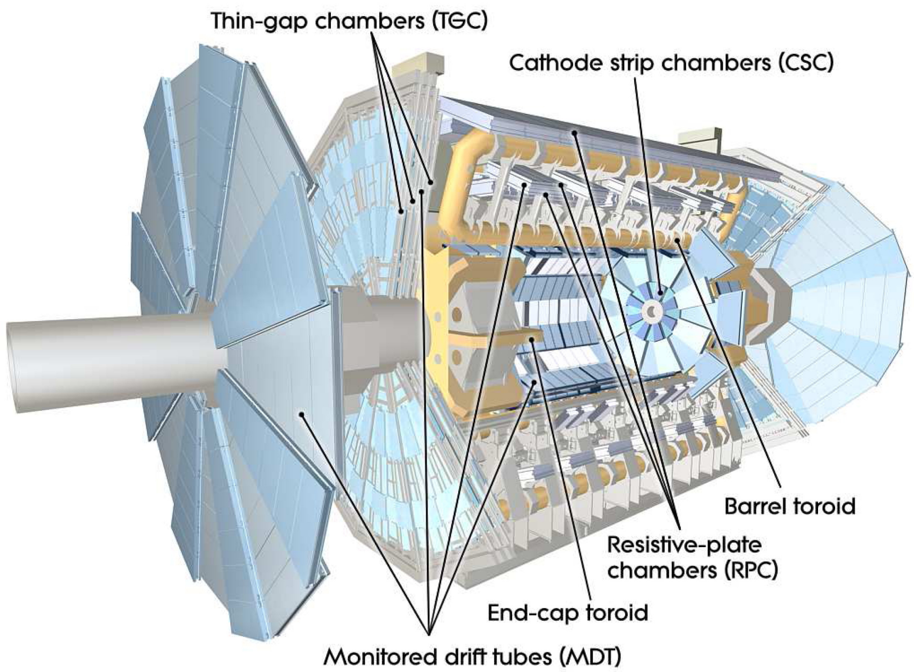


Figure 2.13: Cut-away view of the ATLAS muon system.

The muon spectrometer consists of four subsystems: the Monitored Drift Tubes (MDT), the Cathode Strip Chambers (CSC), the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC). The first two provide precision tracking, while the last two provide triggering capabilities. The main parameters of the muon chambers are listed in Table 2.3

System	Coverage	Number of Chambers	Number of Channels	Function
MDT	$\|\eta\| < 2.7$	1150	354000	Precision tracking
CSC	$2.0 < \|\eta\| < 2.7$	32	31000	Precision tracking
RPC	$\|\eta\| < 1.05$	606	373000	Triggering, second coordinate
TGC	$1.05 < \|\eta\| < 2.7$	3588	318000	Triggering, second coordinate

Table 2.3: The main parameters of the subsystems conforming the muon spectrometer.

2.2.3.1 Precision tracking chambers. MDTs and CSCs.

The MDT chambers are made of three to eight drift tube layers, operated at an absolute pressure of 3 bar and achieving an average resolution of 80 μm per tube, or about 35 μm per chamber. The MDTs perform the precision momentum measurement in the pseudorapidity range $|\eta| < 2.7$.
The CSCs are multiwire proportional chambers with cathode planes segmented into strips in orthogonal directions. This allows the measurement of both coordinates from the induced-charge distribution. The resolution of a chamber is 40 μm in the bending plane and about 5 mm in the transverse plane. They cover the pseudorapidity region $2 < |\eta| < 2.7$.

2.2.3.2 Triggering chambers. RPCs and TGCs.

The precision-tracking chambers are complemented by a system of fast trigger chambers capable of delivering track information within a few tens of nanosecond. The RPCs (TGCs) were chosen in the barrel region $|\eta| < 1.05$ (in the end-caps $1.05 < |\eta| < 2.4$) for this purpose. The spread of the delivered signals is in the range 15-25 ns. They also measure one coordinate in the bending plane, in the η direction, and one in the non-bending azimuthal plane.
The purpose of the precision-tracking chambers is to determine the coordinate of the track in the bending plane. After matching of the MDT and trigger chamber hits in the bending plane, the trigger chamber's coordinate in the non-bending plane is adopted as the second coordinate of the MDT measurement. This method assumes that in any MDT/trigger chamber pair a maximum of one track per event be present, since with two or more tracks the η and φ hits cannot be combined in an unambiguous way.

2.2.4 The ATLAS Trigger System

The ATLAS Trigger system consists of three consecutive levels of event selection: level 1 (L1), level 2 (L2) and Event Filter (EF), collectively referred to as the High Level Trigger (HLT). It reduces the event rate to about 200 Hz from the nominal bunch crossing rate of 40 MHz. Fig. 2.14 shows a flow chart of the three trigger levels.

2.2.4.1 Level 1

The L1 triggers are hardware based and use coarse resolution information from dedicated detectors like the calorimeters and the muon spectrometer to identify high energy activity in the so called regions of interest (RoI). The calorimeter

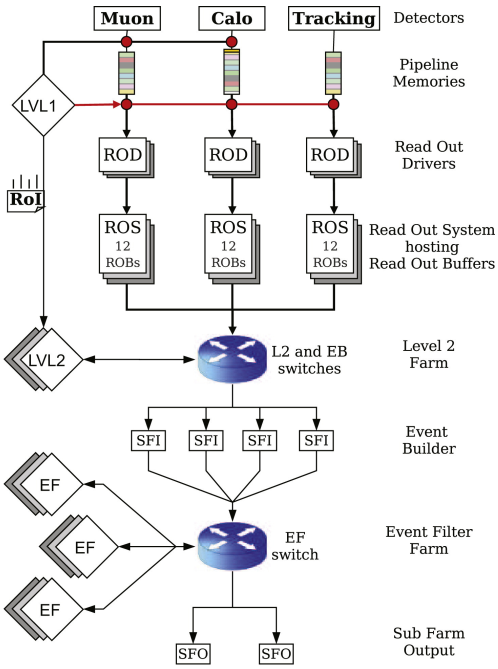


Figure 2.14: Flow chart of the ATLAS Trigger system [38]

selection (L1Calo) is based on the information provided by all the calorimeter subsystems, being consistent with high-E_{T} objects such as electrons, photons, hadronically decaying τ-leptons or jet clusters, as well as events with large $E_{\mathrm{T}}^{\mathrm{miss}}$. The L1Calo is designed to work with 7000 analogue trigger towers of reduced granularity ($\Delta\eta \times \Delta\varphi = 0.1 \times 0.1$ in the central region, but larger at high $|\eta|$), see Fig. 2.9. It sends the signal to the L1 computer farm approximately 1.5 μs after the event occurs.

The L1 muon trigger is based on the information provided by the muon chambers with triggering capabilities in the barrel and end-cap regions, namely the RPCs and the TGCs respectively. The algorithm is based on the coincidence of aligned hits in different trigger stations, tracking the trajectory of a muon. The low-p_{T} trigger is capable of tracking muons with a threshold range of about 6–9 GeV, while the high-p_{T} trigger has a threshold which varies from 9 to 35 GeV.

2.2.4.2 Level 2 and Event Filter

The L2 triggers make use of fast software based online data reconstruction algorithms. The L2 algorithms also use information which is not available at the L1 level, such as the Inner Detector tracks. The L2 trigger has access to partial event data mainly from the L1 RoIs and decides whether to keep or reject an

event in a few tens of ms, thus reducing the event rate below 3.5 kHz.
The Event Filter triggers use offline data reconstruction algorithms to provide additional rejection to reach the desired 200 Hz rate. The EF has access to the complete data from a given event. The EF selection is made on the basis of combinations of candidate physics objects after the event building step. The average event processing time is about 4 s.
For this analysis, the relevant triggers select events with two oppositely charged leptons.

Chapter 3

Object reconstruction and identification

In this chapter, the performance of the ATLAS detector in terms of physics object reconstruction and identification is briefly discussed, including tracks and vertices, electrons, muons, jets, b-jets and missing transverse energy. These objects will be later used in the physics analyses presented in this thesis.

3.1 Tracks and vertices

Track reconstruction is essential for electron and muon identification, as well as for track-jet studies. It is done using all subsystems in the inner detector, including the pixel detector, the SCT and the TRT. The track reconstruction software, using global-χ^2 and Kalman-filter techniques [39], divides the pattern recognition processing into three stages:

- **Pre-processing:** The raw data from the pixel and SCT detectors are converted into clusters, while the TRT raw timing information is calibrated. The SCT clusters are transformed into space-points, combining the cluster information from opposite sides of the SCT modules.

- **Track-finding:** Track seeds are formed by combining space-points in the pixel detector and the first SCT layer. These seeds are then extended throughout the SCT to form track candidates, which are then fitted and subject to quality cuts. The selected tracks are then extended into the TRT to associate the drift-circle information. Finally, the extended tracks are refitted including the information from all three subsystems.

- **Post-processing:** A vertex-finding algorithm is used to reconstruct primary vertices, followed by the reconstruction of photon conversions and secondary vertices. The vertex candidates are selected by minimsing a χ^2 function which depends on the (x, y, z) position as well as on the track momenta $\vec{p}_i$. The primary vertex is defined as the one maximising the quantity $\sum_i (p_{\mathrm{T}}^i)^2$, where the index i runs over all tracks originating from it.

3.2 Electrons

In order to achieve physics goals such as the measurement of the $H \to 4\ell$ or that of top-quark properties, a high efficiency electron reconstruction is required. The electron reconstruction starts with a clustering algorithm maximising the energy containment and follows with differents sets of cuts to separate the high jet background from the electron signal.

3.2.1 Calorimeter-seeded reconstruction

Electron (and photon) reconstruction starts with a three-step clustering algorithm, referred to as the sliding-window method, described in detail in Ref. [40].

- **Tower building:** The η–φ space in the calorimeter is divided into a grid of towers with size $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$, with the constraint $|\eta| < 2.5$.

- **Seed finding:** A window of 5×5 EM cells is moved across the tower grid defined in the previous step. The sum of the transverse energy of each cell within the window is required to be greater than 3 GeV and to be a local maximum. If these two conditions are fulfilled, then the window is selected as a precluster with η and φ calculated as the energy-weighted average of the position of all cells within a new 3×3 window around the center of the sliding window. Precluster duplicates are removed.

- **Cluster formation:** Once the seeds are defined, a layer-by-layer clustering algorithm is ran by considering all 3×7 (Barrel) and 5×5 (Endcap) rectangles centred in a given seed position. The algorithm starts with the middle layer, using as seed the precluster position computed in the previous step. The strip layer is done in the second place, taking as input the middle-layer barycenter. Finally, the Presampler and back layers are included using the strip and middle layer barycenters as seeds.

This cluster is said to be matched to a track when the extrapolation of this track from the inner detector to the LAr calorimeter is within $\Delta\eta \times \Delta\varphi = 0.05 \times 0.10$ of the cluster position. If the ratio $E/|\vec{p}|$ between the cluster energy and the track momentum is smaller than 10, both the track and the cluster are combined to form the electron candidate. Its energy is calculated as the weighted average

between the cluster energy and the track momentum. Its angular coordinates η and φ correspond to those from the track.
The electron reconstruction efficiency is high: Approximately 93% of true isolated electrons with $E_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ are selected as candidates. The 7% inefficiency is understood to arise from the large amount of material in the inner detector, and therefore is η-dependent.

3.2.2 Electron identification

The major goal of the ATLAS electron identification algorithm is to separate the electron signal from the huge multijet background arising from QCD mediated processes. This is achieved by considering different sets of cuts based on several variables. Three electron identification categories referred to as Loose, Medium and Tight [41], are built using this information

3.2.2.1 Loose cuts

The loose set of cuts makes use of the different shower shapes for jets and electrons, as well as their different level of penetration in the hadronic calorimeter. They involve a cut in the lateral shower width, defined as

$$w_{\eta,2} = \sqrt{\frac{\sum_i E_i \eta_i^2}{\sum_i E_i} - \left(\frac{\sum_i E_i \eta_i}{\sum_i E_i}\right)^2} \tag{3.1}$$

The sum is calculated in a window of 3×5 cells centred at the cluster position, with E_i the cell energy and η_i its pseudorapidity. The loose cuts provide a background rejection factor of about 500 and a high efficiency selection.

3.2.2.2 Medium cuts

The medium cuts provide a jet rejection factor of 5000. They include information from both the calorimeter and the inner detector. They provide a good rejection of the neutral pion decays to two photons, $\pi^0 \to \gamma\gamma$. In the case of π^0 decays, the calorimeter is capable of resolving a second maximum with energy $E_{\mathrm{max},2}$ in a window of size $\Delta\eta \times \Delta\varphi = 0.125 \times 0.2$ around the cell with maximum energy deposit. The calorimetric variables also include the total shower width, defined as

$$w_{\mathrm{stot}} = \sqrt{\left[\sum_i E_i (i - i_{\mathrm{max}})^2\right] \sum_i E_i} \tag{3.2}$$

In this case, the index i runs over all strips in a window of $\Delta\eta \times \Delta\varphi = 0.0625 \times 0.2$, and i_{max} is the index of the maximum-energy strip. The tracking variables include the number of pixel and SCT hits as well as the transverse impact parameter, defined as the transverse distance of closest approach to the primary vertex. A track-cluster matching within $|\Delta\eta_{\mathrm{ct}}| < 0.01$ is also applied.

3.2.2.3 Tight cuts

The tight cuts provide a rejection factor of 50000 thanks to the tightening of the cuts defining the track-cluster spatial matching ($|\Delta\varphi| < 0.02$, $|\Delta\eta| < 0.005$) as well as the ratio E/p. Electrons from photon conversions are also discarded.

3.3 Muons

The ATLAS muon system is designed to provide momentum measurement of muons in the pseudorapidity region $|\eta| < 2.7$ with a relative resolution better than 3% over a wide p_{T} range and up to 10% at transverse momenta close to 1 TeV. This is important for the reconstruction of many physics processes such as $t\bar{t}$ production treated in this thesis, new physics searches like $Z' \to \mu\mu$ or multilepton SUSY searches.
The ATLAS muon reconstruction algorithms combine the information from the muon specrometer with that from the inner detector and to a lesser extent from the calorimeter leading to different muon classes, namely:

3.3.1 Standalone muons

The standalone muon algorithm reconstructs the muon trajectory using only the information provided by the muon spectrometer. The parameters of the track at the interaction point are determined by extrapolation to the perigee using the Staco [42] or Muid [43] algorithms. In the Staco case, the extrapolation algorithm is called Muonboy [42], while the Moore [44] algorithm is used on the Muid side. In this extrapolation, energy losses in the calorimeter are taken into account. The Muonboy algorithm does this correction by estimating the amount of material crossed by the muon depending on its pseudorapidity, while Moore also includes calorimetric measurements.
Muons produced in hadron decays in the calorimeter, such as π and K decays, are normally reconstructed in the standalone scheme. They give an estimation of these kind of backgrounds when the matching to the inner detector case is performed.
Standalone muons are mainly used to extend the muon acceptance to the pseudorapidity range $2.5 < |\eta| < 2.7$ which is not covered by the inner detector.

3.3.2 Combined muons

Combined muons are formed from the succesful combination of a muon spectrometer track and an inner detector track.

- Segment-tagged muons (ST): a track in the ID is classified as a muon if, once extrapolated to the muon spectrometer, it is associated with at least one track segment in the MDT or CSC chambers. ST muons are used to increase the acceptance in cases where low transverse momentum muons crossed only one layer of the MS chambers.
- Calorimeter-tagged muons (Calo-Tag): A track in the ID is identified as a muon if it can be associated to an energy deposit in the calorimeter consistent with a minimum ionizing particle. This type of muon has the lowest putity.

3.4 Jets

Jets play an important role in the analyses presented in this thesis. As already discussed in Section 1.5, jets are collimated showers of hadrons, defined in an attempt to reconstruct the kinematic properties of the partons from which they originate. The primary objects used in ATLAS for jet building are topological clusters, constructed from calorimeter cells and used as input four-momenta for the anti-k_t algorithm. Then, further calibrations are needed to correct the jet energy measurement for calorimetric non-compensation, leakage and other detector effects.

3.4.1 Topological clusters

The basic idea of the topological clustering algorithms [40] is to combine groups of cells with an amount of energy well above the noise threshold. Topological clusters are seeded by calorimeter cells with energy $|E| > 4\sigma$, where σ is the cell by cell RMS of both the electronic and pile-up noise. Neighbouring cells are added if $|E| > 2\sigma$ and clusters are formed with the help of an iterative procedure as sketched in Fig. 3.1.

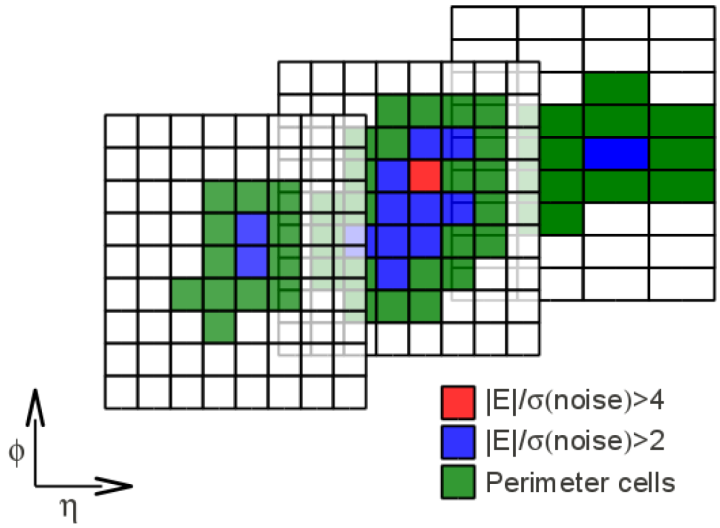


Figure 3.1: Sketch of the formation of a 3-dimensional topological cluster attending to the signal to noise ratio of the constituent calorimeter cells.

Once the topological clusters are built as described above, one can use these clusters as inputs to the jet algorithms described in Section 1.5 and then apply the jet calibration procedure (EM clusters), or one can calibrate their energy individually before any attempt to reconstruct the jets. This second procedure is called the local cluster weighting (LCW) calibration [45, 46]. In the LCW method, clusters are classified into two types, electromagnetic or hadronic, depending on their energy, depth in the calorimeter and cell energy density. Three weights are then applied separately to account for different effects:

- The calorimeter cells in the clusters are weighted according to the cluster energy and the energy density of the cell, thus taking into account the response to hadrons in the calorimeter using test beam data.
- The cluster is weighted according to the energy deposits in the cluster neighbourhood and to the longitudinal depth of the cluster barycenter λ_c, thus taking into account energy deposits not contained within the cluster.
- Finally, the cluster is weighted according to its energy and the fraction of energy in each layer of the calorimeter, thus accounting for energy deposited in the calorimeter dead material.

The depth of the cluster barycenter is characterized by the distance λ_c from the front of the calorimeter to the center of the shower. The center of the shower has spatial coordinates given by

$$\langle x_i \rangle = \frac{\sum_k E_k x_i^{(k)}}{\sum_k E_k} \tag{3.3}$$

The index i runs over the three spatial coordinates (x, y, z), and the index k runs over all cells within the clusters with positive energy $E_k > 0$. After these weighting techniques, the resulting energy of the clusters is compared to the energy before the weighting (EM scale). Fig. 3.2 shows the cluster average response $\langle E_{\text{calib}}/E_{\text{EM}} \rangle$ as a function of the cluster energy and pseudorapidity for the three different weighting procedures described above to correct for the low hadronic response, energy deposits outside the topological cluster and non-active material in the calorimeter. The jets used for these studies have transverse momenta above 20 GeV.

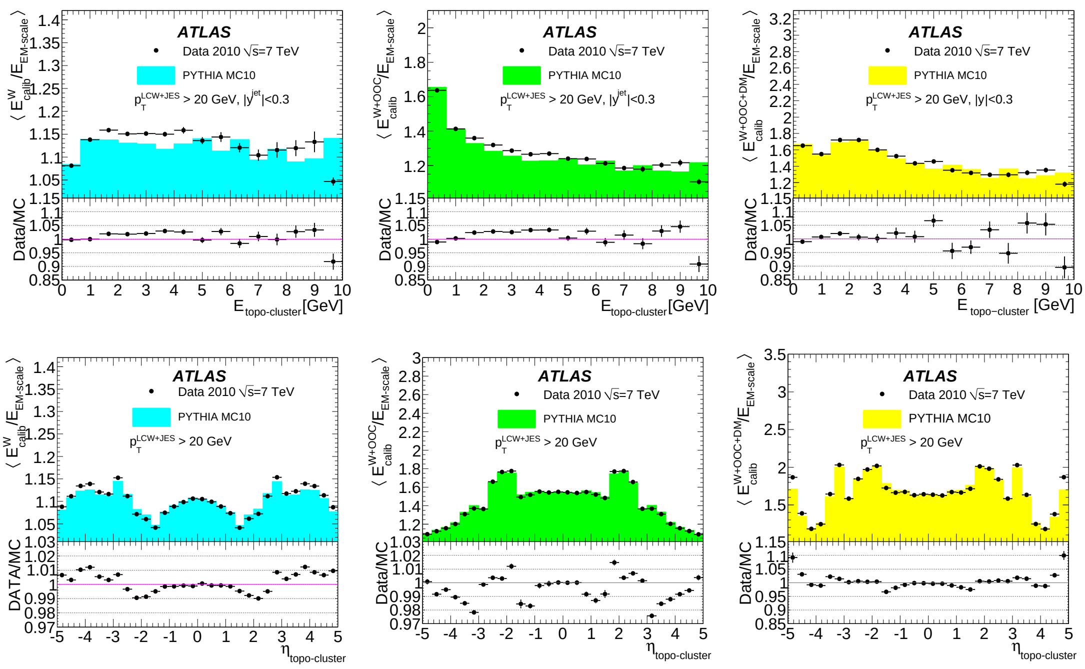


Figure 3.2: Average value of the topocluster response $\langle E_{\text{calib}}/E_{\text{EM}} \rangle$ for hadronic response (left column), out-of-cluster weights (center column) and dead material weights (right column). The top row presents the results as a function of the cluster energy for jets with $|y| < 0.3$, while in the bottom row results are presented as a function of the cluster pseudorapidity [46].

The results above show that the agreement between the data and the Monte Carlo simulation for the average cluster responses to each weighting procedure is within 5%, and it is better for low energy and central pseudorapidity clusters.

3.4.2 Jet calibration

Jets seeded from the EM or LCW topological clusters described above are then reconstructed using the anti-k_t sequential recombination algorithm described in Section 1.5. Once the jets four-momenta $(E, p_{\mathrm{T}}, \eta, \varphi)$ are reconstructed, a jet energy calibration algorithm restores the jet energy to that of the corresponding particle level jet. This is done following a three-step process [47]

3.4.2.1 Pile-up and origin correction

The pile-up correction is designed to remove the contamination due to both in-time and out-of-time pile-up interactions. This corrections is MC based and taks into account the number of primary interaction vertices N_{PV}, which is a measure of the in-time pile-up, and the average number of interactions per bunch crossing μ, which is a measure of the out-of-time pile-up.

3.4.2.2 Calibration from MC simulations

The calibration of the energy of a given jet is performed by multiplying it by simple correction factors. To this end, it is useful to define the jet response function as

$$\mathfrak{R}^{\mathrm{EM\ (LCW)}} = \frac{E_{\mathrm{reco}}^{\mathrm{EM\ (LCW)}}}{E_{\mathrm{truth}}} \tag{3.4}$$

This response function is the inverse of the correction factors, calculated in bins of the jet energy and pseudorapidity. Figure 3.3 shows the jet response for jets built from both EM and LCW clusters in the full η_{det} range of the detector.

3.4.2.3 In situ corrections

The last step of the jet calibration is performed by deriving an in situ correction using relative calibrations between the central and forward rapidity regions. This is referred to as the η intercalibration and it is described in Ref. [47]. The jet calibration is also tested in data using the jet momentum balance techniques in several physics processes such as $Z/\gamma^* + 1$ jet, γ + jet.

3.4.3 Uncertainties in the JES calibration procedure

The main source of uncertainty due to the calibration procedure is derived using in situ measurements of the jet balance in event samples for Z/γ^* + jet, γ + jet and inclusive jet production. For jets with $p_{\mathrm{T}} > 1$ TeV, the uncertainty is estimated using the calorimeter response to single pions [48]. These, together with

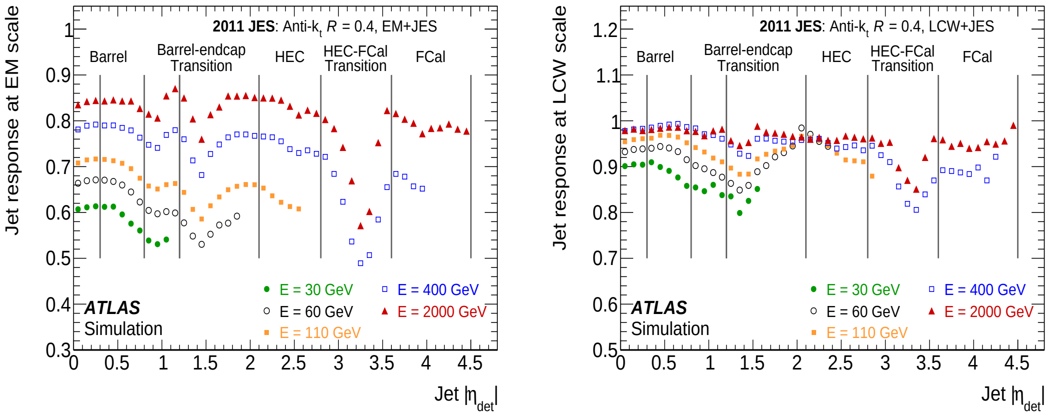


Figure 3.3: Average value of the jet response as a function of the pseudorapidity $|\eta_{\rm det}|$ for several values of the jet energy E for both the EM (left) and the LCW (right) calibration schemes [47].

the uncertainty in the η-intercalibration mentioned above define the baseline JES uncertainty.Additional uncertainties due to the different response of the calorimeter to gluon and quark jets, and the limited knowledge of the gluon fraction in the jet sample are also included in the total JES uncertainty.

3.4.4 The JVF algorithm

In order to reject jets stemming from pile-up interactions, an algorithm based on the relative amount of the jet transverse momentum due to charged tracks ($\rm trk_{k,l}$) associated to the primary vertex (PV) to that of all possible vertices ($\rm vtx_n$) has been developed. This is called the 'Jet Vertex Fraction' algorithm or JVF [49] and is defined for a given jet and PV as:

$$\mathrm{JVF} = \frac{\sum_k p_{\rm T}(\mathrm{trk_k}, \mathrm{PV})}{\sum_n \sum_l p_{\rm T}(\mathrm{trk_l}, \mathrm{vtx_n})} \tag{3.5}$$

Jets from the hard scattering process pointing to the primary vertex are characterised by large values of the JVF while those coming from pile-up interactions are characterized by small values.

3.5 Identification of jets containing B-hadrons

The identification of jets containing hadrons composed of at least one b-quark, or simply b-tagging, plays a key role in ATLAS. This is dictated by the great

importance of processes like $H \to b\bar{b}$ or $t \to Wb$.
The identification of jets containing B-hadrons relies on several properties of these jets. Due to the large lifetime τ of these hadrons, their flight distance with respect to the primary interaction vertex can be resolved in a displaced vertex. Additionally, the average impact parameter of the tracks within one of these jets is greater than that of the tracks within a light jet. Figure 3.4 illustrates these properties. Note that there is a large variety of b-tagging algorithms, each of them with its own particularities. Figure 3.5 shows the light-jet rejection factor versus the b-jet tagging efficiency for some of these taggers, together with the b-jet tagging efficiency as a function of the jet p_{T} for the MV1 tagger which is the one used in this thesis.

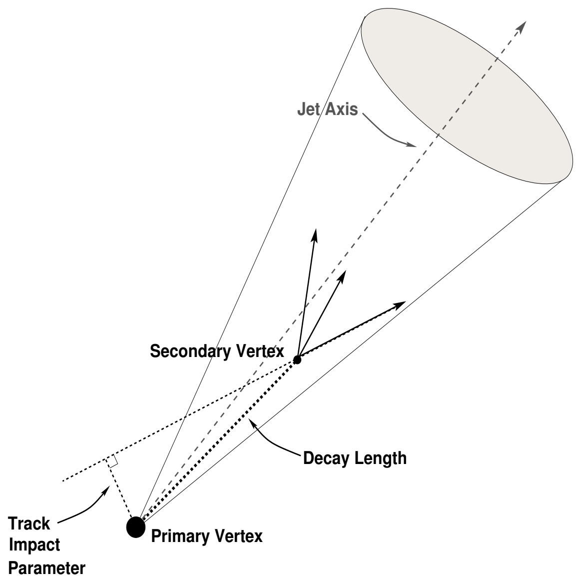
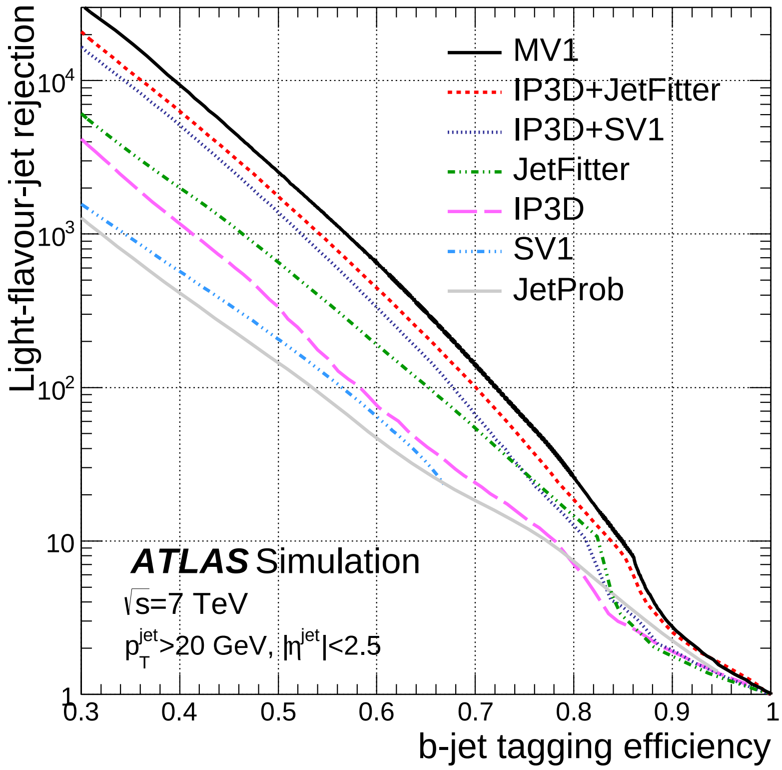


Figure 3.4: Sketch of a jet containing a B-hadron, showing the largest impact parameter of the tracks within the jet as well as the displaced secondary vertex from the B-hadron decay.



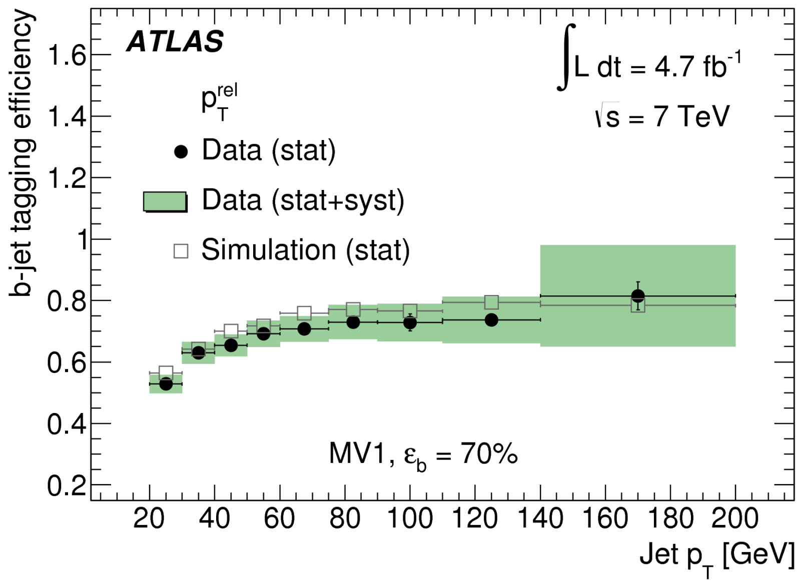


Figure 3.5: Light-jet rejection as a function of the b-tagging efficiency for several b-tagging algorithms (top) and efficiency versus jet p_{T} (bottom) for MV1 [50].

3.5.1 Track selection

The properties of the tracks used for b-tagging are descibed in Ref. [51]. This selection is based on the following quality requirements for these tracks:

- At least seven precision hits (pixel or silicon micro-strip hits) on the track.
- At least two hits in the pixel detector, one or more of which must be in the so-called B-layer, the innermost layer of the pixel tracker.
- The transverse momentum of the track must be above 1 GeV.
- The transverse impact parameter has to fulfill $|d_0| < 1$ mm.
- The longitudinal impact parameter has to fulfill $|z_0| \sin\theta < 1.5$ mm.

The tracks selected following the above criteria are then associated to jets depending on their distance ΔR to the axis of a given jet. The association cut is varied depending on the jet p_T, because high-p_T are more collimated, in order to reduce tracks stemming from the underlying event or pile-up interactions which would reduce the discrimination. The cut varies from $\Delta R = 0.45$ at 20 GeV to 0.25 for more energetic jets around 150 GeV. To avoid ambiguities, one track is associated univocally to one jet, the one closest in ΔR.
The transverse and longitudinal impact parameters d_0 and z_0 of the selected tracks are often signed. If the extrapolation of the track crosses the jet axis in front of the primary interaction vertex, its impact parameter is considered as positive, and otherwise negative. The discriminating variable used for b-tagging purposes is the impact parameter significance, defined as the signed impact parameter divided by its error. Figure 3.6 shows the significances for both the transverse and longitudinal impact parameters of tracks associated to jets.

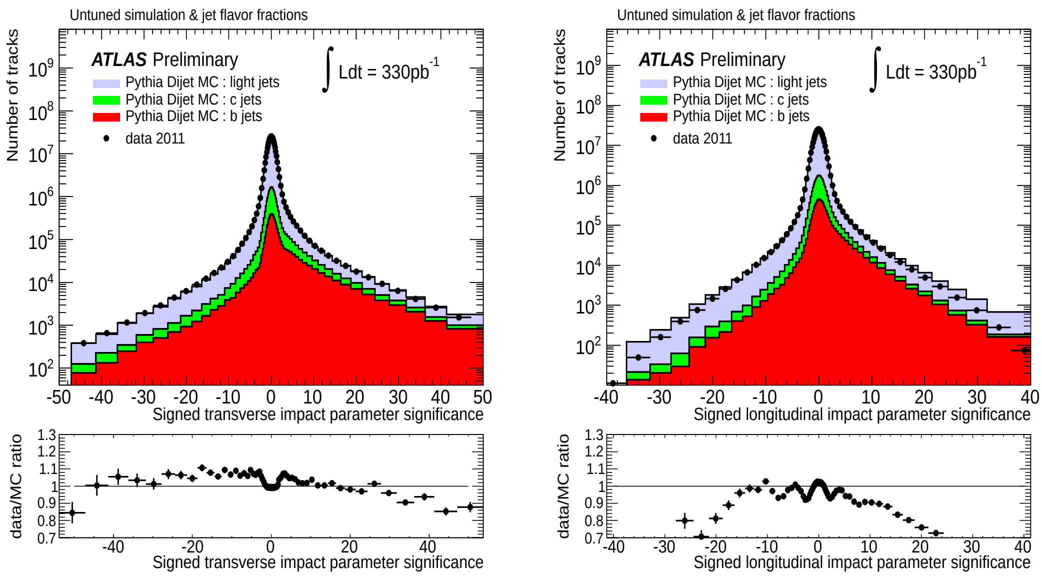
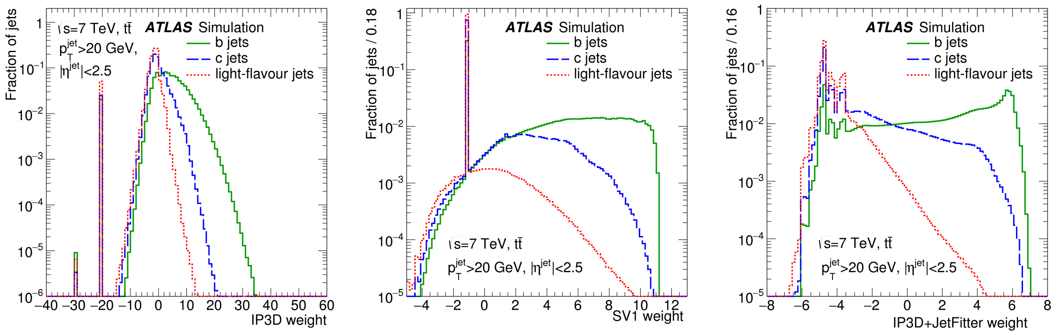


Figure 3.6: Data to MC comparison of the transverse impact parameter significance (left) and the longitudinal impact parameter significance (right) for tracks fulfilling the b-tagging quality criteria and associated to jets. The Monte Carlo predictions are separated in jets arising from the hadronisation of light quarks and gluons (blue), charm quarks (green) and bottom quarks (red) [51].

3.5.2 The MV1 algorithm

MV1 is a b-tagging algorithm [50] used for ATLAS analysis during Run 1. Distributions of the three MV1 input variables (the IP3D and SV1 discriminants as well as the sum of the IP3D and JetFitter discriminants) are shown in Figure 3.7, for b-jets, c-jets, and light-flavour jets in simulated $t\bar{t}$ events.



(a) IP3D (b) SV1 (c) IP3D+JetFitter

Figure 3.7: Distribution of the IP3D (a), SV1 (b) and IP3D+JetFitter (c) weights, for b, c and light-flavour jets. These three weights are used as inputs for the MV1 algorithm. The spikes at $w_{IP3D} \simeq -20$ and -30 correspond to pathological cases where the IP3D weight could not be computed, due to the absence of good-quality tracks. The spike at $w_{SV1} \simeq -1$ corresponds to jets in which no secondary vertex could be reconstructed by the SV1 algorithm, and where discrete probabilities for a b and light-flavour jet not to have a vertex are assigned. The irregular behaviour in $w_{IP3D+JetFitter}$ arises because both the w_{IP3D} and the $w_{JetFitter}$ distribution (not shown) exhibit several spikes. Figure taken from Ref. [50].

The MV1 neural network is a perceptron with two hidden layers consisting of three and two nodes, respectively, and an output layer with a single node which holds the final discriminant variable. The implementation used is the MLP code from the TMVA package [52]. The training relies on a back-propagation algorithm and is based on two simulated samples of b-jets (signal hypothesis) and light-flavour jets (background hypothesis). Most of the jets are obtained from simulated $t\bar{t}$ events and their average transverse momentum is around 60 GeV. To provide jets with higher p_{T} for the training, simulated dijet events with jets in the 200 GeV $< p_{\mathrm{T}} <$ 500 GeV range are also included. As in the case of the JetFitter neural network, since the tagging performance depends strongly on the p_{T} and, to a lesser extent, on the η of the jet, biases may arise from the different kinematic spectra of the two training samples (of light-flavour and b-jets). To reduce this effect, weighted training events are used. Each jet is assigned to a category defined by a coarse two-dimensional grid in (p_{T},η) with four bins in η and ten bins in p_{T}. Jets in the same category are given the same weight, defined according to the overall fraction of all jets in this category, and the jet category is fed to the network as an additional input variable.

3.6 Missing transverse energy

Some SM particles such as neutrinos or some SUSY-predicted eigenstates such as the neutralino are not expected to leave a physics signature in any of ATLAS subsystems, because they interact very weakly with the matter of the detector. However, their presence in physics events can be deduced from the fact that the total momentum in the transverse plane has to be null due to the momentum conservation. This can be used to define the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ [53] in events where the transverse momentum balance does not compensate. The x and y components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ can be expressed as the sum of two terms, one from the calorimeter and the other from the muon system

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss,calo}} + E_{x(y)}^{\mathrm{miss},\mu} \tag{3.6}$$

The missing transverse energy and its azimuthal angle are then defined as the usual modulus and argument of a 2-dimensional vector

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \tag{3.7}$$

$$\varphi^{\mathrm{miss}} = \arctan\left(\frac{E_y^{\mathrm{miss}}}{E_x^{\mathrm{miss}}}\right) \tag{3.8}$$

Chapter 4

The importance of jet shapes for tagging purposes.

Recent measurements by the ATLAS Collaboration of b- and light jet-shapes in $t\bar{t}$ final states at the LHC, show clear evidence that b-jets are wider. There are also indications that b-jets in $t\bar{t}$ final states are narrower than those produced inclusively. This could be interpreted as due to the two different colour structures considered. Multivariate analysis techniques are applied on MC generated samples to illustrate possible applications of these observations in two study cases:

- measurement of $|V_{ts}|$ in $t\bar{t}$ events and
- separation of $H \to b\bar{b}$ in Higgs associated production with a Z from the continuum background i.e. $Z + b\bar{b}$.

The lessons to be learned is that jet shape measurements show more potential in the former case than in the latter. These results were presented at ICHEP 2014 in Valencia [54].

4.1 Introduction

The $t\bar{t}$ final states in the dileptonic mode at the LHC are a copious and clean source of b-jets, while the single lepton mode offers the possibility to study light jets through $W \to u\bar{d}, c\bar{s}$ decays. Integrated jet shapes, in a cone of radius $r \leq R = 0.4$ around the jet axis, defined as:

$$\Psi(r) = \frac{p_T(0,r)}{p_T(0,R)} \tag{4.1}$$

where $p_T(0,r)$ is the scalar sum of the transverse momenta within that cone, have been recently measured by the ATLAS Collaboration [55, 56]. These measurements show clear evidence that b-jets are wider than light jets, see Fig.4.1 (left). This can be understood as due to the b-quark being heavier than the light quarks and decaying sequentially via $b \to c \to s$.

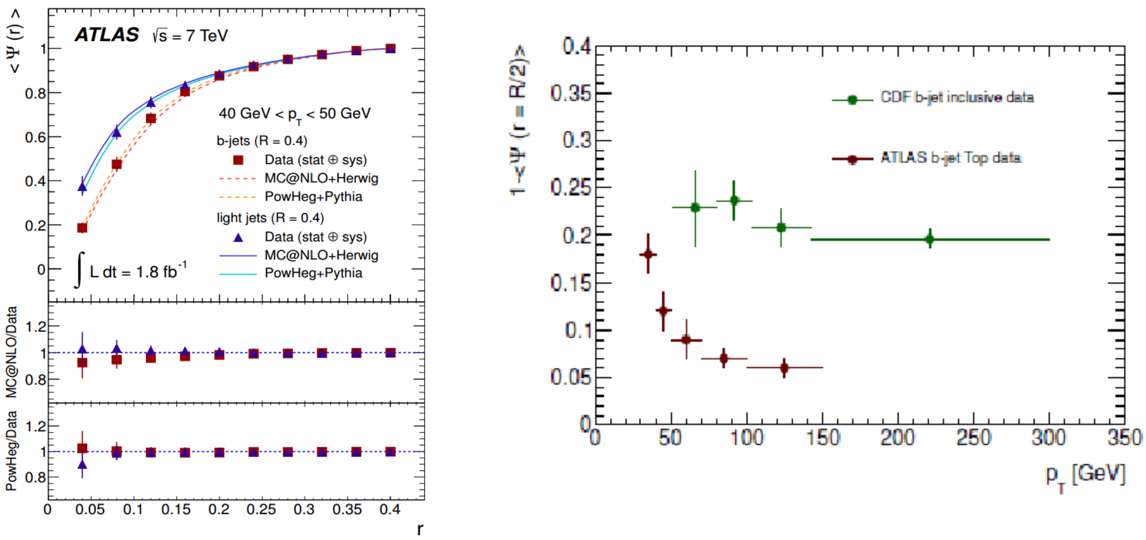


Figure 4.1: Average integrated b- and light-jet shapes measured by ATLAS (left) and energy flow outside half of the jet cone, ATLAS and CDF comparison (right) [55].

Furthermore, by comparing the energy flow outside half of the jet cone for these b-jets in ATLAS $t\bar{t}$ final states with those inclusively measured by CDF [57] one is tempted to conclude that the former are narrower than the latter, see Fig.4.1 (right). This could be due to the different colour flows in the two different final states investigated. It will be interesting to check these ideas within one single

LHC experiment, thus avoiding comparisons between LHC and Tevatron experiments which have different systematic uncertainties and make use of different jet algorithms and radii.
The authors in [58] have recently advocated the use of a new measure, pull, defined as

$$\vec{t} = \sum_{i \in Jet} \frac{p_T^i |r_i|}{p_T^{jet}} \vec{r_i} \tag{4.2}$$

with $\vec{r_i} = (\Delta y_i, \Delta\phi_i)$ being the distance in the (rapidity,azimuth) plane between a given cell in a jet and the jet axis. This variable can be used to discriminate between jets originating from a color singlet, like $W \to u\bar{d}$ or $H \to b\bar{b}$, and those coming from colour multiplets, see Fig. 4.2

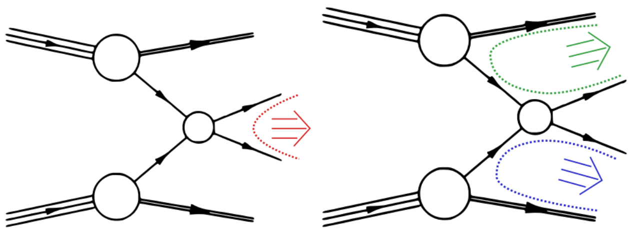


Figure 4.2: Color connections for a color singlet (left) and a color multiplet (right) [58].

The purpose of this analysis is to investigate possible applications of jet shape measurements in two study cases:

- the feasibility to measure $|V_{ts}|$ in $t\bar{t}$ final states when one $t \to Wb$ and the second one goes to $t \to Ws$
- the possibility to discriminate $pp \to ZH$, with the Z boson decaying to electron or muon pairs and the $H \to b\bar{b}$, from the main background $pp \to Z + b\bar{b}$

4.2 Multivariate-Analysis Classification

As explained above, to test the usefulness of jet shapes for tagging purposes, two study cases are analysed. Tagging is nothing but classification. In order to get the goals of these analyses, a classification multivariate-analysis (MVA) technique is needed. This section explains how this works and which method was used.
Classification multivariate-analysis (MVA) techniques are "learning" algorithms based on the use of a certain amount of variables in order to distinguish 'wanted' sort of events (signal) from 'unwanted' events (background). These type of "learning" algorithms have been used in the field of data mining, and are widely popular in high energy physics as well.
To perform the classification multivariate-analysis, the software package Tool kit for Multivariate Analysis (TMVA) [52] in the ROOT framework has been used. In particular one technique based on boosted decision trees has been tried as it will be explained below.

4.2.1 Boosted Decision Trees (BDT)

A decision tree is a binary tree structured classifier where repeated yes/no decisions are taken on one variable at a time until a stop criterion is fulfilled. The phase space is divided in regions classified as signal or background, depending on the majority of training events ending up in the final leaf node. The node splitting is performed on the variable which, at this node, gives the best separation between signal and background. This is an easily interpretable classifier, as shown in the Fig.4.3 below. It is similar to rectangular cuts, but instead of a hypercube signal-like, several signal-like or background-like hypercubes in phase space are considered. Nevertheless, a single decision tree is too sensitive to statistical fluctuations in the training sample.
The boosting of a decision tree extends the concept from one to many trees thus forming a forest. The trees are produced with the same training sample, and are combined in a single classifier as a weighted average of all the clasiffier trees. The weights are calculated in relation with the missclasified events of the training sample. This is performed with an Adaptative Boost. The boosting can be combined with bagging, i.e. using a fraction of the sample (randomly selected) in each tree. The boosting and bagging procedures lead to a greater stability with respect to statistical fluctuations in the training sample and enhance the classification performance of a single tree.

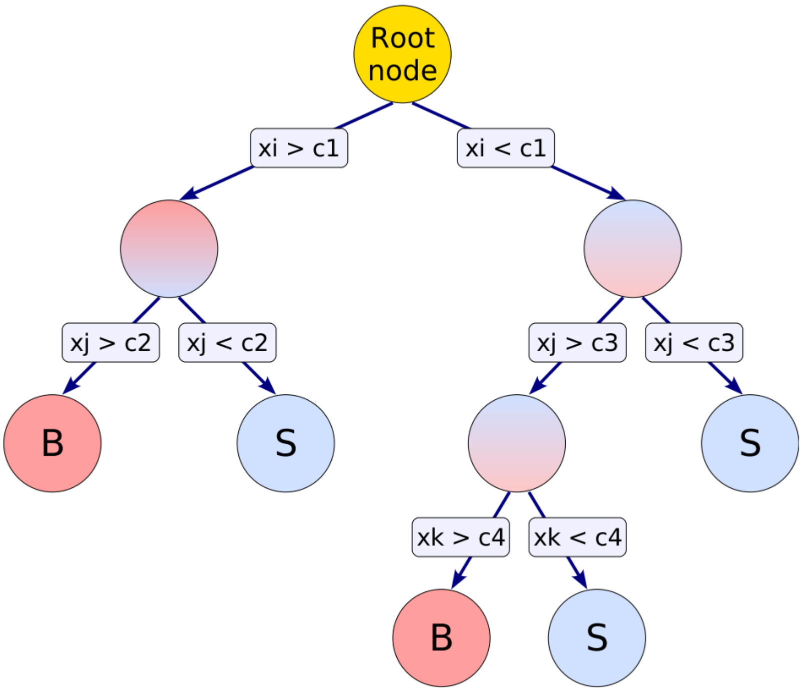


Figure 4.3: Scheme of how a decision tree works [52].

This method has been selected because it is transparent, easily understable, robust and it does not need many changes of the default parameters to reach high level performance in most problems.

4.3 Prospects of measuring the CKM matrix element $|V_{ts}|$

Transitions between quark generations through weak charged interactions are ruled by the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix V_{CKM} [12, 13]:

$$V_{CKM} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{4.3}$$

In the Wolfenstein parametrisation [14]

$$V_{CKM} \simeq \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda(1+iA^2\lambda^4\eta) & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2(1+i\lambda^2\eta) & 1 \end{pmatrix} \tag{4.4}$$

where A, λ, ρ and η are the Wolfenstein parameters.
As $A \sim 1$ and $\lambda = \sin\theta_C \approx 0.23$, from direct measurements and overall fits, it is easy to realize that the CKM matrix is nearly diagonal, and the transitions between generations are more suppressed, the further appart they are. Since the CKM matrix defines a change of basis, it has to be unitary, as long as the SM has only three quark generations. Therefore, to test the CKM matrix unitarity, means to test the SM.
Measurements of single top production cross sections at the Tevatron and the LHC have provided first direct measurement of the dominant CKM matrix element $|V_{tb}|$. A very recent determination by the CMS Collaboration reads $|V_{tb}| = 0.998 \pm 0.038(exp) \pm 0.015(syst)$ yielding $|V_{tb}| \geq 0.92$ at $95\% C.L.$ [59]. Similar limits can be obtained from b-tagged decays of top pair production. They can be compared with indirect determinations based on various loop-induced processes in which the top quark participates as a virtual state such as $B^0 - \bar{B}^0$ mixings, radiative decays $B \to X_s + \gamma$ and the CP violation parameter ϵ_K in the kaon sector. Overall fits of the CKM unitarity yield $|V_{tb}| = 0.999133(44)$ and $|V_{ts}| = 0.0407 \pm 0.001$ so that $|V_{ts}|^2/|V_{tb}|^2 = 1.7 \times 10^{-3}$ [6]. Nevertheless, the $|V_{ts}|$ matrix element has not been measured directly yet. To achieve this, top decaying events are needed, but the decay channel $t \to Ws$ is very suppressed. The LHC with its high energy and luminosity opens a window to precision studies of top decayss, so the direct measurement of the $|V_{ts}|$ matrix element can be explored.
In Ref.[60], the prospects of measuring the CKM matrix element $|V_{ts}|$ have been discussed. The idea is to study $t\bar{t}$ final states in the process $pp \to t\bar{t} + X$ with subsequent decays $t \to Wb, Ws$. The goal is to discern between $pp \to t\bar{t} \to WbWs$, which will be the signal process, from $pp \to t\bar{t} \to WbW\bar{b}$, which will be

the dominant background, see Fig. 4.4. Cross-sections for top pair production at the LHC have been calculated at NNLO accuracy. A typical estimate at $\sqrt{s} = 14\,TeV$ is $\sigma(pp \to t\bar{t}+X) = 874^{+14}_{-33}\,pb$ at $\sqrt{s} = 14\,TeV$ [61]. This translates into a yield of 10^7 top pair events for an integrated luminosity of 10 fb^{-1}. The PDF's used are those in Ref. [62].

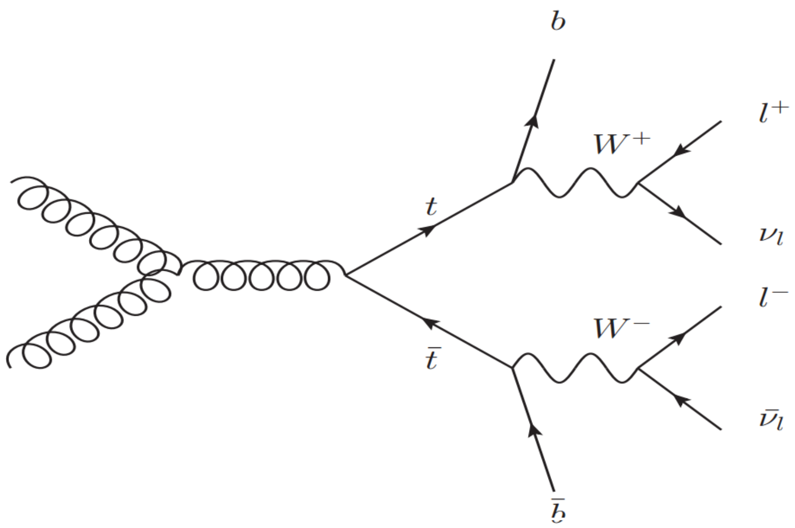


Figure 4.4: Feynman diagram of the process $pp \to t\bar{t} \to WbW\bar{b}$.

The jet profiles in the signal process $t \to Ws$ and the dominant background from $t \to Wb$ are analysed. To that end:

- the energy and transverse momentum distributions of neutral strange particles, K^0_S and Λ^0's, which are expected to be leading (non-leading) in s-(b-) initiated jets, see Fig. 4.5
- the spectrum of soft leptons, see Fig. 4.6,
- the presence (absence) of secondary vertices in b- (s-) initiated jets, see Fig 4.7

are studied.

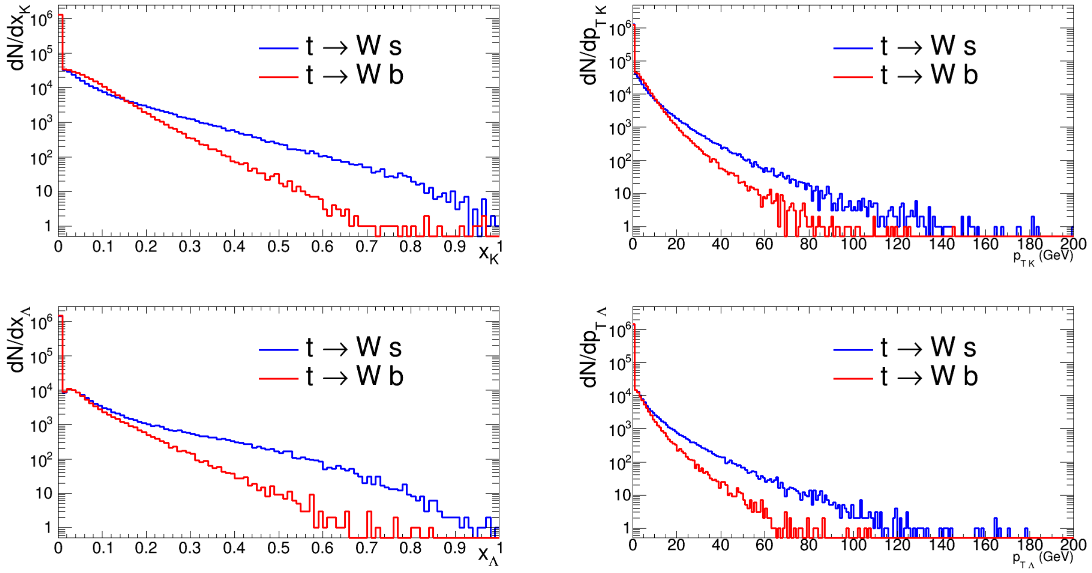


Figure 4.5: Energy fraction and p_T distributions for K^0 and Λ particles inside b- and s-jets.

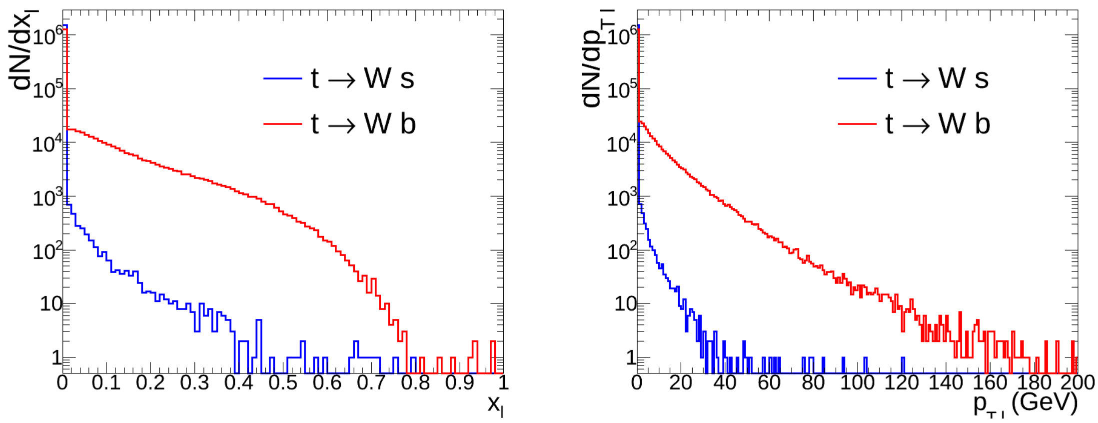


Figure 4.6: Energy fraction and p_T distributions for leptons inside b- and s-jets.

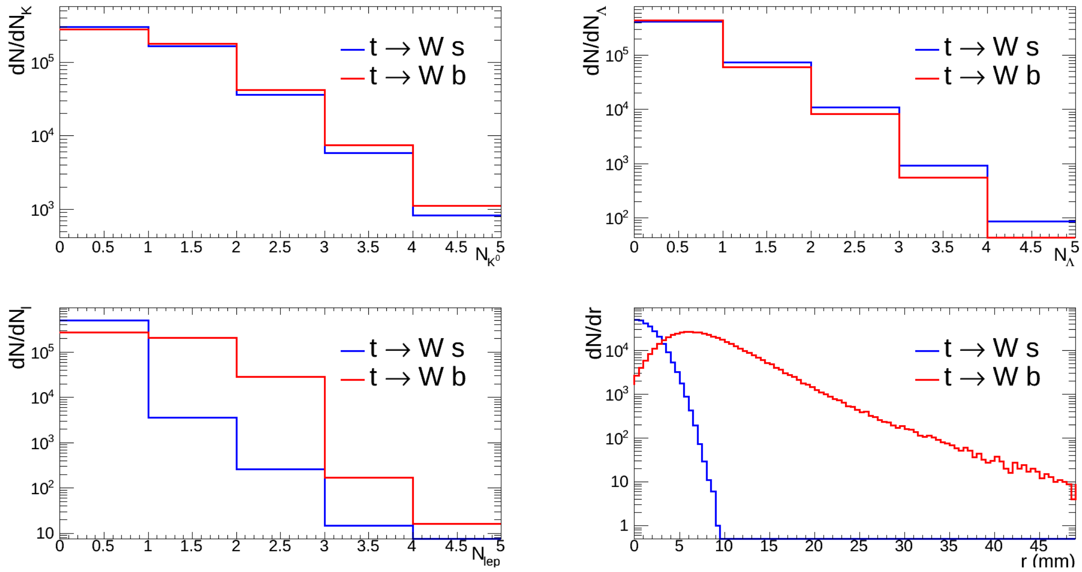


Figure 4.7: Multiplicity distributions for K_S^0 and Λ particles, and leptons inside b- and s-jets. In adition, the secondary vertex flight distance distributions for b- and s-jets.

The authors of Ref. [60] used the PYTHIA.6 Monte Carlo [23] to generate 10^6 events for the process $pp \to t\bar{t}X$, followed by the decay chains $t \to W^+b, W^+s$ and $\bar{t} \to W^-\bar{b}, W^-\bar{s}$. The $W^\pm$ are forced to decay leptonically to reduce the jet activity associated to hadronic W decays. This corresponds to an integrated luminosity of 10 fb^{-1}. The top branching ratios to Wb and Ws were artificially set equal to 50%. With this generated sample they obtained the distributions discussed above which were fed into a TMVA [52] classifier. A BDT algorithm was used for this purpose. For a rejection of the background channel $pp \to t\bar{t} \to WbW\bar{b}$ of 0.1%, the tagging efficiency of the signal $pp \to t\bar{t} \to WbWs$ was found to lie in the range $5-10\%$. This means that $0.05 \times 2 \times 1.7 \times 10^{-3} \times 10^6 = 170$ signal events are expected, with a background of $10^{-3} \times 10^6 = 10^3$ events, giving a significance of $170/\sqrt{1000}$ i.e. more than 5σ.

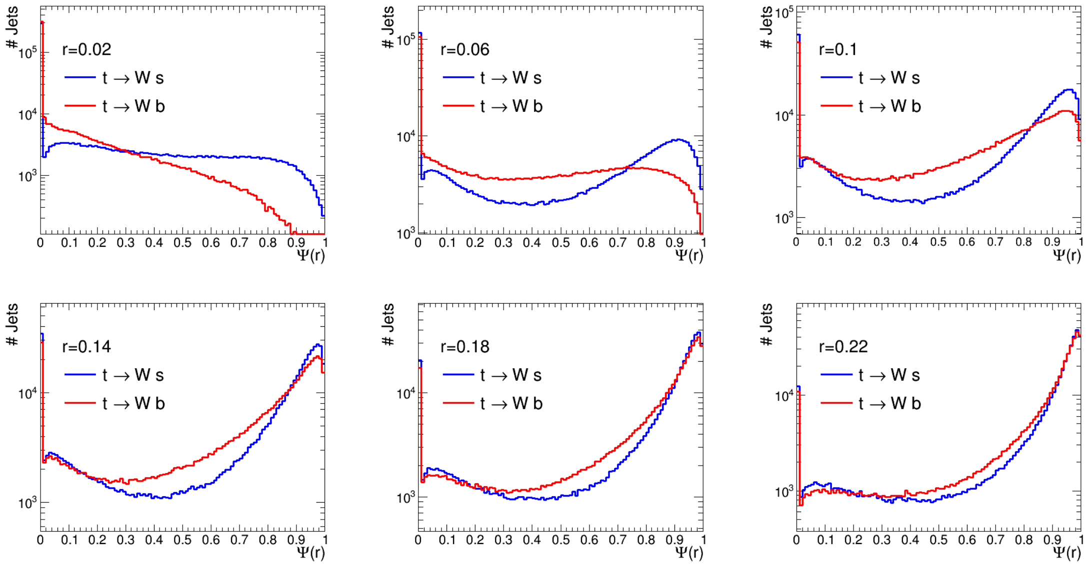


Figure 4.8: Integrated shapes for b- (red) and s-jets (blue) generated with Pythia6.

Next, the benefit of including b- and s-jet shapes in the MVA is investigated. To this end, the exercise in [60] is repeated but now including additional variables associated to the b- and s-jet shapes. Fig.4.8 shows the expected integrated jet shapes for b- and s-jets in the process $pp \to t\bar{t} \to WbWs$. There is a clear difference between these two shapes as discussed in the introduction. This is supported by recent measurements by the ATLAS Collaboration [55]. Notice that this difference is more marked for smaller values of r. The fact that the transverse spectra of both jet types are very similar lead us to expect that these differences will be also maintained when jets are binned in $p_{\rm T}$. This is illustrated in Fig.4.9 where the average integrated shapes are plotted in several $p_{\rm T}$ bins. The main features exhibited by the data can be summarized by stating that the differences between b- and s-jets are more significant at low values of r and diminish with increasing $p_{\rm T}$.

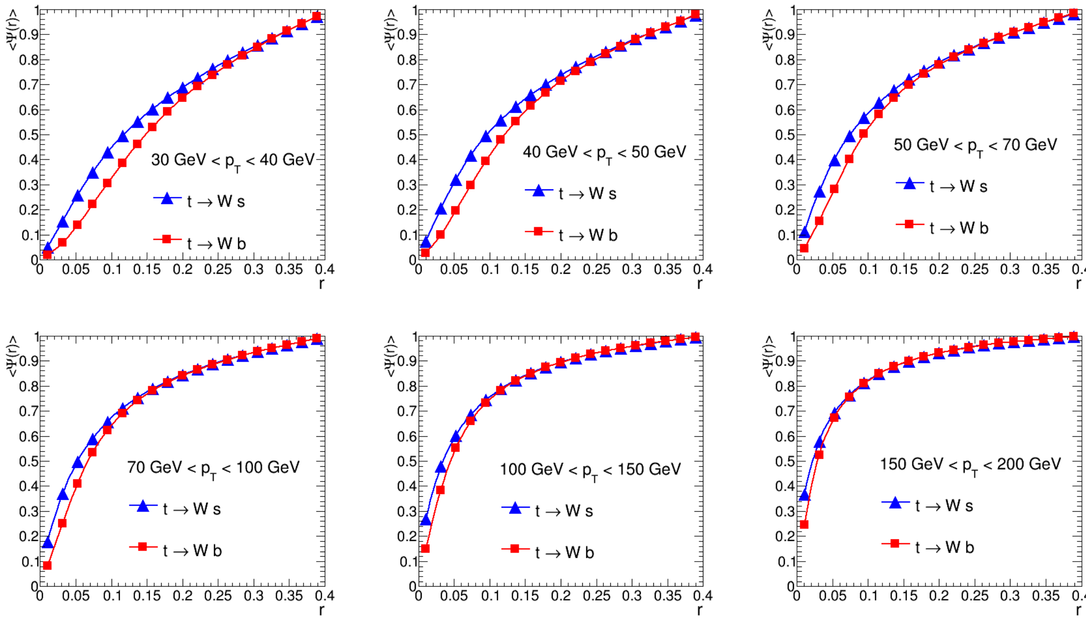


Figure 4.9: Average integrated b- (red squares) and s-jet (blue triangles) shapes with PYTHIA6.

In order to investigate the discrimination power associated to the jet shapes data, a two step multivariate analysis based on a BDT classifier is carried out. To begin with, the jet shapes data alone are fed in the BDT classifier. The corresponding ROC curves are shown in the left hand side of Fig. 4.10. Notice that one can achieve a non-negligible 90% rejection of the background process, i.e. $pp \to t\bar{t} \to WbW\bar{b}$, with an efficiency of the signal, $pp \to t\bar{t} \to WbWs$, of roughly 40%. In a second step, the ROC curves obtained using the variables in [60] with and without adding the jet shape information are compared. Clearly inclusion of the latter improves the discrimination power of the BDT algorithm, as illustrated in the right hand side of Fig.4.10. For high background rejection the improvement exceeds a factor of two. Fig. 4.11 shows the BDT distributions for signal and background, with the training and test samples of the MC simulation. This figure shows the case in which all variables have been used in the multivariate-analysis, including the jet shapes. 100000 signal (background) events are used for training the BDT, and another 100000 signal (background) events for testing the performance. For all the case, 1600 trees have been training.

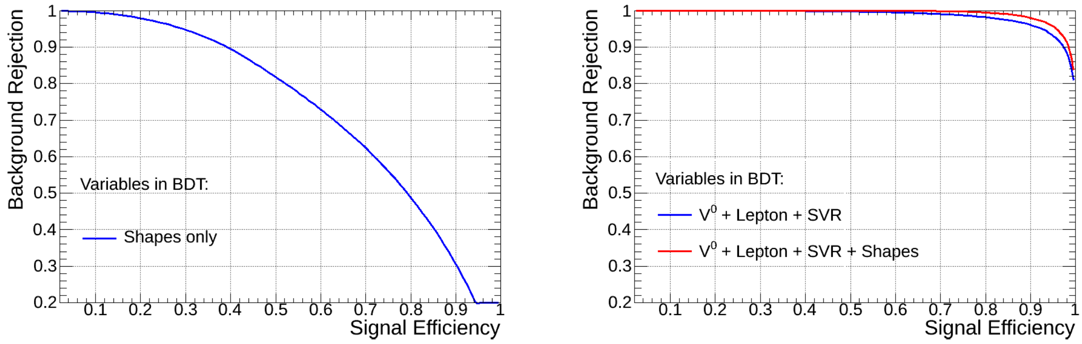


Figure 4.10: ROC curves obtained from shapes data only for $pp \to t\bar{t} \to WbWs$, signal and $pp \to t\bar{t} \to Wb W\bar{b}$, background, left, and in addition with all the distributions described in the text [60], right

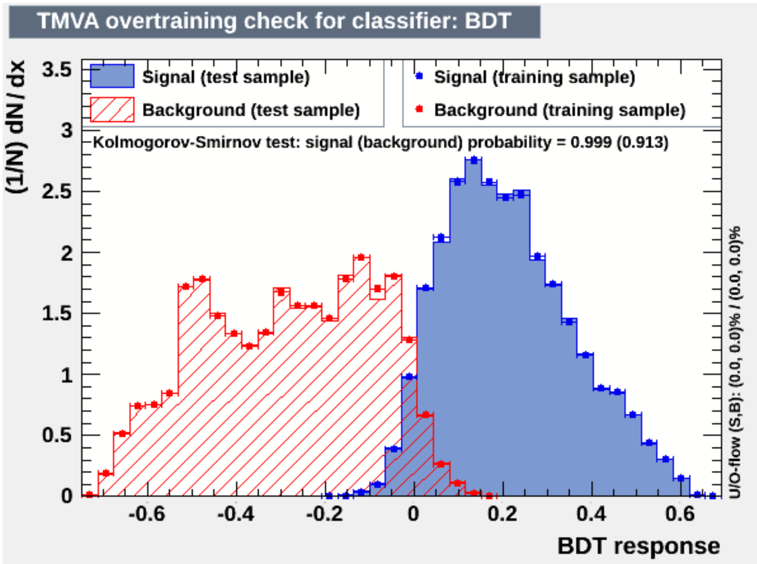


Figure 4.11: BDT distributions obtained from jet shapes in addition with all the distributions described in the text [60]. Training and test sample results are shown to check for overtraining efects.

4.4 Associated Higgs production with a Z boson

Although the Higgs decaying to gauge boson pairs, WW, ZZ, $\gamma\gamma$ has been observed by both the ATLAS and CMS collaborations [63, 64], direct evidence for the coupling to fermions pairs has only very recently been observed, the associated production $t\bar{t}H$ [65, 66] and the decay channels $H \to b\bar{b}$ [67, 68] and $H \to \tau\tau$ [69, 70]. The focus is on associated production with a gauge boson, Z or W, to reduce the QCD background of $b\bar{b}$ production. In the case of $pp \to ZH$ with the subsequent decay $H \to b\bar{b}$, from now on the signal process, the main background is due to $pp \to Z + b\bar{b}$.
Prompted by the observation that b-jets in the signal process are coming from a colour singlet while those in the background are coming from colour triplets or octets [58], see Fig. 4.2 in Section 4.1, one could also expect b-jet shapes in the former to be somewhat narrower than those in the latter. In the following the aim is to investigate whether b-jet shapes have the potential to improve on the separation between signal and background processes discussed above.
The first step it is to look at relevant *kinematic* quantities associated to the b-jets in these two processes, $pp \to ZH \to l^+l^-b\bar{b}$ and $pp \to Z + b\bar{b} \to l^+l^- + b\bar{b}$, like

- transverse momenta of the leading (subleading) b-jet i.e. p_{T1} (resp.p_{T2}) as well as their rapidities,
- their separation in rapidity, azimuth or (η, ϕ) plane, i.e. $\Delta\eta, \Delta\phi, \Delta R$, see Fig. 4.12
- similar variables now giving the separation between the Z boson and the $b\bar{b}$, see Fig. 4.13
- the polar angle of the leading b-jet in the rest frame of the $b\bar{b}$ pair (see Fig. 4.14), which is expected to be flat in the case of Higgs boson decays, see Fig. 4.12.

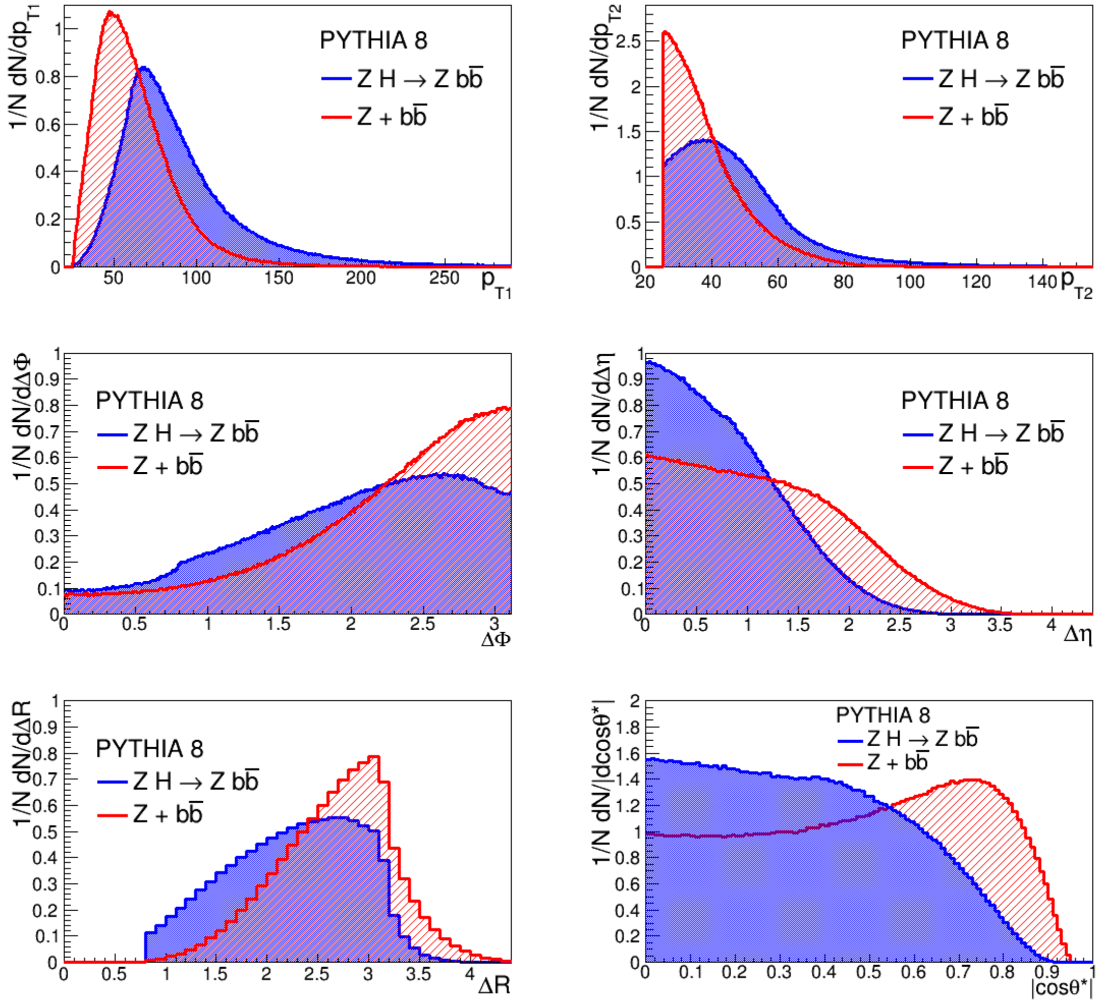


Figure 4.12: Relevant kinematic variables used to discriminate $pp \to ZH(b\bar{b}$ from $pp \to Z + b\bar{b}$.

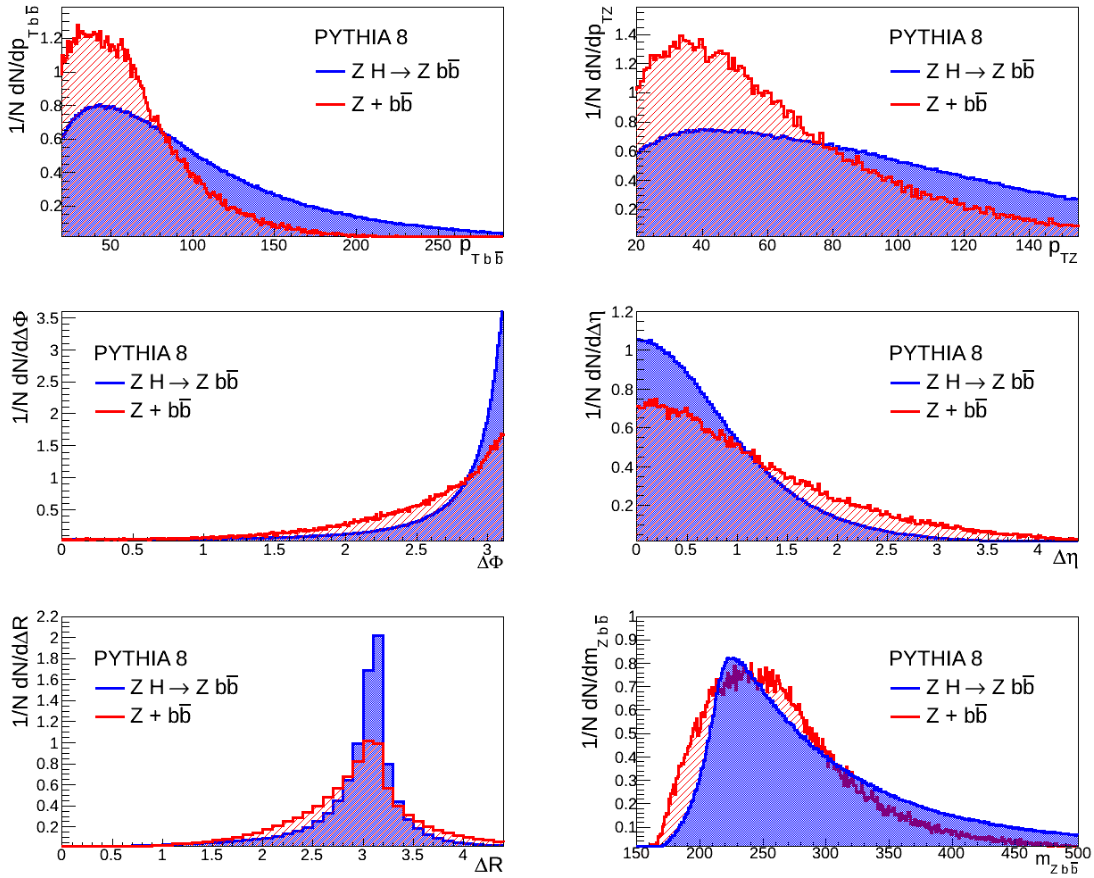


Figure 4.13: Relevant kinematic variables used to discriminate $pp \to ZH(b\bar{b}$ from $pp \to Z + b\bar{b}$.

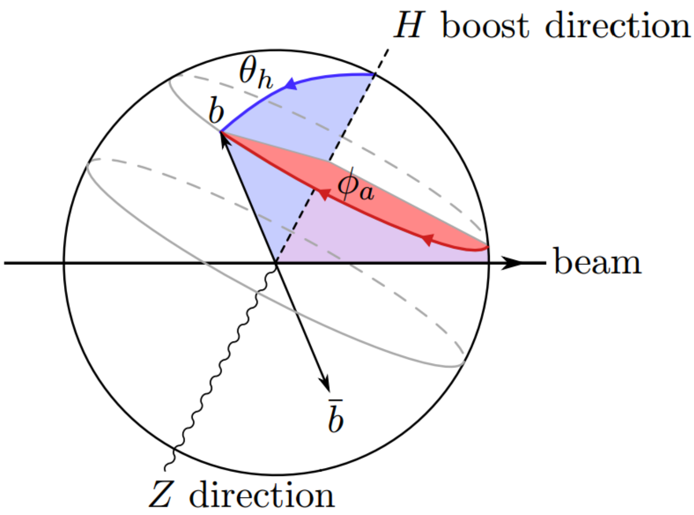


Figure 4.14: Helicity angle θ for $H \to b\bar{b}$.

Notice that the mass of the $b\bar{b}$ pair has not been considered. These distributions have been obtained with the ALPGEN Monte Carlo generator [29] coupled to PYTHIA8 [71], and using FASTJET [34]for jet recosntruction. The Z boson has been forced to decay to e^+e^- pairs and the b-jets, defined with the anti-k_T algorithm with radius R=0.4, are required to have $p_T \geq 25\ GeV$ in order to be in the regime where current b-tagging algorithms are operational. Clearly, the main observation is that b-jets coming from the signal process are harder, in transverse spectra, and better separated in (η, ϕ), than those coming from the background process. These differences were exploited in [72] who performed a multivariate analysis base on a BDT classifier to investigate the feasibility of discriminating $pp \to ZH \to l^+l^-b\bar{b}$ from $pp \to Z + b\bar{b} \to l^+l^- + b\bar{b}$. In fact the authors of Ref.[72] went a step further and also studied the relevance of two new variables which they called pull angles for the leading and subleading b-jets. The variable $\vec{t}$ in Eq. 2 can also be written as

$$\vec{t} = |\vec{t}|(cos\theta_t, sin\theta_t) \tag{4.5}$$

in terms of the pull angle. If the b-jets are coming from a single colour object, the pull angles are expected to peak around $\theta_t = 0$ which is defined on an event-by-event basis as the direction pointing from one b- to another. The discrimination power added by these two variables, see Fig. 4.15 is not that strong as already noted in [72].

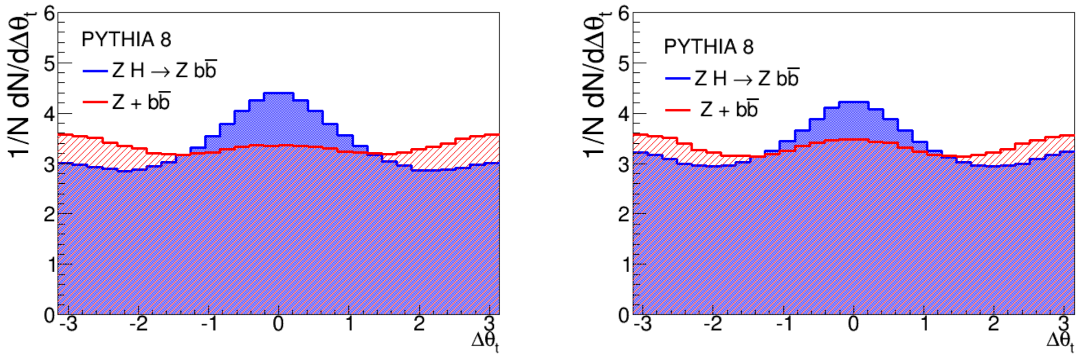


Figure 4.15: Pull variables also used to discriminate $pp \to ZH(b\bar{b})$ from $pp \to Z + b\bar{b}$.

The event shapes variables such as sphericity and aplanarity, have also been investigated. These variables are constructed from eigenvalues of a tensor composed of 3-momenta, with the sum over the four primary objects (the two b-jets, and the two leptons from the Z boson). This tensor is defined as follows:

$$\text{SphericityTensor} \equiv \frac{1}{\sum_i |\vec{p_i}|^2} \sum_i \begin{pmatrix} p_x p_x & p_x p_y & p_x p_z \\ p_y p_x & p_y p_y & p_y p_z \\ p_z p_x & p_z p_y & p_z p_z \end{pmatrix} \tag{4.6}$$

The eigenvalues are computed, ordered, and normalized $\lambda_1 \geq \lambda_2 \geq \lambda_3$ with $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The event shapes are defined as [73]:

- Sphericity: $S = \frac{3}{2}(\lambda_2 + \lambda_3)$, with $0 \leq S \leq 1$, and $S \approx 0$ for a 2-jet event, and $S \approx 1$ for an isotropic one.
- Aplanarity: $A = \frac{3}{2}\lambda_3$, with $0 \leq A \leq 1/2$, and $A \approx 0$ for a planar event, and $A \approx 1/2$ for an isotropic one.

The signal is expected to be less spherical (more planar), than the background, due to the back-to-back configuration of the ZH event, and the b-jets from the Higgs boson are expected to be closer (more boosted). These effects can be observed in Fig. 4.16

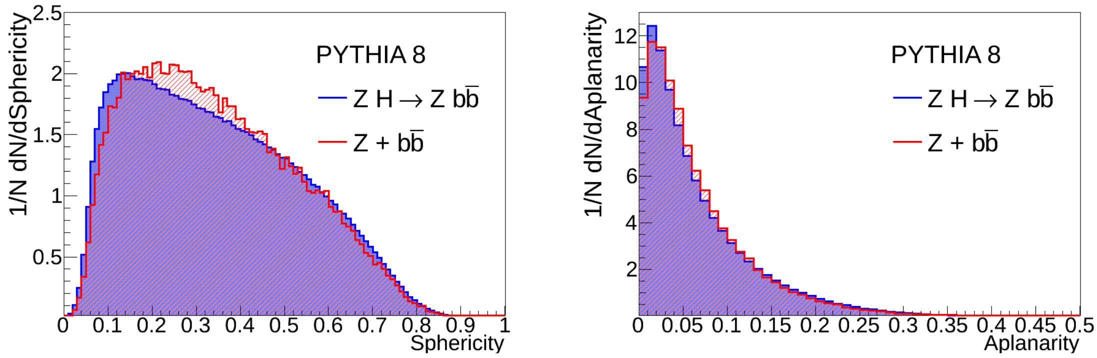


Figure 4.16: Sphericity and Aplanarity distributions for $pp \to ZH(b\bar{b})$ (signal)and $pp \to Z + b\bar{b}$ (background).

Triggered by these considerations, the potential of b-jet shapes in discriminating these two processes is studied. The differential jet shape $\rho(r)$ in an annulus of inner radius $r - \Delta r$ and outer radius r from the axis of a given jet is defined as

$$\rho(r) = \frac{1}{\Delta r} \frac{p_{\mathrm{T}}(r - \Delta r, r)}{p_{\mathrm{T}}(0, R)} \tag{4.7}$$

where $r \leq R$ and $p_{\mathrm{T}}(r_1, r_2)$ is the scalar sum of the p_{T} of the clusters within radii r_1 and r_2.
Some distributions for $\rho(r)$ are shown in Fig. 4.17 for b-jets in the signal and background process.

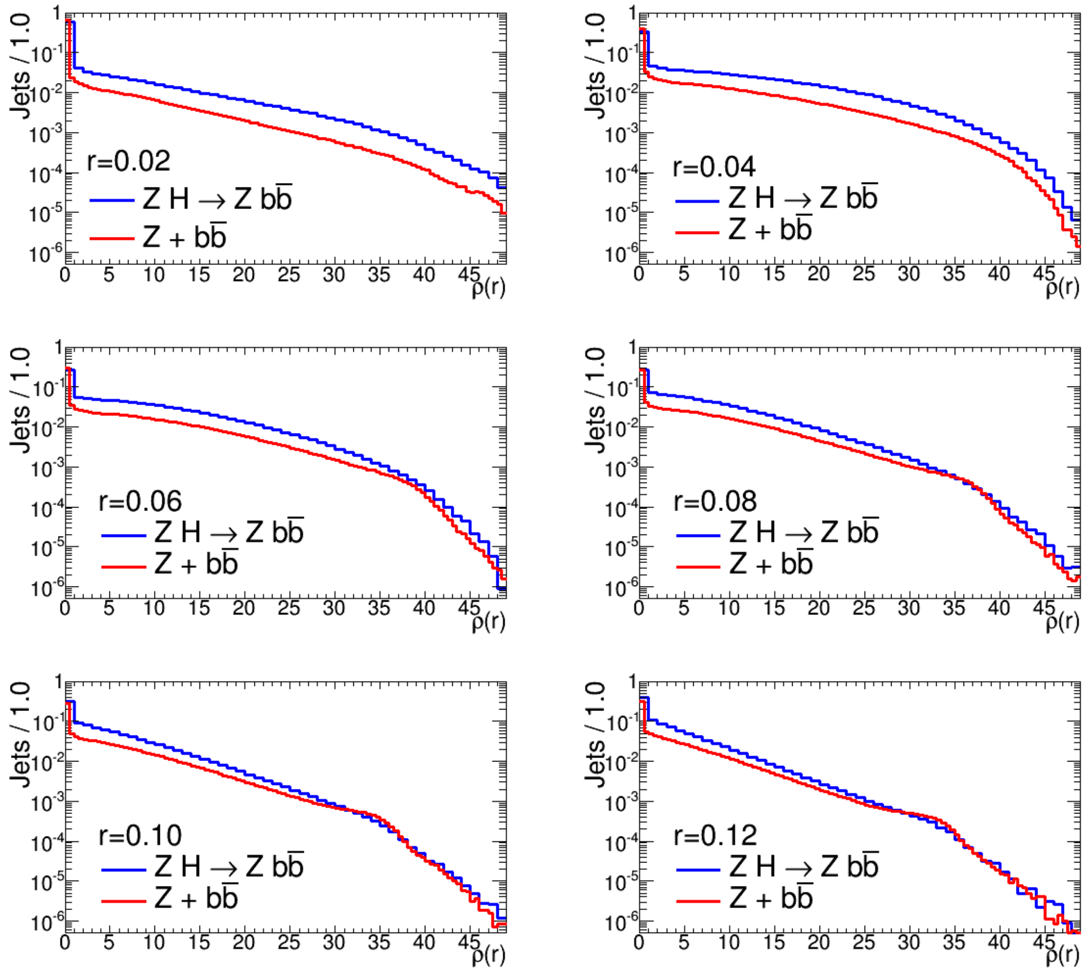


Figure 4.17: Differential b-jet shapes in the signal, $pp \to ZH(b\bar{b})$, and background, $pp \to Z + b\bar{b}$, processes.

The value $\Delta r = 0.02$ has been used in Fig.4.17 and only differential shapes up to $r = 0.12$ have been fed into a TMVA classifier, the reason being that for larger values of r the differences are washed away. The MVA proceeds in three steps. Firstly, only the kinematic variables discussed above are considered, in a second step the pull and event shapes information is added, and finally in a third step the b-jet shapes are included. The resulting ROC curves for signal and background are displayed in Fig. 4.18. All three ROC curves are very similar with marginal improvement after inclusion of pull and shapes information.

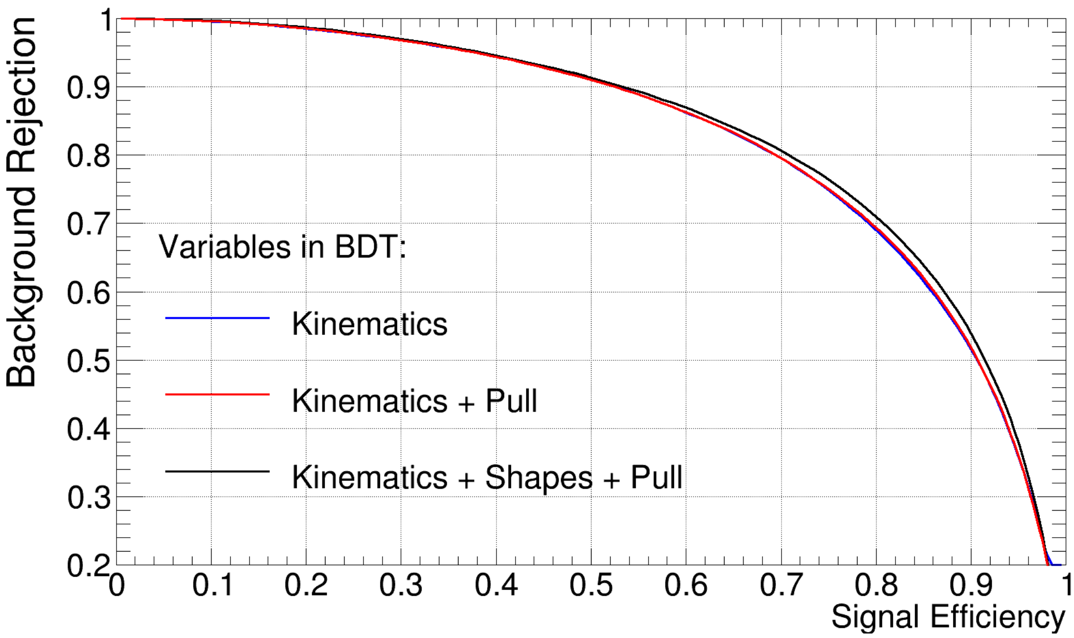


Figure 4.18: ROC curves for the separation between $pp \rightarrow ZH(b\bar{b})$ and $pp \rightarrow Z + b\bar{b}$.

This result was somehow expected from the shapes distributions, where the differences between signal and background processes were not so significant. This is even more understandable, when looking at the jet shapes averaged over p_{T} bins. The higher the p_{T}, the less different the jet shapes are. In this case the p_{T} distributions of the b-jets are the most powerful discriminant, so the first cuts in the trees will be on these variables, thus reducing the potential of jet shapes to discriminate.

4.5 Summary, conclusions and outlook

It has been shown that jet shapes data have a potential to discriminate b- from light-jets. This can be of help for improving b-tagging algorithms which are mainly based on secondary vertices and charged particle impact parameter measurements. The separation power of the jet shapes alone are almost competitive with standard b-taggers, so the combination with other b-tagging common variables should produce a fairly strong discriminator.
In the case of the prospects for measuring $|V_{ts}|$, the inclusion of jet shape variables in a MVA leads to significant improvements of the signal efficiency for high background rejection figures. However, this analysis has some weaknesses, such as not to take into account detector effects, or not to use a jet algorithm in the simulations. Therefore,it is foreseen to repeat the analysis with a simulated sample including full dectector reconstruction, to check the real potential of this discrimation method.
For the separation of $pp \to Z(l^+l^-)H(b\bar{b})$ from $pp \to Z(l^+l^-) + b\bar{b}$, jet shapes are not so powerful as the previous case. They barely increase the discrimation power when is adding to the classifier. This is due to the fact that the kinematical variables are very strong discriminants, and the jet shapes are only an small correction, as the event shape (sphericity or aplanarity) and color connection (pull) variables.

Chapter 5

Measurement of neutral strange particle production in $t\bar{t}$ dileptonic events.

In this chapter, the first measurements of K^0_S and Λ^0 production in $t\bar{t}$ final states will be discussed. They are based on a data sample with integrated luminosity of 4.6 fb^{-1} from proton-proton collisions at a center of mass energy of 7 TeV, collected in the year 2011 with the ATLAS detector at the Large Hadron Collider. Neutral strange particles are separated in three classes, depending on whether they are associated with a jet, b-tagged or not, or not associated with a selected jet. The aim is to look for differences in their main kinematic distributions. A comparison with several Monte Carlo simulations using different hadronisation and fragmentation schemes, colour reconnection proposals and different tunes for the underlying event has been made. The conclusion of this study is that neutral strange particle production from $t\bar{t}$ dileptonic events is well described by current Monte Carlo models for K^0_S and Λ^0 production within jets, but not for those produced outside jets. This study has been recently approved for publication by the ATLAS collaboration and published by EPJC [74].

5.1 Introduction and motivation

Neutral strange particle production has been studied in collider experiments using e^+e^- [75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85], pp [86, 87, 88, 89, 90], $p\bar{p}$ [91, 92, 93], ep [94, 95] and heavy-ion collisions [?] as well as in fixed target experiments [96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109]. These measurements provide interesting tests of theoretical jet fragmentation functions [110] and can be used to validate and tune empirical parameters used in the Monte Carlo (MC)

models. Since the mass of the strange quark is comparable to the QCD scale parameter Λ_{QCD}, perturbative calculations cannot be performed.
The accuracy of these models is very important for constraining the underlying event effects in the high-p_{T} collisions investigated at the LHC. In particular the ratio s/u, giving the supression of strange to non-strange mesons in the hadronic final states, turns out to be larger in pp collisions than in e^+e^- annihilation.
It was suggested [111] that this suppression factor would be significantly larger, or even tend to 1, in nucleus collisions due to the many strings produced within a limited phase space which may interact giving rise to the formation of 'colour ropes'. Recent data from RHIC [**?**] tend to support these ideas and do show that neutral strange meson production is enhanced as is the Λ/K^0_{S} ratio.
In pp collisions at LHC energies, many overlapping strings due to multiparton interactions are also expected to come into play, so that also here larger rates for strange meson and baryon production are expected [112]. This effect may be center of mass (c.m.) energy dependent.
Furthermore, in Ref.[60, 54], the prospects of measuring the CKM matrix element $|V_{ts}|$ have been discussed.This matrix element is determined from overall fits, imposing unitarity of the CKM matrix, to be $|V_{ts}|^2/|V_{tb}|^2 = 1.6 \times 10^{-3}$ [6]. The idea is to study $t\bar{t}$ dileptonic final states in the process $pp \to t\bar{t} + X$ with subsequent decays $t \to Wb, Ws$. The goal is to discern between $pp \to t\bar{t} \to WbWs$, which will be our signal process,from $pp \to t\bar{t} \to WbW\bar{b}$, which will be the dominant background. In order to perform this separation a multi-variate analysis classifier is trained with MC simulated events with the PYTHIA6 generator [23]. Some powerful discriminants for this are the kinematic distributions, such as the energy, p_{T}, energy fraction or transverse flight distance, as well as the multiplicity of K^0_{S} and Λ particles inside b- and s- jets. As shown in Figure 4.5, these distributions are different for b- and s- jets (jet definition can be found in the reference paper), due to the different processes which lead to the neutral strange particle production in the two environments. Therefore, as this method to measure the CKM matrix element $|V_{ts}|$ depends on the accuracy of the model description of these processes, a comparison of the neutral strange particle spectra inside jets, b-tagged or not, with current MC expectations is needed, and this is the goal of this analysis. This may give us confidence in using MVA techniques in a future determination of this matrix element.
Studies of neutral strange particle production at the LHC have been carried out using minimum bias events at low luminosities. The aim of this note is to extend them to $t\bar{t}$ production which is known to be a copious source of high transverse momentum jets, in particular b-jets. In doing so two cases are considered depending on whether the neutral strange particles are:

- embedded in jets, b-tagged or not

- non associated with jets

For those in the first category the focus is to see whether current Monte Carlo fragmentation models describe some of the kinematical features as discussed in Figure 1. For those in the second category, one of the main points of interest will be to see how well is the suppression factor γ_s i.e. the s/u ratio tuned to describe the data. Figure 5.1 shows the dependence of the K_S^0 multiplicity inside b-jets and outside any jet with γ_s for $\gamma_s = 0.3, 0.4, 0.5$.

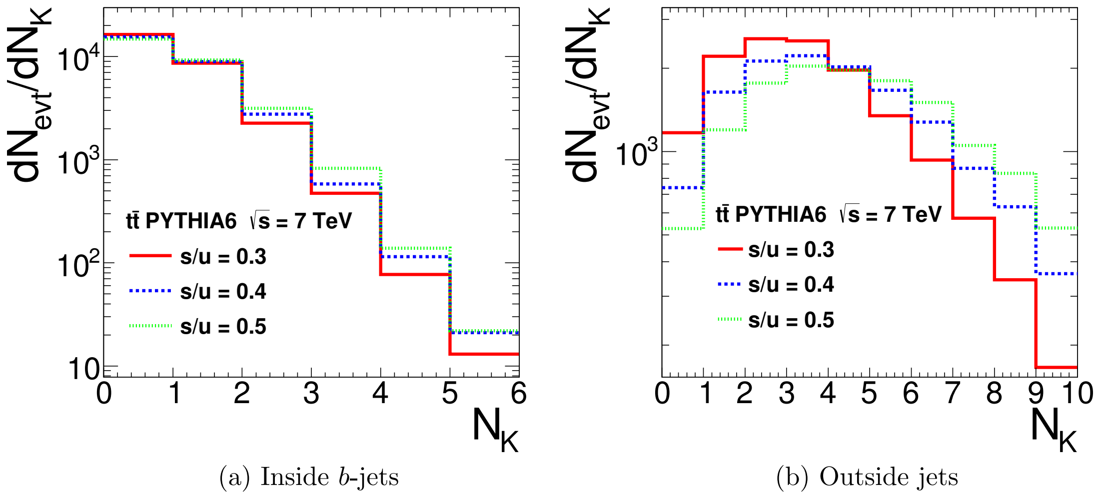
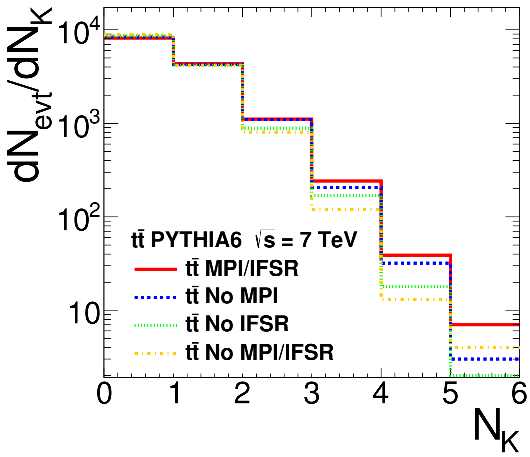
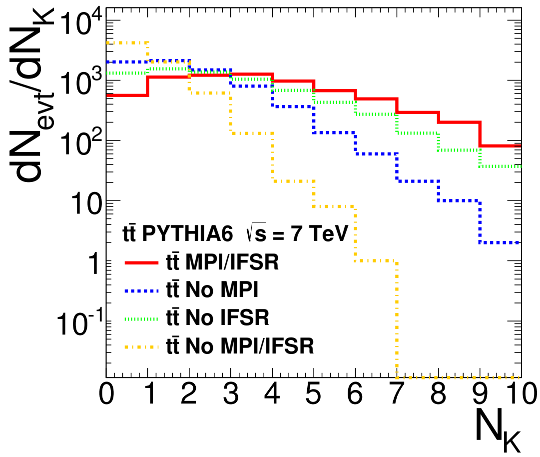


Figure 5.1: K_S^0 multiplicity distributions inside b-jets (left) and outside any jet (right) for three different values of $\gamma_s = 0.3, 0.4, 0.5$, as in [60].

Furthermore, neutral strange particle production within jets in top-quark decays exhibit little sensitivity to initial-state radiation effects, different PDF choices or underlying-event effects. In contrast, neutral strange particle production outside jets is more sensitive to details of the parton shower i.e. initial and final state radiation, fragmentation scheme and multiparton interactions, as illustrated in Figure 5.2. As a trivial remark, when MPI and I/FSR are turned off, neutral strange particles are still produced as a result of colour string hadronisation.



(a) Inside b-jets



(b) Outside jets

Figure 5.2: K^0_S multiplicity distributions inside b-jets (left) and outside any jet (right), simulated with the LO-PYTHIA6 generator with and without MPI and IFSR, as in [60].

This analysis has been carried out on a $t\bar{t}$ events sample with 4.6 fb^{-1} integrated luminosity collected with the ATLAS detector at $\sqrt{s}$ =7 TeV, i.e. the whole 2011 run, which is less affected by pile-up than the 2012 sample.
This note is organized as follows. Section 2 gives a brief description of the ATLAS detector. Section 5.2 is devoted to the Monte Carlo samples used. Section 5.3 explains the data sample and the event selection criteria. Section 5.4 is dedicated to the neutral strange particles reconstruction and selection, as well as the background subtraction procedure. Section 5.5 shows the results at the detector level compared to the POWHEG+PYTHIA6 and the MC@NLO+HERWIG simulations. Section 5.6 dwells on the efficiency correction calculations and the statistical error propagation. Section 5.8 shows the results corrected at the particle level compared to POWHEG+PYTHIA6. In section 5.9 a comparison between the corrected data and different MC models at the particle level is presented, thus checking the model-dependence of the neutral strange particle production in these events. Finally Section 5.10 is left for a summary, conclusions and a discussion of our plans for the future.

5.2 Monte Carlo samples

Monte Carlo generators are used in which $t\bar{t}$ production is implemented with matrix elements calculated up to NLO accuracy. The generated events are then passed through a detailed GEANT4 simulation [113] of the ATLAS detector.
The baseline $t\bar{t}$ MC sample was produced with the next-to-leading-order (NLO) generator POWHEGBOX (referred to hereafter as POWHEG) [28, 114, 115] for the matrix element calculation with CTEQ66 NLO PDF. The parton shower and hadronisation processes are implemented with PYTHIA6 [23] with CTEQ6L [116] parton distribution functions(PDFs). PYTHIA6 orders the parton shower by p_{T} and uses the Lund string fragmentation scheme [18]. The underlying event effects are considered with the PERUGIA2011C tune [24], while pile-up contributions are accounted for with the PYTHIA6 AMBT2B minimum bias (MB) tune. The strangeness suppression factor $\gamma_{\mathrm{s}} = s/u$ is taken at its default value i.e. $\gamma_{\mathrm{s}} = 0.3$ in AMBT2B tune, while $\gamma_{\mathrm{s}} = 0.2$ in the PERUGIA2011C tune.
Additional MC samples are used to check the hadronisation model dependence of K^0_{S} and Λ production. They are based on MC@NLO+HERWIG [27, 25] which orders the parton showers by angular separation and uses the cluster hadronisation model [19] and CT10 NLO [117] PDFs. Multi-parton interactions are simulated using JIMMY [?] with the AUET2 tune, while pile-up contributions are accounted for with the PYTHIA6 AMBT2B minimum bias (MB) tune.
The data corrected for detector effects, are also compared with events from other MC generators at partcle level, without detector simulation:

- the SHERPA 2.1.1 generator [118]. SHERPA, which uses a different approach than previous generators for matrix element calculation up to NLO accuracy with the CT10 PDFs, as well as for parton shower implementation. SHERPA uses $\gamma_{\mathrm{s}} = 0.4$.

- POWHEG with the NNPDF3.0 NLO PDF set [119], interfaced to Pythia8 [71] with the NNPDF2.3 LO PDF set and the A14 tune [120] for the parton shower, hadronisation and UE modelling.

- POWHEG interfaced to HERWIG7 (v7.1)[121] with the NNPDF3.0 NLO PDF set and H7UE tune, as default, for the parton shower, hadronisation and UE modelling.

- MADGRAPH5 aMC@NLO generator (referred to hereafter as aMC@NLO) [122] interfaced to HERWIG7 as before.

- the leading-order (LO) ACERMC generator [123] interfaced to PYTHIA6, with different tunes such as the PERUGIA2011C or the TUNEAPRO for parton showering and hadronisation, as well as different colour recnonnection (CR) schemmes.

Background samples were generated for the production of W and Z bosons in association with jets, including heavy flavours, using the ALPGEN [29] generator with the CTEQ6L PDFs [116], and interfaced with HERWIG and JIMMY. The same generator is used for the diboson backgrounds, WW, WZ and ZZ, while MC@NLO has been used for the simulation of the single-top background in the Wt-channel.
The MC simulated samples are normalised to the corresponding cross-sections. The $t\bar{t}$ signal is normalised to the cross-section calculated at approximate next-to-next-to-leading order (NNLO) using the HATHOR package [61], while for the single top production cross-section, the calculations in Ref. [124] were used. The W and Z plus jets cross-sections are taken from ALPGEN [29] with additional NNLO K-factors as given in Ref. [125].
The simulated events are weighted such that the distribution of the number of interactions per bunch crossing in the simulated samples matches that of the data. This is performed with a toolkit package within the ROOT framework, namely PILEPREWEIGHTING-00-02-03. The statistics in the MC samples considered in this analysis exceeds by more than an order of magnitude that of the data, see Table 5.1 below.

Table 5.1: Summary table with MC samples information.

MC	Level	PDF	UE tune	N_{events}
POWHEG+PYTHIA6	reco	CTEQ66 NLO	PERUGIA2011C	145644
MC@NLO+HERWIG	reco	CT10 NLO	JIMMY-AUET2	206266
SHERPA 2.1.1	particle	CT10 NLO		9391929
POWHEG+PYTHIA8	particle	NNPDF3.0 NLO	A14	∼7000
POWHEG+HERWIG7	particle	NNPDF3.0 NLO	H7UE	∼50000
aMC@NLO+HERWIG7	particle	NNPDF3.0 NLO	H7UE	∼50000
ACERMC+PYTHIA6	particle	CTEQ6L	PERUGIA/	∼1800000
			TUNEAPRO	∼1700000

5.3 Data sample, trigger and event selection

The data sample used in this analysis corresponds to an integrated luminosity of 4.6 fb^{-1}, collected in 2011. The uncertainty of the integrated luminosity is 1.8% [126].
In order to reduce the jet activity from hadronic $W^{\pm}$ decay channels, the $t\bar{t}$ dileptonically decaying sample is used in this analysis. These events were selected as in Ref. [127, 55]. The events were triggered by an inclusive high-p_{T} electron or muon EF trigger with $p_{\mathrm{T}}(e) > 20$ GeV and $p_{\mathrm{T}}(\mu) > 18$ GeV. Events are required to have at least one primary vertex, with five or more tracks with $p_{\mathrm{T}}^{\mathrm{track}} \geq$ 150 MeV. If there are more than one primary vetex, the one maximizing $\sum p_{\mathrm{T}}^2$ is chosen.
Tracks used in the primary vertex reconstructions are required to fulfill the following criteria:

- Each track must have at least 4 (6) hits in the SCT (pixel+SCT) detectors
- $p_{\mathrm{T}} > 400$ MeV
- $|d_0| < 0.4$ cm
- $\sigma(d_0) < 0.5$ cm
- $\sigma(z_0) < 1$ cm

Then two isolated charged leptons (e or μ) are required with $E_{\mathrm{T}}(e) > 25$ GeV and $p_{\mathrm{T}}(\mu) > 20$ GeV, where $E_{\mathrm{T}} = E_{\mathrm{cluster}} \sin(\theta_{\mathrm{track}})$. At least one of them must match with the corresponding trigger object. The isolation is required by summing the transverse energy deposited around the electron in the calorimeter, or the transverse momentum around the muon in the tracking detector. Cosmic muons are rejected by a veto on back-to-back pairs with transverse impact parameter relative to the beam axis $|d_0| > 0.5$ mm. The leptons are required to have opposite charges. The invariant mass of two electrons, or two muons, is required to be greater than 15 GeV, to reject Drell-Yan lepton pairs background, and to be 10 GeV away from the Z boson mass.
Jets are reconstructed with the anti-k_t algorithm [35] with radius parameter $R = 0.4$. The input objects to the jet algorithm, both for data and detector level simulation, are topological clusters in the calorimeter. These clusters are seeded by calorimeter cells with $|E_{\mathrm{cell}}| > 4\sigma$, with σ the RMS of the noise. Neighbouring cells are added and clusters are formed following an iterative procedure. The baseline calibration for these clusters corrects their energy to the electromagnetic energy scale, which is established using test beam measurements for electrons, pions and muons in the electromagnetic and hadronic calorimeters

[40, 128, 129]. Effects due to non-compensation, energy losses in the dead material, shower leakage, as well as inefficiencies in energy clustering and jet reconstruction have to be taken into account. This is done by matching calorimeter jets with MC particle jets in bins of $|\eta|$ and p_{T}. This is called the jet energy scale (JES), thoroughly discussed in [130]. It is different for b- and light-jets since they have different particles composition. See [131] for more details and for a discussion of JES uncertainties.

The selected events should have at least two jets which are required to satisfy:

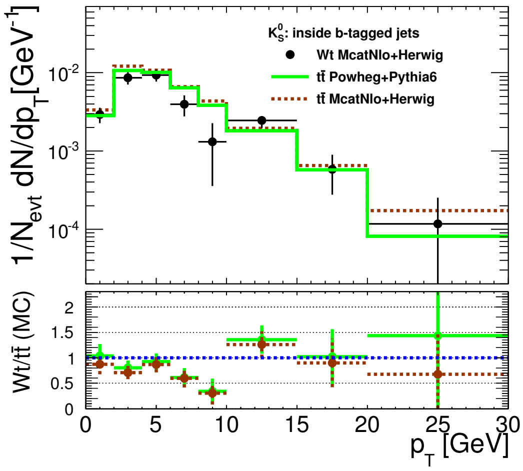
- $|\eta| < 2.5$
- $p_{\mathrm{T}} > 25$ GeV

Both "bad" and "ugly" jets are removed using cleaning cuts defined in [132]. In addition jets are required to satisfy $|JVF| > 0.75$ in order to minimize pile-up effects. At least one of the jets must be tagged as a b-tagged jet, using a multivariate-analysis (neural-network) b-tagging algorithm based on reconstruction of secondary vertices and three dimensional impact parameter information. For this purpose, the MV1 algorithm, has been used at a working point with b-tagging efficiency of 70%. Jets with an overlap with an accepted electron are removed if their separation is $\Delta R < 0.2$. The electron is removed if $0.2 < \Delta R < 0.4$. The muon is removed if the separation is $\Delta R < 0.4$ to a jet.
Finally, the missing transverse energy is required to be $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV for the ee and $\mu\mu$ channels, and for the $e\mu$ channel the requirement is $H_{\mathrm{T}} > 130$ GeV, where H_{T} is the scalar sum of the transverse momentum of the muon plus the transverse energy of the electron and the transverse momentum of the jets. A sample of 6926 $t\bar{t}$ events fulfils all these selection criteria. MC studies [55] indicate that the background contamination in the sample after event selection is $\sim 6\%$ and is dominated by single top events. Background contamination from processes as $W+$ jets or $Z+$ jets, with the electroweak gauge bosons decaying leptonically, is almost negligible. The expected composition from these MC studies can be seen in detail in the Table 5.2 below, where 'Other EW' means $W+$ jets and diboson (WW,WZ and ZZ) contributions. The percentages for signal and background processes quoted in Table 5.2 are in rough agreement with those quoted in [55], with small differences due to the different b-tagging algorithms used. They are also in agreement with a more recent analysis [133] which uses similar selection cuts and b-tagging algorithm and working point as this one.
Since this analysis is focused on kinematics of neutral strange particles in top events, no attempt is made to remove the single top background contamination. Figure 5.3 below shows that the neutral strange particle spectra in the Wt

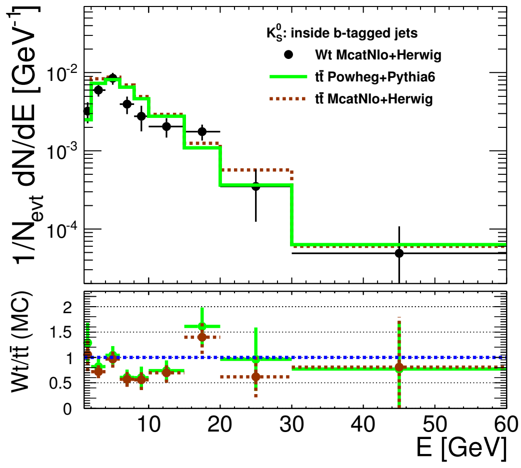
background is within statistical uncertainties the same as in the *ttbar* dileptonic sample.

Table 5.2: Expected sample composition from MC studies after applying the selection criteria.

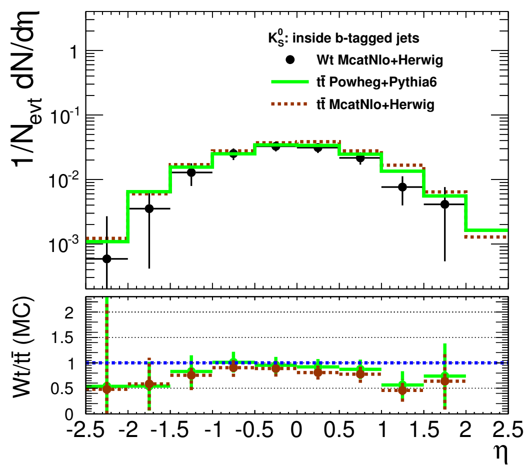
Process	N_{evt}	Percentage (%)
$t\bar{t}$ dileptonic	6860±80	93.9
Single top	300±20	4.1
Z +jets	77±9	1.1
Dibosons	61±8	0.9
Predicted	7300±90	
Observed	6926	



(a) $p_{\rm T}$



(b) Energy



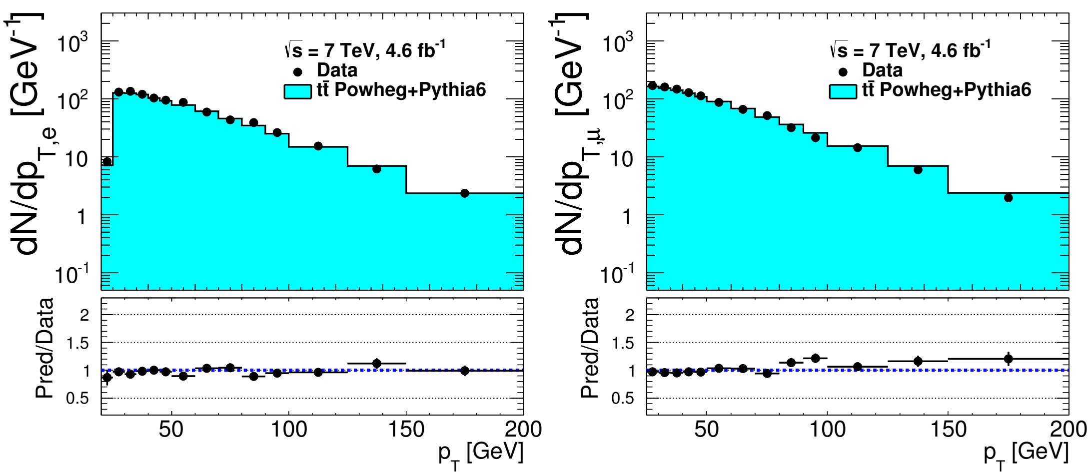
(c) Pseudorapidity

Figure 5.3: Comparison of $p_{\rm T}$, energy and pseudorapidity spectra, for K_S^0's associated with b-tagged jets, for the Wt background simulated with the MC@NLO+HERWIG generator as compared to $t\bar{t}$ dileptonic simulated with the POWHEG+PYTHIA6 and MC@NLO+HERWIG generators.

The dileptonic $t\bar{t}$ events can be divided in three groups according to the leptons arising from the W decays, namely ee, e μ and $\mu\mu$. These groups include contributions from the τ-lepton decay channels where the τ-lepton decays leptonically, $\tau \to \ell\nu_\ell\nu_\tau$. The expected composition along with the data sample composition are stated in Table 5.3, below. These results do not fit expectations from braching ratios $\mathcal{B}(W^\pm \to \ell^\pm\nu) \simeq 11\%$. Nevertheless, this can be understood as due to the higher reconstruction efficiency for muons and the looser selection cuts for the $e\mu$ channel, which has a larger acceptance since neither lepton pair mass constraints nor $E_{\rm T}^{\rm miss}$ cuts are applied. In Figure 5.4 the transverse momentum distributions of these leptons are shown along with a comparison with $t\bar{t}$ MC expectations.

Table 5.3: $t\bar{t}$ dileptonic (expected) composition for data (POWHEG+PYTHIA6 MC) sample. It is to be noticed that e/μ makes reference to the final reconstructed lepton, but these can come from $W^{\pm} \to \tau^{\pm} \to \ell^{\pm}$.

	Expected		Data	
Process	$N_{\rm evt}$	%	$N_{\rm evt}$	%
$t\bar{t} \to e^+e^-$	693	10.1	741	10.7
$t\bar{t} \to e\mu$	4626	67.4	4710	68.0
$t\bar{t} \to \mu^+\mu^-$	1545	22.5	1475	21.3



(a) electron $p_{\rm T}$ (b) muon $p_{\rm T}$

Figure 5.4: Comparison of the lepton trigger $p_{\rm T}$ distributions for data and POWHEG+PYTHIA6 MC predictions, for electrons (left) and muons (right).

In order to shed light on the jet structure in the hadronic final states under investigation, the normalized jet multiplicities as well as their $p_{\rm T}$ spectra are shown in Figure 5.5. The following features can be observed:

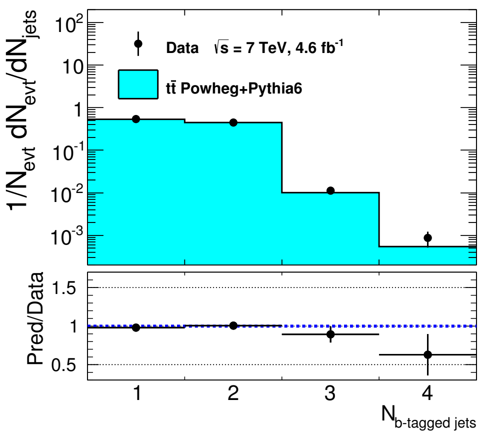
i) the jet activity is indeed very limited since 94% of the final states contain at most 4 jets

ii) about 54% (45%) of the events contain one (two) b-tagged jet(s)

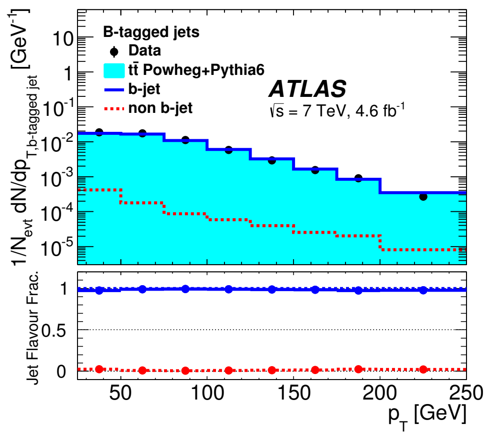
iii) the mean p_{T} of the b-tagged (non b-tagged) jets is 73.4 ± 0.4 GeV (63.2 ± 0.4 GeV)

iv) MC expectations are in very good agreement with data for both jets p_{T} and multiplicities, see [55] for more details.

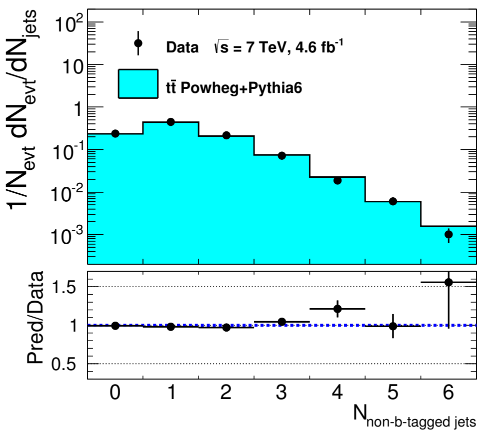
From MC expectations, the b-tagged jet sample has a purity of roughly 99%, while the non b-tagged jet one, contains a fraction of 25% b-jets which are untagged by the MV1 algorithm. These fractions have been calculated by matching reco level jets, b-tagged or not, to their corresponding particle level jets which will be properly defined in forthcoming Sections. These fractions are found to be largely independent of whether non $t\bar{t}$ backgrounds, as given in Table 5.2, are considered or not. This is illustrated in Figure 5.5.



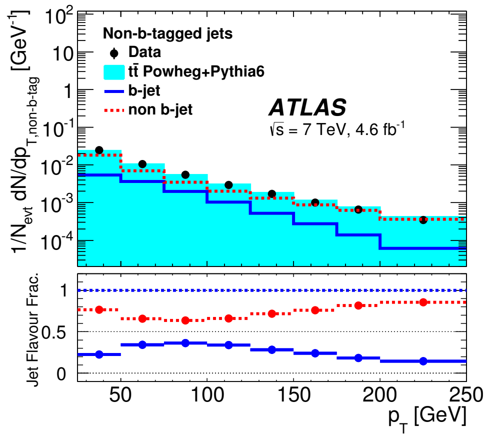
(a) N_{jet} b-tagged



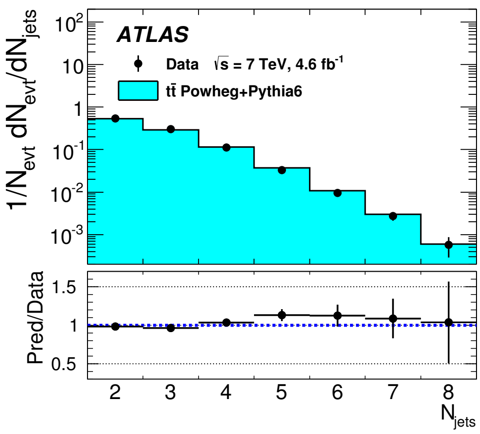
(b) b-tagged jet p_{T}



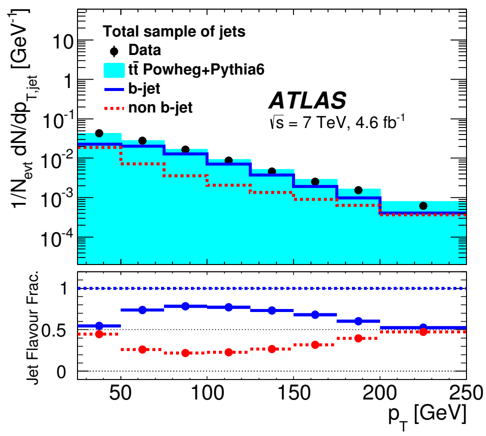
(c) N_{jet} non b-tagged



(d) Non b-tagged jet p_{T}



(e) N_{jet} total



(f) Jet p_{T}

Figure 5.5: Jet multiplicities (left) and p_{T} (right) spectra, for: b-tagged jets (top), non b-tagged jets (medium) and total sample (bottom), in data as compared with POWHEG+PYTHIA6 MC. The expected b-jet flavour fractions for (b) b-tagged, (d) non-b-tagged and the (f) total sample of jets as a function of jet p_{T} are also compared with data. These distributions are normalised to the total number of selected events in data or MC predictions.

5.4 K^0_S and Λ reconstruction: background subtraction

5.4.1 Neutral particle reconstruction

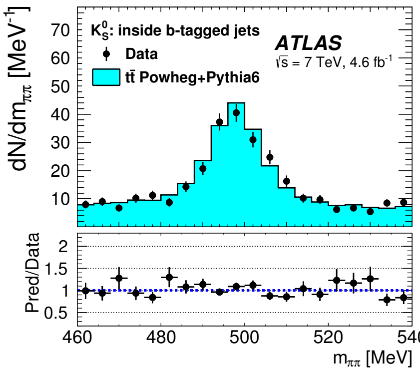
Neutral strange hadrons are reconstructed in the $K^0_S \to \pi^+\pi^-$ ($\Lambda \to p\pi$, $\bar{\Lambda} \to \bar{p}\pi^+$)[1] decay mode by identifying two tracks originating from a displaced vertex, thus profitting from the long lifetimes of neutral kaon mesons (lambda baryons) with $c\tau_0 \approx 2.7$ cm ($c\tau_0 \approx 7.9$ cm).
Tracks are recontructed within the $|\eta| < 2.5$ acceptance of the ID (see Ref. [134, 135]). The K^0_S (Λ) candidates are oppositely charged track pairs with $p_T >$ 100 (500) MeV and at least two silicon hits fitted to a common vertex. For K^0_S reconstruction, the pion mass is assumed for both tracks, while the proton and pion mass are assumed for the Λ case. Further requirements on these candidates, see [86] for more details, are:

- The χ^2 of the two-track vertex fit is required to be less than 15 (with 1 degree of freedom).
- The transverse flight distance (R_{xy}), the distance between the K^0_S (Λ) decay point and the reconstructed primary vertex (jet secondary vertex if the K^0_S (Λ^0) is inside a jet) has to be 4 mm $< R_{xy} <$ 450 mm (17 mm $< R_{xy} <$ 450 mm). This assures the reconstruction of the tracks inside the silicon part of the detector.
- The angle between the K^0_S (Λ) momentum vector and the K^0_S (Λ) flight direction (line connecting decay vertex to primary vertex, or secondary vertex if the K^0_S (Λ) is inside a jet) has to satisfy $\cos\theta_K > 0.999$ ($\cos\theta_\Lambda >$ 0.9998).

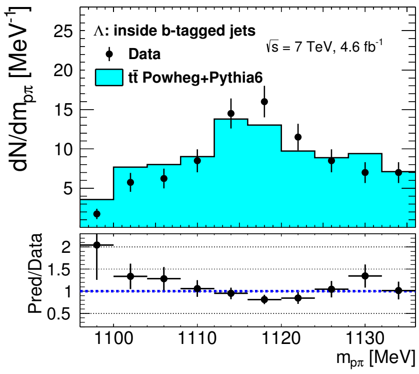
The K^0_S (Λ) candidates that fulfil these conditions, are then separated in three classes: inside a b-tagged jet, inside a non b-tagged jet and outside any jet. To this end the separation ΔR, in the $\eta - \phi$ plane, between the K^0_S (Λ) line of flight and the jet axis is computed. For $\Delta R < 0.4$, a K^0_S (Λ) is associated with a jet, b-tagged or not. Otherwise it is classified as outside any jet. There are no cases of a K^0_S (Λ) inside two different jets simultaneausly.
The mass distributions of the candidates fulfilling all these conditions are shown in Figure 5.6. These mass distributions exhibit a resonance structure, centered around the nominal K^0_S (Λ) mass, with flat tails extending at both sides. The MC

[1]For Λ and $\bar{\Lambda}$ decays, the track with the higher p_T is assigned the proton mass and the other track is assigned the pion mass. In the following and due to statistics Λ will refer to the sum of Λ and $\bar{\Lambda}$ particles.

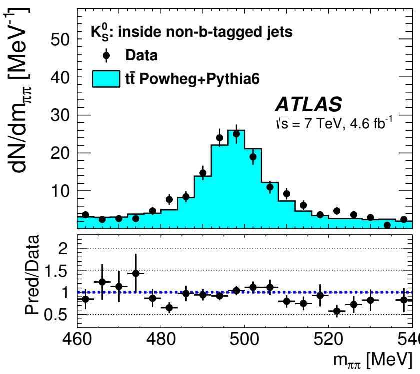
expectations agree fairly well with the data but for K_S^0 (Λ) outside jets, where it falls short by some 40%.



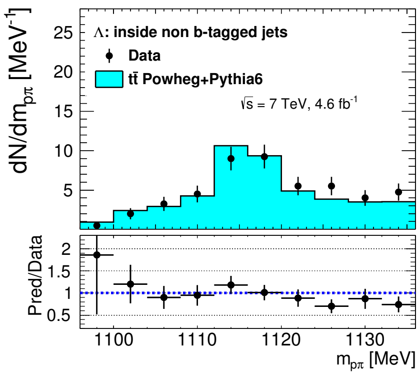
(a) b-tagged jets



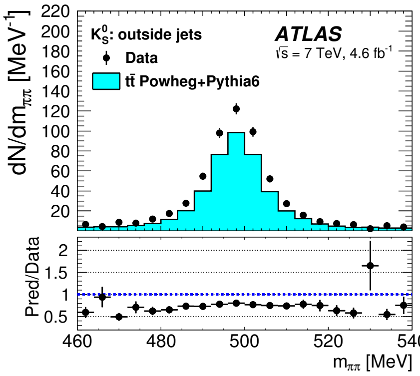
(b) b-tagged jets



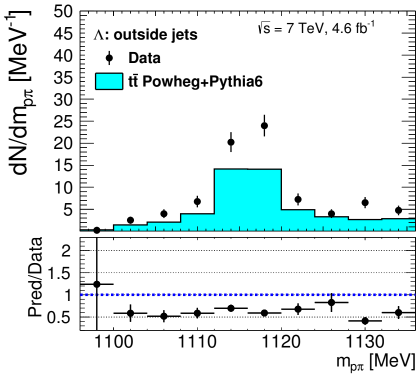
(c) Non b-tagged jets



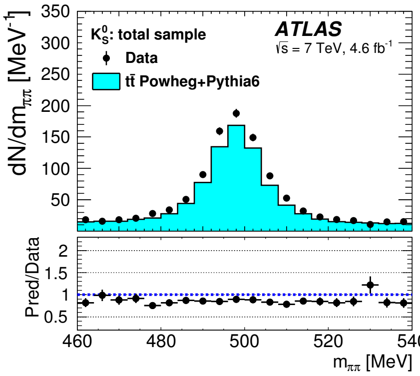
(d) Non b-tagged jets



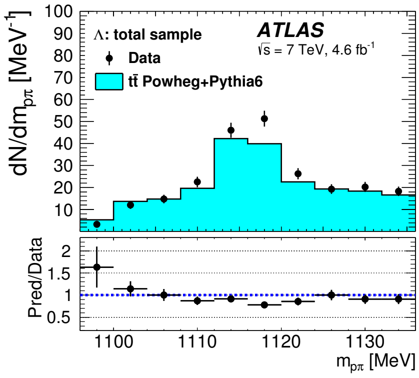
(e) Outside jets



(f) Outside jets



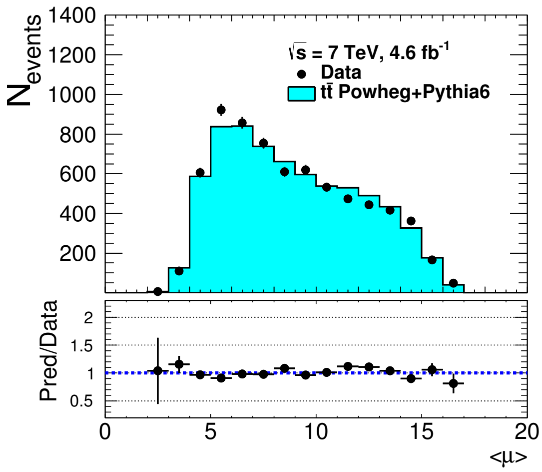
(g) All sample



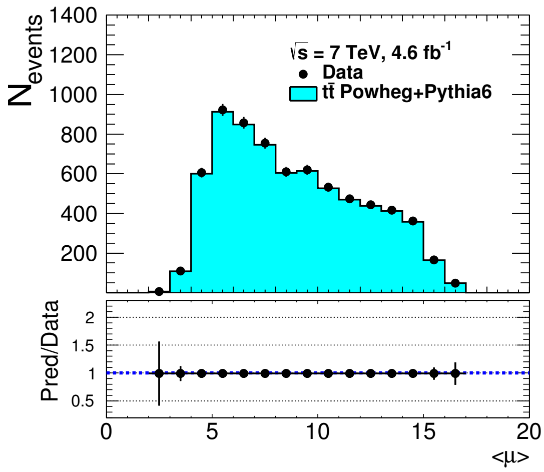
(h) All sample

Figure 5.6: The measured K^0_S (left) and Λ (right) mass distributions along with Powheg+Pythia6 MC simulations.

Due to the fact that pile-up conditions are different for data and MC samples, the latter has to be reweighted on an event-by-event basis depending on the average number of interactions per bunch crossing, $< \mu >$. The ATLAS toolkit package PILEUPREWEIGHTING-00-02-03 has been used to implement this. In Figure 5.7 the $< \mu >$ distributions are shown before and after the reweighting, where the agreement with the data is perfect as expected. The average $< \mu >$ of the data sample is 8.84.



(a) Before reweighting



(b) After reweighting

Figure 5.7: Comparison of the $< \mu >$ distributions for data and POWHEG+PYTHIA6 MC, before (left) and after (right) applying pile-up reweighting.

5.4.2 Background subtraction

In order to get rid of combinatorial background reconstructed as K_S^0 (Λ), a sideband subtraction can be used in the reconstructed mass distribution. The signal range [480–520] MeV ([1106–1126] MeV) is considered for K_S^0 (respectively Λ) production. The background ranges are considered to be [460–480] and [520–540] MeV ([1096–1106] and [1126–1136] MeV) for K_S^0 (respectively Λ) production. Events in the sidebands are given negative weights. The sideband subtraction is applied to both data and MC samples at reco level.

The results of this simple sideband subtraction procedure, used as a baseline, has been cross-checked by fitting the mass distributions to a gaussian centered at the nominal mass plus a flat (linear) shape for the K_S^0 (respectively Λ) background. Variations on the background shape, (flat/linear), have been considered, and the results agree within less than 10% of the statistical uncertainties. Figure 5.8 shows the results of the fits for the total sample of K_S^0 and Λ candidates.

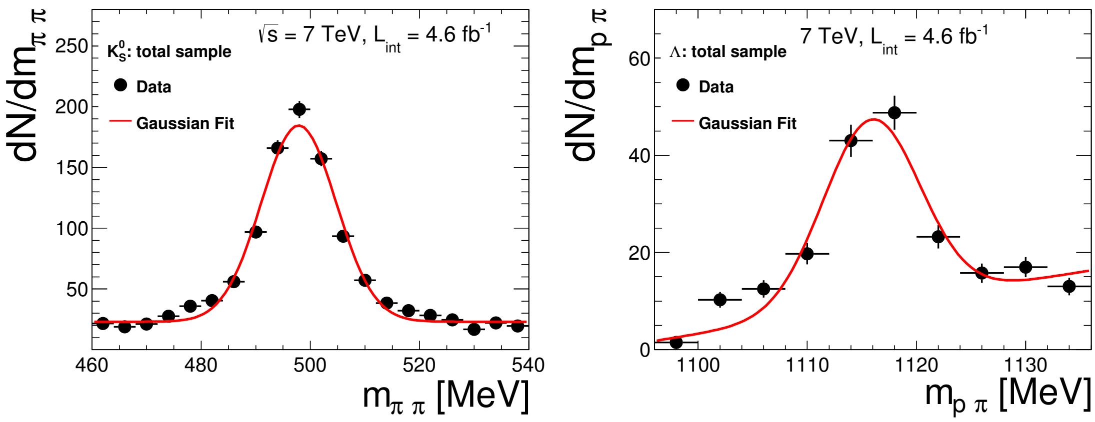
The resulting numbers of signal and background events are found in good agreement with the previous ones within statistical uncertainties. The resulting K^0_S and Λ masses from fits to the total $\pi\pi$ and $p\pi$ mass distributions are 497.8±0.2 MeV and 1115.8±0.3 MeV, in perfect agreement with the PDG values. The K^0_S (Λ) widths from the fits are 6.83±0.03 MeV (4.16±0.04 MeV), therefore the signal ranges chosen for the background subtraction contains 99% (95%) of the signal. The numbers of K^0_S and Λ particles reconstructed in the data after sideband background subtraction for the three classes considered as well as for the total sample are shown in Tables 5.4 and 5.5. For completeness, the signal and background estimates (in the signal regions defined above) extracted from the fits are also shown.
The following remark is in order: the χ^2 values of the fits are somewhat poor, namely 55 (38) for 20 (10) K^0_S (Λ) data points. As already noted in [86, 87, 88, 89, 90], fitting to a double gaussian with a common mean value improves considerably the quality of the fit. This exercise has been tried and in fact the resulting χ^2 values are 11 and 16 for K^0_S and Λ particles, giving very reasonable p-values, Figure 5.9. The resulting gaussian mean values coincide with those obtained before and the number of signal and background events are stable within statistical uncertainties.

Table 5.4: The number of K^0_S particles reconstructed in the data after sideband background subtraction, which is used as a baseline, for each class and for the total sample, along with the signal and background estimates in the signal regions as extracted from the fit to a gaussian (double gaussian) plus a flat background, as a cross-check. All errors are statistical only.

	Reco	Fit Gaussian		Fit Mod. Gaussian	
Class	Signal	Signal	Back.	Signal	Back.
Inside b-tagged jets	530 ± 34	540 ± 38	309 ± 18	574 ± 81	276 ± 21
Inside non b-tagged jets	391 ± 25	416 ± 33	99 ± 12	438 ± 49	74 ± 15
Outside any jet	1837 ± 49	1863 ± 70	216 ± 15	1965 ± 77	135 ± 18
Total sample	2758 ± 69	2777 ± 91	996 ± 32	2987 ± 140	831 ± 37

Table 5.5: The number of Λ particles reconstructed in the data after sideband background subtraction, which is used as a baseline, for each class and for the total sample, along with the signal and background estimates in the signal regions as extracted from the fit to a gaussian (double gaussian) plus a linear background, as a cross-check. All errors are statistical only.

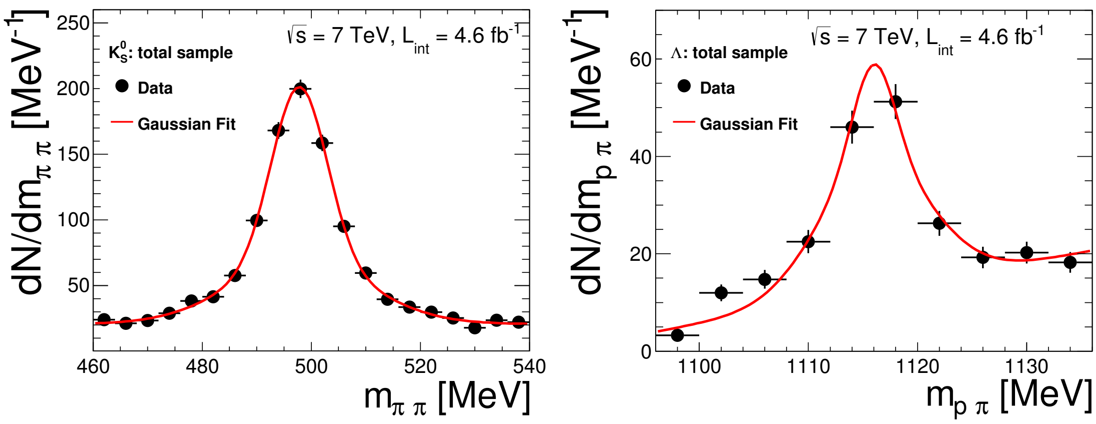
	Reco	Fit Gaussian		Fit Double Gaussian	
Class	Signal	Signal	Back.	Signal	Back.
Inside b-tagged jets	115 ± 19	165 ± 24	88 ± 37	166 ± 54	69 ± 30
Inside non b-tagged jets	65 ± 14	82 ± 18	53 ± 17	85 ± 39	47 ± 14
Outside any jet	183 ± 18	185 ± 23	61 ± 13	201 ± 46	53 ± 14
Total sample	363 ± 31	405 ± 40	254 ± 42	445 ± 87	216 ± 39



(a) K^0_S mass fit

(b) Λ mass fit

Figure 5.8: Invariant mass distribution for the total sample of K^0_S (left) and Λ (right) candidates (black dots), in the data, along with their fitting functions (red line).

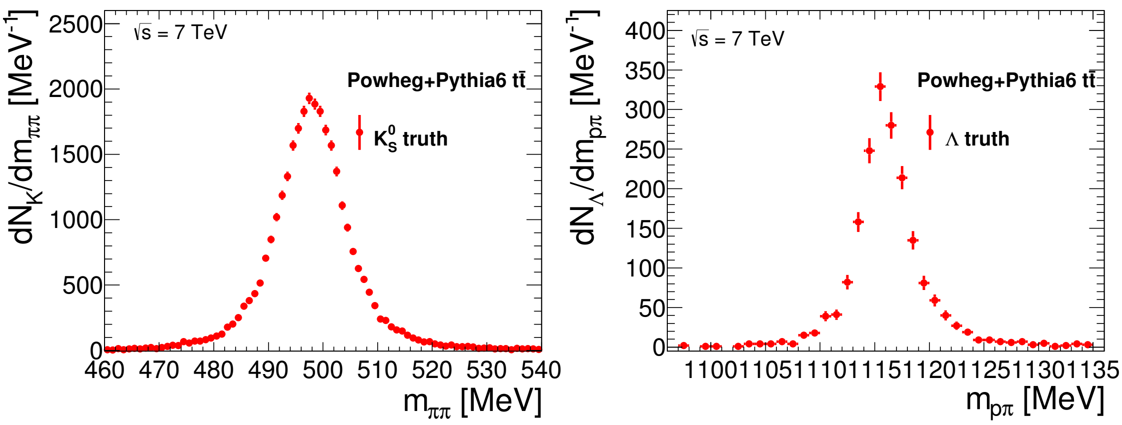


(a) K_S^0 mass fit (b) Λ mass fit

Figure 5.9: Invariant mass distribution for the total sample of K_S^0 (left) and Λ (right) candidates (black dots), in the data, along with their fitting functions (red line) using a double gaussian with a common mean.

5.4.2.1 Validation of the background subtraction

As a further check of the validity of the background subtraction discussed in the previous subsection, Figure 5.10 shows the Powheg+Pythia6 $\pi^+\pi^-$ and $p\pi$ detector-level mass distributions for K_S^0 and Λ candidates which can be matched at particle level.



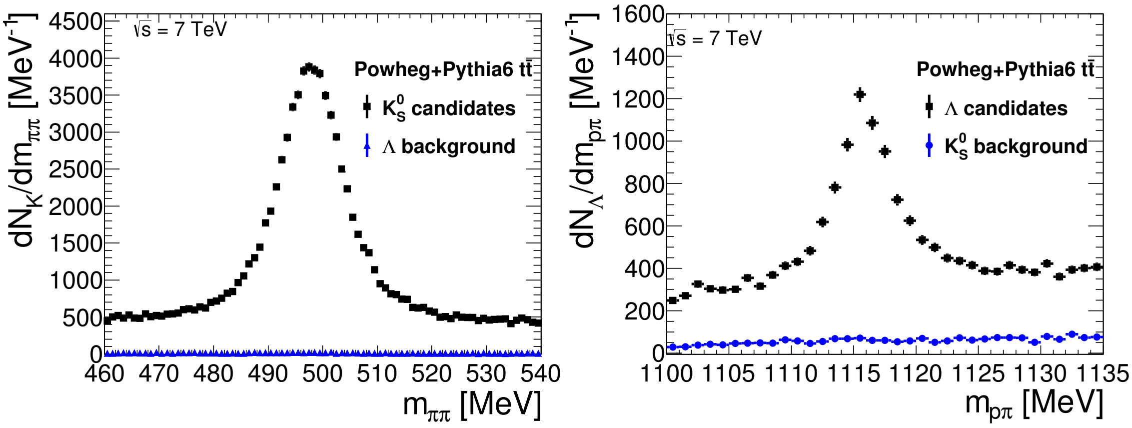
(a) K_S^0 mass (b) Λ mass

Figure 5.10: Invariant mass distribution for the total sample of K_S^0 (left) and Λ (right)candidates in the Powheg+Pythia6 sample at detector level, which found a match at particle level.

This Figure shows that 5 % (4 %) of the K_S^0 (Λ) leak into the considered sideband regions. In order to check the uncertainty due to this effect, the analysis has been

repeated redefining the signal mass ranges to [475–525] MeV ([1104–1128] MeV) for K^0_S (respectively Λ) production. The new background ranges considered are [450–475] and [525–550] MeV ([1092–1104] and [1128–1140] MeV) for K^0_S (respectively Λ) production. Thus the mass ranges have been extended by 25 % (20 %). With these new signal and sideband region definitions, the percentage of signal K^0_S (Λ) leaking into the sidebands is much smaller than 1 %. The effects of these variations after unfolding are discussed in Section 5.6.1.2.
To finish this subsection, a MC study was performed to check the Λ (K^0_S) background in the K^0_S (Λ) candidate mass distributions. This is done by matching the candidates at detector level to their corresponding truth particle. This is illustrated in Figure 5.11 which shows that:

- the Λ background under the K^0_S peak is negligible (in the order of 1 %).
- the K^0_S background under the Λ mass distribution is roughly 10 % and equally distributed between signal and sideband regions.



(a) K^0_S mass

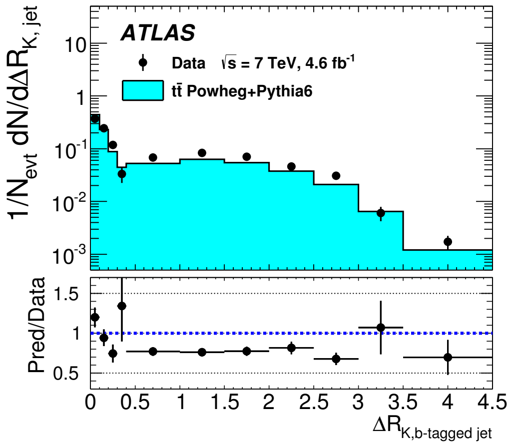
(b) Λ mass

Figure 5.11: Invariant mass distribution for the total sample of K^0_S (left) and Λ (right) candidates (black) in the Powheg+Pythia6 sample at detector level, along with the Λ (left) and K^0_S (right) background (blue).

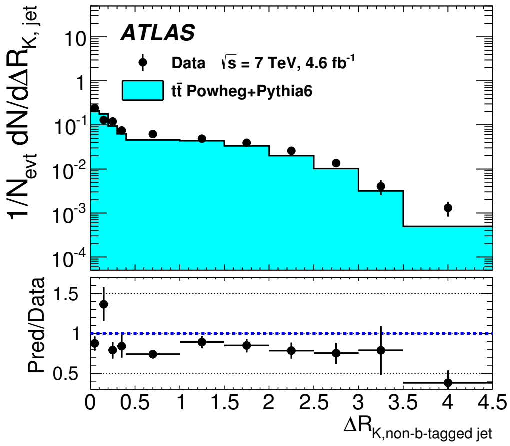
5.4.3 Association with jets

In order to gain further insight into the procedure used to associate neutral strange particles to jets, the angular separation in the (η,ϕ) plane between each K^0_S and its nearest jet is shown in Figure 5.12. Two cases are considered depending on whether the nearest jet is b-tagged or not.
Clearly these distributions show that the K^0_S density is significantly higher in the vicinity of the jet axis, b-tagged or not, than farther away by approximately one

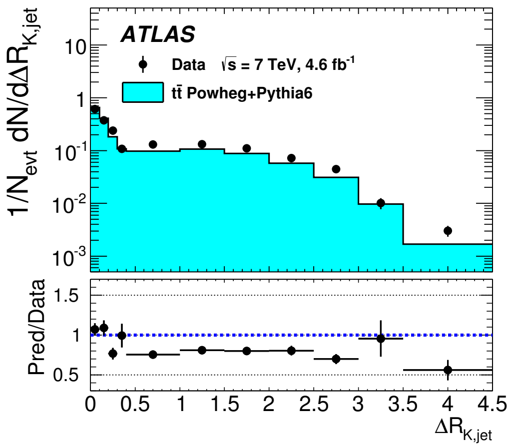
order of magnitude. The phase space outside the jet cone is much larger and the K_S^0 density is roughly constant between $0.4 < \Delta R < 3.0$ and falls sharply beyond that limit. Furthermore, the MC description of the data is fair within the jet cone, but underestimates the data outside the jet regions. This is also observed in the neutral strange particles mass distributions above.



(a) ΔR K_S^0 to b-tagged jet



(b) ΔR K_S^0 to non b-tagged jet



(c) ΔR K_S^0 to any jet

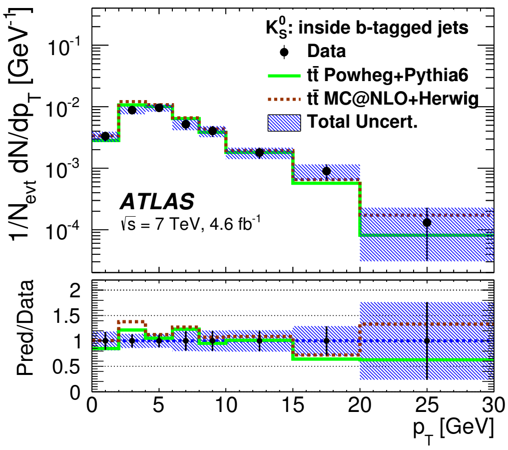
Figure 5.12: The K_S^0 angular separation (ΔR) to the nearest jet, depending on whether the jet is b-tagged (top left) or not (top right), as well as for all jets (bottom left), in data as compared with POWHEG+PYTHIA6 MC.

5.5 Results at detector level

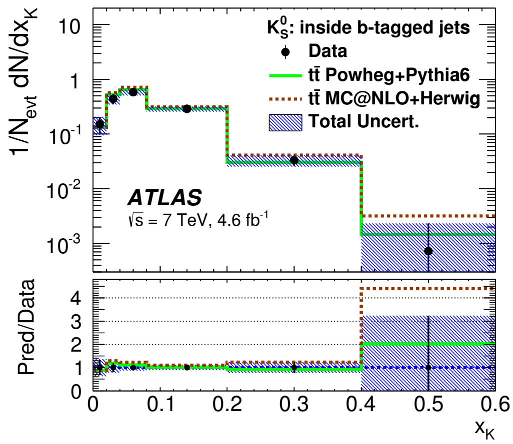
As already discussed in the introduction, neutral strange particle production is studied in terms of simple and generally accepted kinematic variables, like the transverse momentum p_{T}, the energy E, the pseudorapidity η, the transverse flight distance R_{xy} and the multiplicity N. For K^0_{S} and Λ production inside jets, also the energy fraction, $x_{K,\Lambda} = \frac{E_{K,\Lambda}}{E_{jet}}$, is considered .
To this purpose, the reconstructed K^0_{S} and Λ mass distributions have been obtained in different bins of the variables discussed above, and the number of background subtracted candidates have been determined, as discussed in the previous Section. They have been normalized to the total number of events in the dileptonic $t\bar{t}$ sample.

5.5.1 K^0_{S} production at detector level

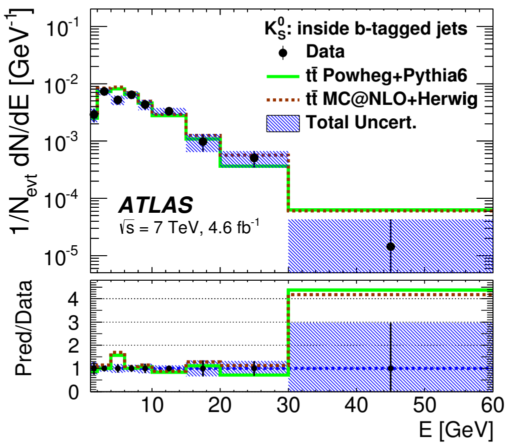
These distributions for K^0_{S} production are displayed in Figures 5.13 to 5.16. Three different classes are considered i.e. K^0_{S} inside b-tagged jets, non b-tagged jets and outside any jet. For completeness the distributions for the total sample are also shown. They have been compared to two broad MC models namely Powheg+Pythia6 and MC@NLO+Herwig. Error bars (blue striped area) give account for statistical (total, i.e. statistical plus detector level sytematics added in quadrature) uncertainties in the data while the light blue shaded area in the MC to data ratios give account for statistical uncertainties in the data.
Approximately two thirds of the total sample of K^0_{S}'s are not associated with jets. The K^0_{S} multiplicity inside b-tagged jets is similar to that inside non b-tagged jets. The K^0_{S} spectra inside jets do not seem to strongly depend on whether the jets are b-tagged or not. On the other hand, K^0_{S}'s not associated with jets are softer in p_{T} or energy than those embedded in jets, b-tagged or not, and the pseudorapidity is flatter.
The gross features of the data are fairly well described by both MCs with the only exception of single K^0_{S} production i.e. K^0_{S}'s which are not associated with any jet. Here the MC expectations fall below the data by roughly 30 %.



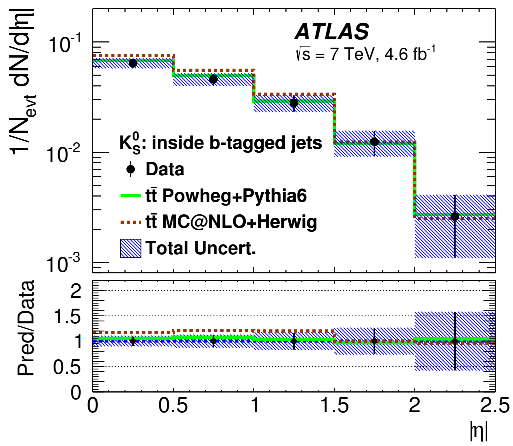
(a) p_{T} distribution



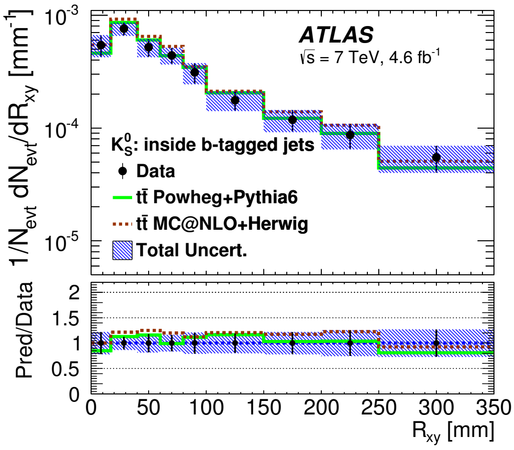
(b) Energy fraction distribution



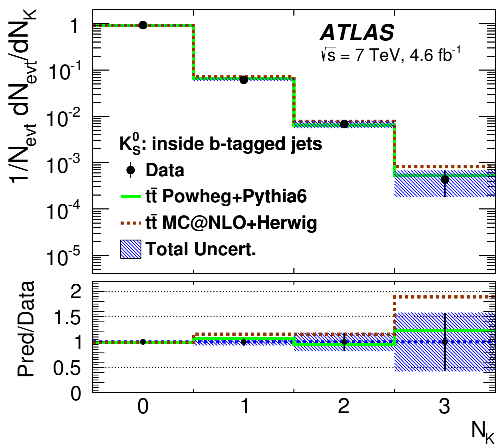
(c) Energy distribution



(d) $|\eta|$ distribution

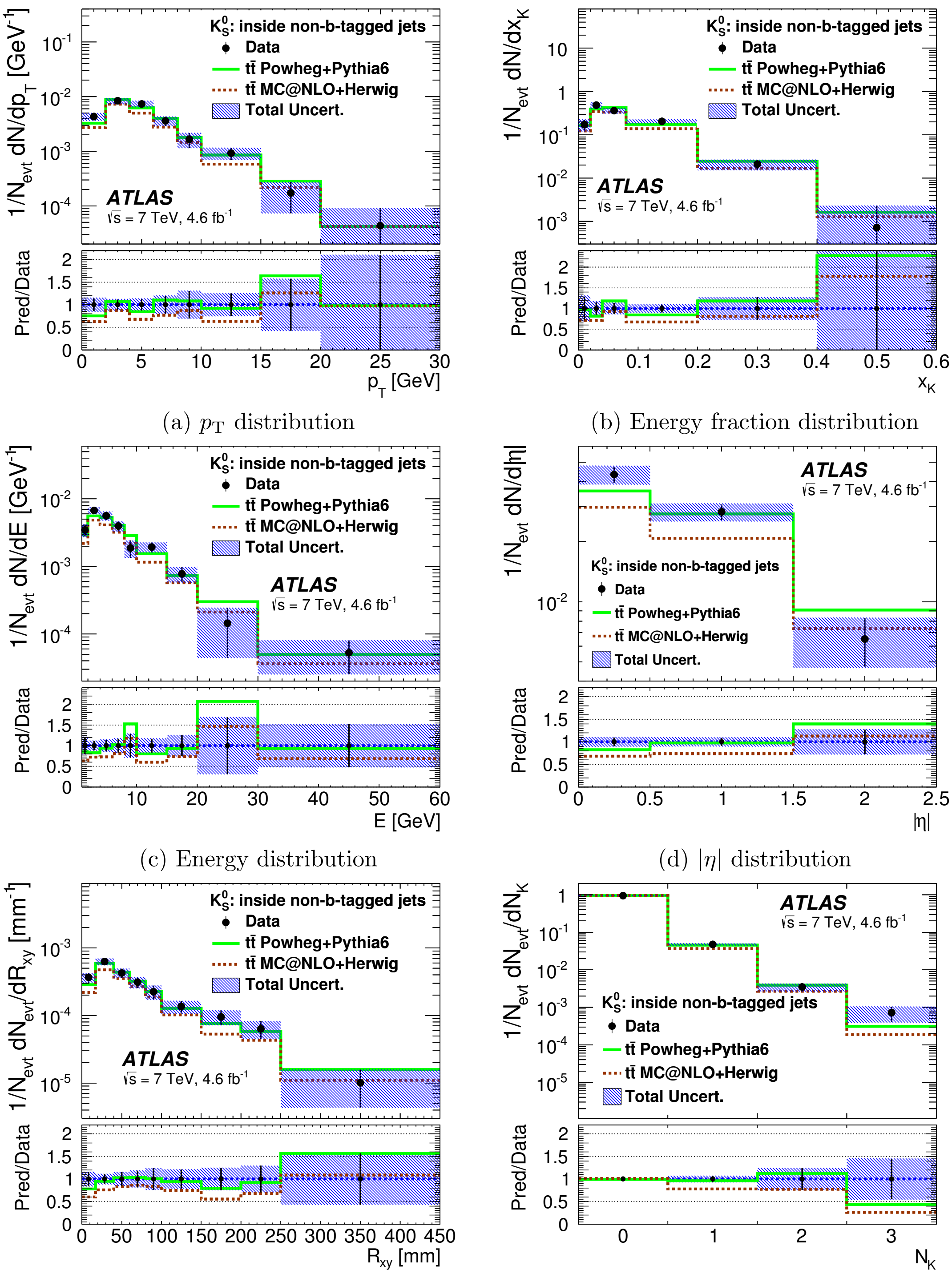


(e) Transverse flight distance distribution



(f) Multiplicity distribution

Figure 5.13: Kinematic characteristics for K^0_{S} production inside b-tagged jets, for data and detector level MC simulated with the POWHEG+PYTHIA6 and MC@NLO+HERWIG generators. Total uncertainties are represented by the shaded area.



(a) p_T distribution

(b) Energy fraction distribution

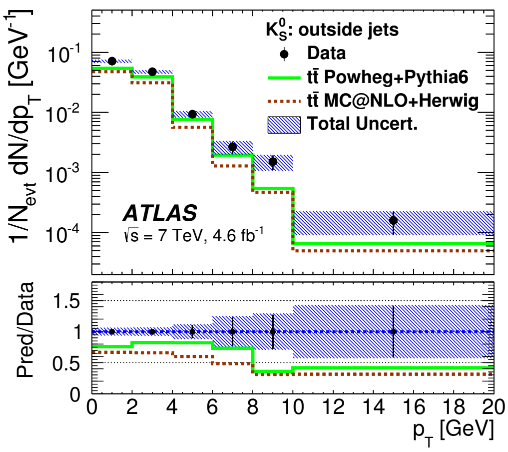
(c) Energy distribution

(d) $|\eta|$ distribution

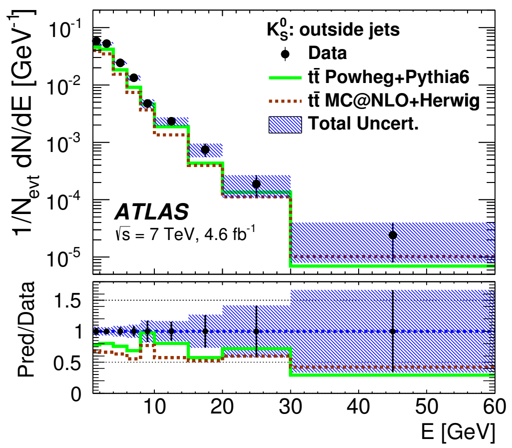
(e) Transverse flight distance distribution

(f) Multiplicity distribution

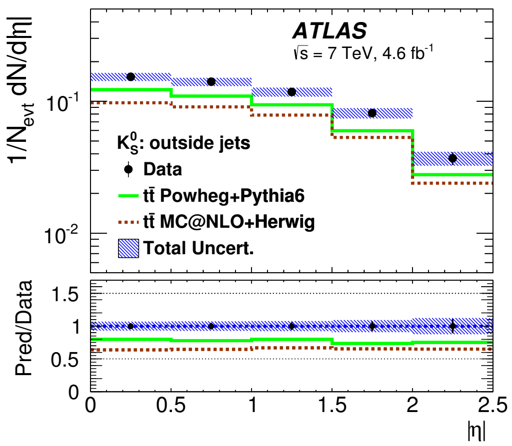
Figure 5.14: Kinematic characteristics for K^0_S production inside non b-tagged jets, for data and detector level MC simulated with the POWHEG+PYTHIA6 and MC@NLO+HERWIG generators. Total uncertainties are represented by the shaded area.



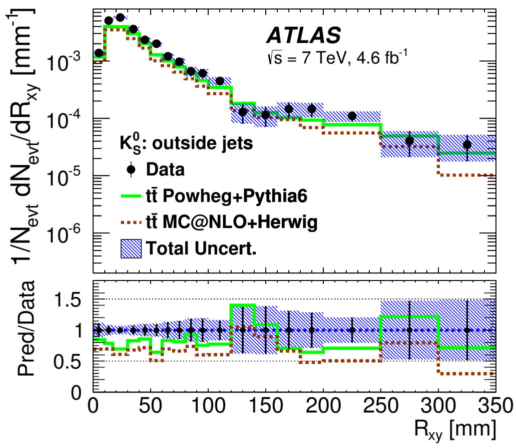
(a) p_T distribution



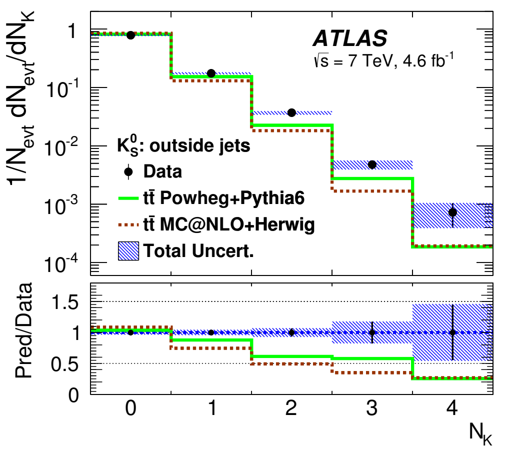
(b) Energy distribution



(c) $|\eta|$ distribution

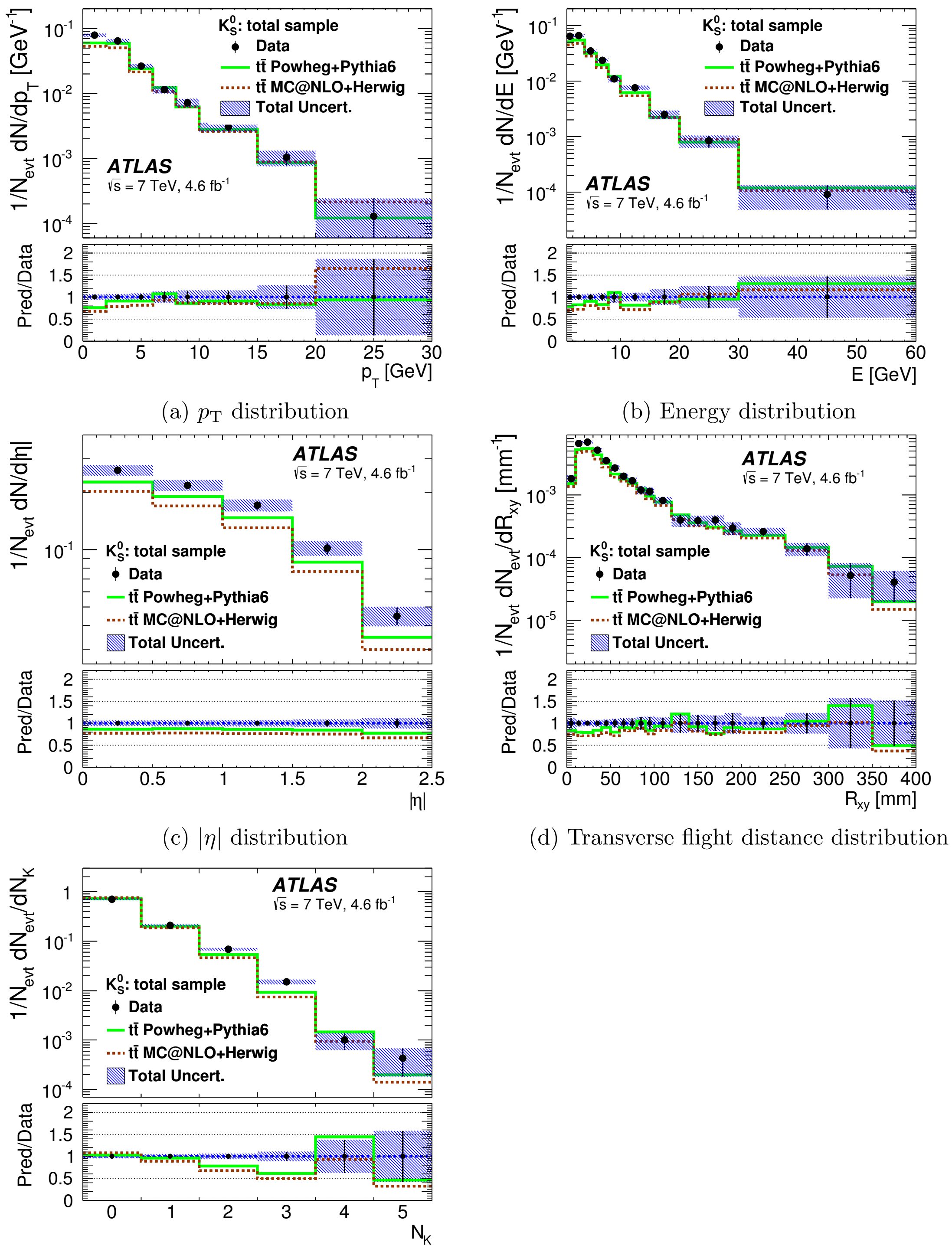


(d) Transverse flight distance distribution



(e) Multiplicity distribution

Figure 5.15: Kinematic characteristics for K^0_S production not associated with jets, for data and detector level MC simulated with the Powheg+Pythia6 and MC@NLO+Herwig generators. Total uncertainties are represented by the shaded area.



(a) p_{T} distribution

(b) Energy distribution

(c) $|\eta|$ distribution

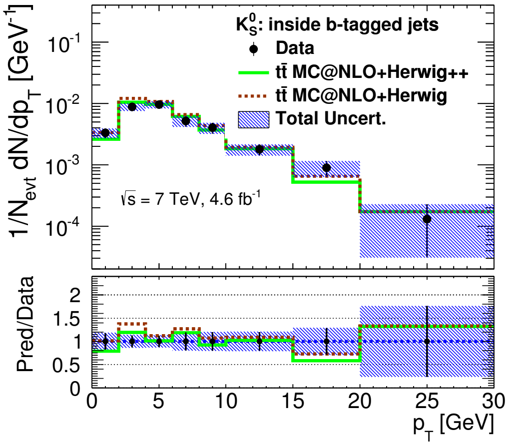
(d) Transverse flight distance distribution

(e) Multiplicity distribution

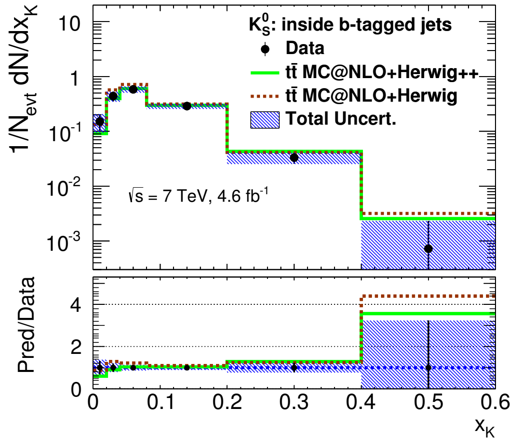
Figure 5.16: Kinematic characteristics for the total K_S^0 production, for data and detector level MC simulated with the POWHEG+PYTHIA6 and MC@NLO+HERWIG generators. Total uncertainties are represented by the shaded area.

5.5.1.1 Comparisons with Herwig++

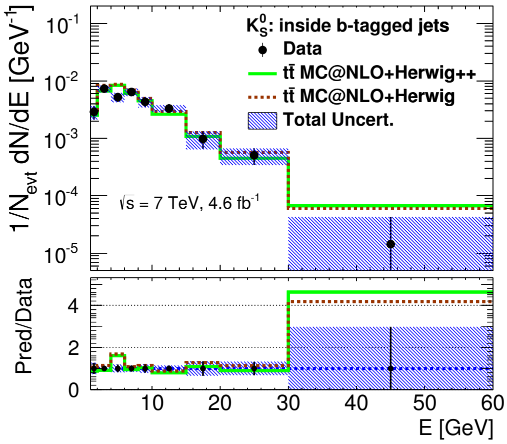
The same distributions for K^0_S production are displayed in Figures 5.17 to 5.20 along with a comparison with MC@NLO+HERWIG++ predictions [136]. The data description of this new generator is very similar to the old MC@NLO+HERWIG.



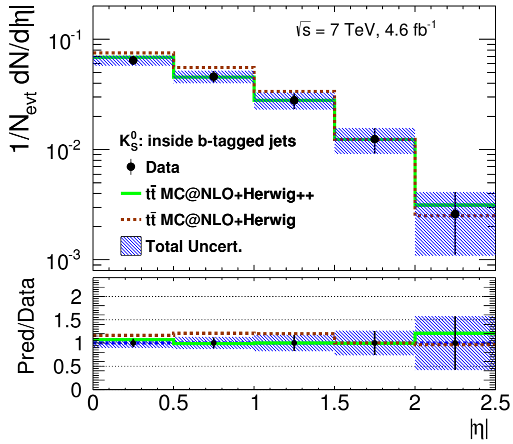
(a) p_T distribution



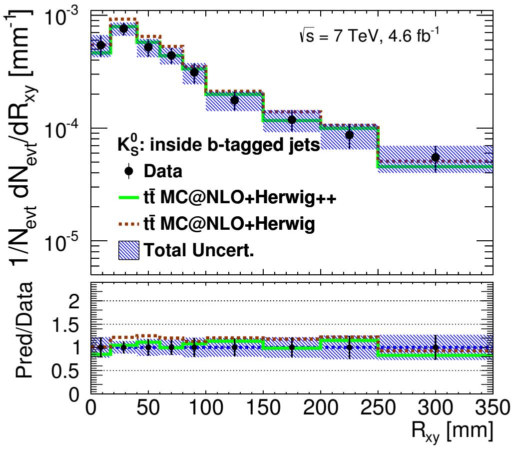
(b) Energy fraction distribution



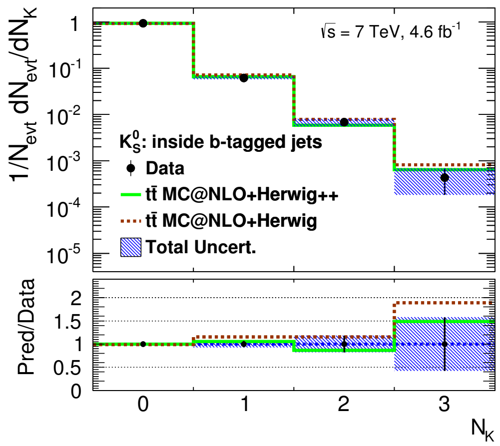
(c) Energy distribution



(d) $|\eta|$ distribution

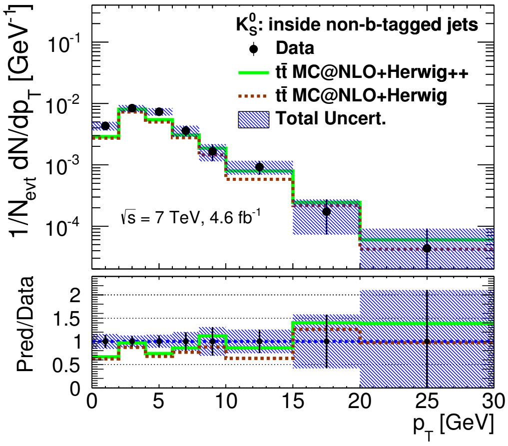


(e) Transverse flight distance distribution

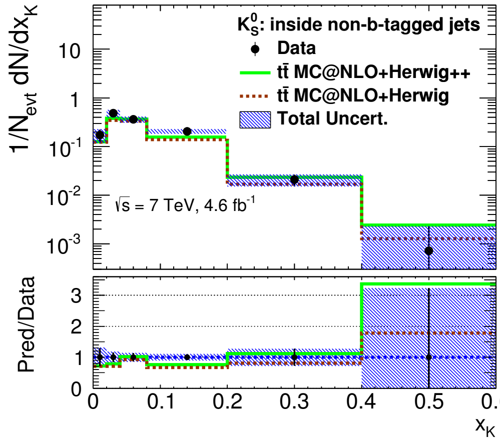


(f) Multiplicity distribution

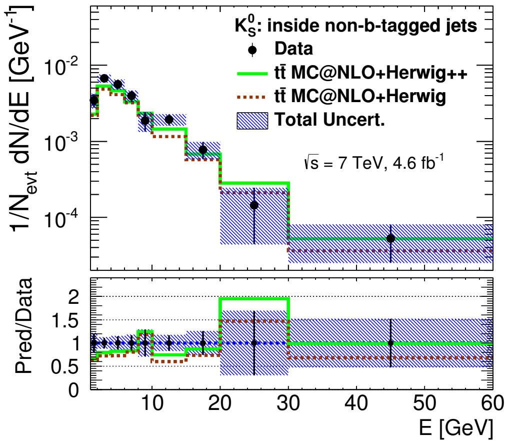
Figure 5.17: Kinematic characteristics for K_S^0 production inside b-tagged jets, for data and detector level MC simulated with the MC@NLO+HERWIG++ and MC@NLO+HERWIG generators. Total uncertainties are represented by the shaded area.



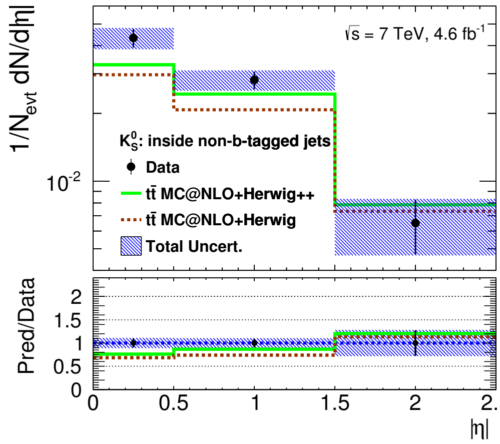
(a) p_{T} distribution



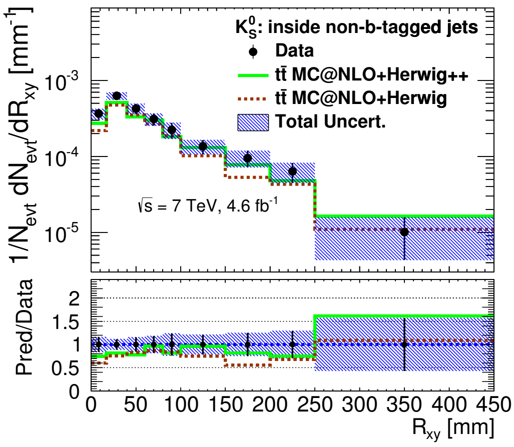
(b) Energy fraction distribution



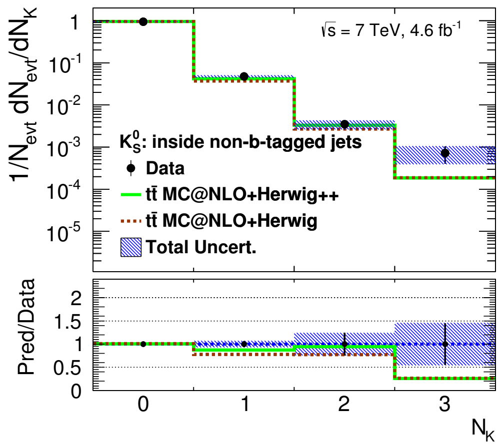
(c) Energy distribution



(d) $|\eta|$ distribution

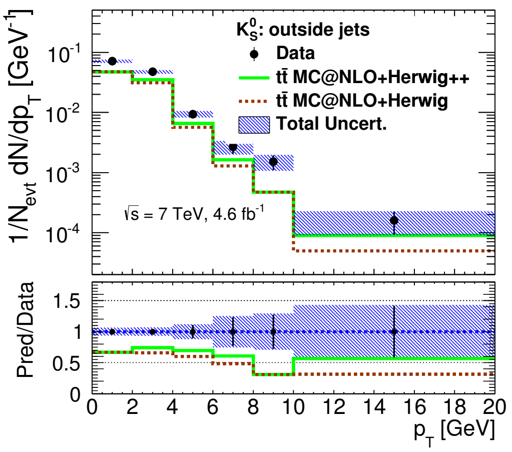


(e) Transverse flight distance distribution

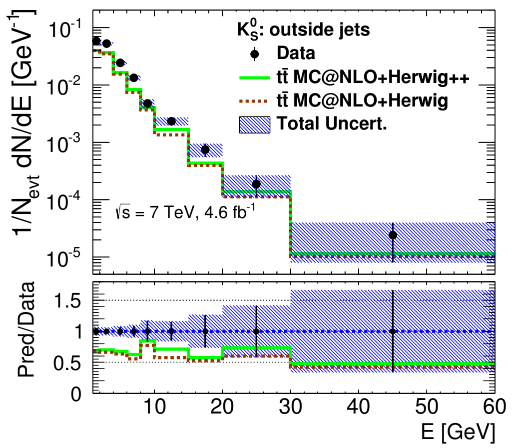


(f) Multiplicity distribution

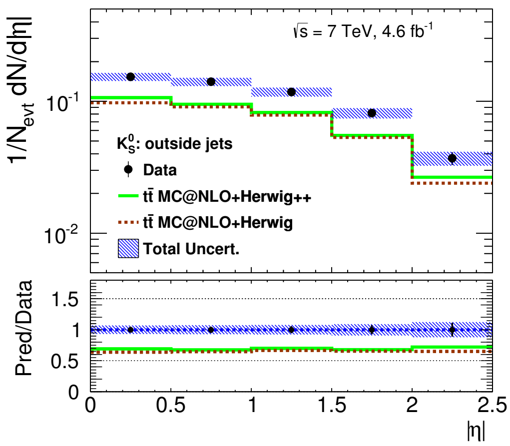
Figure 5.18: Kinematic characteristics for K_S^0 production inside non b-tagged jets, for data and detector level MC simulated with the MC@NLO+Herwig++ and MC@NLO+Herwig generators. Total uncertainties are represented by the shaded area.



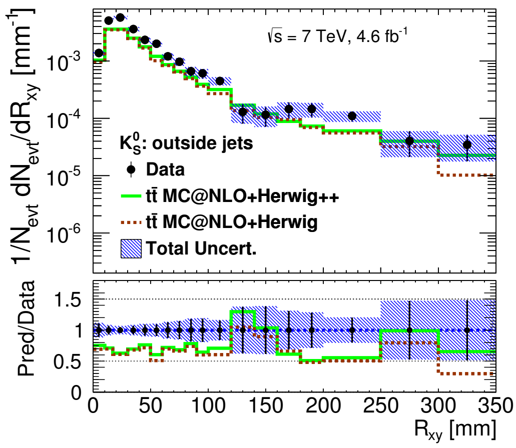
(a) p_{T} distribution



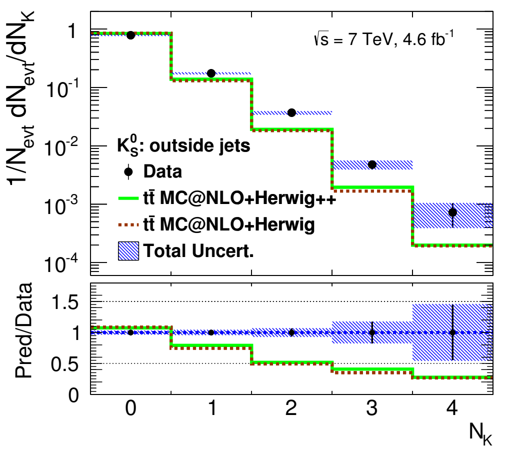
(b) Energy distribution



(c) $|\eta|$ distribution

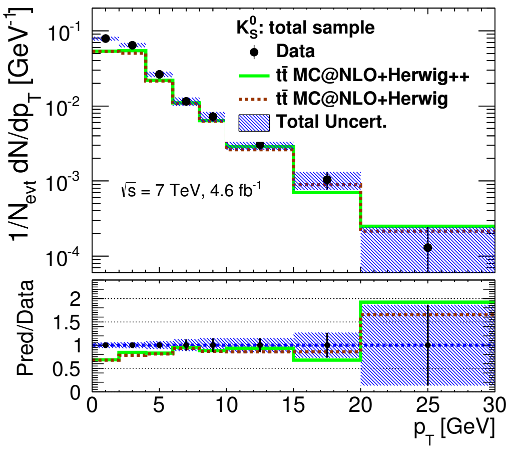


(d) Transverse flight distance distribution

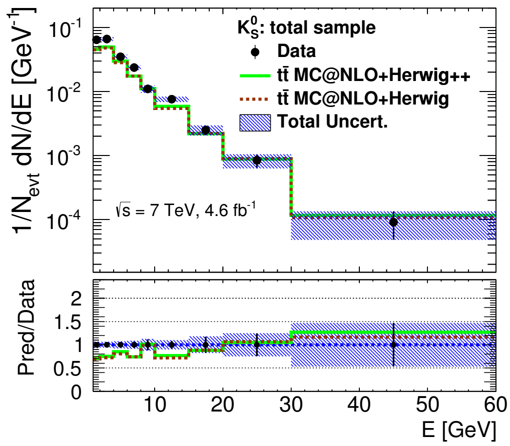


(e) Multiplicity distribution

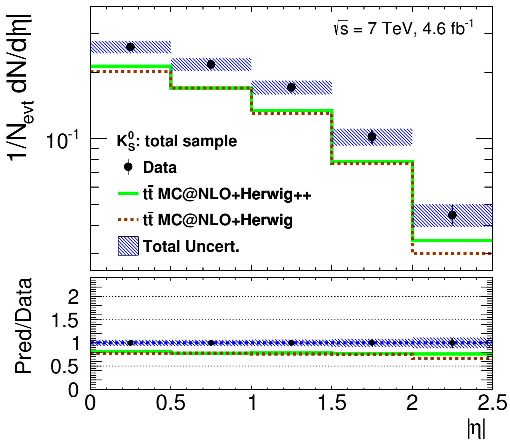
Figure 5.19: Kinematic characteristics for K^0_{S} production not associated with jets, for data and detector level MC simulated with the MC@NLO+Herwig++ and MC@NLO+Herwig generators. Total uncertainties are represented by the shaded area.



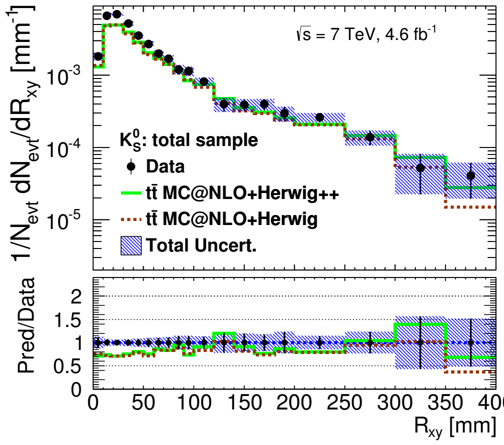
(a) p_{T} distribution



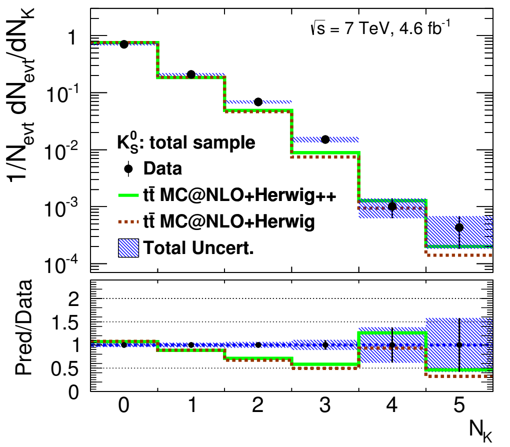
(b) Energy distribution



(c) $|\eta|$ distribution



(d) Transverse flight distance distribution

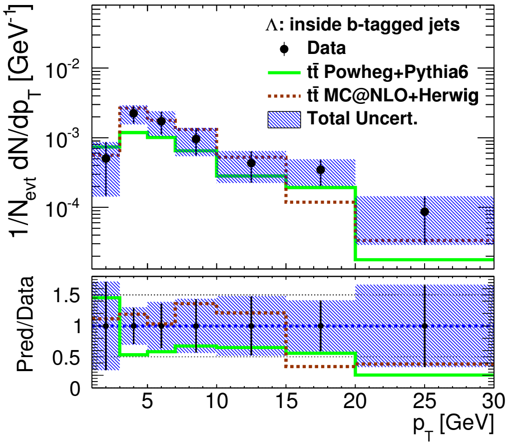


(e) Multiplicity distribution

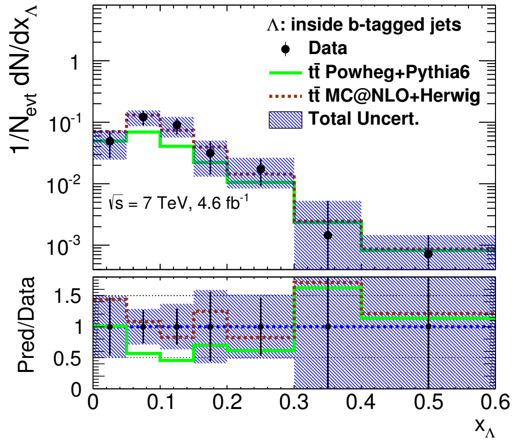
Figure 5.20: Kinematic characteristics for the total K_S^0 production, for data and detector level MC simulated with the MC@NLO+Herwig++ and MC@NLO+Herwig generators. Total uncertainties are represented by the shaded area.

5.5.2 Λ production at detector level

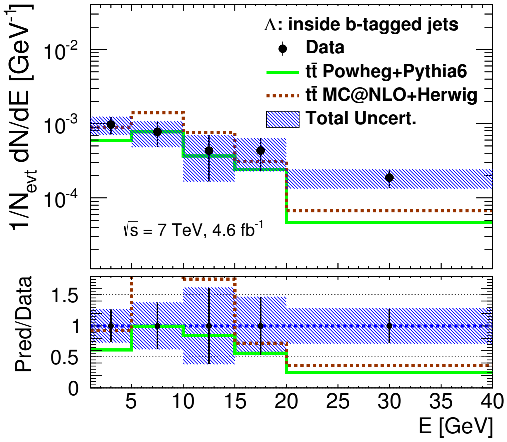
Similar distributions for Λ production are displayed in Figures 5.21 to 5.24.



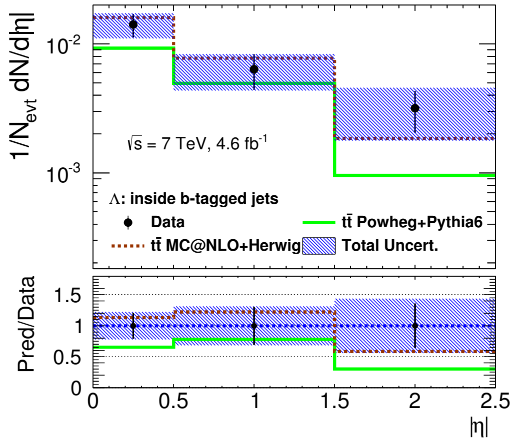
(a) p_{T} distribution



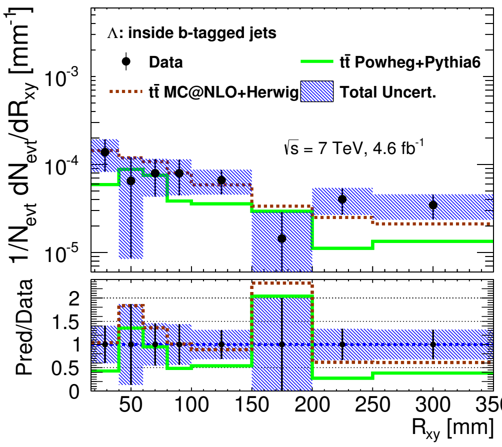
(b) Energy fraction distribution



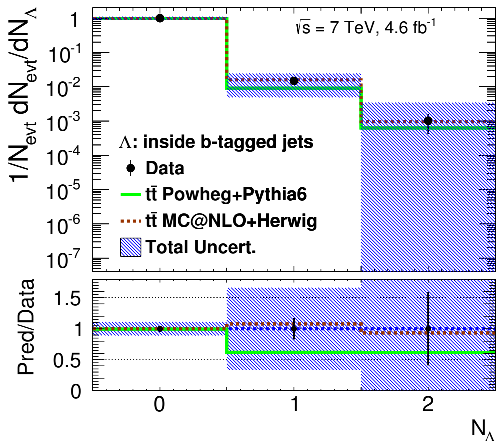
(c) Energy distribution



(d) $|\eta|$ distribution

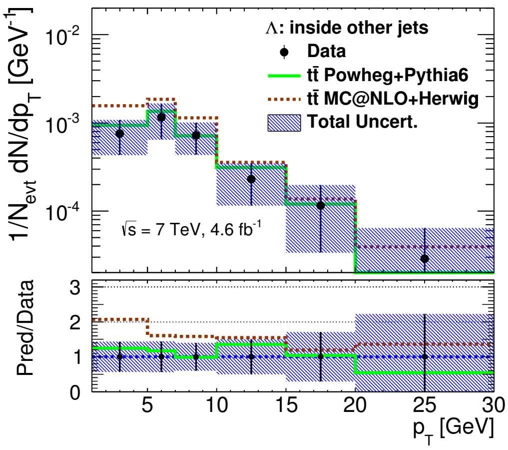


(e) Transverse flight distance distribution

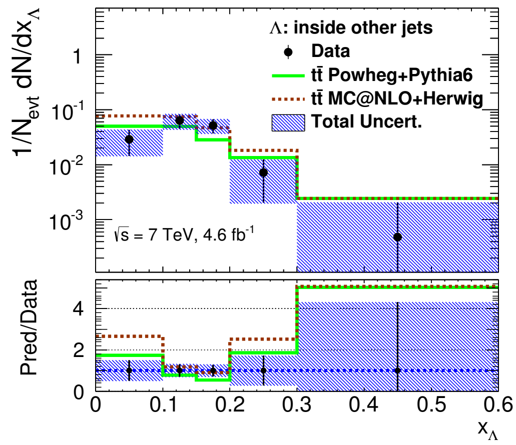


(f) Multiplicity distribution

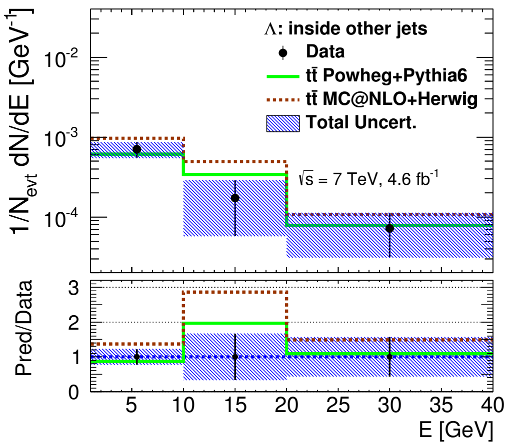
Figure 5.21: Kinematic characteristics for Λ_S^0 production inside b-tagged jets, for data and detector level MC simulated with the Powheg+Pythia6 and MC@NLO+Herwig generators. Total uncertainties are represented by the shaded area.



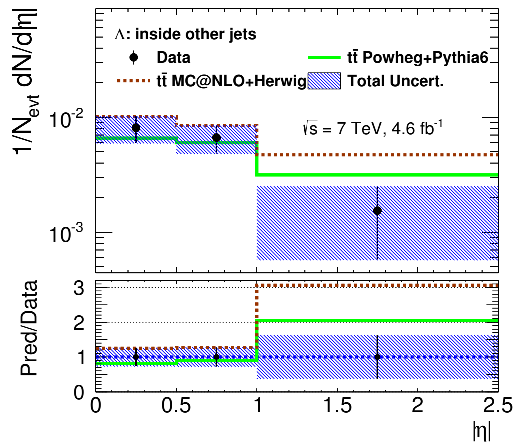
(a) p_{T} distribution



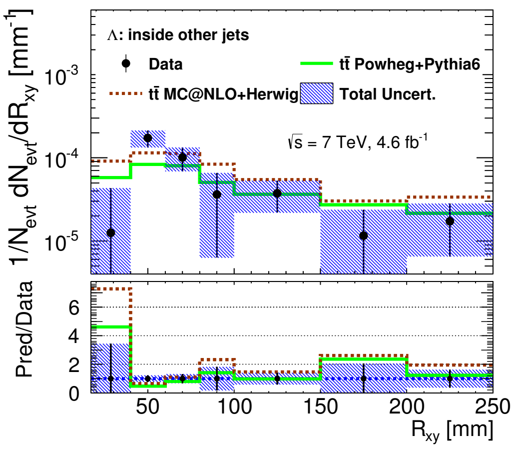
(b) Energy fraction distribution



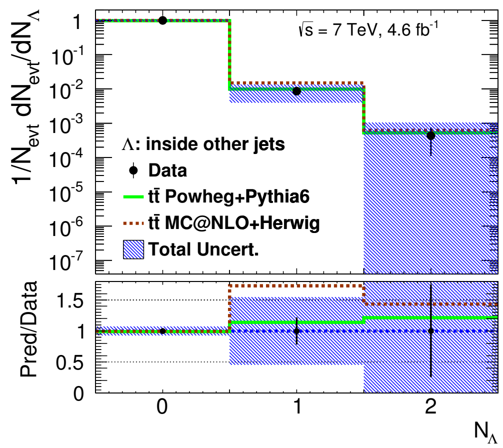
(c) Energy distribution



(d) $|\eta|$ distribution

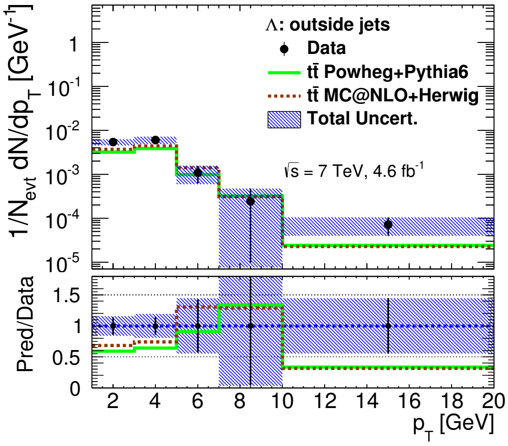


(e) Transverse flight distance distribution

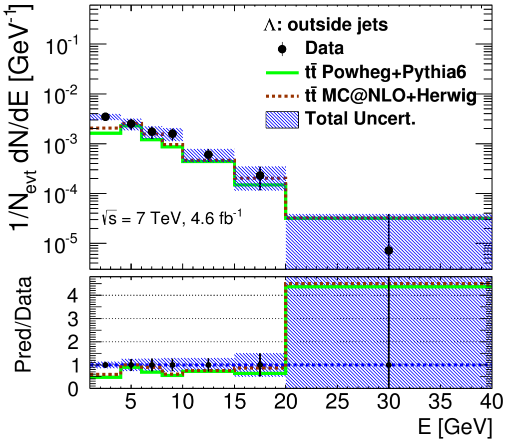


(f) Multiplicity distribution

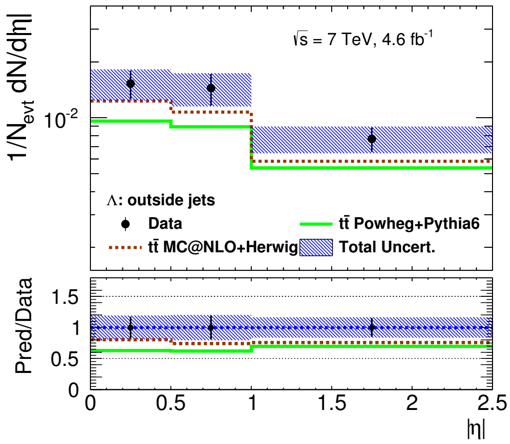
Figure 5.22: Kinematic characteristics for Λ production inside non b-tagged jets, for data and detector level MC simulated with the POWHEG+PYTHIA6 and MC@NLO+HERWIG generators. Total uncertainties are represented by the shaded area.



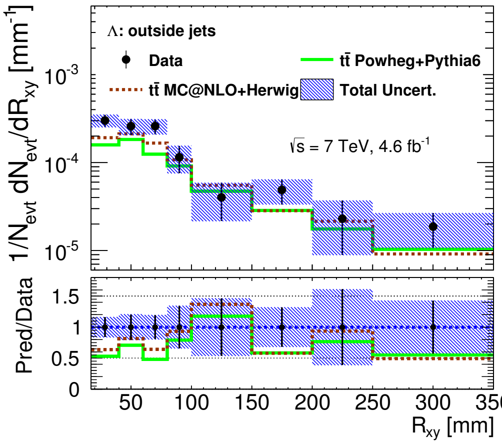
(a) p_{T} distribution



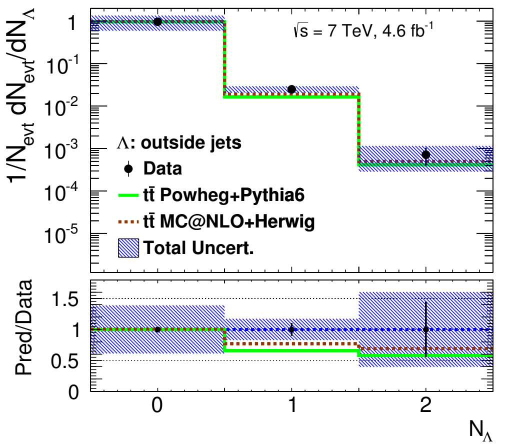
(b) Energy distribution



(c) $|\eta|$ distribution

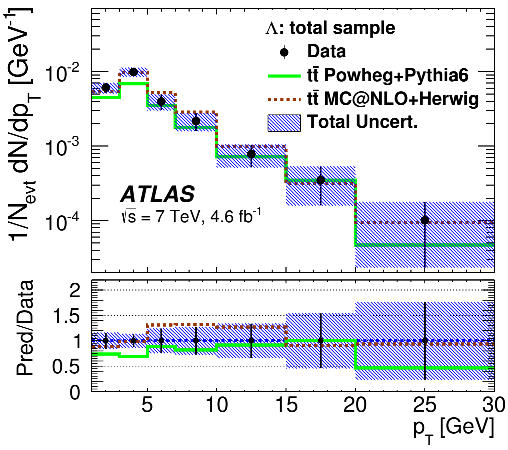


(d) Transverse flight distance distribution

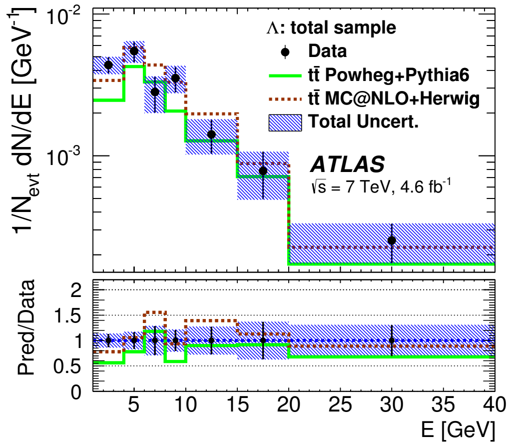


(e) Multiplicity distribution

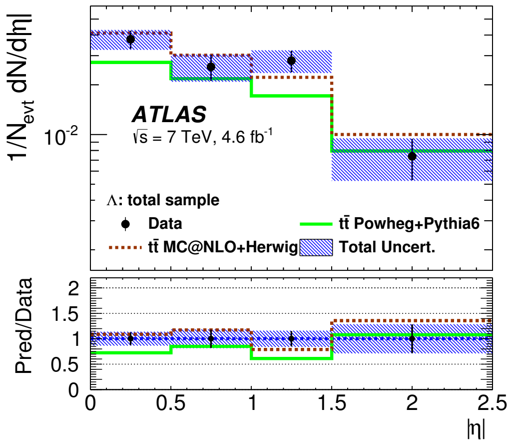
Figure 5.23: Kinematic characteristics for Λ production not associated with jets, for data and detector level MC simulated with the Powheg+Pythia6 and MC@NLO+Herwig generators. Total uncertainties are represented by the shaded area.



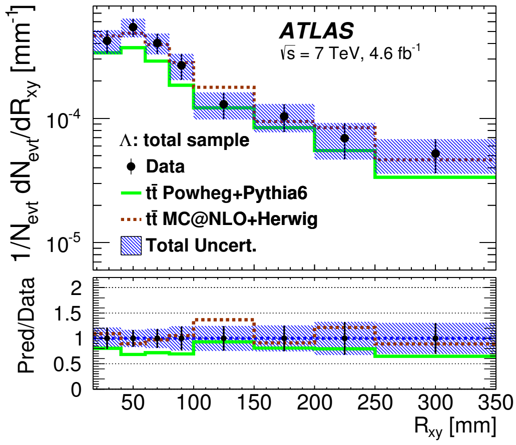
(a) p_{T} distribution



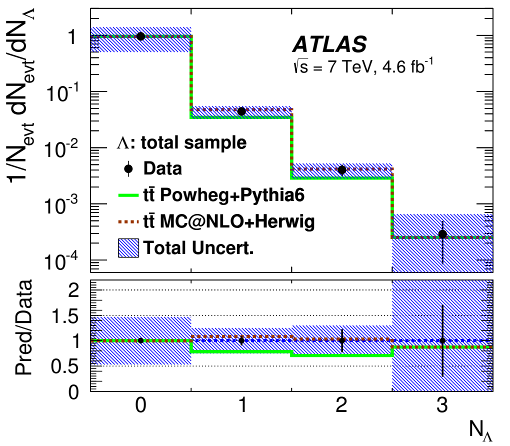
(b) Energy distribution



(c) $|\eta|$ distribution



(d) Transverse flight distance distribution



(e) Multiplicity distribution

Figure 5.24: Kinematic characteristics for the total Λ production, for data and detector level MC simulated with the POWHEG+PYTHIA6 and MC@NLO+HERWIG generators. Total uncertainties are represented by the shaded area.

5.6 Unfolding to the particle level

In order to take into account detector effects, the data are unfolded to the particle level. This allows a direct comparison with theoretical calculations as well as with measurements from other experiments. For kinematic quantities such as the transverse momentum and pseudorapidity, for which migrations are negligible, this is done by computing the reconstruction efficiencies on a bin-by-bin basis (as also in Refs. [75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 94, 95]). For the multiplicity distributions, however, a Bayesian unfolding procedure is applied because the bin-to-bin migrations are relevant.

5.6.1 Efficiency correction as a function of p_{T}, $|\eta|$, energy and energy fraction

The reconstruction efficiencies (ϵ) are calculated by dividing on a bin-by-bin basis each of the distributions (p_{T}, $|\eta|$, energy and energy fraction) at reconstructed and particle levels for each of the three classes considered:

$$\epsilon_i = \frac{\frac{1}{N_{\mathrm{evt}}^{\mathrm{reco}}}\frac{\mathrm{d}N_{\mathrm{K},\Lambda}^{\mathrm{reco}}}{\mathrm{d}x_{\mathrm{i}}}}{\frac{1}{N_{\mathrm{evt}}^{\mathrm{particle}}}\frac{\mathrm{d}N_{\mathrm{K},\Lambda}^{\mathrm{particle}}}{\mathrm{d}x_{\mathrm{i}}}} \tag{5.1}$$

where x_{i} stands for the i-th bin in the variable x, which denotes any of the kinematic variables mentioned above. They are shown in Figures 5.25, 5.26 and 5.27, for each of the classes discussed so far, as well as for the total sample. For neutral strange particles embedded in b-tagged jets, this efficiency correction also includes the b-tagging efficiency.

The particle level distributions are obtained using generator level trigger leptons, i.e. leptons coming from W decays, jets and neutral strange particles (K_{S}^0's and Λ's) in those events fulfilling the selection criteria at reconstructed level. These generator level jets are built using all particles with a lifetime above 10^{-11}s, excluding muons and neutrinos. The kinematical criteria for jets at particle level are the same as the reconstructed level ones, namely $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. The b-jets are defined at the particle level, as those containing a B-hadron, with $p_{\mathrm{T}} > 5$ GeV and $\Delta R < 0.3$ from the jet axis. Particle level K_{S}^0's and Λ's, including their neutral particle decay modes, are required to have status code 1 and to lie within $|\eta| < 2.5$ with an energy $E > 1$ GeV, as no K_{S}^0 candidates are reconstructed below that energy at detector level. Neutral strange particles coming from secondary interactions with the detector material are removed by demanding a barcode smaller than 200000. Similarly than for reconstructed level, the K_{S}^0's (Λ's) at particle level, which fulfill these conditions, are then separated in three classes: inside a b-jet, inside a non b-jet, and outside any jet, depending on whether they are closer than $\Delta R = 0.4$ to a b-jet or non b-jet at particle level. MC studies show that migrations between classes when going from detector to particle level are generally smaller than 5%. For example, K_S^0 candidates which are not associated with any jet at detector level, have a 1% (3%) probability to be classified as embedded in a b-jet (non-b-jet) at particle level. A notable exception is that of K_S^0 candidates inside non-b-tagged jets at detector level, which have a 32% probability to be classified as embedded in a b-jet at particle level. This is due to the b-tagging efficiency, which is included in the reconstruction efficiency as defined above.

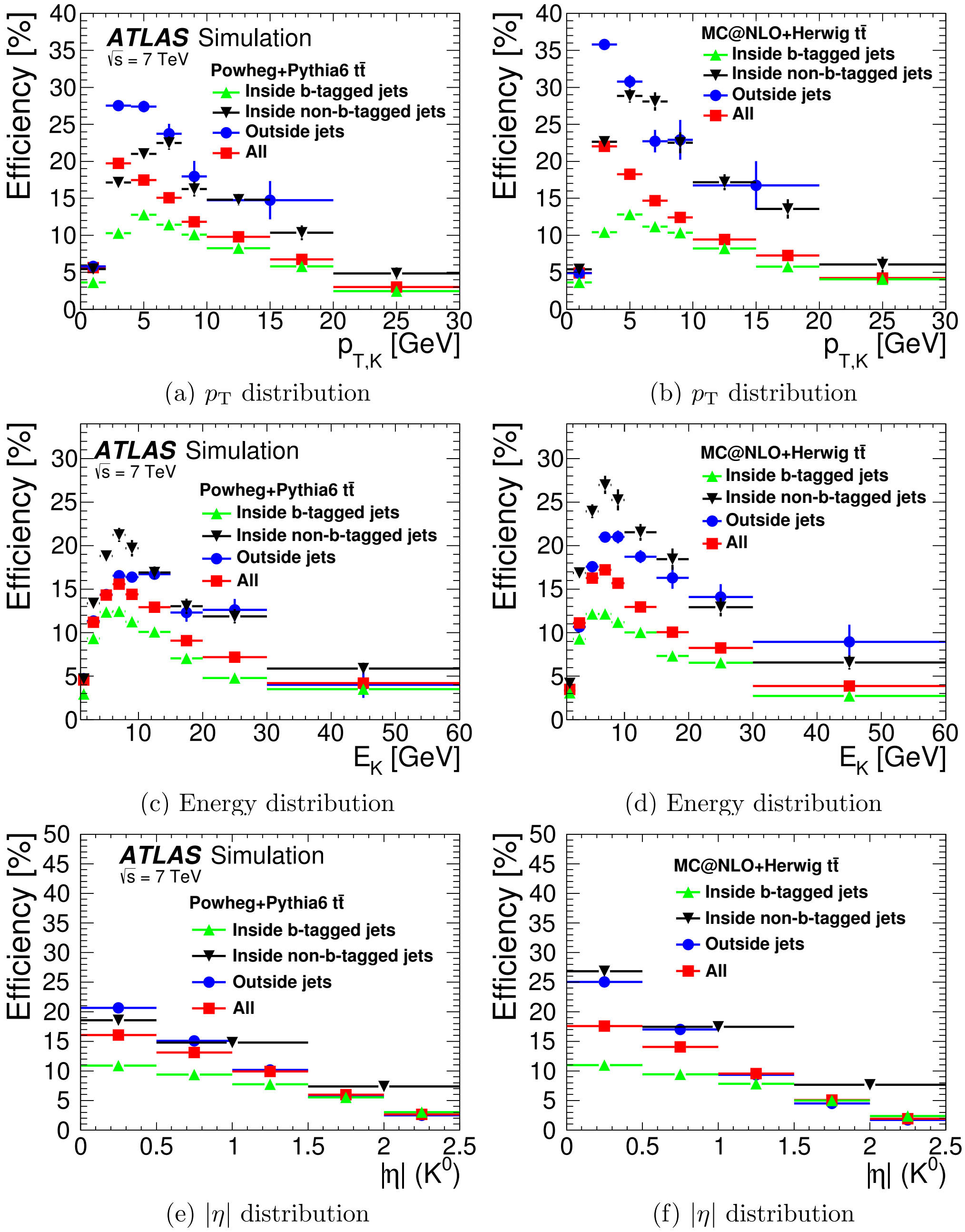
The contribution of non-fiducial events, i.e. events which pass the detector-level selection but are not present at the particle level, introduces a small bias which is taken into account as a systematic uncertainty. More details are given in Section 5.7.

The reconstructed distributions for K_S^0 (Λ) particles in the data sample, are corrected with a weight given by $1/\epsilon_i$, depending on their class. The baseline MC sample used for this purpose is the POWHEG+PYTHIA6 generator. It is to be noticed that since the MC simulation does not include the pile up at the particle level, this efficiency calculation has the additional outcome of correcting for the pile-up effects present at the reconstructed level. This correction would be as accurate as the MC description of the pile-up at reconstructed level, and the possible mismodelling is taken into account as a systematic uncertainty as will it be explained in Section 5.7.

For completeness in Figures 5.25, 5.26 and 5.27 the efficiency calculation using MC@NLO+HERWIG is also shown. Both sets of efficiencies exhibit similar trends and even are comparable for the total sample but do show numerical differences within each of the three classes being considered.

Figure 5.25 shows that the reconstruction efficiency inside b-tagged jets is lower than inside non-b-tagged jets, due to the fact that the average b-tagging efficiency is 70% and the b-jet contamination in the non-b-tagged sample is around 30%. It was checked that the efficiency for K_S^0 reconstruction inside b-jets is independent of whether they are b-tagged or not at detector level. Since the efficiency mainly depends on the transverse flight distance, i.e. the p_T, the peak efficiencies for energy and $|\eta|$ are expected to be smaller as they represent an average over a given p_T spectrum. This has been studied by looking at the p_T distribution of the total K_S^0 sample for two energy bins, i.e. below and above 10 GeV, showing that K_S^0 in a given energy bin are spread in p_T.

The peak efficiency outside jets is higher than inside, and falls more sharply at the distributions tails. This can be understood as due to a greater track density inside jets which makes the K_S^0 (Λ) reconstruction more difficult and the differences in their transverse momentum spectra.. This efficiency outside jets is lower than the one reported in Ref. [86], but that was in a minimum bias sample, and with less pile-up.



(a) p_{T} distribution

(b) p_{T} distribution

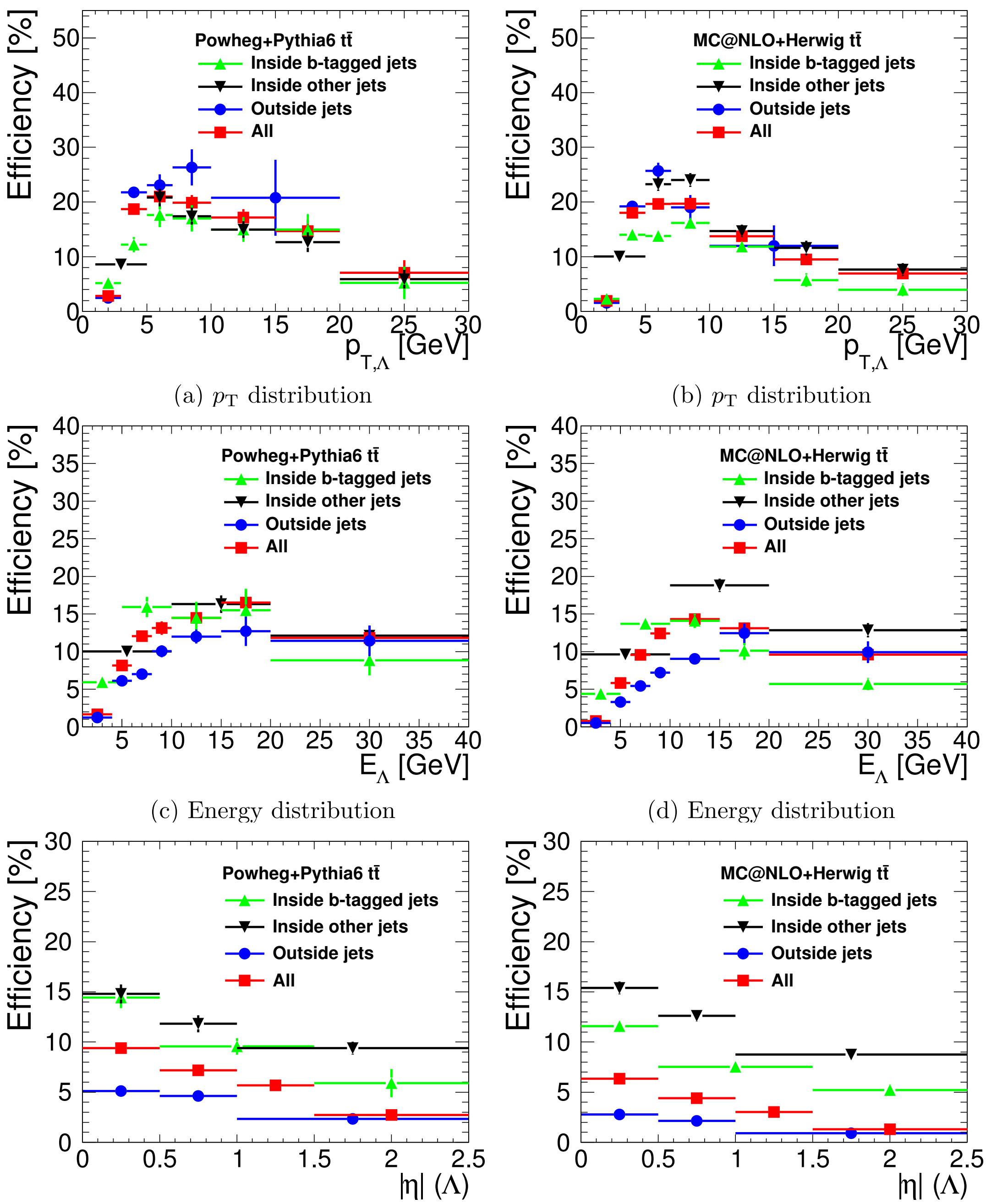
(c) Energy distribution

(d) Energy distribution

(e) $|\eta|$ distribution

(f) $|\eta|$ distribution

Figure 5.25: The K^0_{S} reconstruction efficiency as a function of p_{T}, $|\eta|$ and energy for POWHEG+PYTHIA6 (left) and MC@NLO+HERWIG (right) MC simulations. Four classes are considered, for K^0_{S} inside b-tagged jets (green), inside non b-tagged jets (black), outside any jet (blue), and the total sample (red).



(a) p_{T} distribution

(b) p_{T} distribution

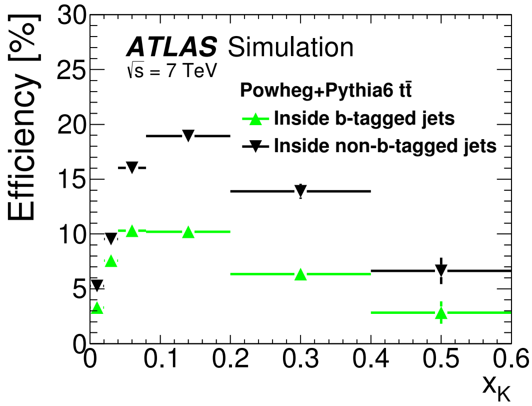
(c) Energy distribution

(d) Energy distribution

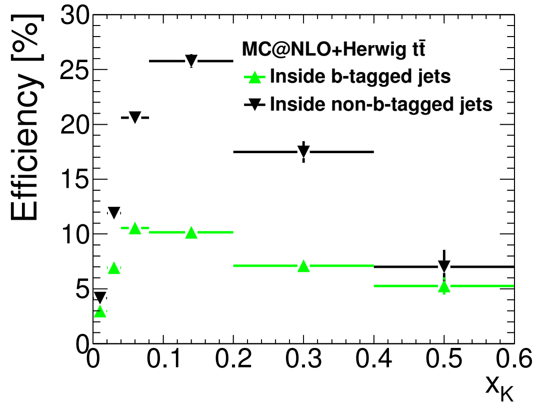
(e) $|\eta|$ distribution

(f) $|\eta|$ distribution

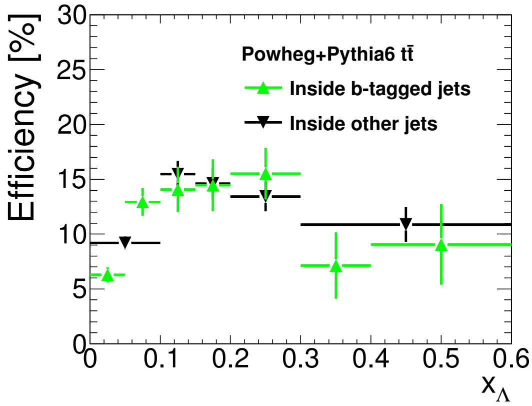
Figure 5.26: The Λ reconstruction efficiency as a function of p_{T}, $|\eta|$ and energy for POWHEG+PYTHIA6 (left), and MC@NLO+HERWIG (right) MC simulations. Four classes are considered, for K^0_{S} inside b-tagged jets (green), inside non b-tagged jets (black), outside any jet (blue), and the total sample (red).



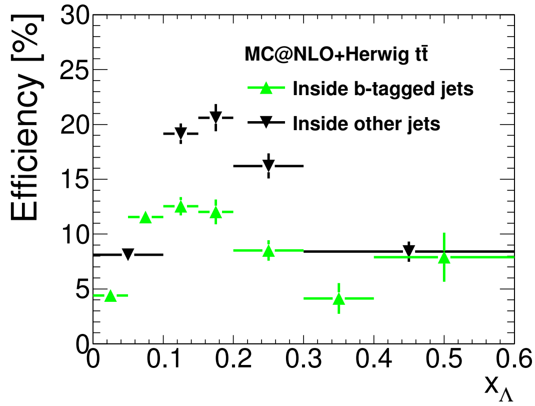
(a) Energy fraction distribution



(b) Energy fraction distribution



(c) Energy fraction distribution

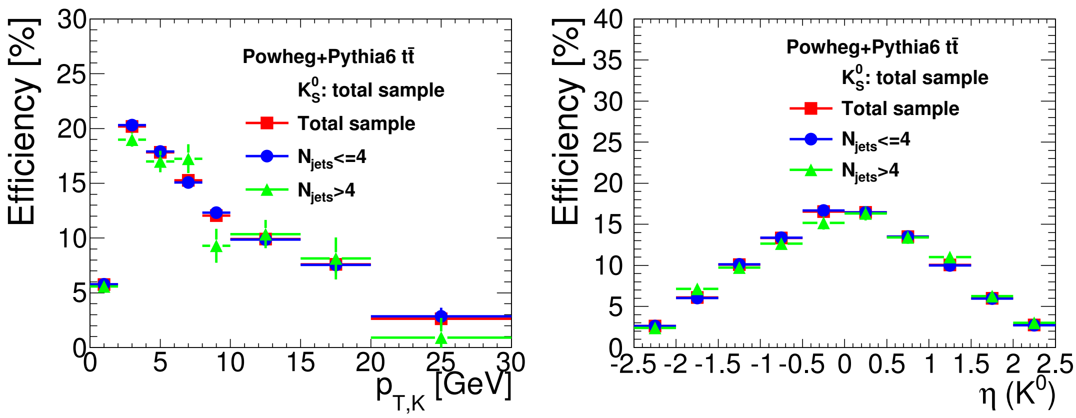


(d) Energy fraction distribution

Figure 5.27: The reconstruction efficiency for the K^0_S (Λ) as a function of the energy fraction for POWHEG+PYTHIA6 (left) and MC@NLO+HERWIG (right) MC simulations. Two classes are considered, for K^0_S (Λ) inside b-tagged jets (green) and inside non b-tagged jets (black).

5.6.1.1 Cross-check on efficiency dependence with jet multiplicity

The question has been raised whether the efficiency corrections for neutral strange particles can be affected by the jet multiplicity. In order to further investigate this issue, these efficiencies have been recalculated for two samples of events depending on whether the jet multiplicity is larger than 4 or not. In Figure 5.28 the total K^0_S efficiency corrections for which statistics are best, are shown as a function of p_T and $|\eta|$ for these two samples as well as for the total sample. These effiencies agree within statistical uncertainties. This could be expected since each additional jet with radius R=0.4 just represents 1.5% of the total available phase space ($|\eta| < 2.5$).

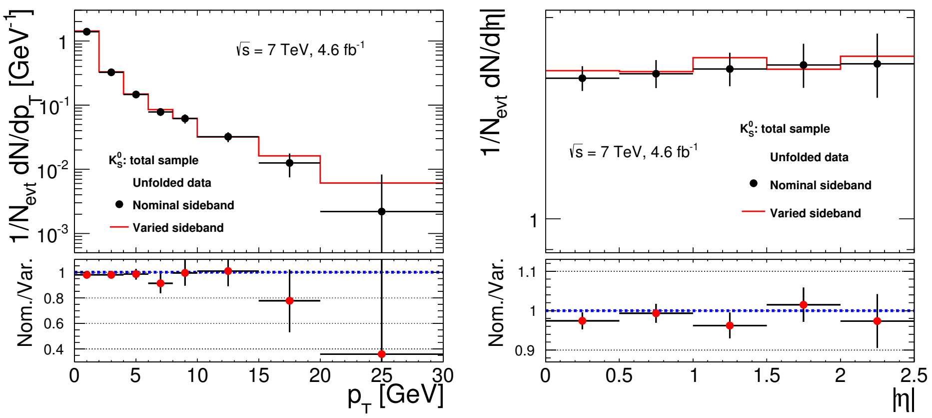


(a) p_T efficicencies (b) η efficiencies

Figure 5.28: The total K^0_S reconstruction efficiency as a function of p_T and $|\eta|$ for Powheg+Pythia6 MC simulation. Three cases are considered, events with more than 4 jets (green), equal or less than 4 jets (blue) and the total sample (red).

5.6.1.2 Dependence on the signal and sideband region definitions

The unfolded K^0_S p_T and η distributions with the extended signal and sidebands definitions discussed in Section 5.4.2.1 are compared in Figure 5.29 with the nominal unfolded distributions. No systematics effects are observed.



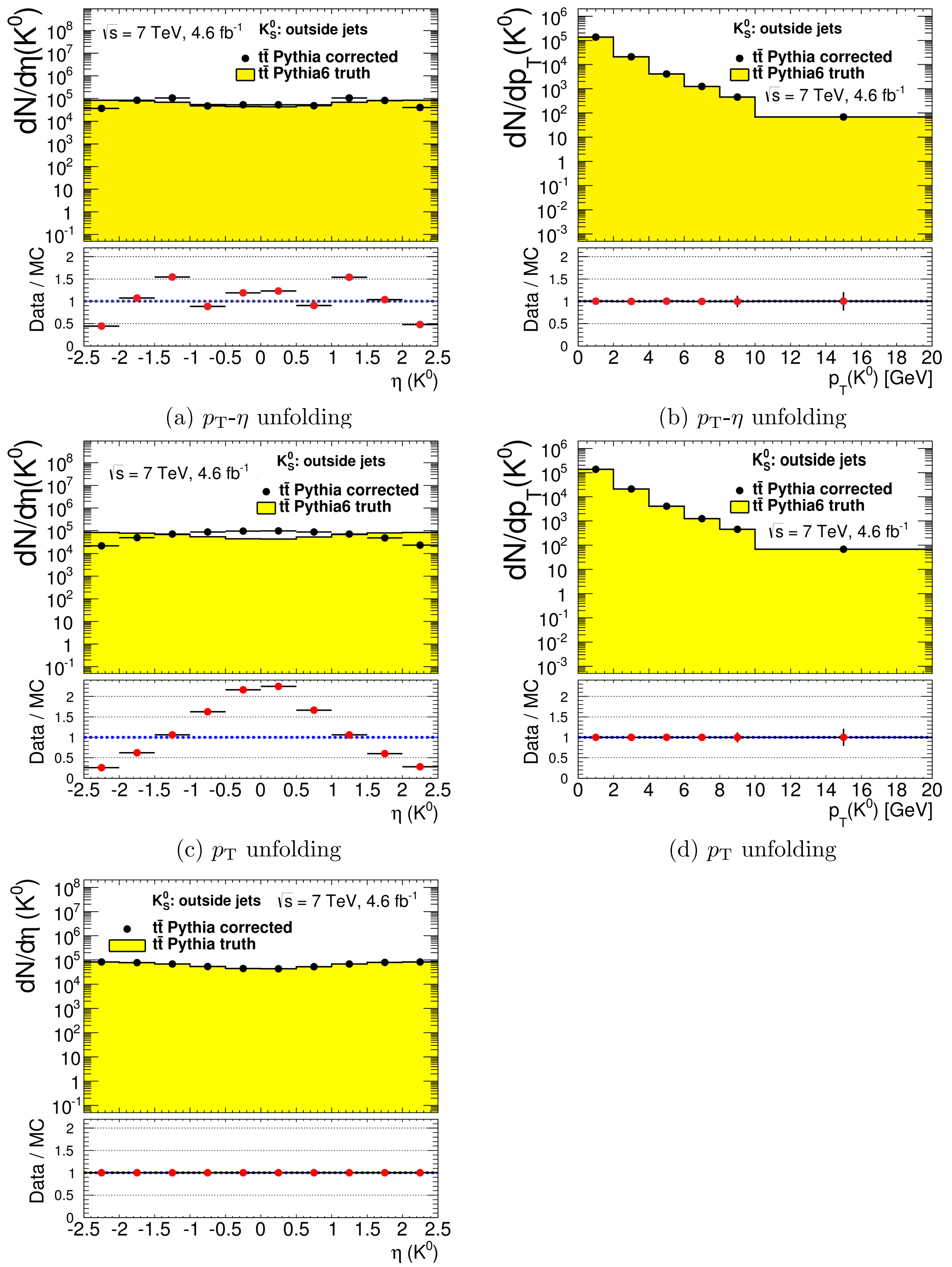
(a) p_{T} efficicencies (b) η efficiencies

Figure 5.29: Comparisons of the total K^0_{S} unfolded p_{T} and $|\eta|$ distributions using the nominal and extended signal and sideband region definitions.

5.6.1.3 Closure tests

As a closure test of the unfolding procedure for the neutral strange particle's transverse momentum and pseudorapidity, a comparison between the full statistics POWHEG+PYTHIA6 sample at particle level and the reco level unfolded with its own efficiencies (1-D) is performed, which have to trivially agree with each other if everything has been properly calculated.
In addition, a 2-D unfolding based on a p_{T}-η binning was also tried, although due to the low statistics only 2 bins in η were possible, namely $|\eta| < 1.1$ and $1.1 < |\eta| < 2.5$. This is illustrated in Figure 5.30 which shows a comparison of the two unfolding methods (1D and 2 D) as applied on the p_{T} and η distributions. The following comments are in order:

- the closure test is successful on the p_{T} distribution, for both methods 1D and 2D, since the binning is common for both unfoldings.
- the closure test on the η distribution is also trivially succesful for the 1D unfolding using the η efficiencies, while the 2D unfolding using the p_{T}-η efficiencies fails due to the fact that in the latter only two bins in η were possible as mentioned above.
- the closure test on say η fails when applied using the p_{T} efficiencies



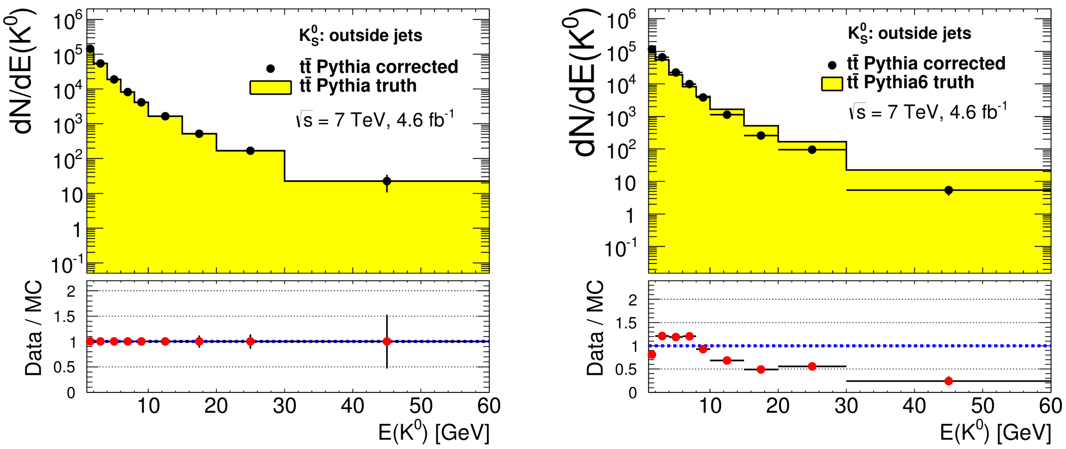
(a) p_T-η unfolding (b) p_T-η unfolding

(c) p_T unfolding (d) p_T unfolding

(e) η unfolding

Figure 5.30: Closure test for the unfolding of p_T (right) and η (left) distributions for Powheg+Pythia6 MC reco K^0_S's not associated with jets. Efficiencies are calculated as a funtion of p_T-η (top), p_T (middle) and η (bottom). The nominal unfolding procedure are the d) and e).

One could be tempted to think that the efficiency for reconstructing neutral strange particles would only depend on transverse momentum and pseudorapidity since these two variables determine the track reconstruction goodness. This is not indeed generally so, as for example for the energy of the neutral strange particles. Figure 5.31 shows that the closure test for this distribution is not completely successful when applied using $p_{\mathrm{T}}(-\eta)$ efficiencies.
This leads to the necessity of either a 1-D direct unfolding for each distribution, or an $N-D$ unfolding, where N is the number of distributions considered in the analysis. This is in general precluded by statistics. This is also the method followed by the ATLAS Minimum Bias analysis [86], where in spite of their far better statistics a 1D unfolding is preferred over a 2D.

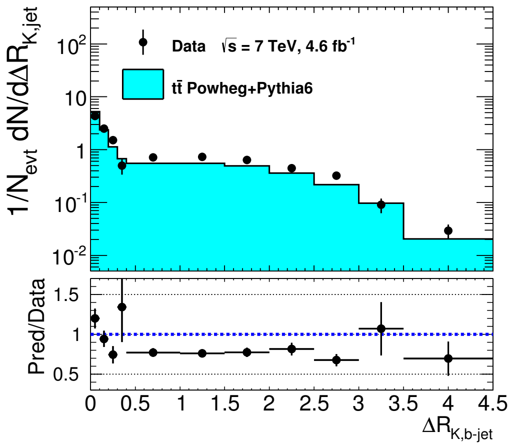


(a) E unfolding

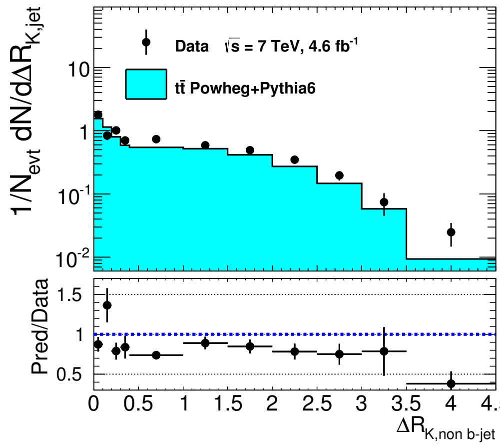
(b) p_{T} unfolding

Figure 5.31: Closure test for the unfolding of the energy distribution for POWHEG+PYTHIA6 MC reco K^0_{S}'s not associated with jets. Efficiencies are calculated as a funtion of energy (left) and p_{T} (right). The nominal unfolding procedure is the one on the left.

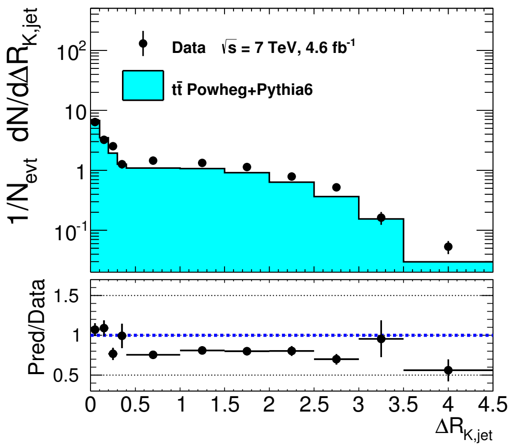
Nevertheless, there are variables which are strongly correlated with p_{T} and η for which a closure test with $p_{\mathrm{T}}(-\eta)$ efficiencies should work successfully. One such example is the angular separation (ΔR) of each K^0_{S} to the nearest jet. This has been shown at reconstructed level in Section 5.4.3, see Figure 5.12. The directly unfolded distributions are presented in Figure 5.32 below along with a comparison to MC expectations at particle level. They show similar features to those discussed at reco level in Section 5.4.3.



(a) ΔR K^0_S to b-tagged jet



(b) ΔR K^0_S to non b-tagged jet



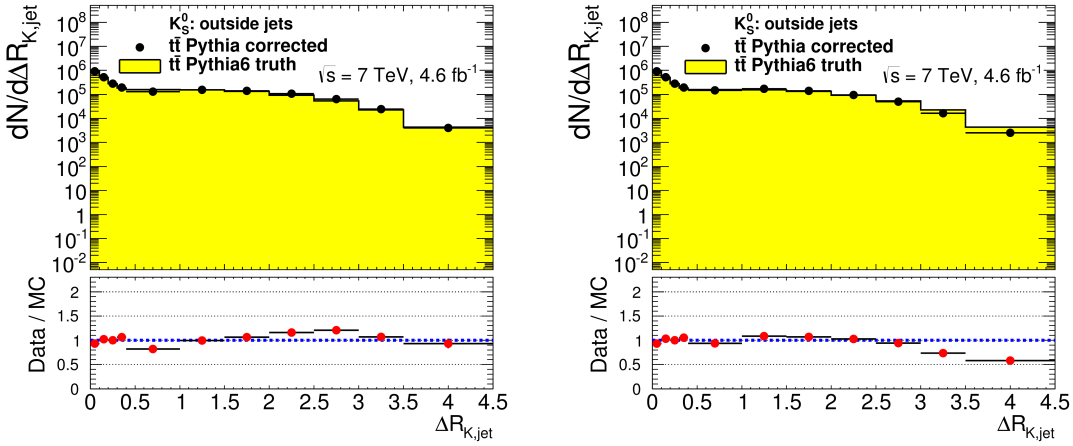
(c) ΔR K^0_S to any jet

Figure 5.32: The K^0_S angular separation (ΔR) to the nearest jet, depending on whether the jet is b-tagged (top left) or not (top right), as well as for all jets (bottom left), in unfolded data as compared with POWHEG+PYTHIA6 MC at particle level.

Figure 5.33 shows the results of a closure test for the ΔR distribution. The POWHEG+PYTHIA6 MC at particle level is compared to the corresponding distribution at reco level unfolded using a 1-D (p_T) and 2-D (p_T-η) efficiencies. Though some residual effects are still presented, the results of the closure test are reasonable and they can be summarised as follows:

- for the part inside jets ($\Delta R < 0.4$), both methods give similar results
- for $0.4 < \Delta R < 3$, the 1-D unfolding gives ratios nearer to unity
- for the tail ($\Delta R > 3$), the 2-D unfolding performs better, this could be explained since the information from the η efficiencies, in particular at large

η values, becomes more relevant when the K_S^0 are far away from any jet, which are mostly centrally produced



(a) p_{T}-η unfolding

(b) p_{T} unfolding

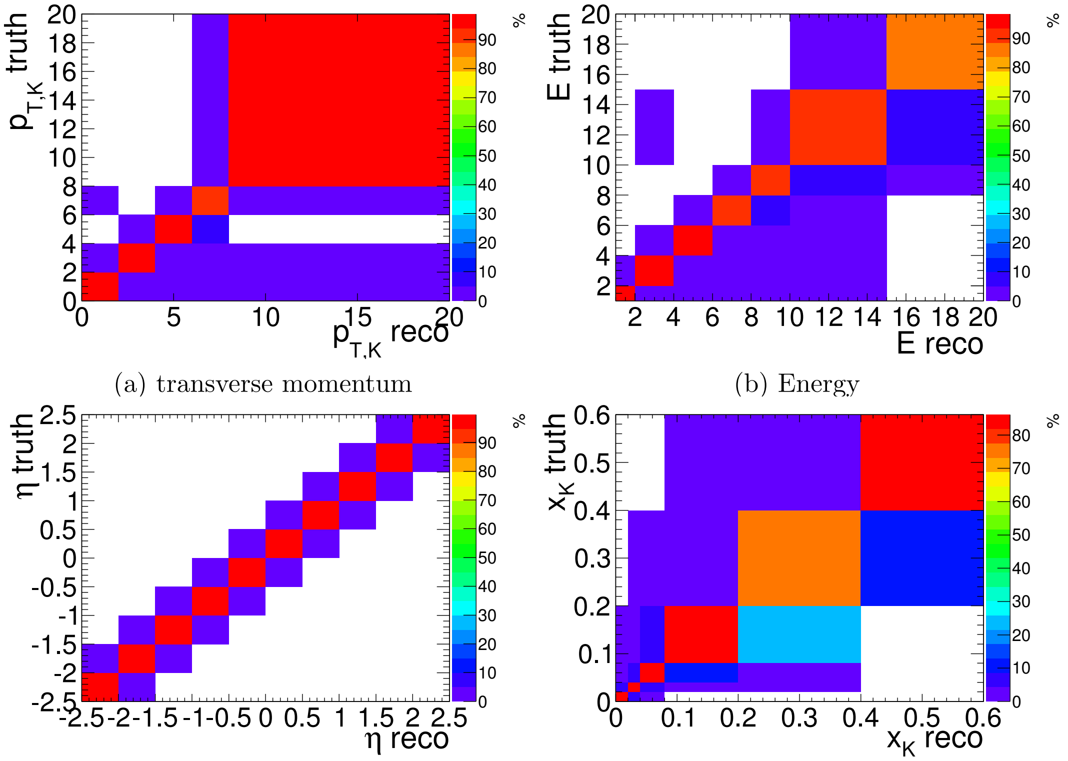
Figure 5.33: Closure test for the unfolding ofK_S^0 angular separation (ΔR) to the nearest jet for POWHEG+PYTHIA6 MC reco. Efficiencies are calculated as a funtion of p_{T}-η (left) and p_{T} (right).

5.6.1.4 Cross-check: efficiency vs. Bayes

The unfolding based on the efficiency calculations disscused in this subsection relies on the assumption that the neutral strange particles, once reconstructed, are very well measured. As an alternative a Bayesian unfolding [137] as implemented in the RooUnfold programme [138] was tried. The method numerically calculates the inverse of the migration matrices for each of the distributions under study. The POWHEG+PYTHIA6 MC sample is used to determine these migration matrices by matching detector-level and particle-level K_S^0 that are in the same class and have an angular separation $\Delta R < 0.01$. The resulting matrices exhibit a very pronounced diagonal correlation. This is illustrated in Fig 5.34, which shows the K_S^0 not associated with jets, p_{T}, energy and pseudorapidity transfer matrices, as well as the energy fraction (x_K) for K_S^0 associated with non b-tagged jets. The number of iterations is chosen such that the residual bias, evaluated through a closure test as discussed in Section 5.7, is within a tolerance of 1% for the statistically significant bins as in Ref. [139].
The results ($N^{\mathrm{part},i}$) of this Bayesian unfolding procedure are given by:

$$N^{\mathrm{part},i} = \sum_j N_{\mathrm{detec},j} \times \epsilon_{\mathrm{detec},j} A^{\mathrm{part},i}_{\mathrm{detec},j} / \epsilon^{\mathrm{part},i}$$

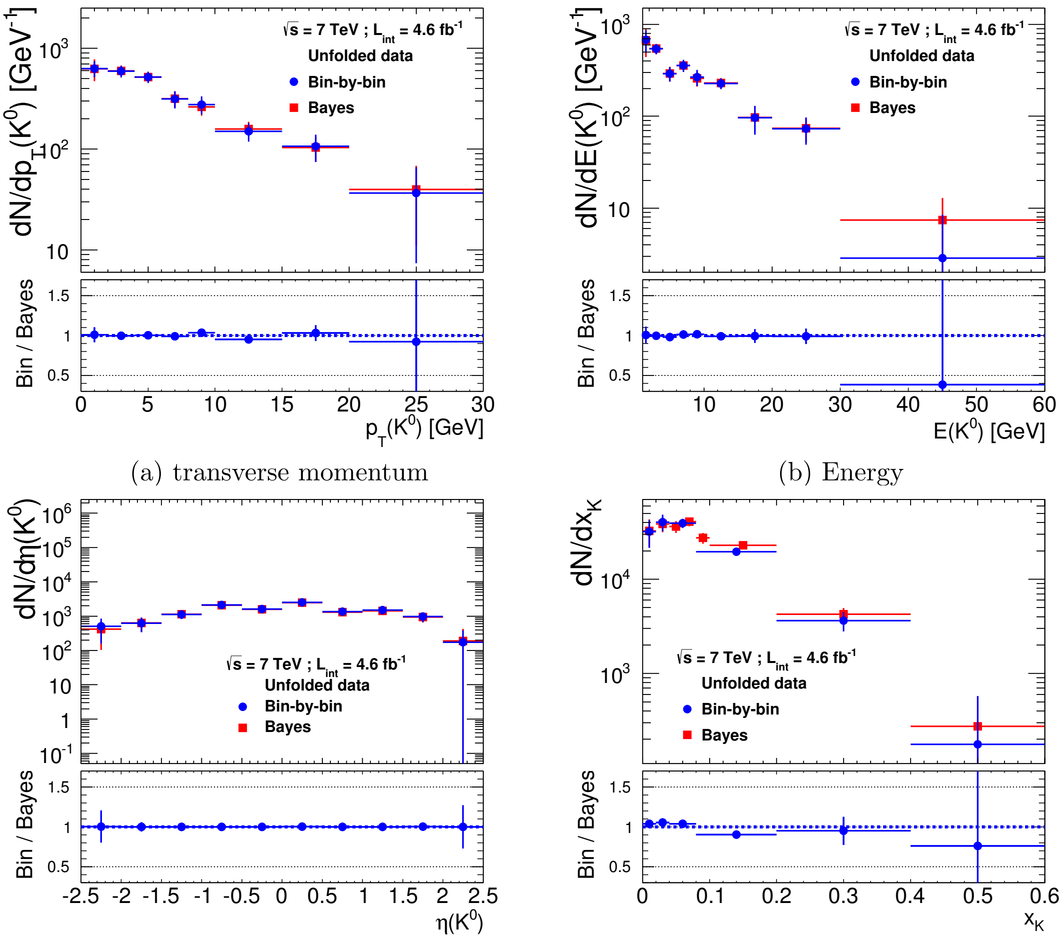
where i and j are the particle and detector level bin indices, respectively, $N_{\text{detec},j}$ is the data result at detector level, $A^{\text{part},i}_{\text{detec},j}$ is the migration matrix refined through iteration as explained above, and $\epsilon_{\text{detec},j}$ and $\epsilon^{\text{part},i}$ the matching efficiencies for detector and particle level neutral strange particles.



(a) transverse momentum
(b) Energy
(c) transverse momentum
(d) Energy

Figure 5.34: Transfer matrices for K^0_S, not associated with jets, p_T (top left) and energy (top right), η (bottom left) and x_K (bottom right) for K^0_S associated with non b-tagged jets. These matrices are obtained with the POWHEG+PYTHIA6 generator.

The results obtained with the Bayesian unfolding are found to be in very good agreement within statistics with the bin-by-bin unfolding used before, which by the way was used by the TASSO [81] and ZEUS [94]collaborations. This is even so for the energy fraction (x_K). This is illustrated in the Figures 5.35 and 5.36 below for K^0_S inside b-jets and not associated with jets, respectively, where the ratio between the unfolded distributions using bin-by-bin and Bayes are presented. The statistical uncertainty in the ratios have been calculated using the Bootstrap method in order to take into account possible correlations.



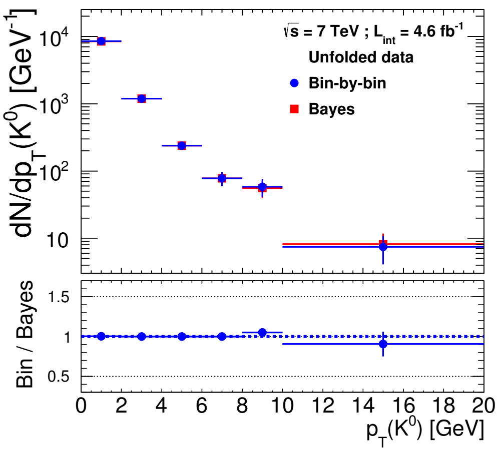
(a) transverse momentum

(b) Energy

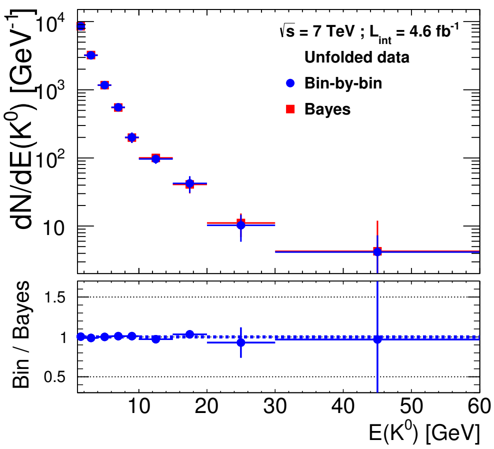
(c) pseudorapidity

(d) Energy fraction

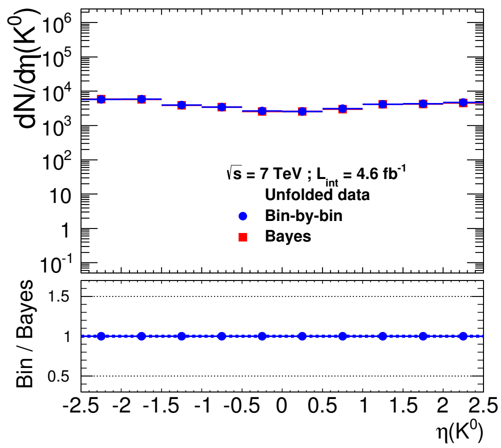
Figure 5.35: Comparison of unfolded distributions by bin-by-bin (blue) and Bayes (red) methods for: (a)$p_{\rm T}$, (b) energy, (c) η and (d) x_K for $K^0_{\rm S}$ associated with b-tagged jets.



(a) transverse momentum



(b) Energy

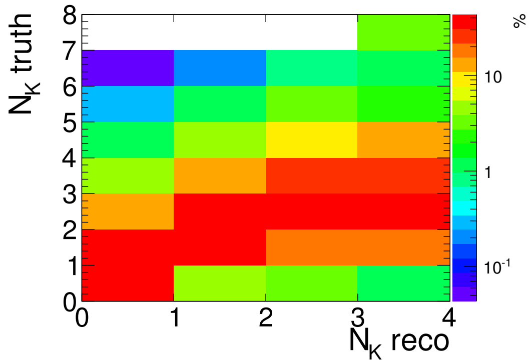


(c) pseudorapidity

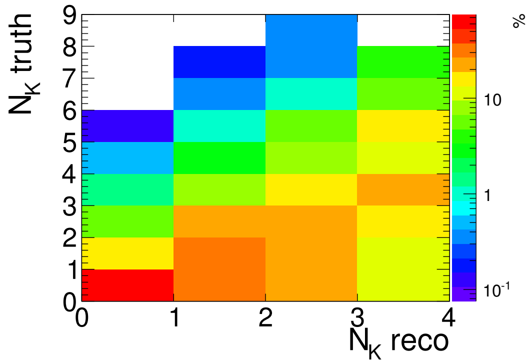
Figure 5.36: Comparison of unfolded distributions by bin-by-bin (blue) and Bayes (red) methods for: (a) p_{T}, (b) energy and (c) η for K^0_{S} not associated with jets.

5.6.2 Bayesian unfolding of the multiplicity

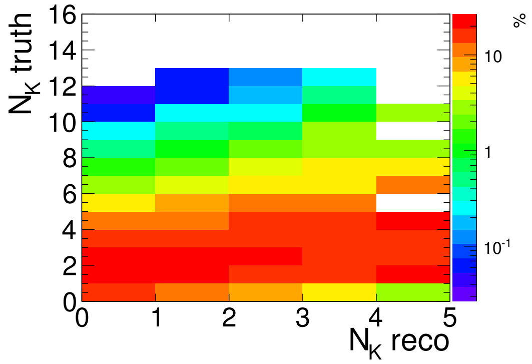
To correct the multiplicity distributions, the previous method based on a bin-by-bin efficiency correction, cannot be applied due to the big migrations between detector and particle levels. Transfer matrices are filled on an event-by-event basis with the multiplicities at detector and particle level from the POWHEG+PYTHIA6 sample. They are shown in Figure 5.37 and Figure 5.38 for K^0_{S} and Λ production respectively. This procedure allows calculation the conditional probabilities that a given reconstructed multiplicity N_{reco} comes from a given particle multiplicity N_{part}. Then the detector level data multiplicity distributions can be unfolded by applying Bayes theorem.



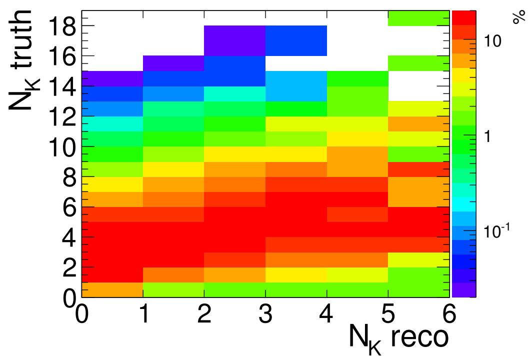
(a) b-jets



(b) Non b-jets

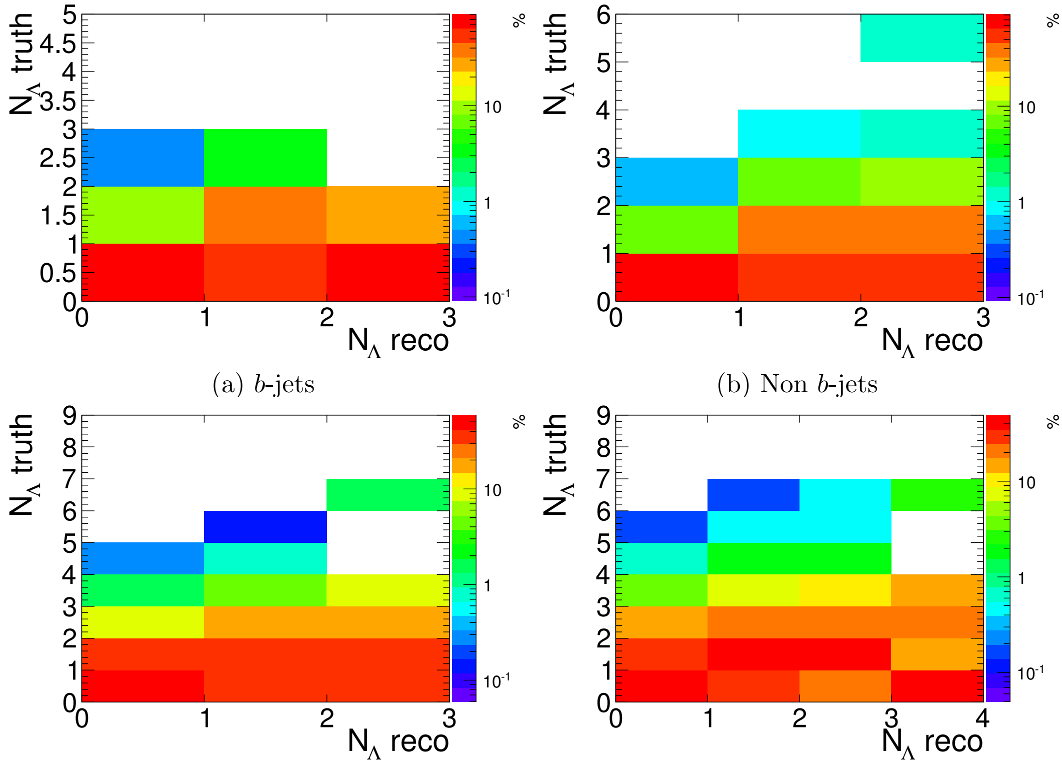


(c) Outside jets



(d) all sample

Figure 5.37: Transfer matrices for K_S^0 production at particle and detector level of the MC simulation, normalized to 100 % for each column. For the four classes discussed in the text.

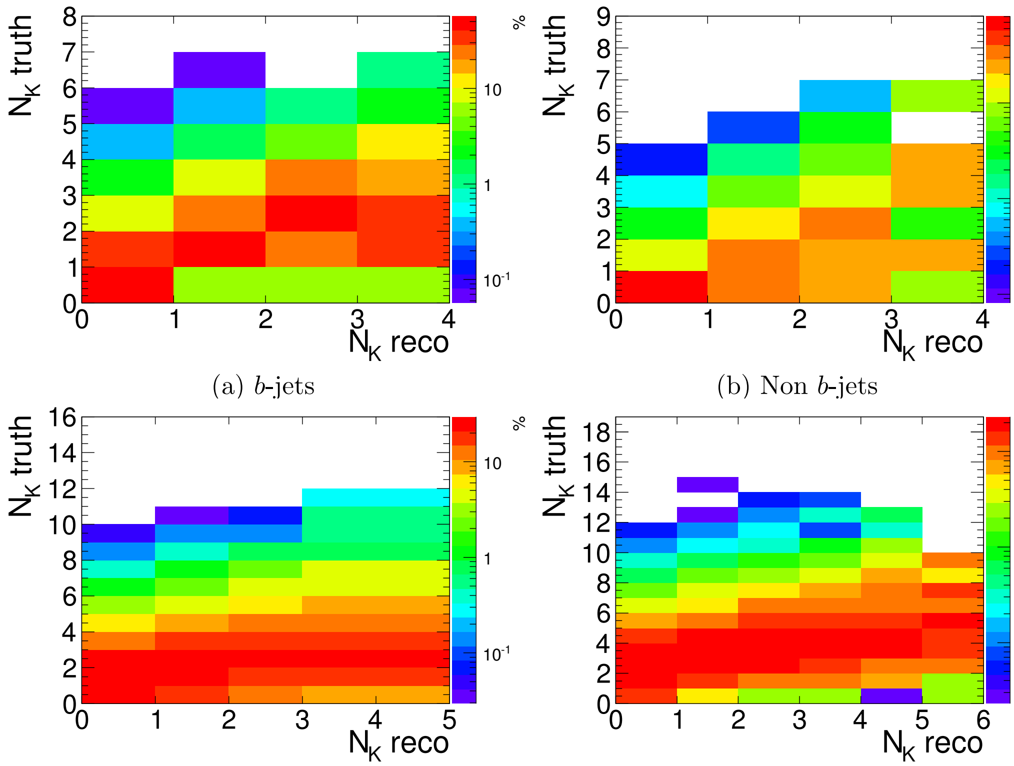


(a) b-jets (b) Non b-jets

(c) Outside jets (d) All sample

Figure 5.38: The tranfer matrices for Λ production at particle and detector level of the MC simulation, normalized to 100 % for each column. For the four classes discussed in the text.

These matrices show significant non-diagonal terms. This is expected in particular for neutral strange multiplicities non associated with jets, which are the largest, since reconstruction efficiencies are of the order of 25% and at particle level invisible decays are being considered. Therefore, one could think that the unfolding with these matrices is not trustworthy. In order to cross-check on this, another set of multiplicity transfer matrices has been calculated by considering at particle level only the visible decays ($K^0_S \to \pi^+\pi^-$). They are shown in Figure 5.39. The importance of nondiagonal terms is reduced as expected, so that this second set is chosen to perform the unfolding procedure.



(a) b-jets (b) Non b-jets

(c) Outside jets (d) all sample

Figure 5.39: Transfer matrices for K_S^0 production at particle and detector level of the MC simulation, normalized to 100 % for each column. For the four classes discussed in the text. At particle level only visible decays i.e. $K_S^0 \to \pi^+\pi^-$ have been considered here.

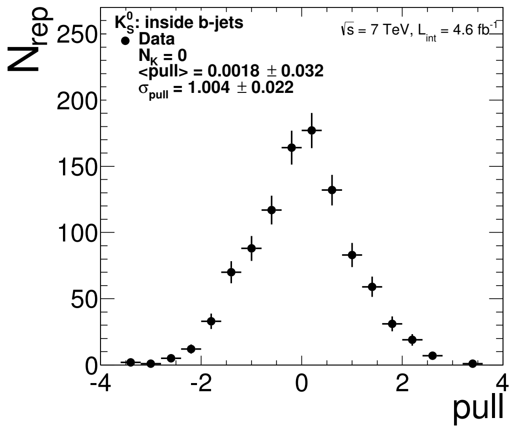
The bayesian unfolding has been done with the iterative method as implemented in the RooUnfold tool [138, 137]. This method attempts to calculate numerically the inverse of the transfer matrices for the multiplicity applying the Bayes theorem with an iterative algorithm. It uses the transfer matrices shown above and the data multiplicity at the detector level. The number of iterations has been chosen such that the residual bias, evaluated through a data-driven closure test is within a tolerance of 1% for the statistically significant bins. This is the criteria followed by the ATLAS measurement of the inclusive jet cross-sections [139].

5.6.2.1 Propagation of statistical uncertainty. The bootstrap method

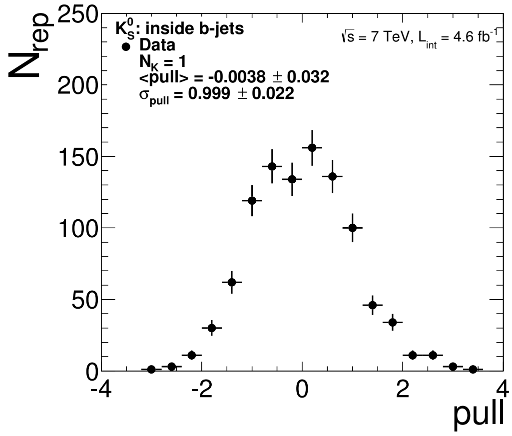
For calculating properly the propagation of the statistical uncertainties from the reconstructed data and the conditional probabilities to the unfolded multiplicity distributions the Bootstrap generator tool is used. It is based on a Poisson weight generator to create 1000 replicas of the transfer matrices, and a statistically independent generator to create another 1000 replicas of the reconstructed data multiplicity distributions. The RMS of the replicas is taken as the statistical uncertainty of the unfolded distribution.

To check that these statistical uncertainties are well behaved, i.e. Gaussian like, pulls have been calculated for each bin of the multiplicity distribution, as the ratio between the difference of a replica content and the nominal one to the satatistical error in that bin. As long as the errors are Gaussian, the pull for each bin should follow a Normal distribution i.e. a Gaussian centered at zero with standard deviation equal to one.

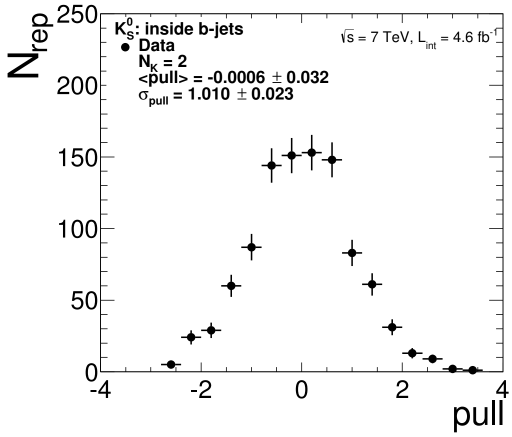
These pull distributions for the replicas of each bin are shown in Figures 5.40 to 5.43. They are found to conform to the expectations discussed above. In fact, the RMS of these distributions deviate from unity at most by 1-2 %, while the mean values deviate from zero an order of magnitude less than their statistical uncertainty of 0.03.



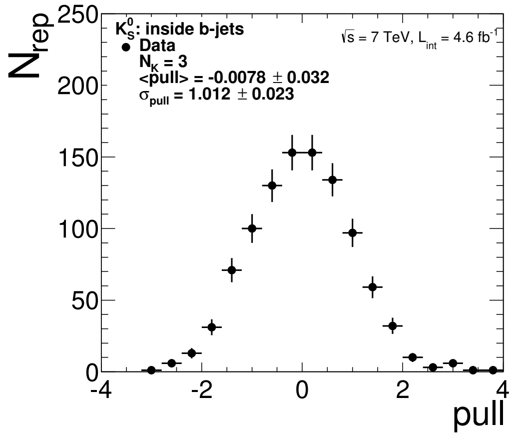
(a) $N_K = 0$



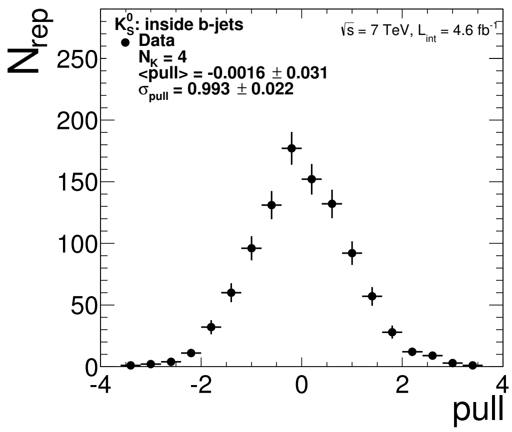
(b) $N_K = 1$



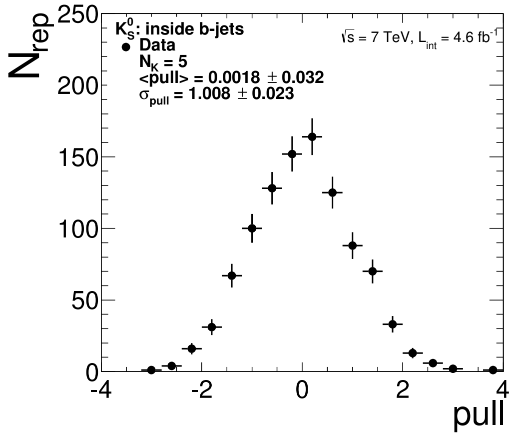
(c) $N_K = 2$



(d) $N_K = 3$

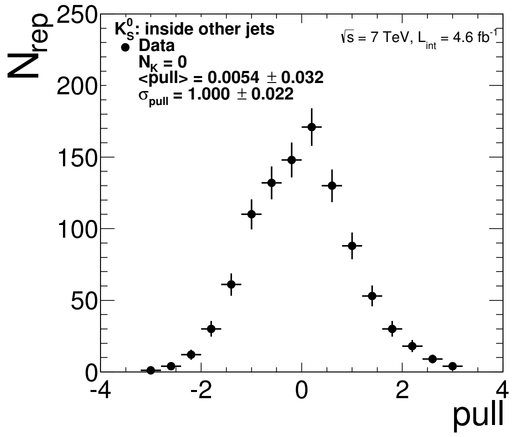


(e) $N_K = 4$

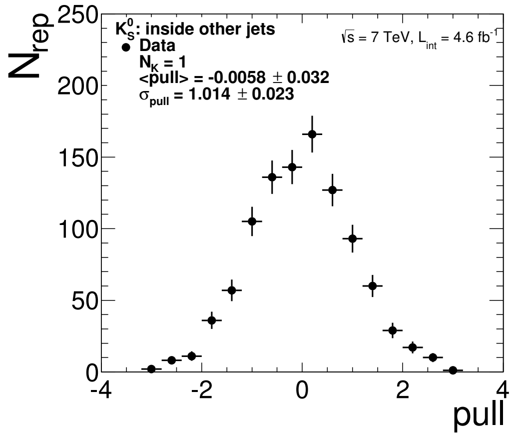


(f) $N_K = 5$

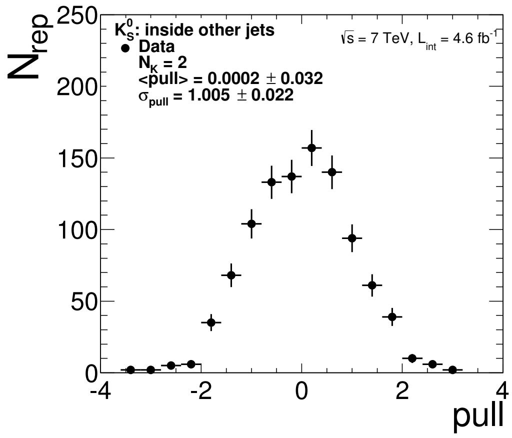
Figure 5.40: The pulls of unfolded multiplicity for K^0_{S} associated with b-jets, calculated by the bootstrap method.



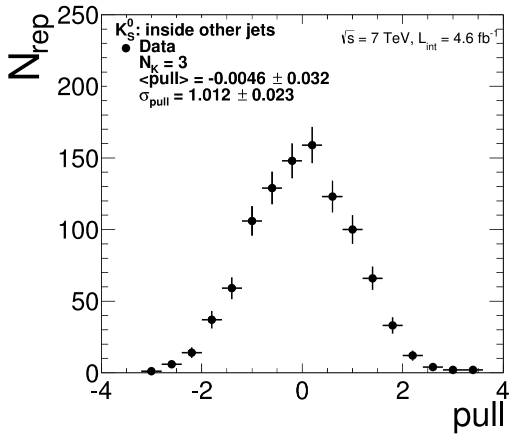
(a) $N_K = 0$



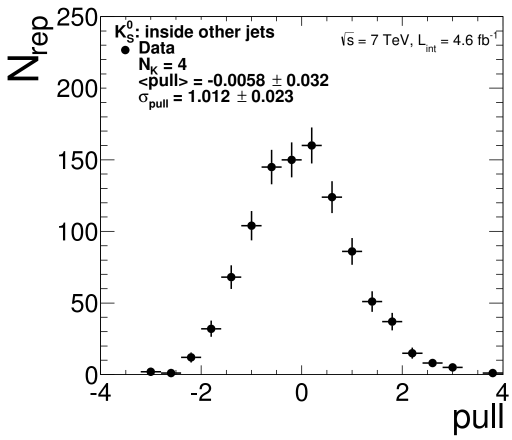
(b) $N_K = 1$



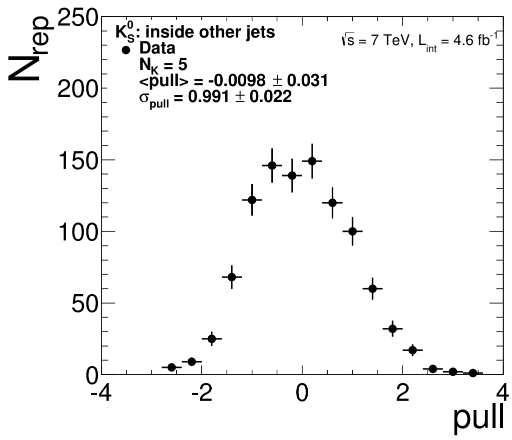
(c) $N_K = 2$



(d) $N_K = 3$

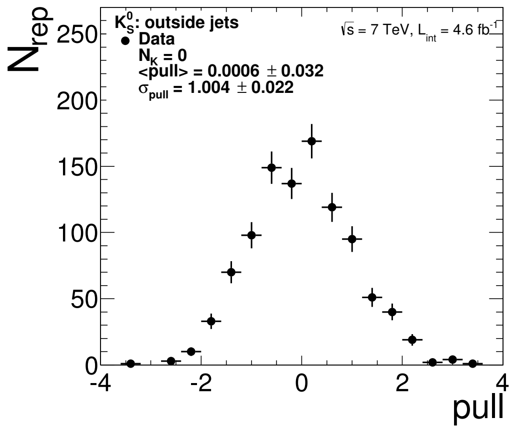


(e) $N_K = 4$

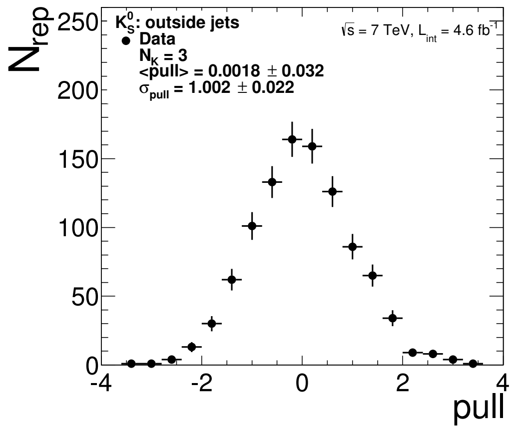


(f) $N_K = 5$

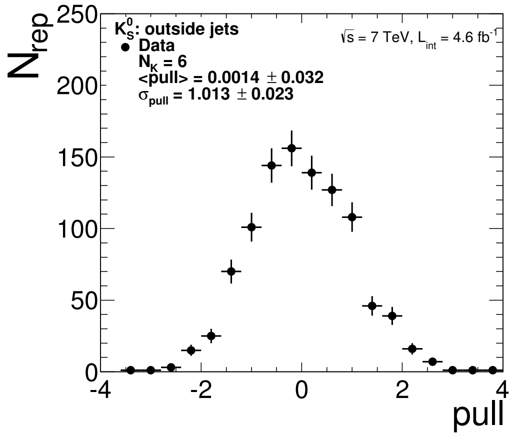
Figure 5.41: The pulls of unfolded multiplicity for K_S^0 associated with non b-jets, calculated by the bootstrap method.



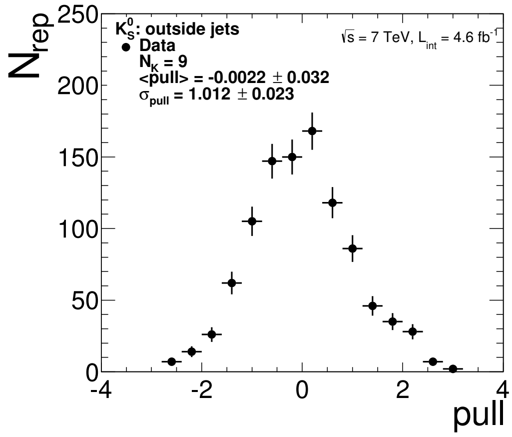
(a) $N_K = 0$



(b) $N_K = 3$

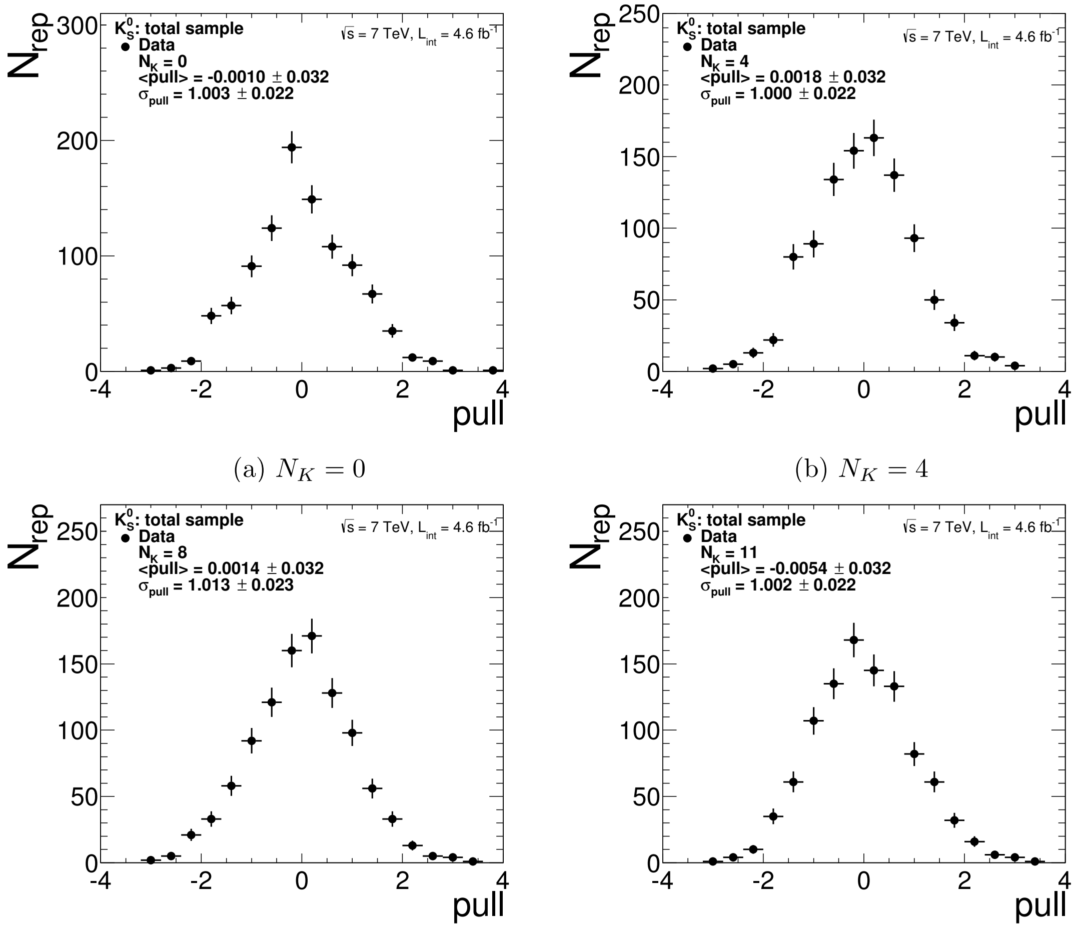


(c) $N_K = 6$



(d) $N_K = 9$

Figure 5.42: The pulls of unfolded multiplicity for K_S^0 not associated with jets, calculated by the bootstrap method. Four bins are shown, for $N_K = 0, 3, 6$ and 9.



(a) $N_K = 0$

(b) $N_K = 4$

(c) $N_K = 8$

(d) $N_K = 11$

Figure 5.43: The pulls of unfolded multiplicity for the total sample of K_S^0, calculated by the bootstrap method. Four bins are shown, for $N_K = 0, 4, 8$ and 11.

5.7 Systematic uncertainties

The main aim of this note is to study the general characteristics of neutral strange particle production in $t\bar{t}$ final states, like the average number of K^0_S's or Λ's, and not their cross-sections. Thus, many systematic uncertainty sources considered in top cross-section measurements, in particular those related to the lepton and b-jet scale factors, do not apply here.
In this Section the following systematic uncertainties will be discussed, namely:

- the systematic uncertainty related to the choice of a MC generator for the unfolding
- the systematic uncertainty related to how well the MC model for the pile-up describes the data at reconstructed level
- the systematic uncertaity related to the jet energy scale (JES) and resolution (JER)
- the b-tagging efficiency working point
- the systematic uncertainty related to the unfolding non-closure
- the systematic uncertainty related to experimental uncertainties due to tracking, GEANT version, dead material and magnetic field, which are reported in [86, 140]
- the systematic uncertainty due to out of fiducial events.

As it is shown below, the main source of uncertainty is the choice of the MC generator for the unfolding, which turns out to be of the order of 20 %.

5.7.1 MC generator uncertainties

As it was explained in the previous Section, the POWHEG+PYTHIA6 MC generator is chosen to perform the efficiency correction for the kinematical distributions and the bayesian unfolding for the multiplicities. Therefore the result of the unfolding depends to some extent on the choice of MC generator to perform it. This is so because the reconstruction efficiencies for neutral strange particles depend on their transverse spectra, and different Monte Carlos exhibit differences in these transverse momentum profiles.
This relative systematic uncertainty, $\sigma^{\mathrm{MC}}_{\mathrm{syst}}$, is estimated as the difference between the unfolded data using POWHEG+PYTHIA6 (nominal MC) and MC@NLO+HERWIG (as a comparison), see Section 5.2 for more details on the MC simulations, divided by the data unfolded by Pythia. This is expressed in the Eq. 5.2 below. This

equation is used to calculate the systematic uncertainty on a bin-by-bin basis for each distribution and each class as discussed in previous Sections.

$$\sigma^{\mathrm{MC}}_{\mathrm{syst}} = \frac{|Data^{\textsc{Pythia6}}_{\mathrm{Corrected}} - Data^{\textsc{Herwig}}_{\mathrm{Corrected}}|}{Data^{\textsc{Pythia6}}_{\mathrm{Corrected}}} \tag{5.2}$$

Here $Data^{\mathrm{Gen.}}_{\mathrm{Corrected}}$ is the data at reconstructed level divided by the recontruction efficiency calculated with the MC generator (Gen. = PYTHIA6/HERWIG):

$$Data^{\mathrm{Gen.}}_{\mathrm{Corrected}} = Data_{\mathrm{reco}} \frac{MC^{\mathrm{Gen.}}_{\mathrm{part}}}{MC^{\mathrm{Gen.}}_{\mathrm{reco}}} = \frac{Data_{\mathrm{reco}}}{\epsilon^{\mathrm{Gen}}} \tag{5.3}$$

where, the efficiency ϵ^{Gen} is defined as:

$$\epsilon^{\mathrm{Gen.}} = \frac{MC^{\mathrm{Gen.}}_{\mathrm{reco}}}{MC^{\mathrm{Gen.}}_{\mathrm{part}}} \tag{5.4}$$

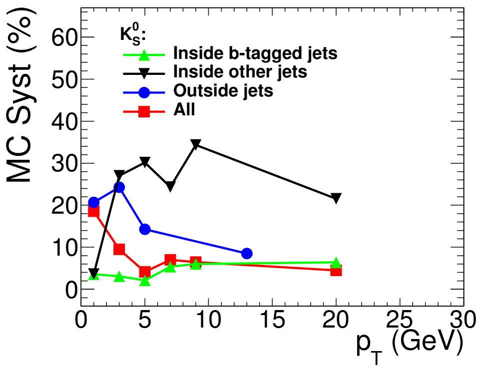
Therefore substituting 5.3 in 5.2 one obtains:

$$\sigma^{\mathrm{MC}}_{\mathrm{syst}} = \frac{Data_{\mathrm{reco}} |\frac{MC^{\textsc{Pythia6}}_{\mathrm{part}}}{MC^{\textsc{Pythia6}}_{\mathrm{reco}}} - \frac{MC^{\textsc{Herwig}}_{\mathrm{part}}}{MC^{\textsc{Herwig}}_{\mathrm{reco}}}|}{Data_{\mathrm{reco}} \frac{MC^{\textsc{Pythia6}}_{part}}{MC^{\textsc{Pythia6}}_{\mathrm{reco}}}} = |1 - \frac{\epsilon^{\textsc{Pythia6}}}{\epsilon^{\textsc{Herwig}}}| \tag{5.5}$$

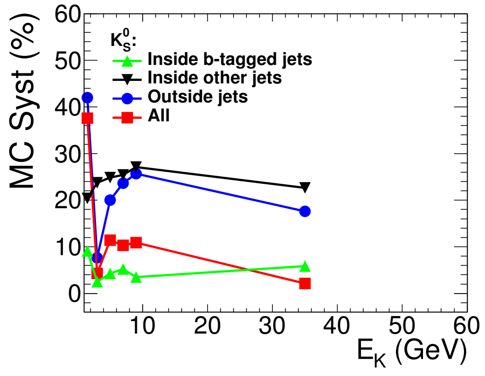
On the other hand, another standard option is to estimate this uncertainty, $\sigma^{\mathrm{MC}}_{\mathrm{syst2}}$, as the relative difference between the nominal MC POWHEG+PYTHIA6 at reconstructed level, unfolded using efficiencies from the MC@NLO+HERWIG MC generator, $MC^{\textsc{Pythia6}}_{\mathrm{Corr.(\textsc{Herwig})}}$, and the Pythia MC at particle level. In Eq. 5.6, applying the definitions of efficiency and correction as above, one obtains:

$$\begin{aligned}\sigma^{\mathrm{MC}}_{\mathrm{syst2}} &= \frac{|MC^{\textsc{Pythia6}}_{Corr.(\textsc{Herwig})} - MC^{\textsc{Pythia6}}_{\mathrm{part}}|}{MC^{\textsc{Pythia6}}_{\mathrm{part}}} = \\ &= \frac{|MC^{\textsc{Pythia6}}_{\mathrm{reco}} \frac{MC^{\textsc{Herwig})}_{\mathrm{part}}}{MC^{\textsc{Herwig})}_{\mathrm{reco}}} - MC^{\textsc{Pythia6}}_{\mathrm{part}}|}{MC^{\textsc{Pythia6}}_{\mathrm{part}}} = |\frac{\epsilon^{\textsc{Pythia6}}}{\epsilon^{\textsc{Herwig}}} - 1|\end{aligned} \tag{5.6}$$

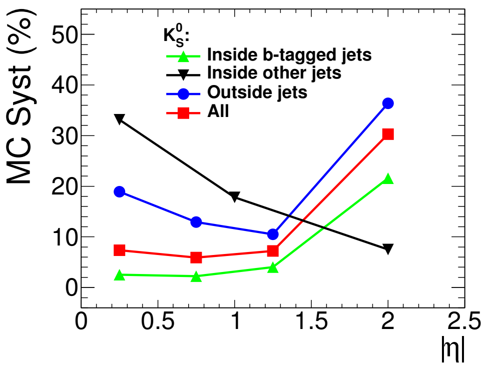
Clearly both approaches give the same results.
In spite of the fact that the statistics in the POWHEG+PYTHIA6 and the MC@NLO+HERWIG samples are far better than that in the data, when computing $\sigma^{\mathrm{MC}}_{\mathrm{syst}}$, with the binning shown in Section 5.5, one is still dominated by statistical fluctuations. To get rid of them $\sigma^{\mathrm{MC}}_{\mathrm{syst}}$ has been calculated with a new coarser binning and the statistical uncertainty has been added to it. The outcome of this procedure for every distribution can be found in appendix A.1.
In Figures 5.44 and 5.45 the values of the relative sytematic uncertainties for all the K^0_S and Λ distributions and classes are shown.



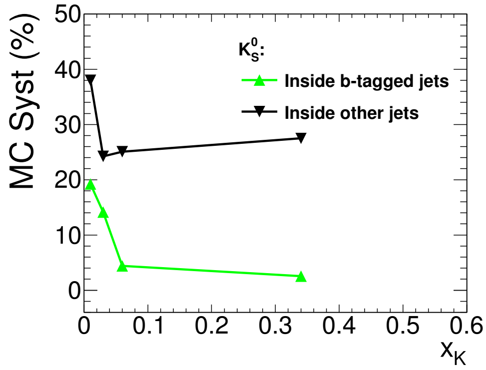
(a) p_{T} distribution



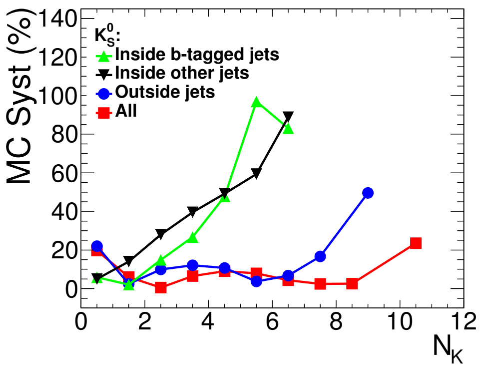
(b) Energy distribution



(c) $|\eta|$ distribution

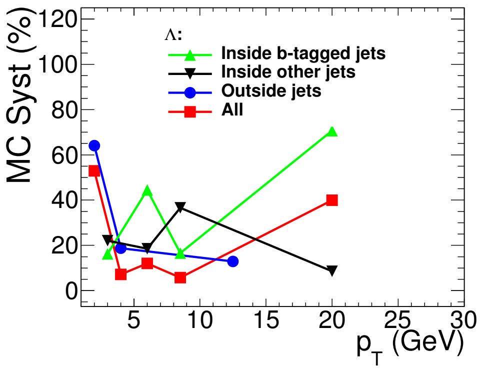


(d) Energy fraction distribution

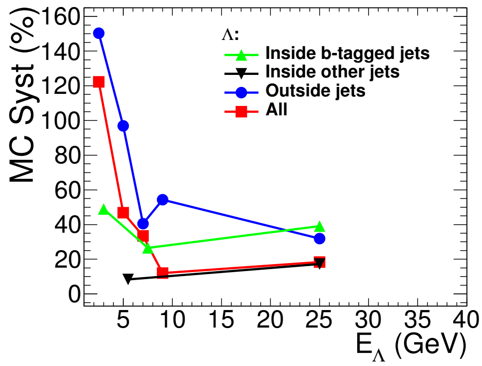


(e) Multiplicity distribution

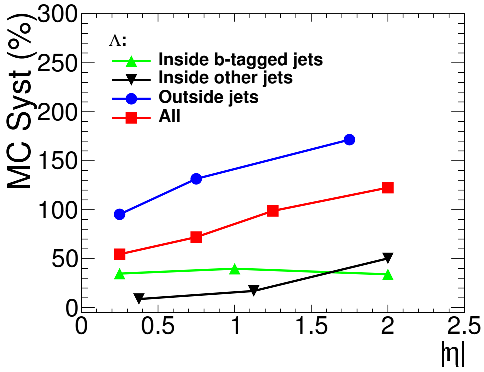
Figure 5.44: The K_{S}^0 systematic uncertainties due to the MC generator, POWHEG+PYTHIA6 (nominal MC) vs MC@NLO+HERWIG, as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K_{S}^0 inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).



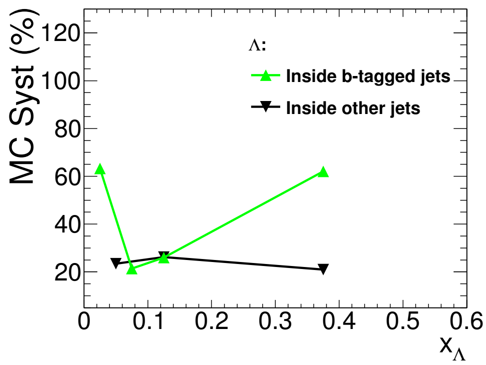
(a) p_{T} distribution



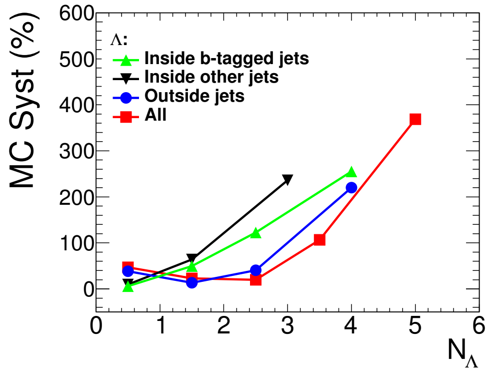
(b) Energy distribution



(c) $|\eta|$ distribution



(d) Energy fraction distribution

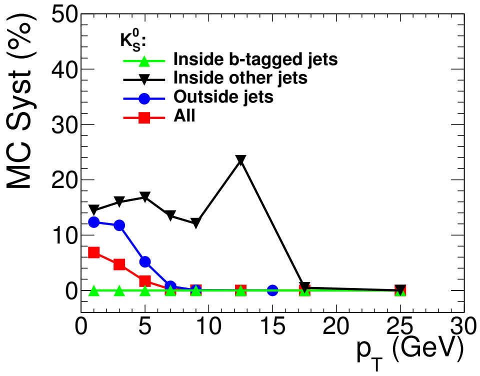


(e) Multiplicity distribution

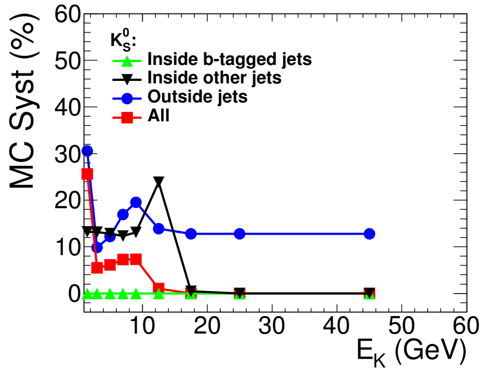
Figure 5.45: The Λ systematic uncertainties due to the MC generator, POWHEG+PYTHIA6 (nominal MC) vs MC@NLO+HERWIG, as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for Λ inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

The systematic uncertainties due to the MC generator used for the unfolding are dominant among all the uncertatinties investigated in the coming subsections. An attempt is made to separate the effect due to the matrix element (ME) calculation from the parton shower (PS) and hadronisation scheme. In a first step, the difference between the unfolded data using Powheg+Pythia6 (nominal MC) and Powheg+Herwig (as a comparison) is calculated, see Figure 5.46. Here the only difference lies in the PS and hadronisation scheme, since they both use the same ME. As it can be seen, Figure 5.46, these uncertainties are of similar order than those shown in Figure 5.44, in which both PS and ME in the generator changed.

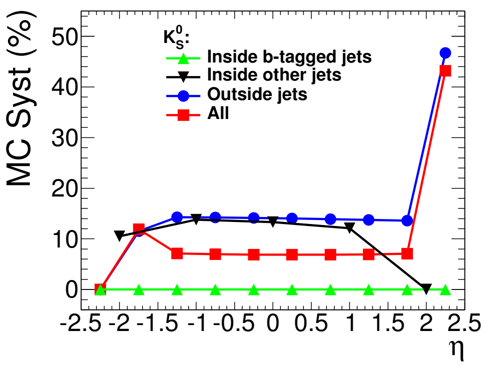
In a second step, the relative difference between the unfolded data has been also obtained by using the MC@NLO+Herwig and Powheg+Herwig. Now the only difference lies the ME since they both use the same PS and hadronisation scheme. As shown in Figure 5.47, uncertainties due to ME are lower than the ones due to PS only. There is here a caveat, since for this comparison of different ME generators, also the reconstruction is different, as AtlasfastII is used for Powheg+Herwig, and full simulation with Geant4 for MC@NLO+Herwig, so the uncertainties shown in this Figure are not due to ME only. As a further cross-check to see the importance of reconstruction simulations, a comparison with two Powheg+Pythia6 samples, one with ATLASFastII and another with Geant4, is shown in Figure 5.48. These uncertainties are similar to those in the previous Figure, so the uncertatinties due to ME only are even smaller than shown in Figure 5.47. This is a good hint that the main uncertainty comes from the PS and hadronisation scheme and not from the ME calculation, as expected.



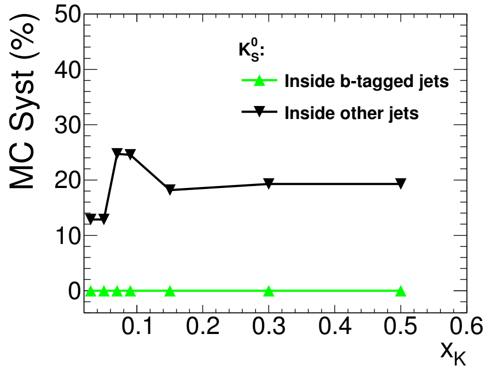
(a) p_{T} distribution



(b) Energy distribution

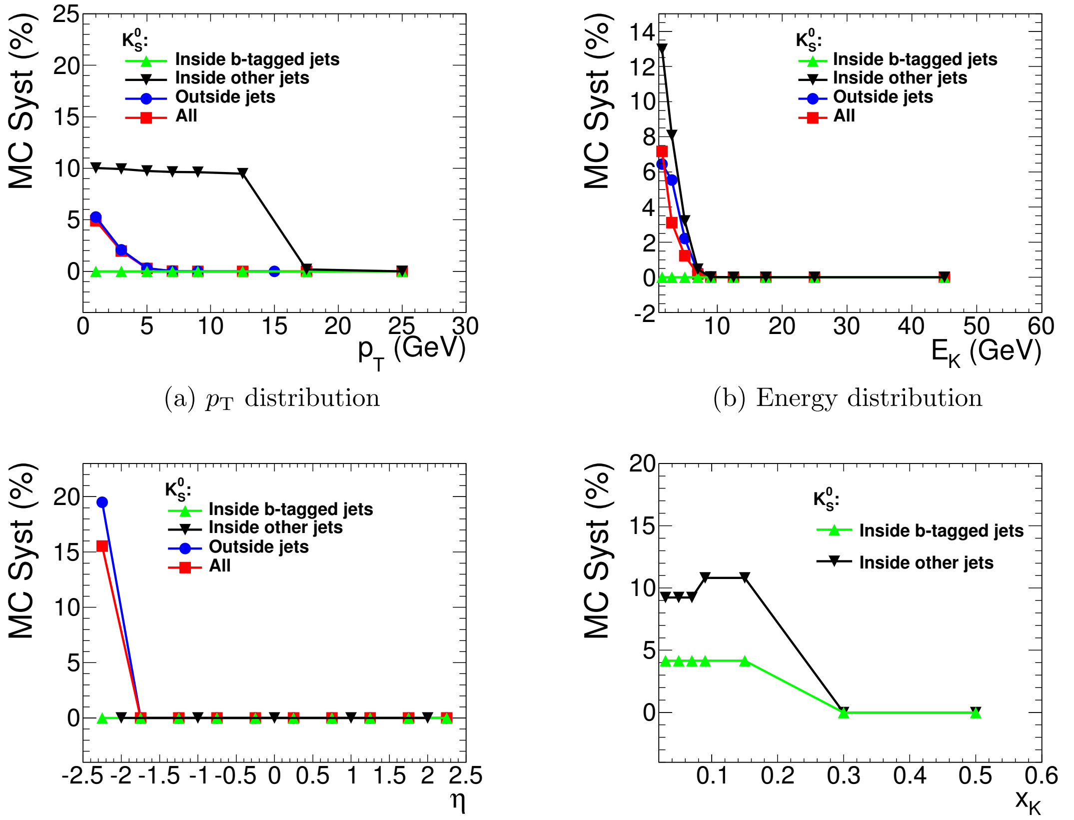


(c) $|\eta|$ distribution



(d) Energy fraction distribution

Figure 5.46: The K_{S}^0 systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, and energy fraction (x_K). Four classes are considered, for K_{S}^0 inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red). These uncertainties have been obtained by comparing unfolded data using POWHEG+PYTHIA6 (nominal MC) and POWHEG+HERWIG.



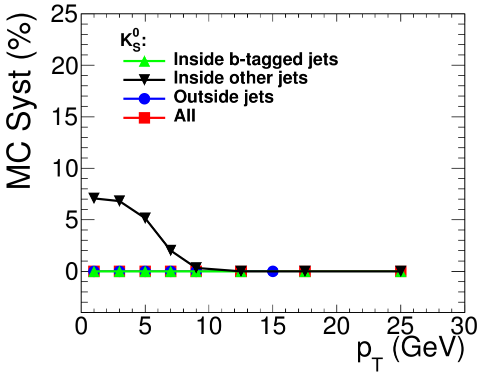
(a) p_{T} distribution

(b) Energy distribution

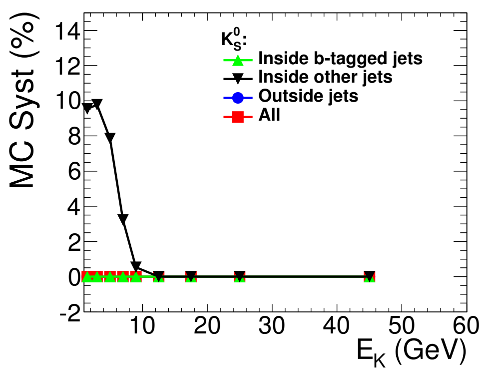
(c) $|\eta|$ distribution

(d) Energy fraction distribution

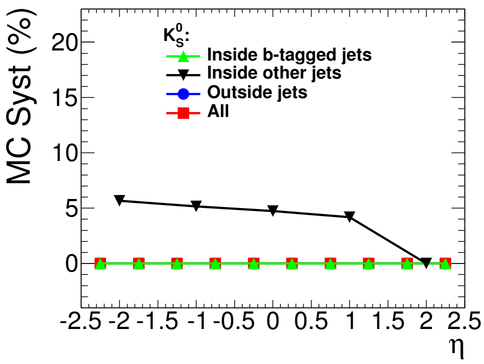
Figure 5.47: The K^0_{S} systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, and energy fraction (x_K). Four classes are considered, for K^0_{S} inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red). These uncertainties have been obtained by comparing unfolded data using MC@NLO+HERWIG and POWHEG+HERWIG, to see the effects due to different ME.



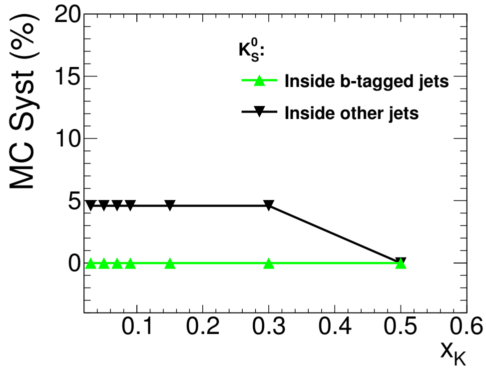
(a) p_{T} distribution



(b) Energy distribution



(c) $|\eta|$ distribution



(d) Energy fraction distribution

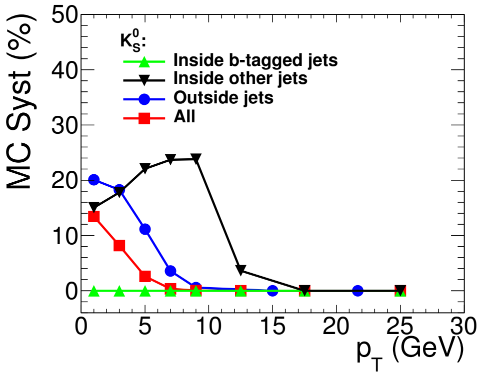
Figure 5.48: The K_{S}^0 systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, and energy fraction (x_K). Four classes are considered, for K_{S}^0 inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red). These uncertainties have been obtained by comparing unfolded data using POWHEG+PYTHIA6 with ATLASFastII and GEANT4, to see the effects due to different reconstruction simulations.

5.7.1.1 Unfolding to visible vs. invisible decays

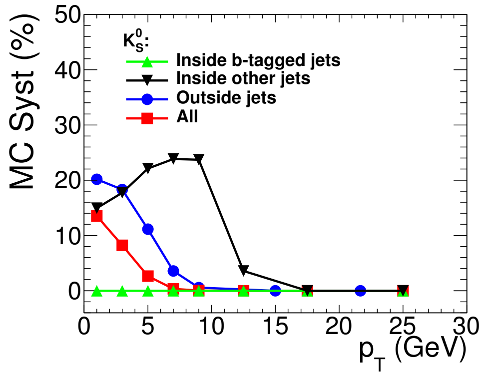
As shown in Eq. 5.5, the systematic uncertainties due to the MC modelling should be the same irrespectively of whether only visible decays are used at the particle level or not. This is so because due to isospin (SU(2)) symmetry there are no differences in the kinematical spectra for say $K_{\mathrm{S}}^0 \to \pi^0\pi^0$ and $K_{\mathrm{S}}^0 \to \pi^+\pi^-$.

5.7.1.2 Bin-by-bin vs. Bayes unfolding

The systematics for this dominant source, which is due to the choice of MC in the unfolding, are really similar no matter which method is used. This is shown in the Figure 5.49 below, Bayes (right) and bin-by-bin (left).



(a) Bin-by-bin unfolding

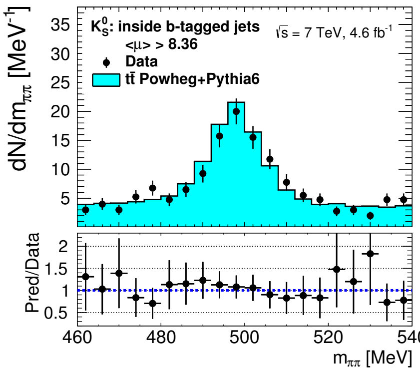


(b) Bayes unfolding

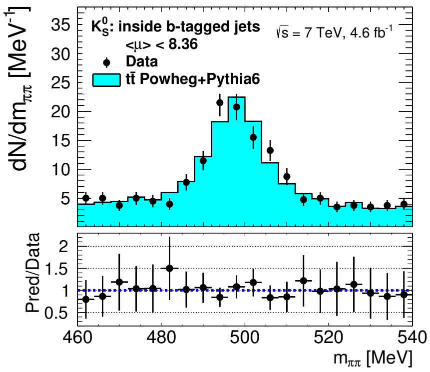
Figure 5.49: The K_S^0 systematic uncertainties due to the MC generator, Powheg+Pythia6 (nominal MC) vs MC@NLO+Herwig, as a function of p_T for bin-by-bin (left) and Bayes (right) unfolding. Four classes are considered, for K_S^0 inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

5.7.2 Pile-up uncertainties

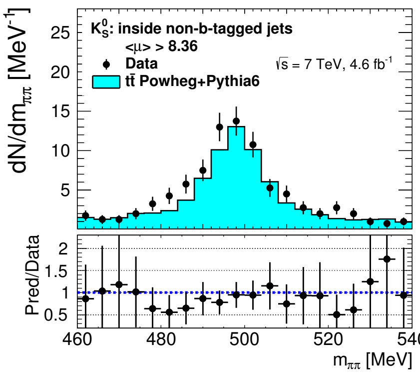
As a first attempt to check how well the MC pile-up modelling describes the data, plots as in Figure 5.6 showing the K_S^0 and Λ candidates mass distributions are repeated but dividing the sample in two halves with average number of interactions per bunch crossing, $< \mu >$, higher or lower than the median ($< \mu >= 8.36$).



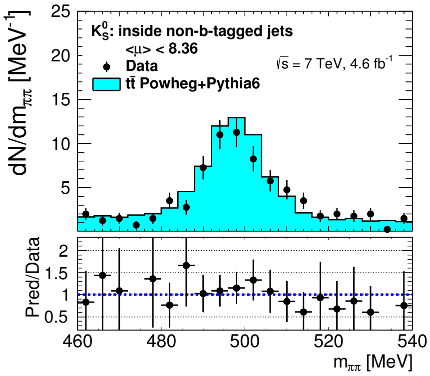
(a) b-jets



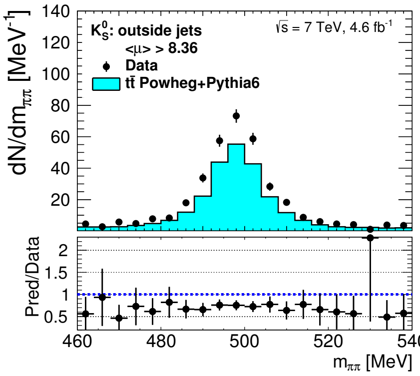
(b) b-jets



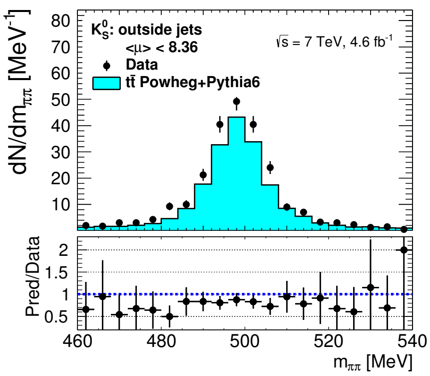
(c) Non b-tagged jets



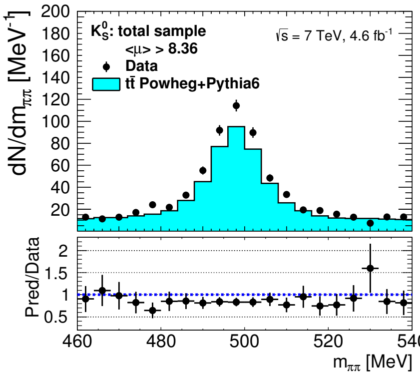
(d) Non b-tagged jets



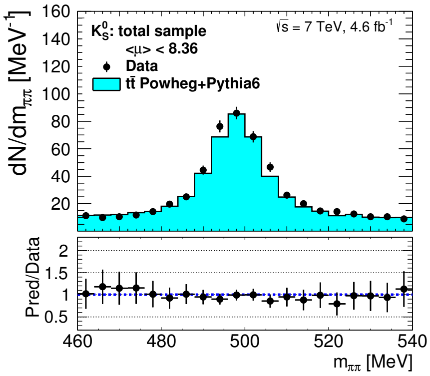
(e) Outside jets



(f) Outside jets

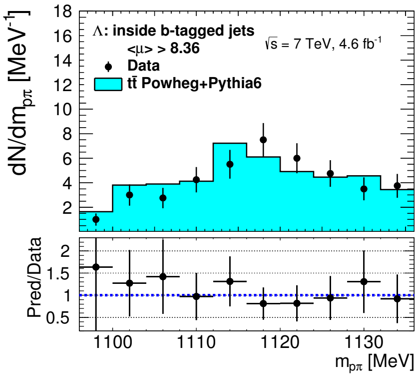


(g) All sample

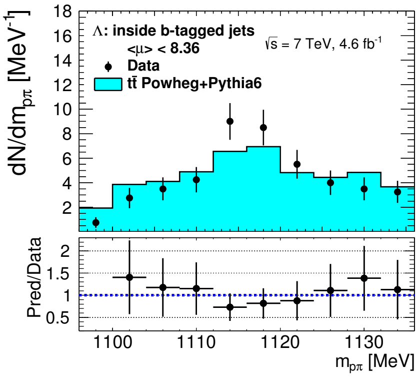


(h) All sample

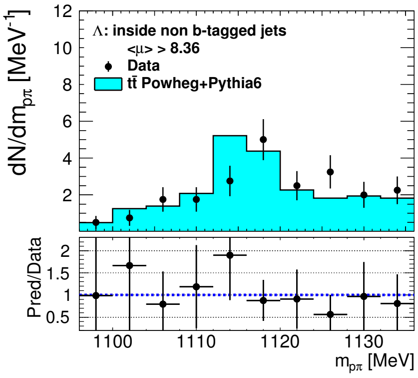
Figure 5.50: The measured K_S^0 candidates mass distributions for the high (left) and low pile-up (right) subsamples, along with Powheg+Pythia6 MC simulation. The mass distributions are scaled to a bin width of 4 MeV.



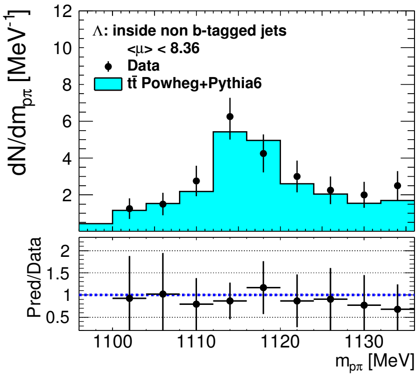
(a) *b*-jets



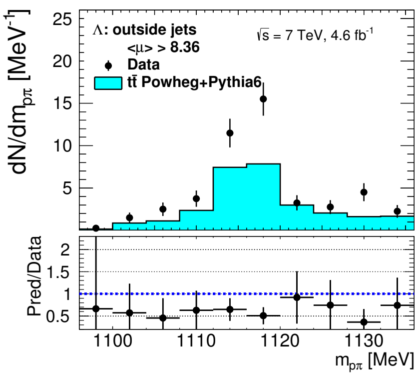
(b) *b*-jets



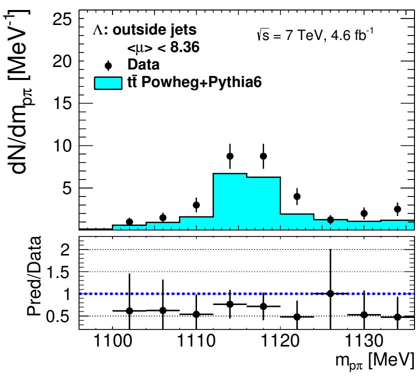
(c) Non *b*-tagged jets



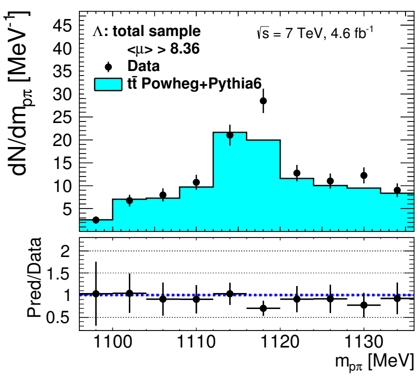
(d) Non *b*-tagged jets



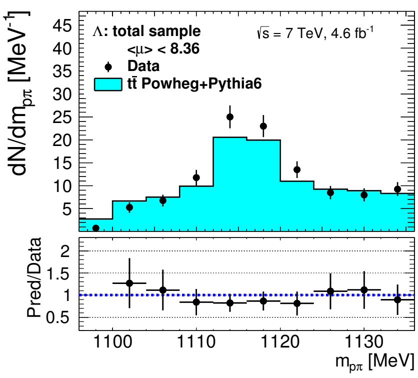
(e) Outside jets



(f) Outside jets



(g) All sample



(h) All sample

Figure 5.51: The measured Λ candidates mass distributions for the high (left) and low pile-up (right) subsamples, along with POWHEG+PYTHIA6 MC simulation. The mass distributions are scaled to a bin width of 4 MeV.

In Figures 5.50 and 5.51 comparisons of the high and low pile-up subsamples are shown. It is to be noticed that for neutral strange particles associated with jets (b-tagged or not) the number of K^0_S's (Λ's) is equally distributed between the two subsamples. This is not the case for neutral strange particles not associated with jets for which the high pile-up subsample shows a higher neutral strange particle multiplicty. Furthermore, the disagreement between data and MC samples for K^0_S's (Λ's) production not associated with jets is slightly bigger in the high pile-up than in the low pile-up subsample.
The K^0_S and Λ yields at reconstructed level for the high and low pile-up subsamples, as well as for the total sample, are shown in Tables 5.6 and 5.8 (5.7 and 5.9) for data (POWHEG+PYTHIA6 MC). Note that the yields are mostly unaffected by the pile-up, within the statistical uncertainties, for K^0_S's and Λ's associated with jets, b-tagged or not, while for the not associated with jets, the yields increase with the pile-up ($< \mu >$). Thus only pile-up systematics to the latter category have been considered.

Table 5.6: K^0_S yields at the reconstructed level data, for the high and low pile-up subsamples, as well as for the total sample.

	High PU	Low PU	Total
Class	$< n_K > \pm(stat)$	$< n_K > \pm(stat)$	$< n_K > \pm(stat)$
Inside b-jets	0.072 ± 0.005	0.081 ± 0.005	0.077 ± 0.004
Inside non b-tagged jets	0.061 ± 0.004	0.052 ± 0.004	0.056 ± 0.003
Outside any jet	0.31 ± 0.011	0.220 ± 0.009	0.265 ± 0.007
Total sample	0.45 ± 0.015	0.361 ± 0.013	0.403 ± 0.010

Table 5.7: K^0_S yields at the reconstructed level POWHEG+PYTHIA6 MC, for the high and low pile-up subsamples, as well as for the total sample.

	High PU	Low PU	Total
Class	$< n_K > \pm(stat)$	$< n_K > \pm(stat)$	$< n_K > \pm(stat)$
Inside b-jets	0.0794 ± 0.0014	0.0834 ± 0.0013	0.0814 ± 0.0010
Inside non b-tagged jets	0.0543 ± 0.0011	0.0556 ± 0.0011	0.0550 ± 0.0008
Outside any jet	0.2403 ± 0.0023	0.1797 ± 0.0019	0.2091 ± 0.0015
Total sample	0.3823 ± 0.0034	0.3267 ± 0.0030	0.3538 ± 0.0023

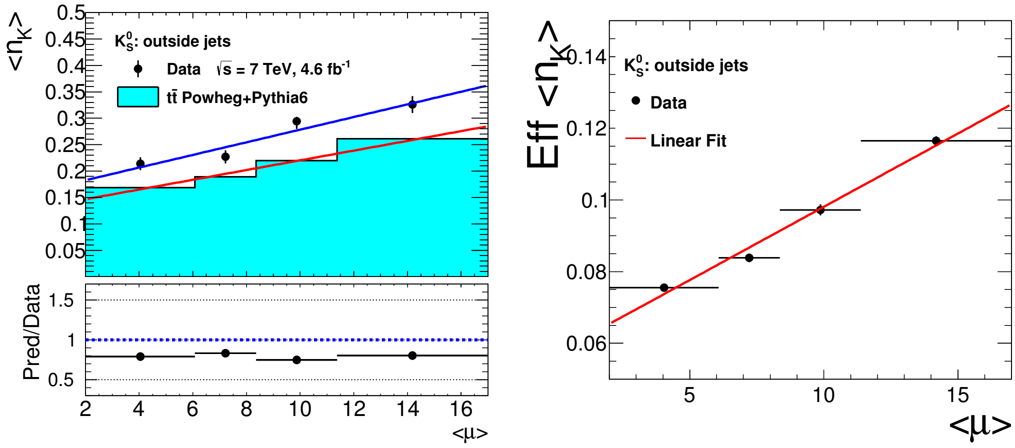
Table 5.8: Λ yields at the reconstructed level data, for the high and low pile-up subsamples, as well as for the total sample.

	High PU	Low PU	Total
Class	$< n_{\Lambda} > \pm(stat)$	$< n_{\Lambda} > \pm(stat)$	$< n_{\Lambda} > \pm(stat)$
Inside b-jets	0.0118 ± 0.0021	0.0214 ± 0.0026	0.0166 ± 0.0016
Inside non b-tagged jets	0.0081 ± 0.0015	0.0107 ± 0.0019	0.0094 ± 0.0012
Outside any jet	0.0303 ± 0.0031	0.0225 ± 0.0026	0.0264 ± 0.0020
Total sample	0.0482 ± 0.0044	0.0578 ± 0.0045	0.0529 ± 0.0031

Table 5.9: Λ yields at the reconstructed level Powheg+Pythia6 MC, for the high and low pile-up subsamples, as well as for the total sample.

	High PU	Low PU	Total
Class	$< n_{\Lambda} > \pm(stat)$	$< n_{\Lambda} > \pm(stat)$	$< n_{\Lambda} > \pm(stat)$
Inside b-jets	0.0111 ± 0.0005	0.0102 ± 0.0005	0.0107 ± 0.0003
Inside non b-tagged jets	0.0102 ± 0.0005	0.0120 ± 0.0005	0.0111 ± 0.0003
Outside any jet	0.0200 ± 0.0006	0.0151 ± 0.0005	0.0175 ± 0.0004
Total sample	0.0437 ± 0.0010	0.0397 ± 0.0009	0.0416 ± 0.0007

Figure 5.52 below shows the outside jets K^0_S multiplicity dependence on the average number of interactions per bunch crossing ($< \mu >$), along with a comparison to the Powheg+Pythia6 MC sample.



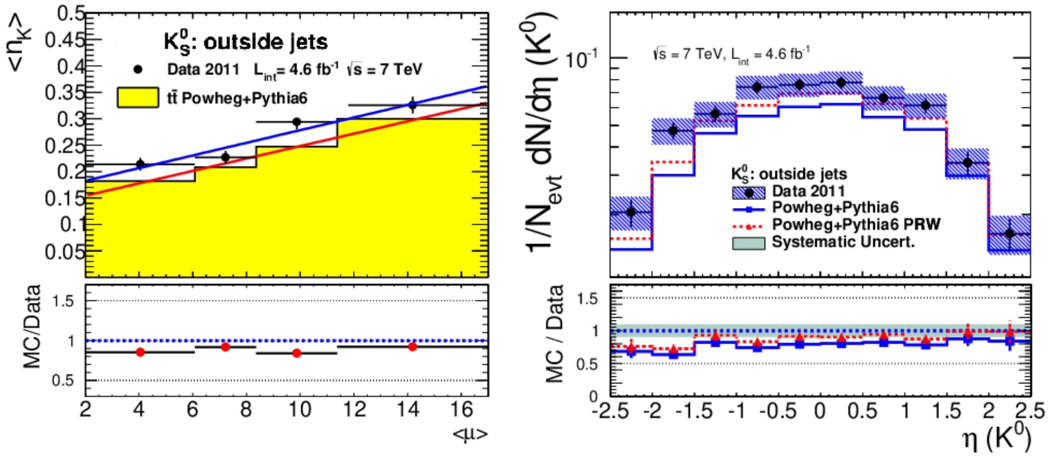
(a) $< n_K >$ fit (b) Efficiency fit

Figure 5.52: K^0_S average multiplicity outside jets (left) and the corresponding efficiency (right) as a function of $< \mu >$ for data and POWHEG+PYTHIA6 MC, along with fits to linear distributions.

The dependence of $< n_K >$ on $< \mu >$ is linear in both data and MC. In fact the slopes of straight lines fitted to them ($< n_K >= a + b < \mu >$), yield the results $b_{MC} = 0.0092 \pm 0.0003$ and $b_{data} = 0.012 \pm 0.002$. The fitted intersect values are $a_{MC} = 0.128 \pm 0.004$ and $a_{data} = 0.16 \pm 0.02$. The fitted values for slope and intersect in the data are larger than those in MC, by one and a half standard deviations. On the other hand, the dependence of the efficiency on $< \mu >$ is also fitted by a straight line resulting in $a_\epsilon = 0.057 \pm 0.002$ and $b_\epsilon = 0.004 \pm 0.0002$. Using these numbers, the data on the mean K^0_S multiplicity outside jets PU corrected (i.e $< \mu >= 1$) and unfolded to the particle level turns out to be $< n_K >= (a_{data} + b_{data})/(a_\epsilon + b_\epsilon) = 0.172/0.061 = 2.82$, which is in good agreement with the resulting $< n_K >$ from unfolded data using the efficiencies as a function of p_{T} as shown in Table 5.12, Section 5.8.
In order to correct for this small mismodelling in the MC, reco level K^0_S's not associated with jets have been reweighted to mimic the data slope, i.e. the $< \mu >$ dependence associated with pile-up. To be more precise, the weights have been defined as:

$$w = \frac{a_{MC} + b_{data} < \mu >}{a_{MC} + b_{MC} < \mu >} \tag{5.7}$$

In order to illustrate the effect of this data-driven pile-up reweighting the comparisons presented in Figures 5.52a and 5.19c are shown again in Figure 5.53, after the procedure discussed above. Clearly, the discrepancies observed in the data-MC comparisons at reco level have been reduced. These discrepancies have not been completely gauged away because the intersect a, which is not related to pile-up, has not been retuned.



(a) $< n_K >$ fit

(b) $|\eta|$ comparisons

Figure 5.53: $K^0_{\rm S}$ average multiplicity outside jets as a function of $< \mu >$ (left) and pseudorapidity distribution (right) for data and POWHEG+PYTHIA6 MC after pile-up reweighting.

As a further test that the reweighting is producing the expected effects, two variations in the procedure have been done, shifting up and down the value for the slope b_{data} in the numerator of Eq. 5.7 by an amount given by the statistical error of the data fit. The results are shown in Figure 5.54 below. Linear fits to these two variations result in the following values for the slopes $b_{MC} = 0.0119 \pm 0.0004$, $b_{MCup} = 0.0136 \pm 0.0004$ and $b_{MCdown} = 0.0099 \pm 0.0004$ in very good agreement with the expectations.

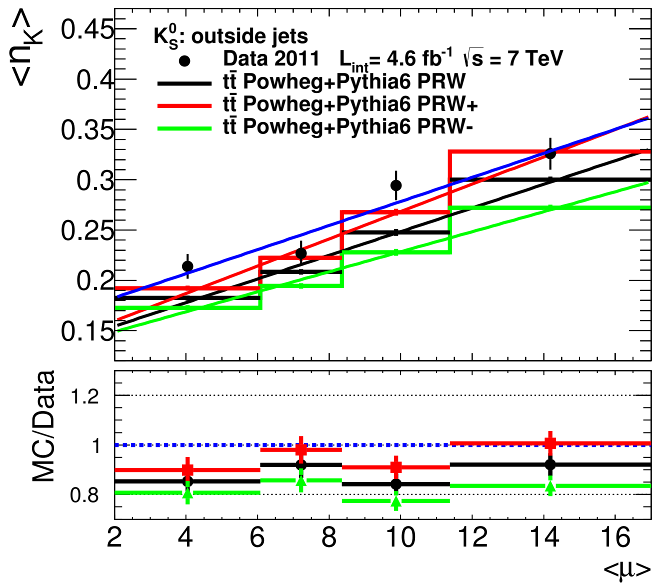
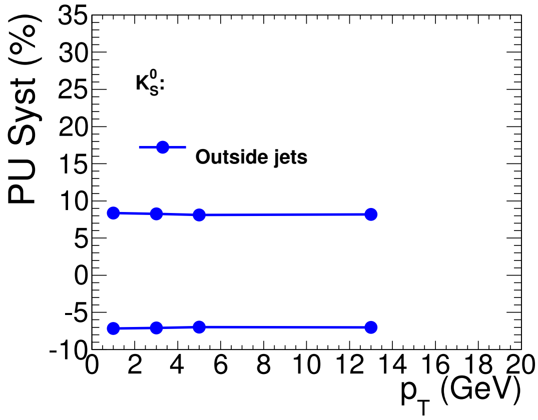
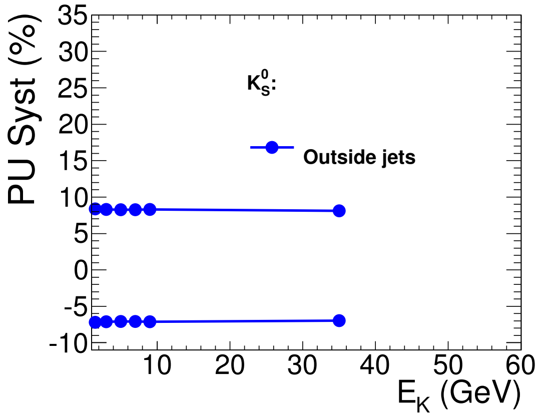


Figure 5.54: K_S^0 average multiplicity outside jets as a function of $< \mu >$ for data and POWHEG+PYTHIA6 MC after pile-up reweighting (PRW) with up and down variations.

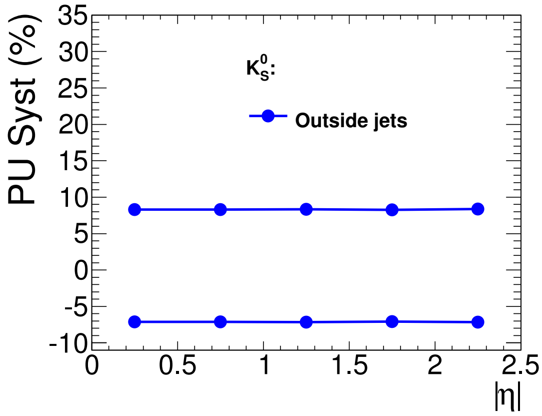
New efficiency sets are calculated using this reweighted reco level MC, including up and down variations on the slope given by the statistical uncertainty on the data fit. This data-driven procedure allows to calculate the systematic uncertainty due to pile-up as the relative difference between the nominal reweighted unfolded distribution and the two up and down variations. The results are shown in Figure 5.55 and found to be of the order of 8% which is smaller than the previous systematics due to the choice of the generator used for the unfolding.



(a) p_{T} distribution



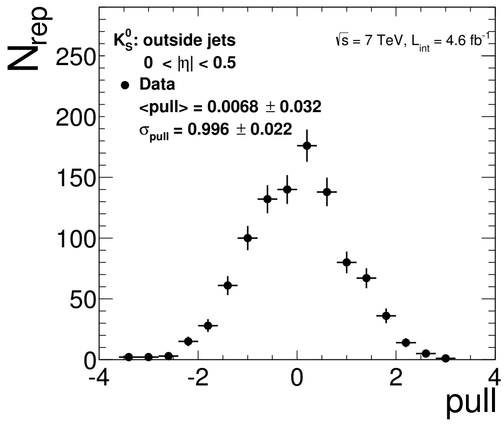
(b) Energy distribution



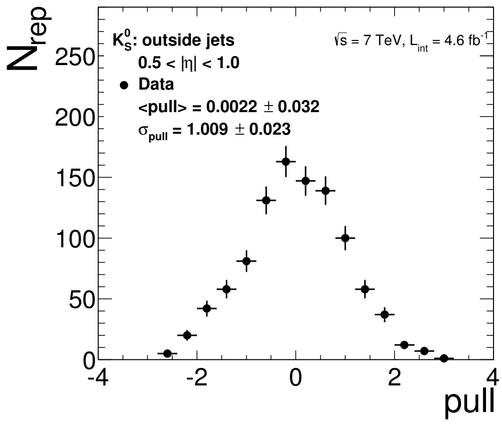
(c) $|\eta|$ distribution

Figure 5.55: The K_S^0 systematic uncertainties due to the pile-up modelling in POWHEG+PYTHIA6 (nominal MC), as a function of p_{T}, $|\eta|$, energy, and multiplicity. Four classes are considered, for K_S^0 inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

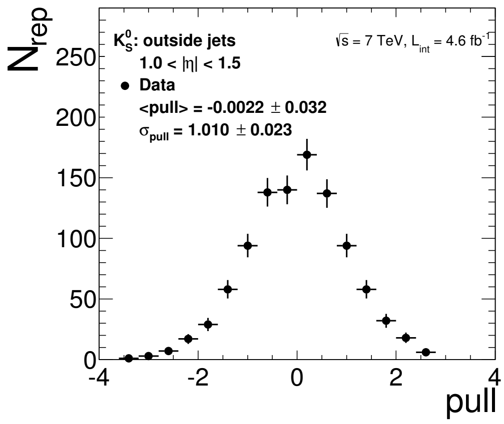
The statistical errors in the systematic uncertatinty have been calculated using the Bootstrap method, they turned out to be small. This is understood since this systematic uncertainty is essenctially the deviation from one of the ratio between two efficiency sets, the nominal and the up and down variations using the same POWHEG+Pythia MC sample. The pulls obtained from the replicas in the Bootstrap procedure are shown in Figure 5.56. They exhibit a normal distribution behaviour as expected. The same comment as at the end of Section 5.6.2.1 applies here.



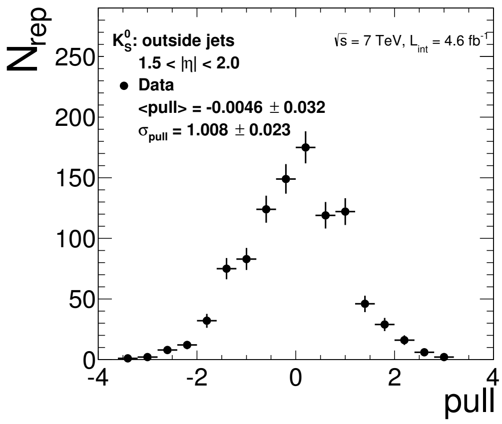
(a) $0 < |\eta| < 0.5$



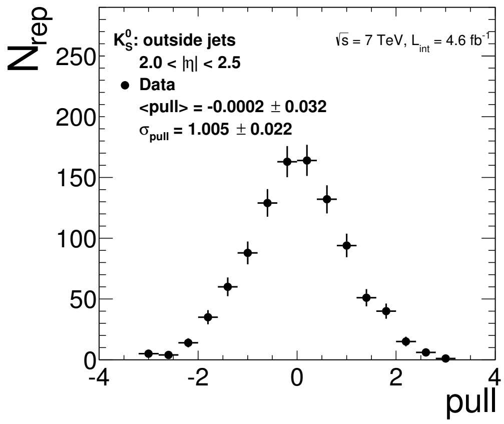
(b) $0.5 < |\eta| < 1$



(c) $1 < |\eta| < 1.5$



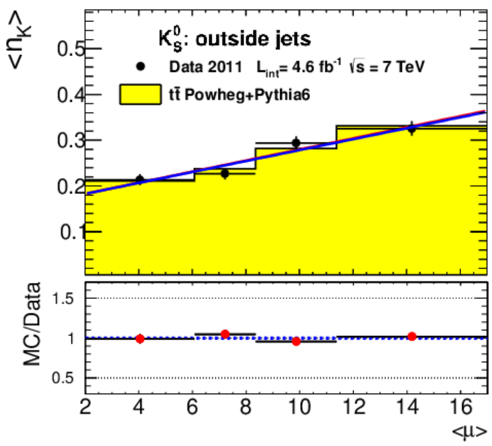
(d) $1.5 < |\eta| < 2$



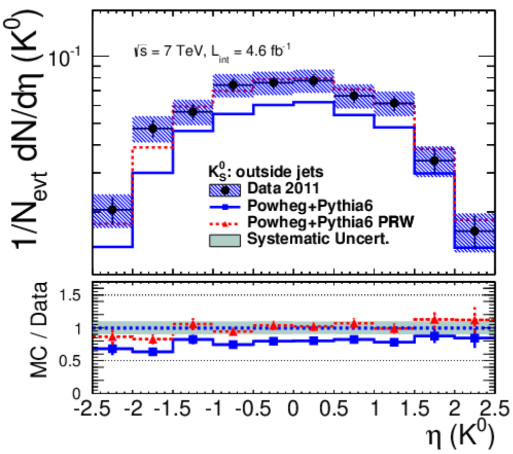
(e) $2 < |\eta| < 2.5$

Figure 5.56: The pulls from replicas obtained for the pseudorapidity K_S^0 systematic uncertainties due to the pile-up modelling in Powheg+Pythia6.

An additional test has been performed by reweighting the MC samples as to match also the intersect of the data fit, changing the a_{MC} by a_{data} in the numerator of Eq. 5.7. The shifted sets of efficiencies are then calculated by just changing the b_{data} as before. The resulting uncertainties are of the order of the previous one, but slightly smaller (around 6%) since the background fraction is reduced as the intersect in data is higher than in MC sample. In Figure 5.57 it is shown that after reweighting the MC samples match almost completely the data.



(a) $< n_K >$ fit

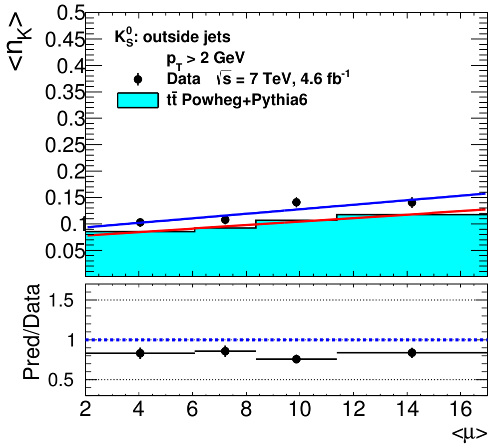


(b) $|\eta|$ comparisons

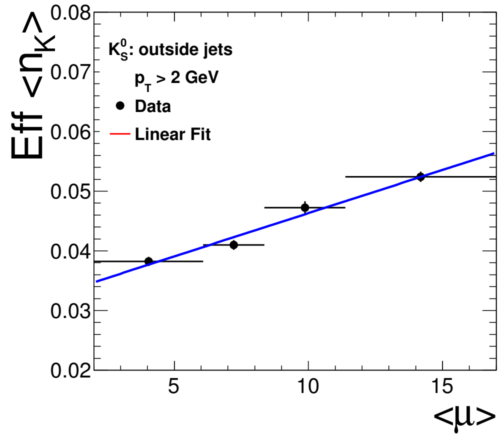
Figure 5.57: K^0_S average multiplicity outside jets as a function of $< \mu >$ (left) and pseudorapidity distribution (right) for data and POWHEG+PYTHIA6 MC after pile-up reweighting, including a_{data}.

As a further cross-check the fits discussed above have been repeated for two subsamples at low and high p_T. The results are shown in Figures 5.58 and 5.59. The following remarks are in order:

- the $< \mu >$ dependence of both subsamples is very different, the low p_T sample being much steeper
- the MC reproduces this feature fairly well
- this gives further support to the procedure described above for the estimation of the pile-up systematics.

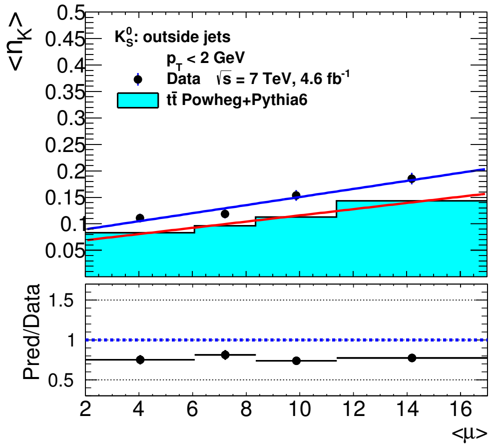


(a) $< n_K >$ fit

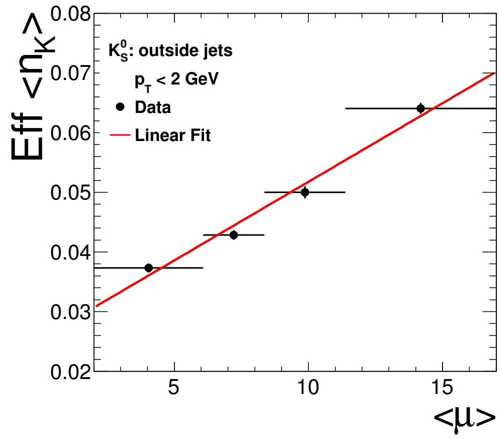


(b) Efficiency fit

Figure 5.58: K^0_S average multiplicity outside jets (left) and the corresponding efficiency (right) as a function of $< \mu >$ for data and POWHEG+PYTHIA6 MC, along with fits to linear distributions for K^0_S with $p_\mathrm{T} > 2$ GeV.



(a) $< n_K >$ fit



(b) Efficiency fit

Figure 5.59: K^0_S average multiplicity outside jets (left) and the corresponding efficiency (right) as a function of $< \mu >$ for data and POWHEG+PYTHIA6 MC, along with fits to linear distributions for K^0_S with $p_\mathrm{T} < 2$ GeV.

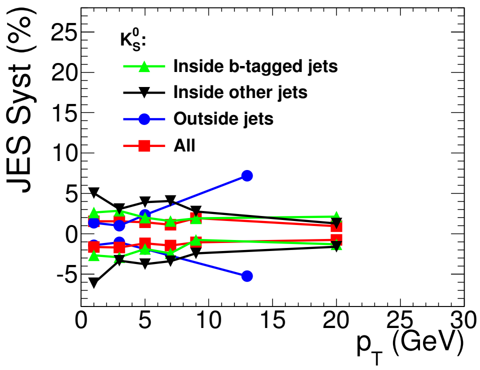
5.7.3 JES and JER uncertainties

In this subsection, the propagation of the jet energy scale (JES) [130] and jet energy resolution (JER) [141] uncertainties to the K^0_S and Λ systematic uncertainties are calculated. These effects will be found smaller than the uncertainties due to the choice of MC for the unfolding, since they affect most of the distributions indirectly through slight changes in the number of jets passing the cuts (see Section 5.3). The only distribution directly affected by these uncertainties is the energy fraction ($x_{K,\Lambda}$). of the scalar sum of the two leading jets transverse momenta ($\Sigma p_{T,b\bar{b}}$).
To propagate the JES to the K^0_S and Λ systematic uncertainties, the energy and transverse momentum of all jets are shifted up and down by the global JES uncertainty (taking into account whether the jet is a b-jet or not) in the POWHEG+PYTHIA6 MC sample at reconstructed level and before applying the cuts. The K^0_S and Λ relative systematic uncertainties due to this shift are asymmetric since the increment in number of jets passing all cuts (Section 5.3) when the JES is added, is different from the reduction when it is subtracted. Therefore, the K^0_S and Λ systematic uncertainties associated with the JES, σ^{JESup}_{syst} ($\sigma^{JESdown}_{syst}$), are calculated on a bin-by-bin basis for each distribution and for each class, as the positive (negative) relative differences between the nominal distribution and shifted ones, more details can be found in Appendix A.3.

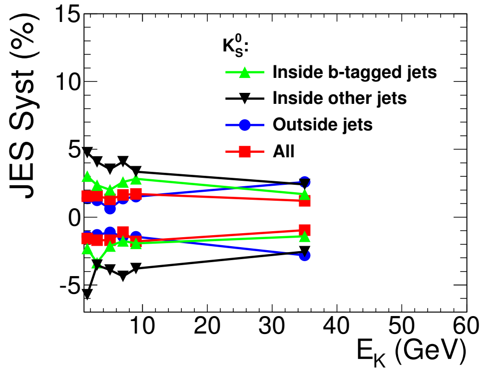
$$\sigma^{JESup}_{syst} = pos[\frac{MC^{Pythia}_{+JES} - MC^{Pythia}_{nominal}}{MC^{Pythia}_{nominal}}, \frac{MC^{Pythia}_{-JES} - MC^{Pythia}_{nominal}}{MC^{Pythia}_{nominal}}] \tag{5.8}$$

$$\sigma^{JESdown}_{syst} = neg[\frac{MC^{Pythia}_{+JES} - MC^{Pythia}_{nominal}}{MC^{Pythia}_{nominal}}, \frac{MC^{Pythia}_{-JES} - MC^{Pythia}_{nominal}}{MC^{Pythia}_{nominal}}] \tag{5.9}$$

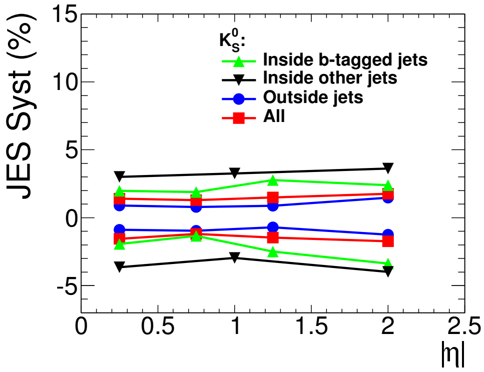
In Figures 5.60 and 5.61 the values of the relative sytematic uncertainties for all the K^0_S and Λ distributions and classes are shown. They are all below 5 %, except for some bins.



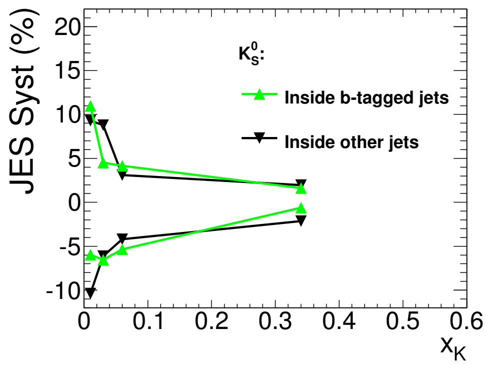
(a) p_{T} distribution



(b) Energy distribution

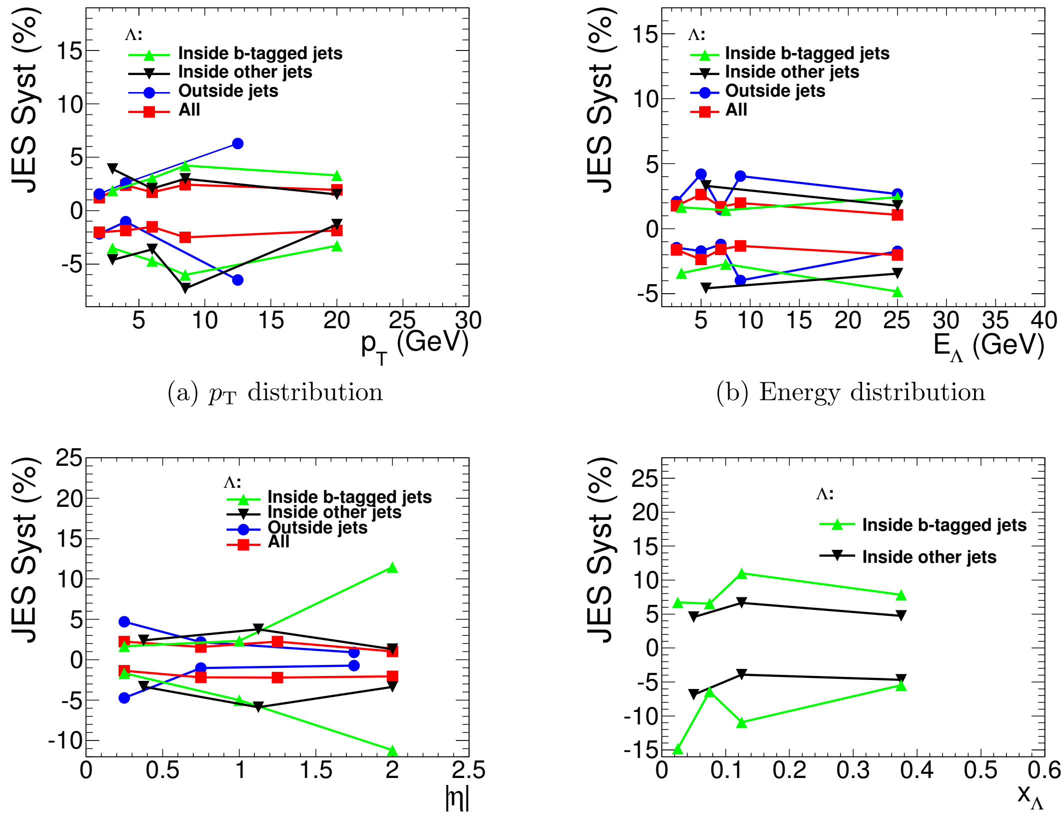


(c) $|\eta|$ distribution



(d) Energy fraction distribution

Figure 5.60: The K^0_{S} systematic uncertainties due to the JES uncertainties as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K^0_{S} inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).



(a) p_T distribution

(b) Energy distribution

(c) $|\eta|$ distribution

(d) Energy fraction distribution

Figure 5.61: The Λ systematic uncertainties due to the JES uncertainties as a function of p_T, $|\eta|$, energy, energy fraction (x_Λ) and multiplicity. Four classes are considered, for Λ's inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

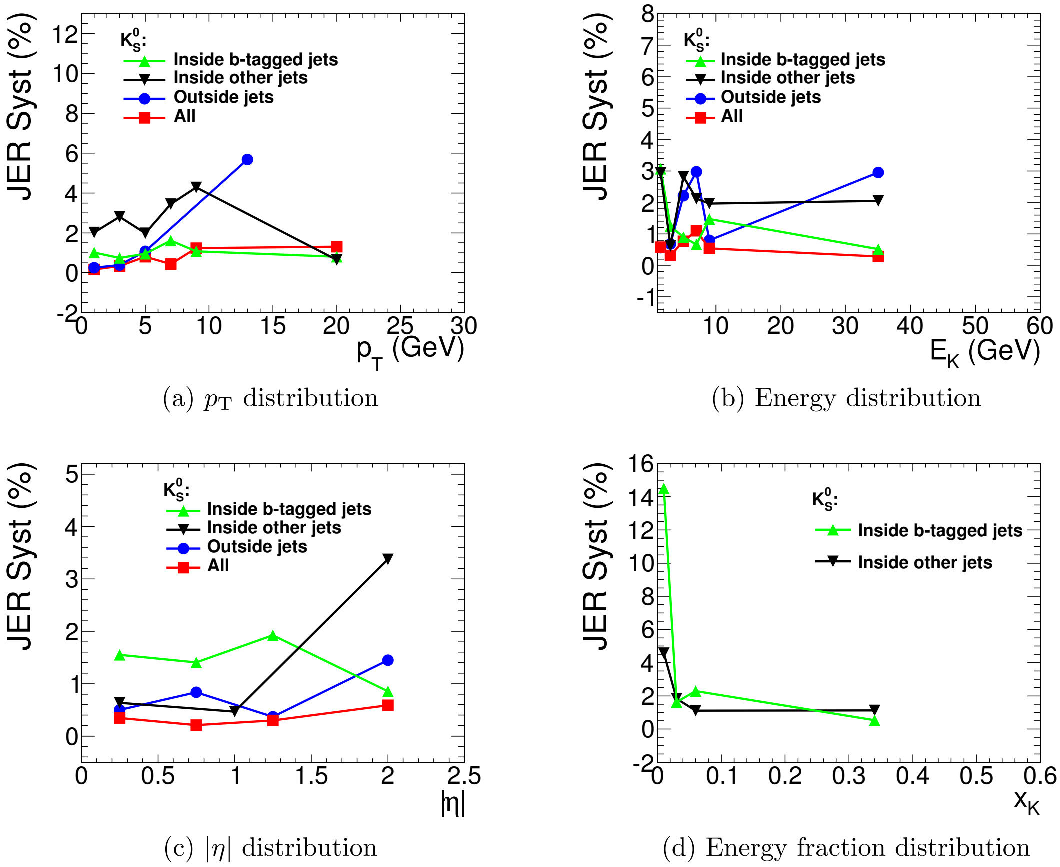
The JER in the ATLAS calorimeter is given by the following equation:

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = \frac{N}{p_\mathrm{T}} \oplus \frac{S}{\sqrt{p_\mathrm{T}}} \oplus C \tag{5.10}$$

where the N,S and C constants are used to parametrize the noise, stochastic and constant terms. To calculate the relative systematics due to this effect, the energy and transverse momentum of all jets, in the Powheg+Pythia6 MC sample at reconstructed level before applying cuts (Section 5.3), are smeared with a Gaussian distribution centered in unity and with standard deviation equal to the JER uncertainty for that jet, as a function of p_T and pseudorapidity [141]. Then the K^0_S and Λ relative systematics (σ^{JER}_{syst}) are obtained on a bin-by-by basis for each distribution and class, as the absolute value of the relative deviation of the smeared from the nominal distributions (Eq. 5.11)

$$\sigma_{syst}^{JER} = \frac{|MC_{smeared}^{Pythia} - MC_{nominal}^{Pythia}|}{MC_{nominal}^{Pythia}} \tag{5.11}$$

In Figures 5.62 and 5.63 the values of the relative systematic uncertainties for the K_S^0 and Λ distributions and classes , are shown, further details can be found in Appendix A.3. They are of the same order as the systematics due to the JER uncertainties.



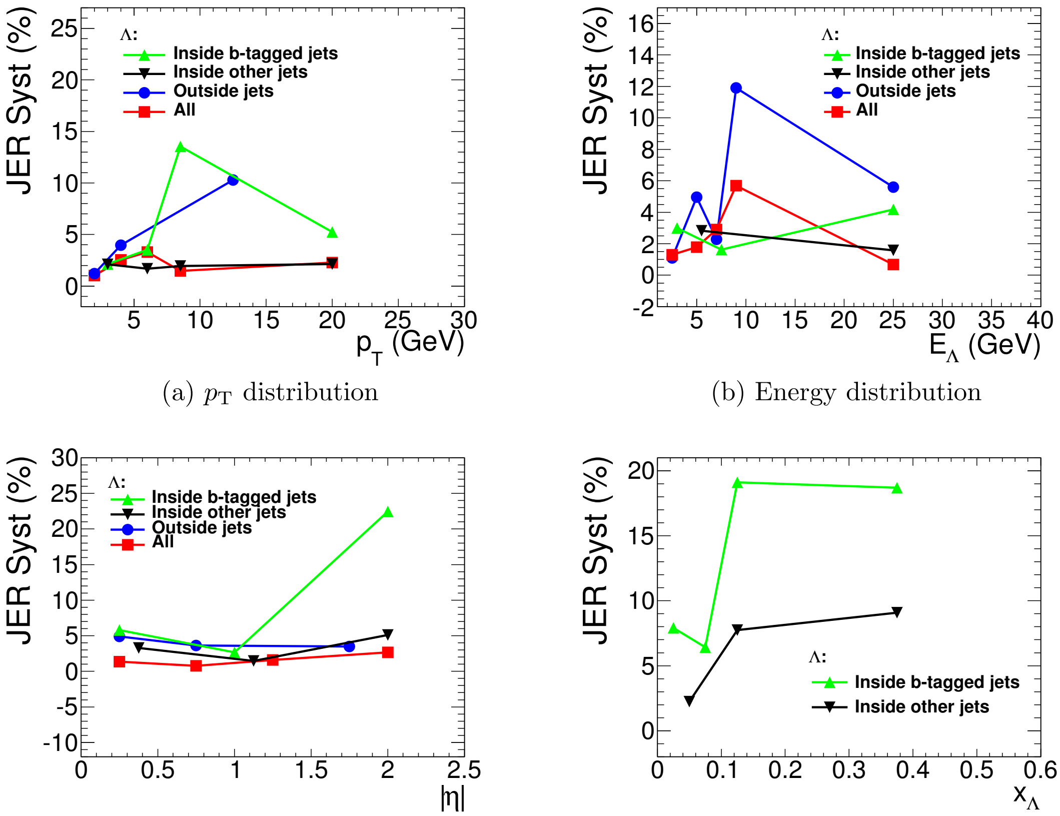
(a) p_T distribution

(b) Energy distribution

(c) $|\eta|$ distribution

(d) Energy fraction distribution

Figure 5.62: The K_S^0 systematic uncertainties due to the JER uncertainties as a function of p_T, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K_S^0 inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).



(a) p_{T} distribution

(b) Energy distribution

(c) $|\eta|$ distribution

(d) Energy fraction distribution

Figure 5.63: The Λ systematic uncertainties due to the JER uncertainties as a function of p_{T}, $|\eta|$, energy, energy fraction (x_{Λ}) and multiplicity. Four classes are considered, for Λ's inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

5.7.4 *b*-tagging uncertainties

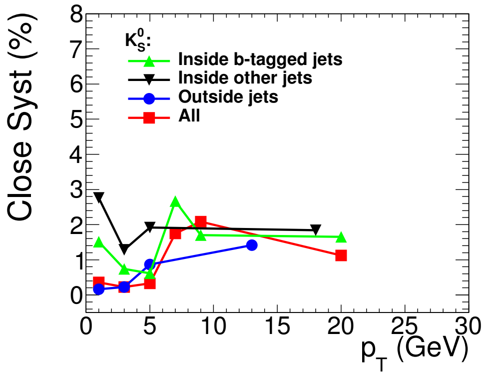
In order to study the systematics induced by the choice of the b-tagging working point, the analysis has been repeated using two additional working points, namely those giving a b-tagging efficiency of 60% and 85%. The main obvious effect of this is that the number of jets classified as b-tagged (or not) is going to change. Nevertheless, since as shown in Fig 5,6 and 14, 15, the distributions for K^0_S and Λ production inside b-tagged jets are very similar to those inside non b-tagged jets, the uncertainties for normalized distributions are expected to be small. In Table 5.10, the average multiplicity for K^0_S production per b-tagged jet (or non b-tagged jet) is shown to be independent from the choice of b-tagging working point within the statistical uncertainty. Therefore, one can conclude that the systematics due to the b-tagging working point are negligible.

Table 5.10: Total number of jets classified as b-tagged or non b-tagged, number of jets containing K^0_S or Λ, and K^0_S (Λ) reconstructed average multiplicities per jet for each class and for the total sample. Three b-tagging working points are considered, namely $\epsilon_b = 70\%$, $\epsilon_b = 60\%$ and $\epsilon_b = 85\%$

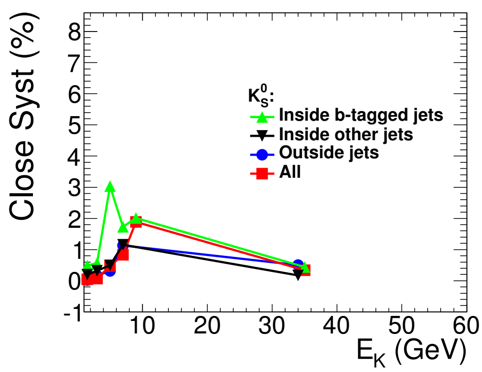
	$\epsilon_b = 70\%$	$\epsilon_b = 60\%$	$\epsilon_b = 85\%$
Number of b-tagged jets	10187	8747	12023
Number of non b-tagged jets	8369	9809	6533
Number of b-tagged jets with K^0_S	490	395	583
Number of non b-tagged jets K^0_S	373	468	280
$< n_K >$ inside b-tagged jets	0.052 $\pm$ 0.003	0.048 $\pm$ 0.004	0.052 $\pm$ 0.003
$< n_K >$ inside non b-tagged jets	0.047$\pm$ 0.003	0.051 $\pm$ 0.03	0.045 $\pm$ 0.003
Number of b-tagged jets with Λ	109	91	128
Number of non b-tagged jets Λ	63	81	44
$< n_\Lambda >$ inside b-tagged jets	0.011 $\pm$ 0.002	0.011 $\pm$ 0.002	0.011 $\pm$ 0.002
$< n_\Lambda >$ inside non b-tagged jets	0.008 $\pm$ 0.002	0.009 $\pm$ 0.003	0.007 $\pm$ 0.002

5.7.5 Unfolding Non-closure

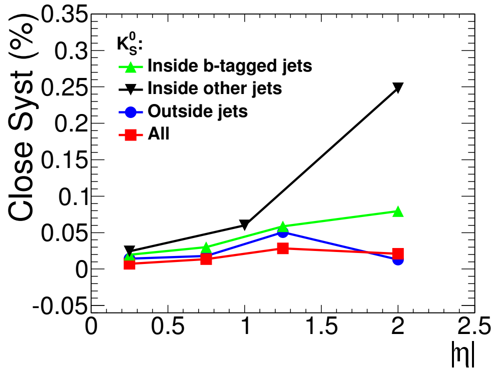
The unfolding non-closure uncertainty is calculated in two steps. In the first step, the p_T, energy, energy fraction, pseudorapidity and multiplicity reco level MC distributions are reweighted to match the data. Then, these reweighted MC distributions are unfolded to the particle level on a bin-by-bin (for the p_T, energy, energy fraction and $|\eta|$ distributions) or Bayesian (for the multiplicity distributions), and compared to the reweighted particle level MC. The latter are obtained directly from the transfer matrices and the reco level weights. This gives a very small uncertainty, as shown below in Figure 5.64 and 5.65.



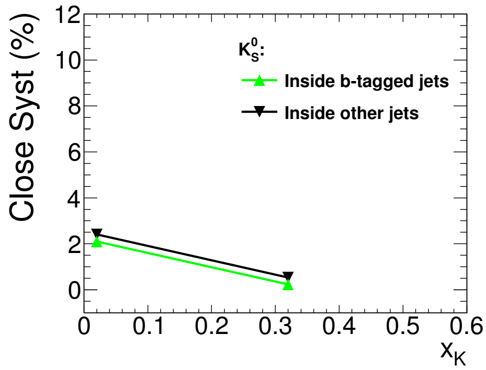
(a) p_{T} distribution



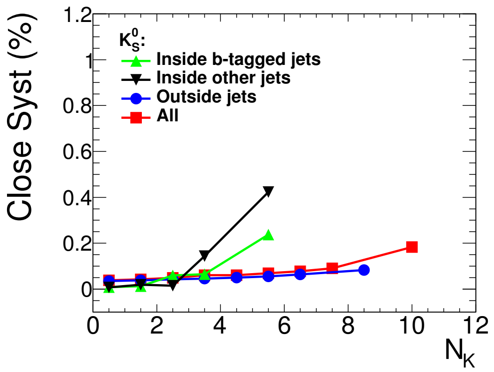
(b) Energy distribution



(c) $|\eta|$ distribution

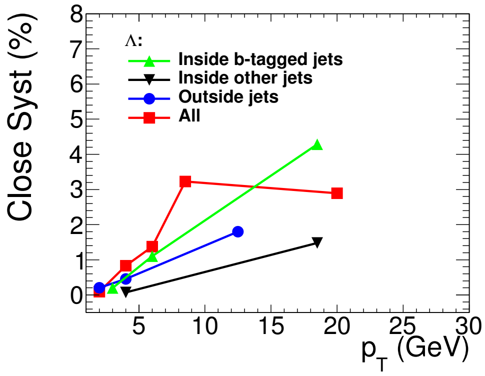


(d) Energy fraction distribution

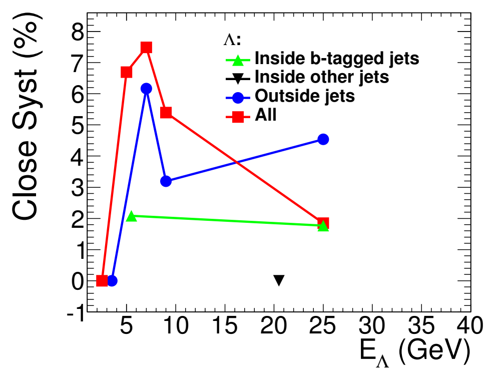


(e) Multiplicity distribution

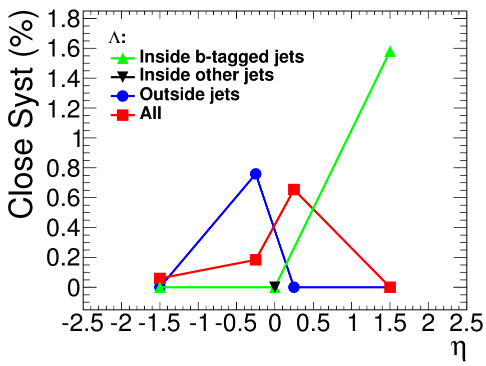
Figure 5.64: The K^0_{S} systematic uncertainties due to the unfolding non-closure, as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K^0_{S} inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).



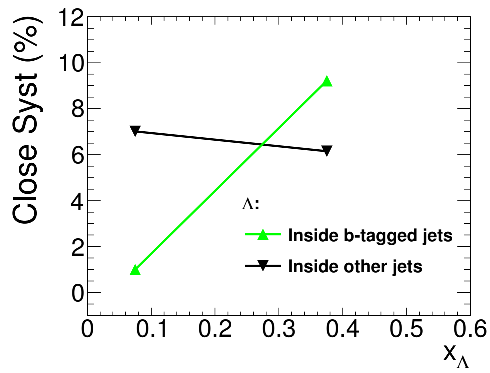
(a) p_{T} distribution



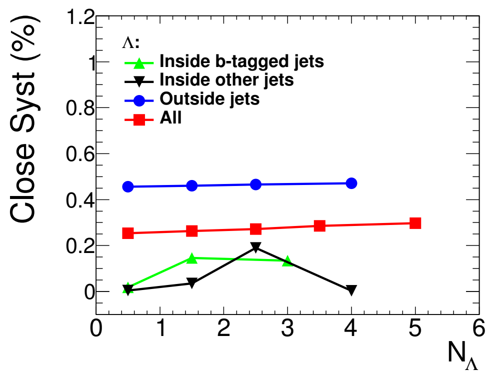
(b) Energy distribution



(c) $|\eta|$ distribution



(d) Energy fraction distribution



(e) Multiplicity distribution

Figure 5.65: The Λ systematic uncertainties due to the unfolding non-closure, as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K^0_{S} inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

5.7.6 Tracking Inefficiencies

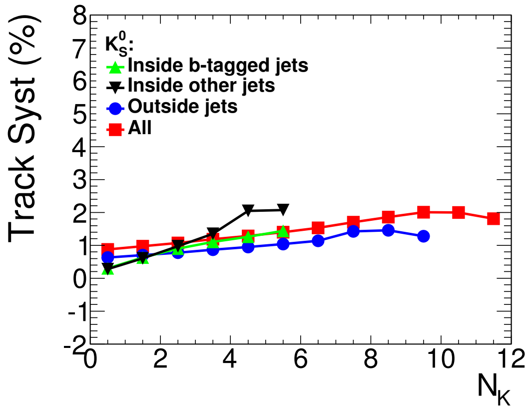
In other to estimate the systematics due to tracking inefficiencies, and taking into account that MC samples for $t\bar{t}$ events with different GEANT versions are not available, use has been made of Table 4 in reference [140] where the systematics due to tracking inefficiencies as a function of pseudorapidity and transverse momentum are given. Typically these are of the order of 2% for tracks in minimum-bias events. The followed procedure is therefore to loop over neutral strange particle candidates, determining the indices of their tracks and deleting them if a random number generator gives a value below the tabulated systematic. The results, obtained by fitting the resulting K^0_S mass distributions before and after the deletion of the tracks, are summarised in Table 5.11 below.

Table 5.11: K^0_S systematics due to tracking inefficiencies for each class.

K^0_S Class	Percentage (%)
Inside b-tagged jets	4.9
Inside non b-tagged jets	4.2
Outside any jet	5.1
Total sample	4.9

Notice that this table has been extended to contain the systematics not only for K^0_S outside jets, but also inside jets, b-tagged or not. In doing so, the validity of the systematics in reference [140], originally derived for minimum-bias events, has been assumed for tracks in a dense environment as given by a jet. In order to check whether there are additional systematics not taken into account in the procedure described above, possible deviations in the MC description of the K^0_S production dependence on the angular separation to the jet-axis (ΔR) and jet transverse momentum have been investigated. The results are shown in Figure 5.67 below.

In order to propagate this reco level systematics to the Bayesian unfolded multiplicity, the procedure is as follows. Firstly, the reco level data multiplicity distributions are shifted down by the amount given in Table 5.11. Then these shifted distributions are unfolded using the nominal transfer matrices, and the resulting relative differences between these and the nominal unfolded distributions are assigned as systematic uncertainty, see Figure 5.66.

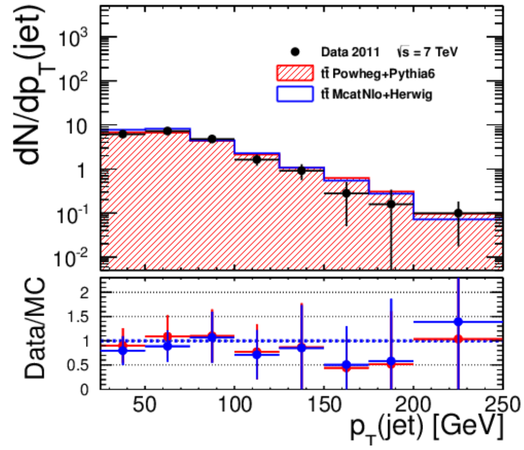


(a) Multiplicity distribution

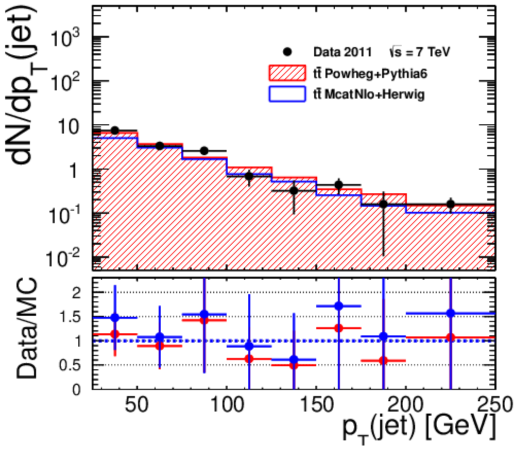
Figure 5.66: K_S^0 unfolded multiplicity systematic uncertainty due to tracking inefficiencies.

The following comments are in order:

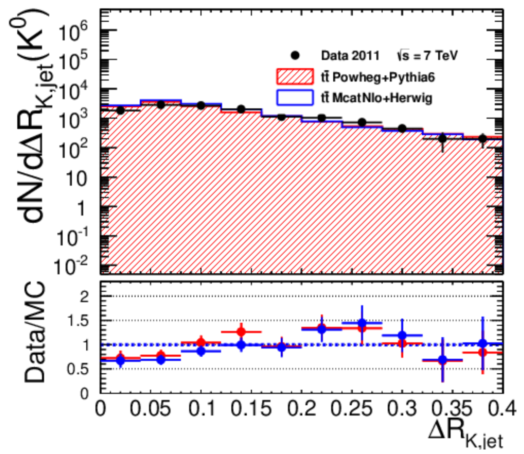
- the systematics due to tracking inefficiencies discussed above are smaller than the statistical errors affecting the K_S^0 signal reconstruction
- the MC description of the K_S^0 production dependence on ΔR and jet transverse momentum is fair and shows no indication of additional systematic effects in jet-like configurations
- therefore they are not considered any further.



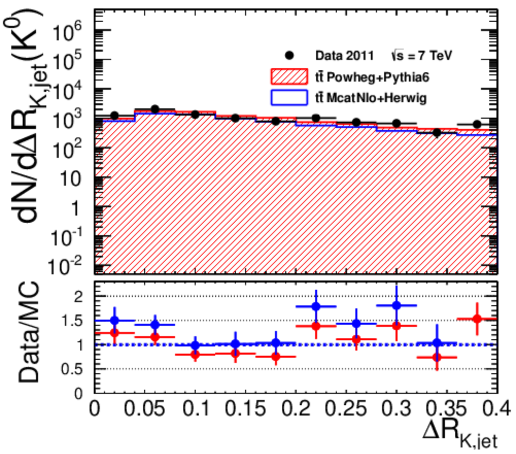
(a) b-tagged jets



(b) Non b-tagged jets



(c) b-tagged jets



(d) Non b-tagged jets

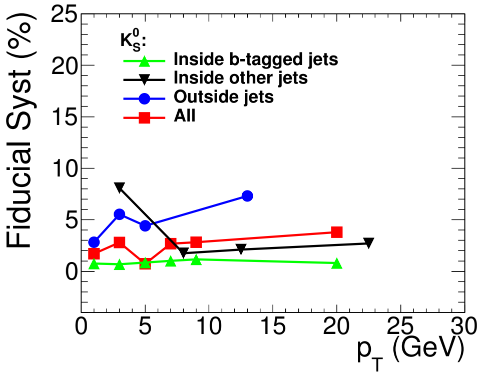
Figure 5.67: K^0_{S} production inside jets, b-tagged (left) or not (right), as a function of jet p_{T} (top) and ΔR (bottom) for data as compared to POWHEG+PYTHIA6 (red) and MC@NLO+HERWIG (blue) MC simulations.

5.7.7 Out of fiducial events

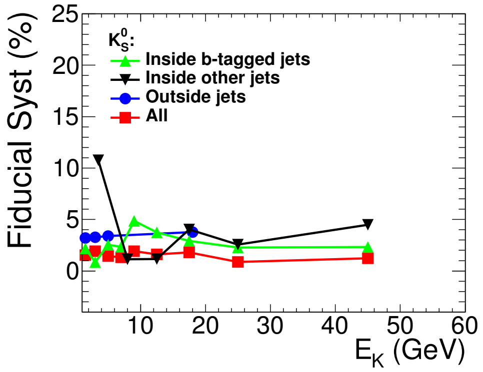
In order to take into account effects due to out of fiducial events,a comparison is made between K^0_{S}'s and Λ's distributions at particle level, for two samples of events selected using the reco or the particle level criteria, using the POWHEG+PYTHIA6 generator. To be more precise, $t\bar{t}$ dileptonic final states at particle level are requested to contain:

- two opposite charge leptons, electrons or muons, coming from W decays, with the same cuts in isolation, pseudorapidity and transverse momentum as applied at reco level
- at least two jets with the same fiducial cuts as those at reco level, and the same lepton-jet overlap removal criteria
- at least one of these jets should be a b-jet with the definition given in subsection 5.6.1.

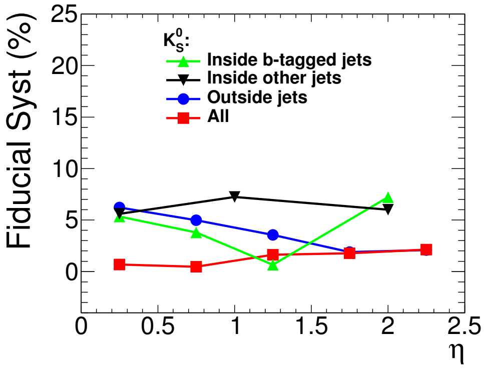
For ee and $\mu\mu$ final states the missing transverse energy is calculated from the neutrinos coming from W decays and the same cuts in the dilepton mass are applied as at reco level. While for $e\mu$ final states the H_{T} variable is defined as at reco level. It should be noted that these last requirements, are essentially irrelevant.
The results of this MC comparison are shown in Figures 5.68 and 5.69 below for the p_{T}, $|\eta|$, energy, energy fraction and multiplicity distribution of K_{S}^{0}'s and Λ's. Typically these systematics due to out of fiducial events are below 5%.



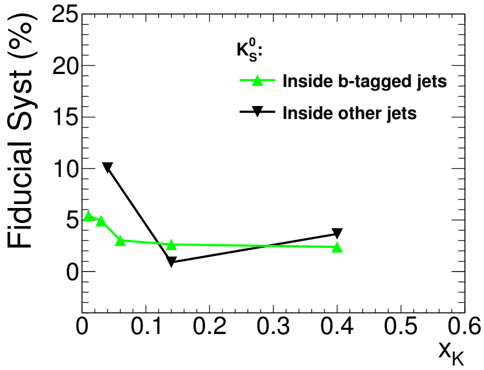
(a) p_{T} distribution



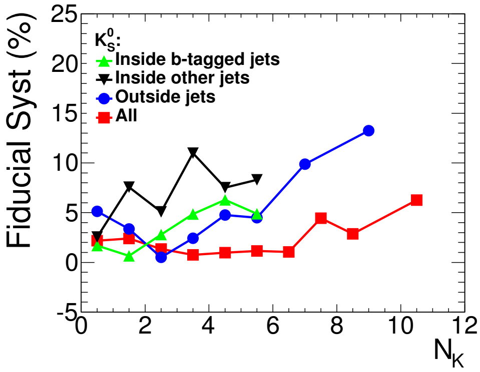
(b) Energy distribution



(c) $|\eta|$ distribution

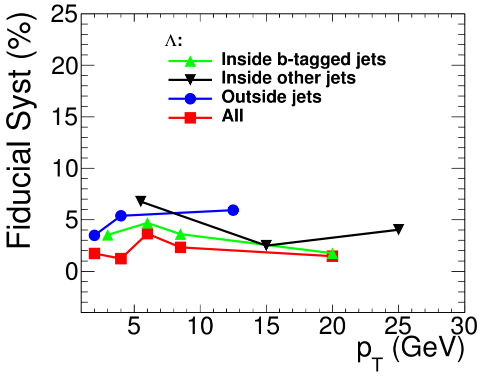


(d) Energy fraction distribution

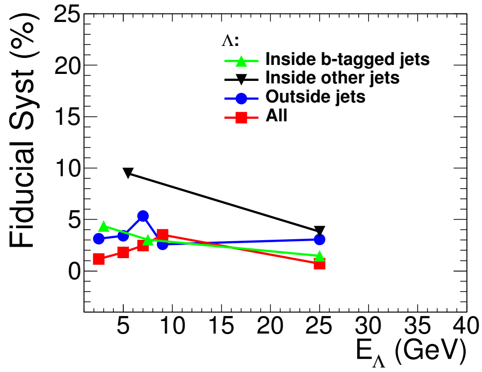


(e) Multiplicity distribution

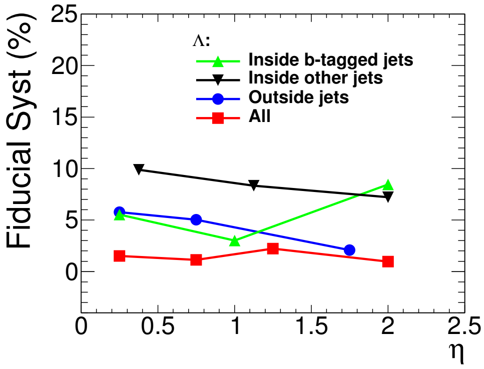
Figure 5.68: The K^0_{S} systematic uncertainties due to out of fiducial events, using Powheg+Pythia6, as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K^0_{S} inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).



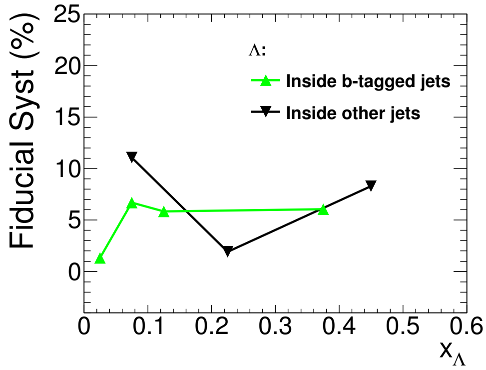
(a) p_{T} distribution



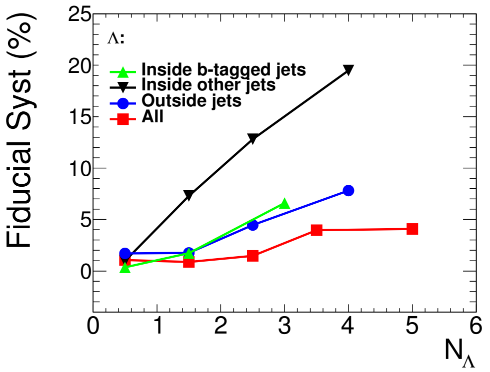
(b) Energy distribution



(c) $|\eta|$ distribution



(d) Energy fraction distribution

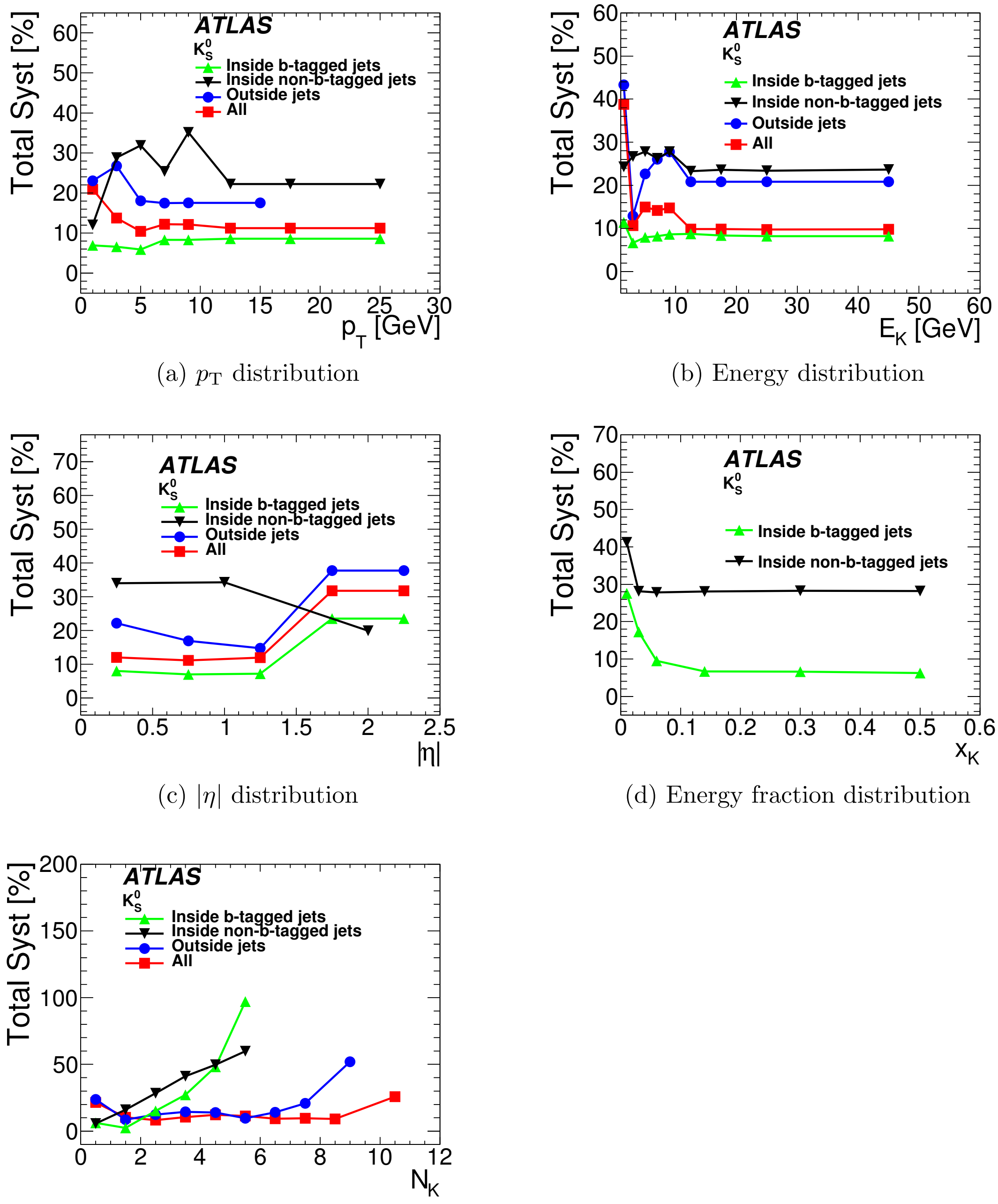


(e) Multiplicity distribution

Figure 5.69: The Λ systematic uncertainties due to out of fiducial events, using POWHEG+PYTHIA6, as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for Λ inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

5.7.8 Total systematics

In Figures 5.70 and 5.71 the values of the relative sytematic uncertainties for all the K^0_S and Λ distributions and classes are shown. These total systematic uncertainty distributions have been calculated as the quadratic sum of the previously mentioned systematics.
As already mentioned in Section 5.6, the systematic uncertainty due to the fiducial definition of the dileptonic $t\bar{t}$ final states is at the per cent level.



(a) p_{T} distribution

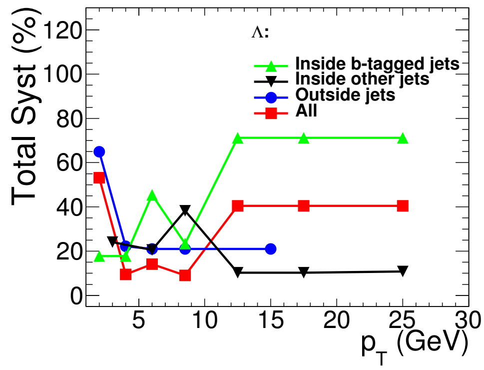
(b) Energy distribution

(c) $|\eta|$ distribution

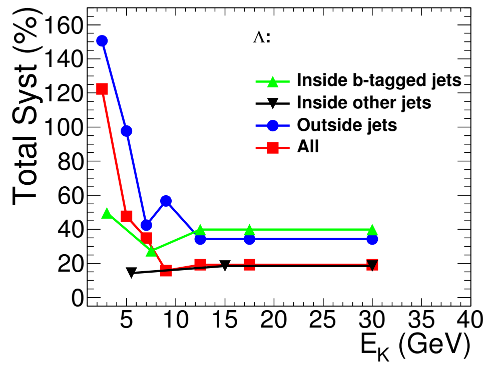
(d) Energy fraction distribution

(e) Multiplicity distribution

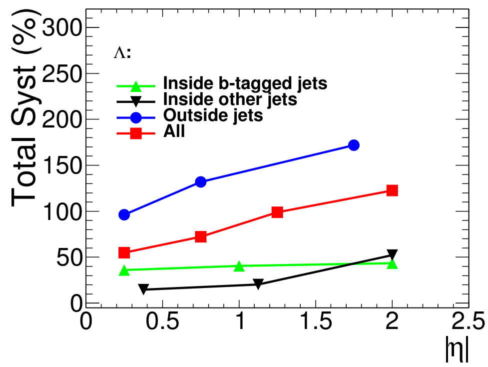
Figure 5.70: The K^0_{S} total systematic uncertainties as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K^0_{S} inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).



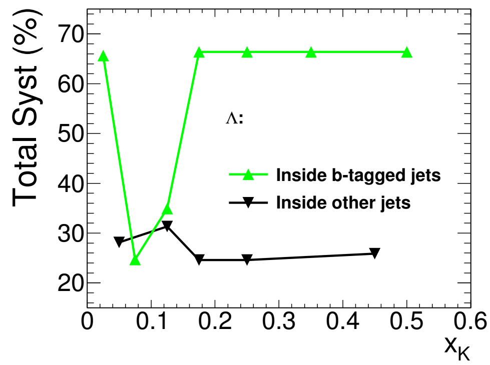
(a) p_{T} distribution



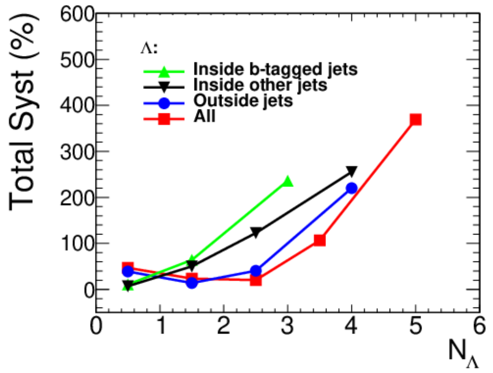
(b) Energy distribution



(c) $|\eta|$ distribution



(d) Energy fraction distribution



(e) Multiplicity distribution

Figure 5.71: The Λ total systematic uncertainties as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K) and multiplicity. Four classes are considered, for K^0_{S} inside b-jets (green), inside non b-tagged jets (black), outside any jet (blue), and all the sample (red).

5.8 Results at the particle level

The K^0_S and Λ yields for corrected data and POWHEG+PYTHIA6 MC at particle level are shown in Tables 5.12 to 5.14, as obtained from the unfolded p_T and multiplicity ($N_{K,\Lambda}$) spectra.

Table 5.12: K^0_S unfolded (particle level) average multiplicities per event for each class and for the total sample.

	Unfolded data from p_T	Unfolded data from N_K
Class	$< n_K > \pm(stat) \pm (syst)$	$< n_K > \pm(stat) \pm (syst)$
Inside b-jets	$0.90 \pm 0.06 \pm 0.007$	$0.91 \pm 0.03 \pm 0.07$
Inside non b-jets	$0.41 \pm 0.03 \pm 0.03$	$0.41 \pm 0.02 \pm 0.08$
Outside any jet	$2.92 \pm 0.09 \pm 0.40$	$2.66 \pm 0.07 \pm 0.12$
Total sample	$4.26 \pm 0.13 \pm 0.32$	$3.98 \pm 0.10 \pm 0.08$

Table 5.13: Λ unfolded (particle level) average multiplicities per event for each class and for the total sample.

	Unfolded data from p_T	Unfolded data from N_Λ
Class	$< n_\Lambda > \pm(stat) \pm (syst)$	$< n_\Lambda > \pm(stat) \pm (syst)$
Inside b-jets	$0.14 \pm 0.03 \pm 0.001$	0.112 ± 0.005
Inside non b-jets	$0.08 \pm 0.02 \pm 0.005$	0.095 ± 0.005
Outside any jet	$0.51 \pm 0.07 \pm 0.07$	0.57 ± 0.01
Total sample	$0.65 \pm 0.07 \pm 0.05$	0.78 ± 0.02

Table 5.14: K^0_S and Λ particle-level POWHEG+PYTHIA6 MC average multiplicities per event for each class and for the total sample.

	MC PYTHIA6 particle	
Class	$< n_K > \pm(stat)$	$< n_\Lambda > \pm(stat)$
Inside b-jets	0.917 ± 0.003	0.095 ± 0.005
Inside non b-jets	0.397 ± 0.002	0.091 ± 0.005
Outside any jet	2.248 ± 0.004	0.313 ± 0.011
Total sample	3.563 ± 0.005	0.499 ± 0.014

The average multiplicities obtained from the p_T distributions, which are unfolded with the efficiency correction only, are found in good agreement within statistical

uncertainties with those obtained from the multiplicity distributions unfolded by the iterative bayesian method with the transfer matrices discussed in Section 5.6.2. This indicates that the unfolding of the multiplicity is trustworthy. As a further cross-check, average multiplicities obtained with the set of multiplicity transfer matrices, where only visible decays at particle level are considered, have been recalculated. When corrected for the branching ratio $K^0_{\rm S} \rightarrow \pi^0\pi^0$, they have been found in good agreement with the ones presented previously.
A complete set of results at particle level can be found in Ref. [142].

5.8.1 $K^0_{\rm S}$ unfolded distributions

The unfolded distributions in the $p_{\rm T}$, E, and $|\eta|$ for $K^0_{\rm S}$ production inside b-jets, inside non b-jets, outside any jet and for the total sample are shown in Figures 5.72 to 5.75. Furthermore, for $K^0_{\rm S}$ production inside jets, the distribution of the energy fraction, $\rm x_K = \frac{E_K}{E_{Jet}}$, is also shown. They are compared to current MC expectations, namely POWHEG+PYTHIA6, MC@NLO+HERWIG, SHERPA, POWHEG+PYTHIA8, POWHEG+HERWIG7 and aMC@NLO+HERWIG7.
In general, the MC distributions reproduce the behaviour of the data sample reasonably well. This is expected since jet fragmentation functions have been studied from $Sp\bar{p}S$ to Tevatron energies, so the MC simulations are tuned fairly well. This is illustrated in the data to MC ratios shown in Figures 5.72 to 5.75, which indeed indicate that these are dominated by statistical rather than systematic uncertainties, in particular in the tails of the distributions.
To be more quantitative on this comparison between data and MC predictions a χ^2 test has been performed for the distributions shown in Figures 5.72 to 5.75. The χ^2 is defined as:

$$\chi^2 = V^T \cdot Cov^{-1} \cdot V \tag{5.12}$$

where V is the vector of differences between MC predictions and unfolded data, and Cov^{-1} denotes the inverse of the covariance matrix.
The covariance matrix has been obtained from the Bootstrap generator tool using one thousand replicas of the corresponding unfolded data distributions. In order to include systematic effects, these replicas have been smeared with Gaussians whose widths are given by the systematic errors considered as uncorrelated.
The following comments are in order:

- On average, the POWHEG+PYTHIA6 or PYTHIA8 and SHERPA generators give a very similar description of the data, while MC@NLO+HERWIG, aMC@NLO+HERWIG7 and POWHEG+HERWIG7 are slightly disfavoured.
- In general, the MC distributions reproduce the $K^0_{\rm S}$ particle spectra inside jets rather well. This is expected since jet fragmentation functions are

studied from $\mathrm{S}p\bar{p}\mathrm{S}$ to Tevatron energies, so the MC simulations are tuned fairly well.

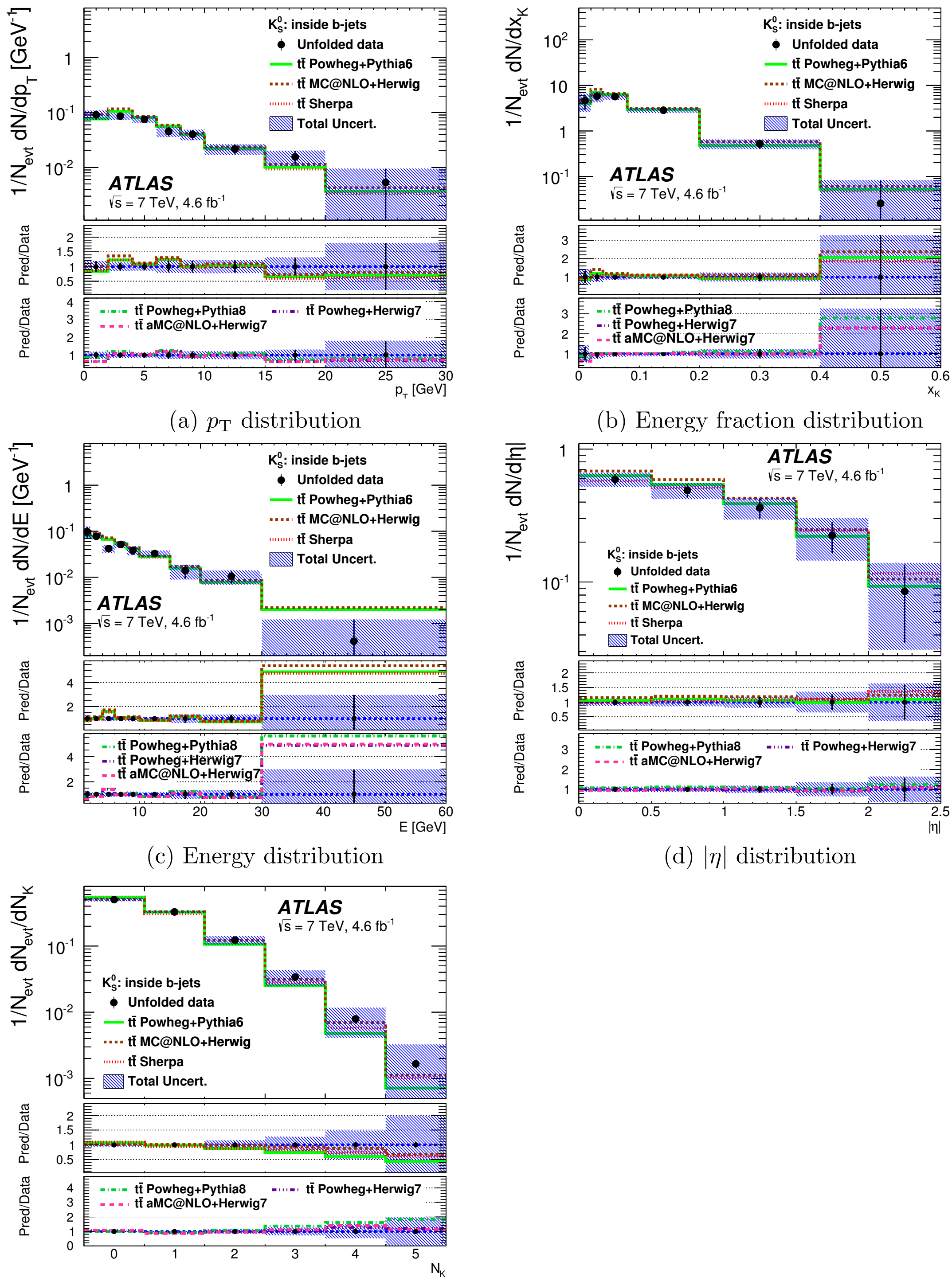
- The spectra for K^0_S production outside jets are reproduced in shape, but are underestimated by approximately 30%. This observation is consistent with a CMS study of strange particle production in the UE [88]. These data could be used to improve the simulation of the UE, especially to tune the $\gamma_\mathrm{s} = s/u$ parameter. This parameter needs to be larger than 0.2, the value used in the PYTHIA6+PERUGIA2011C tune [24]. The PYTHIA8+A14 predictions come closer to the data than the PYTHIA6+PERUGIA2011C ones. This is attributed to the fact that the A14 tune uses γ_s value equal to 0.217, as in the MONASH tune [143], which is 10 % larger than that in the default PYTHIA6+PERUGIA2011C tune. HERWIG+JIMMY and HERWIG7+H7UE gives a somewhat worse description than PYTHIA6+PERUGIA2011C, which indicates the need to also tune the strangeness suppression here or even to use an improved colour reconnection scheme for MPI as suggested in Ref. [144]. SHERPA, which uses strangeness suppression of $\gamma_\mathrm{s} = 0.4$, tends to overestimate the K^0_S yields outside jets.

- The energy and transverse momentum spectra for K^0_S mesons inside b-jets are similar to the spectra for those inside non-b-jets. The spectra for K^0_S mesons produced outside jets are much softer than for those produced in association with a jet.

- The pseudorapidity distributions for K^0_S mesons produced outside jets are constant over a wider central plateau than for those produced in association with a jet.

- Numerical values for unfolded up to non visible decays transverse momentum and pseudorapidity distributions are presented in Tables 5.15 to 5.20

- the results of the χ^2 test discussed above are summarised in Tables from 5.21 to 5.24.



(a) p_{T} distribution

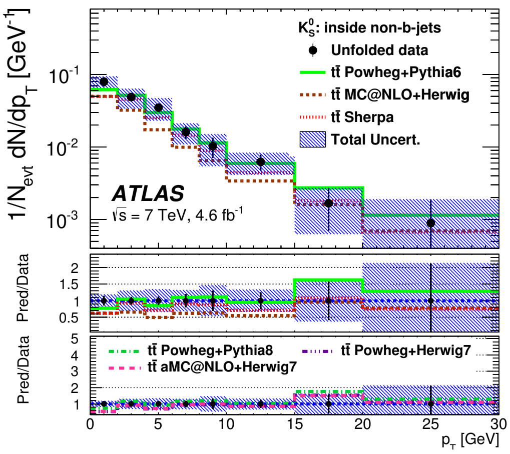
(b) Energy fraction distribution

(c) Energy distribution

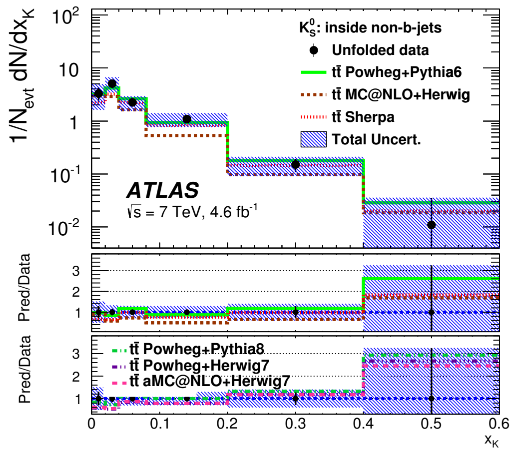
(d) $|\eta|$ distribution

(e) Multiplicity distribution

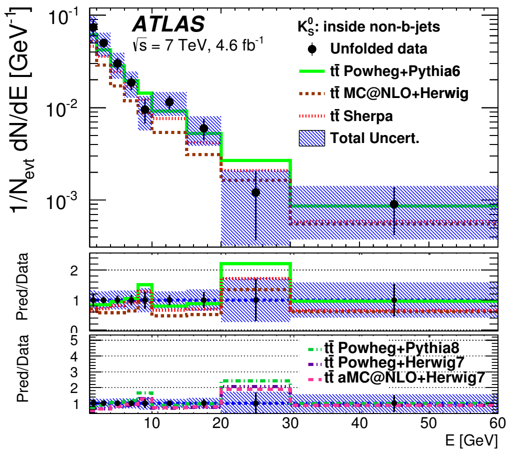
Figure 5.72: Kinematic characteristics for K^0_{S} production inside b-jets, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig, Sherpa, Powheg+Pythia8, Powheg+Herwig7 and aMC@NLO+Herwig7 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



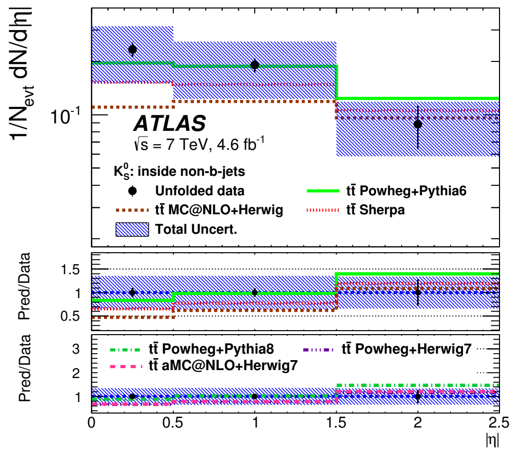
(a) p_{T} distribution



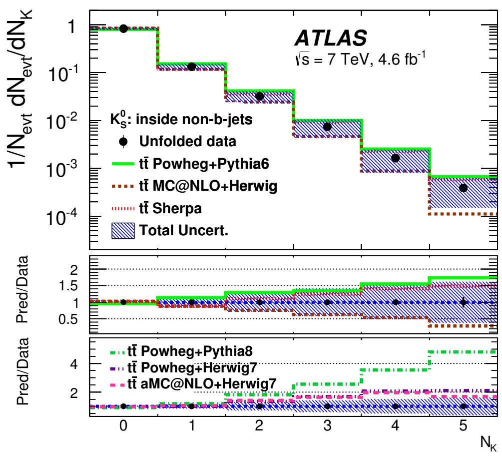
(b) Energy fraction distribution



(c) Energy distribution

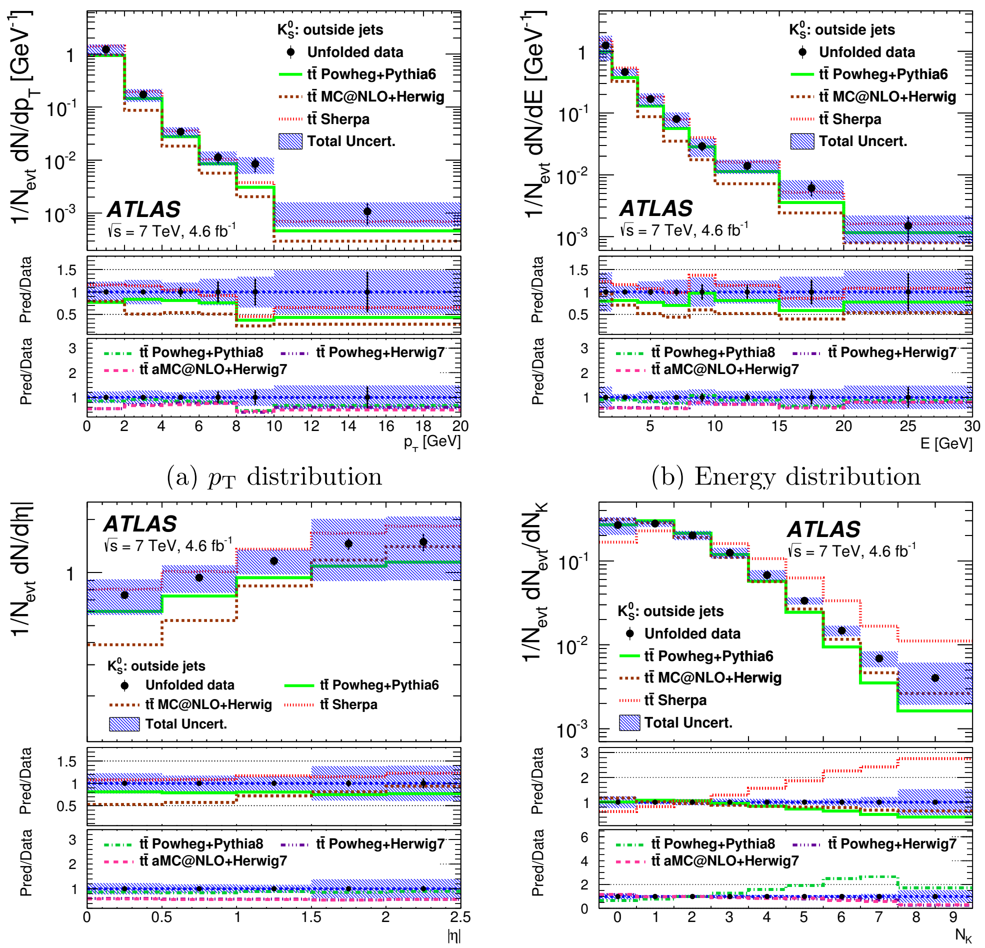


(d) $|\eta|$ distribution



(e) Multiplicity distribution

Figure 5.73: Kinematic characteristics for K^0_S production inside non b-jets, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig, Sherpa, Powheg+Pythia8, Powheg+Herwig7 and aMC@NLO+Herwig7 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



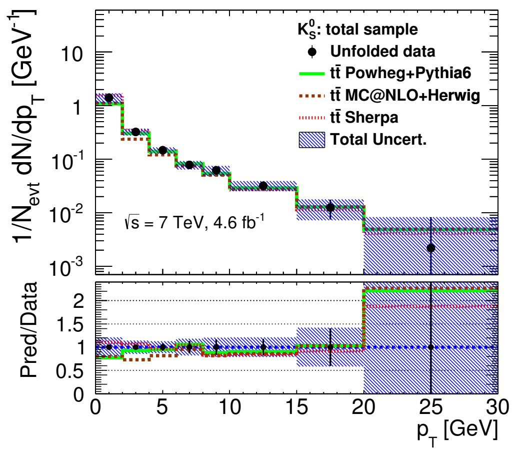
(a) p_{T} distribution

(b) Energy distribution

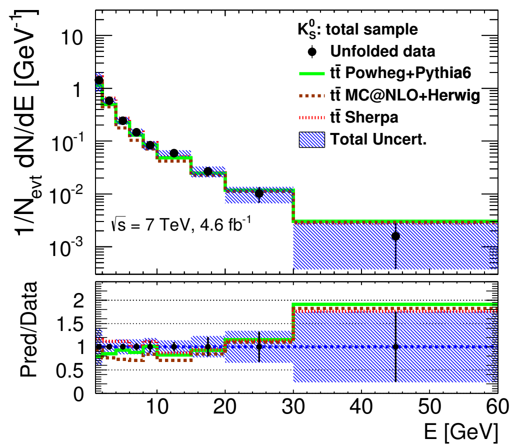
(c) $|\eta|$ distribution

(d) Multiplicity distribution

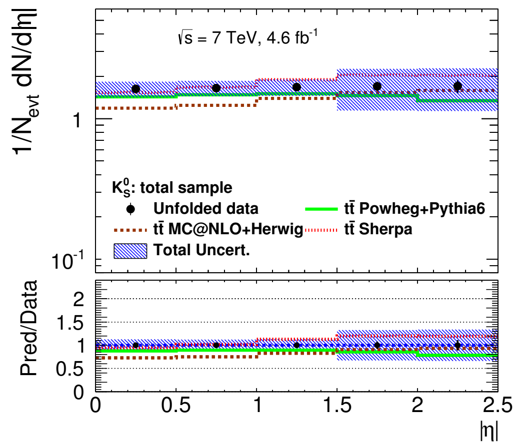
Figure 5.74: Kinematic characteristics for K^0_S production not associated with jets, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig, Sherpa, Powheg+Pythia8, Powheg+Herwig7 and aMC@NLO+Herwig7 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



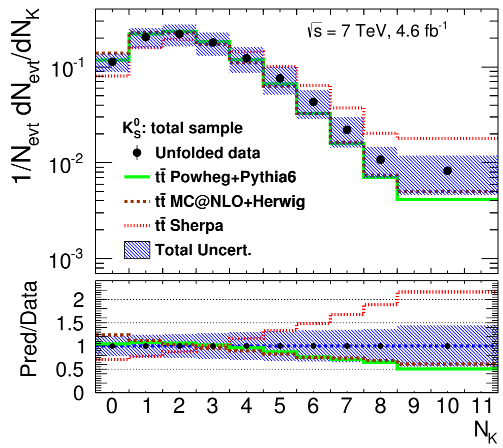
(a) p_{T} distribution



(b) Energy distribution



(c) $|\eta|$ distribution



(d) Multiplicity distribution

Figure 5.75: Kinematic characteristics for the total K_S^0 production, for corrected data and particle level MC simulated with the POWHEG+PYTHIA6, MC@NLO+HERWIG and SHERPA generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.

Table 5.15: Transverse momentum distribution unfolded to particle level for K_S^0 not associated with jets, and including invisible decays, along with the statistical and systematic uncertainties

p_T[GeV]	$\frac{1}{N_{evt}}\frac{dN}{dp_T}$	Stat.	Model	Track	JES	JER	PU	Fid.
(0, 2)	1.22	0.05	0.25	0.06	$^{+0.02}_{-0.02}$	$< 10^{-2}$	0.10	0.03
(2, 4)	0.172	0.009	0.042	0.009	$^{+0.002}_{-0.002}$	0.001	0.014	0.010
(4, 6)	0.035	0.004	0.005	0.002	$^{+0.001}_{-0.001}$	$< 10^{-3}$	0.003	0.002
(6, 8)	0.011	0.003	0.001	0.001	$^{+0.001}_{-0.001}$	0.001	0.001	0.001
(8, 10)	0.0085	0.0026	0.0007	0.0004	$^{+0.0006}_{-0.0004}$	0.0005	0.0007	0.0006
(10, 20)	0.0011	0.0005	0.0001	0.0001	$^{+0.0001}_{-0.0001}$	0.0001	0.0001	0.0001

Table 5.16: Transverse momentum distribution unfolded to the particle level for K_S^0 associated with b-jets, and including invisible decays, along with the statistical and systematic uncertainties.

p_T[GeV]	$\frac{1}{N_{evt}}\frac{dN}{dp_T}$	Stat.	Model	Track	JES	JER	PU	Fid.
(0, 2)	0.091	0.017	0.003	0.004	$^{+0.002}_{-0.002}$	0.001	$< 10^{-3}$	0.001
(2, 4)	0.085	0.012	0.003	0.004	$^{+0.002}_{-0.002}$	0.001	$< 10^{-3}$	0.001
(4, 6)	0.075	0.010	0.002	0.004	$^{+0.001}_{-0.001}$	0.001	$< 10^{-3}$	0.001
(6, 8)	0.045	0.009	0.002	0.002	$^{+0.001}_{-0.001}$	0.001	$< 10^{-3}$	0.001
(8, 10)	0.040	0.008	0.002	0.002	$^{+0.001}_{-<10^{-3}}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$
(10, 15)	0.022	0.004	0.001	0.001	$^{+<10^{-3}}_{-<10^{-3}}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$
(15, 20)	0.0156	0.0046	0.0010	0.0008	$^{+0.0003}_{-0.0002}$	0.0001	$< 10^{-4}$	0.0001
(20, 30)	0.0058	0.0050	0.0004	0.0003	$^{+0.0001}_{-0.0001}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$

Table 5.17: Transverse momentum distribution unfolded to the particle level for K_S^0 associated with non-b-jets, and including invisible decays, along with the statistical and systematic uncertainties.

p_T[GeV]	$\frac{1}{N_{evt}}\frac{dN}{dp_T}$	Stat.	Model	Track	JES	JER	PU	Fid.
(0, 2)	0.080	0.011	0.003	0.003	$^{+0.004}_{-0.005}$	0.002	$< 10^{-3}$	0.007
(2, 4)	0.049	0.006	0.013	0.002	$^{+0.002}_{-0.002}$	0.001	$< 10^{-3}$	0.004
(4, 6)	0.035	0.005	0.011	0.001	$^{+0.001}_{-0.001}$	0.001	$< 10^{-3}$	0.003
(6, 8)	0.016	0.003	0.004	0.001	$^{+0.001}_{-0.001}$	0.001	$< 10^{-3}$	$< 10^{-3}$
(8, 10)	0.010	0.003	0.003	0.001	$^{+<10^{-3}}_{-<10^{-3}}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$
(10, 15)	0.0063	0.0016	0.0014	0.0003	$^{+0.0001}_{-0.0001}$	$< 10^{-4}$	$< 10^{-4}$	0.0002
(15, 20)	0.0016	0.0010	0.0004	0.0001	$^{+<10^{-4}}_{-<10^{-4}}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
(20, 30)	0.0009	0.0011	0.0002	$< 10^{-4}$	$^{+<10^{-4}}_{-<10^{-4}}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$

Table 5.18: Pseudorapidity distribution unfolded to the particle level for K_S^0 not associated with jets, and including invisible decays, along with the statistical and systematic uncertainties.

$\lvert\eta\rvert$	$\frac{1}{N_{evt}}\frac{dN}{d\lvert\eta\rvert}$	Stat.	Model	Track	JES	JER	PU	Fid.
(0.0, 0.5)	0.744	0.039	0.141	0.038	$^{+0.007}_{-0.007}$	0.004	0.062	0.046
(0.5, 1.0)	0.933	0.049	0.120	0.048	$^{+0.007}_{-0.009}$	0.008	0.077	0.046
(1.0, 1.5)	1.162	0.068	0.122	0.059	$^{+0.010}_{-0.008}$	0.004	0.096	0.041
(1.5, 2.0)	1.449	0.110	0.527	0.074	$^{+0.021}_{-0.018}$	0.021	0.120	0.030
(2.0, 2.5)	1.489	0.172	0.541	0.076	$^{+0.019}_{-0.019}$	0.022	0.123	0.031 6

Table 5.19: Pseudorapidity distribution unfolded to the particle level for K^0_S associated with b-jets, and including invisible decays, along with the statistical and systematic uncertainties.

$\lvert\eta\rvert$	$\frac{1}{N_\mathrm{evt}}\frac{\mathrm{d}N}{\mathrm{d}\lvert\eta\rvert}$	Stat.	Model	Track	JES	JER	PU	Fid.
(0.0, 0.5)	0.592	0.053	0.015	0.029	$^{+0.012}_{-0.012}$	0.009	$< 10^{-3}$	0.032
(0.5, 1.0)	0.488	0.056	0.011	0.024	$^{+0.009}_{-0.007}$	0.007	$< 10^{-3}$	0.019
(1.0, 1.5)	0.358	0.061	0.014	0.017	$^{+0.010}_{-0.009}$	0.007	$< 10^{-3}$	0.002
(1.5, 2.0)	0.226	0.059	0.049	0.011	$^{+0.005}_{-0.008}$	0.002	$< 10^{-3}$	0.016
(2.0, 2.5)	0.087	0.053	0.019	0.004	$^{+0.002}_{-0.003}$	$< 10^{-3}$	$< 10^{-3}$	0.006

Table 5.20: Pseudorapidity distribution unfolded to the particle level for K^0_S associated with non-b-jets, and including invisible decays, along with the statistical and systematic uncertainties.

$\lvert\eta\rvert$	$\frac{1}{N_\mathrm{evt}}\frac{\mathrm{d}N}{\mathrm{d}\lvert\eta\rvert}$	Stat.	Model	Track	JES	JER	PU	Fid.
(0.0, 0.5)	0.233	0.023	0.077	0.010	$^{+0.007}_{-0.009}$	0.002	$< 10^{-3}$	0.013
(0.5, 1.5)	0.190	0.017	0.063	0.008	$^{+0.006}_{-0.006}$	0.001	$< 10^{-3}$	0.014
(1.5, 2.5)	0.089	0.024	0.016	0.004	$^{+0.003}_{-0.004}$	0.003	$< 10^{-3}$	0.005

Table 5.21: Values of the χ^2 per degree of freedom and their corresponding p-values, for K^0_S production inside b-jets, for Powheg+Pythia6, MC@NLO+Herwig and Sherpa predictions.

K^0_S inside b-jets $\chi^2/n.d.f.$ (p-value)			
	Pythia6	Herwig	Sherpa
p_T	0.77 (0.63)	1.31 (0.23)	0.80 (0.61)
E	1.81 (0.06)	2.54 (0.007)	1.67 (0.09)
$\lvert\eta\rvert$	0.56 (0.73)	1.44 (0.21)	0.33 (0.90)
x_K	0.40 (0.88)	1.08 (0.37)	0.19 (0.98)
N_K	1.50 (0.19)	0.75 (0.59)	2.33 (0.04)

Table 5.22: Values of the χ^2 per degree of freedom and their corresponding p-values, for K^0_S production inside non b-jets, for Powheg+Pythia6, MC@NLO+Herwig and Sherpa predictions.

K^0_S inside non-b-jets $\chi^2/n.d.f.$ (p-value)			
	Pythia6	Herwig	Sherpa
p_T	0.42 (0.91)	1.32 (0.23)	0.71 (0.69)
E	1.25 (0.26)	1.58 (0.11)	1.12 (0.34)
$\lvert\eta\rvert$	0.90 (0.44)	0.64 (0.59)	0.43 (0.73)
x_K	0.82 (0.55)	1.51 (0.17)	0.78 (0.58)
N_K	1.80 (0.11)	1.45 (0.20)	0.85 (0.52)

Table 5.23: Values of the χ^2 per degree of freedom and their corresponding p-values, for K^0_S production outside jets, for Powheg+Pythia6, MC@NLO+Herwig and Sherpa predictions.

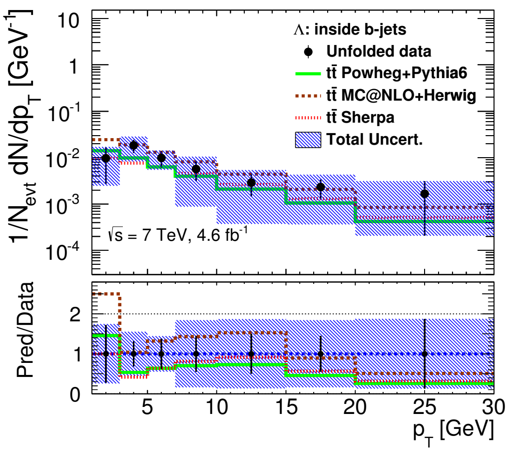
K^0_S outside jets $\chi^2/n.d.f.$ (p-value)			
	Pythia6	Herwig	Sherpa
p_T	0.98 (0.44)	2.64 (0.015)	0.93 (0.47)
E	1.05 (0.40)	2.67 (0.004)	1.14 (0.33)
$\lvert\eta\rvert$	0.94 (0.45)	2.30 (0.04)	1.36 (0.24)
N_K	3.13 (0.002)	2.10 (0.03)	34.17 (0.00)

Table 5.24: Values of the χ^2 per degree of freedom and their corresponding p-values, for total sample of K^0_S production, for POWHEG+PYTHIA6, MC@NLO+HERWIG and SHERPA predictions.

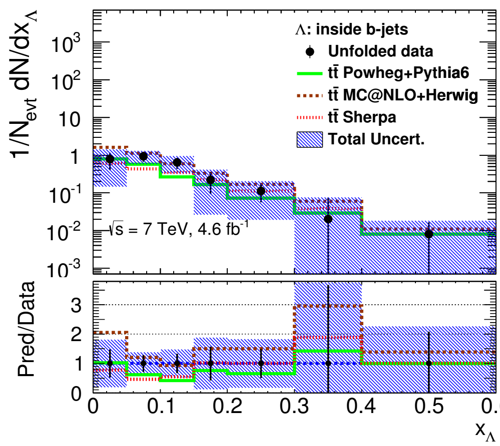
K^0_S total sample $\chi^2/n.d.f.$ (p-value)			
	PYTHIA6	HERWIG	SHERPA
p_T	0.15 (0.99)	0.91 (0.50)	0.47 (0.88)
E	1.28 (0.24)	3.16 (0.00)	1.34 (0.21)
$\|\eta\|$	0.78 (0.56)	2.71 (0.02)	1.09 (0.36)
N_K	10.33 (0.00)	8.77 (0.00)	78.12 (0.00)

5.8.2 Λ unfolded distributions

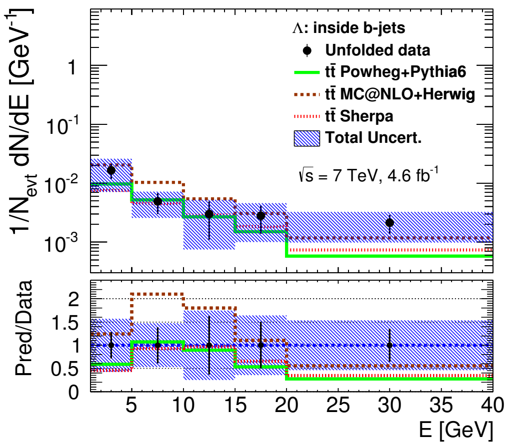
For the sake of completeness, the same distributions as in the previous subsection, are now presented for Λ production in Figures 5.76 to 5.79, along with a comparison to current MC predictions. In spite of the poorer statistics and higher systematics, similar features as for K^0_S production can be observed in these distributions. To be more precise, the p_T and energy distributions for Λ's associated with jets, are harder than those not associated with jets. Nevertheless one can not draw any conclusion from the unfolded multiplicity for lambdas.
Numerical values for unfolded up to non visible decays transverse momentum and pseudorapidity distributions are presented in Tables 5.25 to 5.26.
The results of a χ^2 test for the comparison between unfolded data and MC predictions are summarised in Table 5.27.



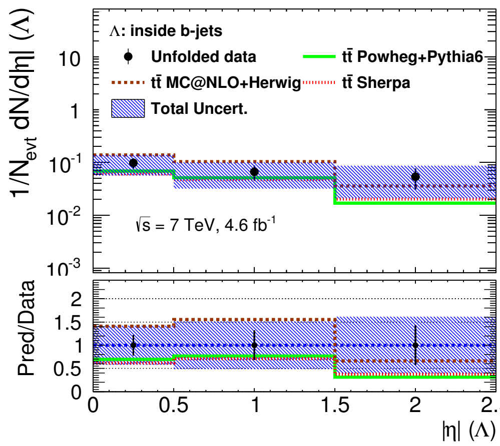
(a) p_{T} distribution



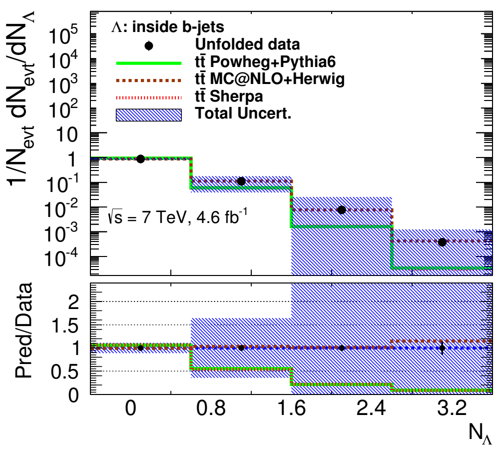
(b) Energy fraction distribution



(c) Energy distribution

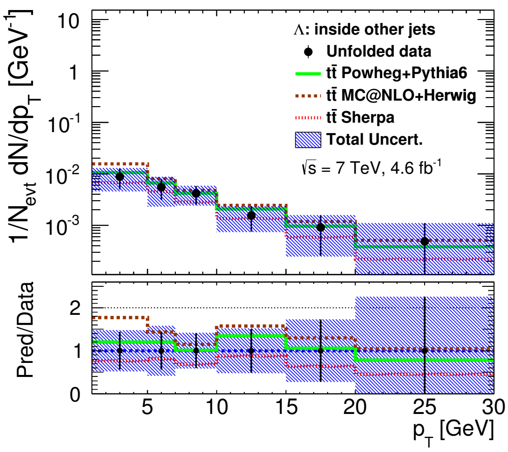


(d) $|\eta|$ distribution

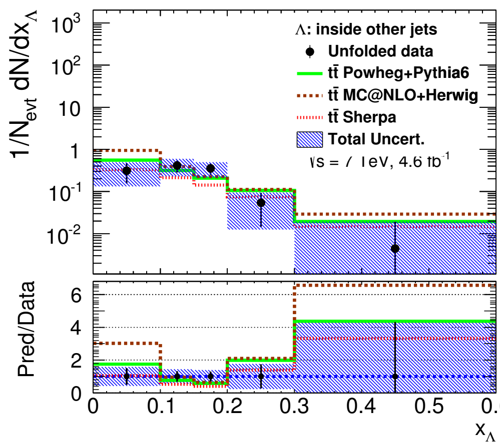


(e) Multiplicity distribution

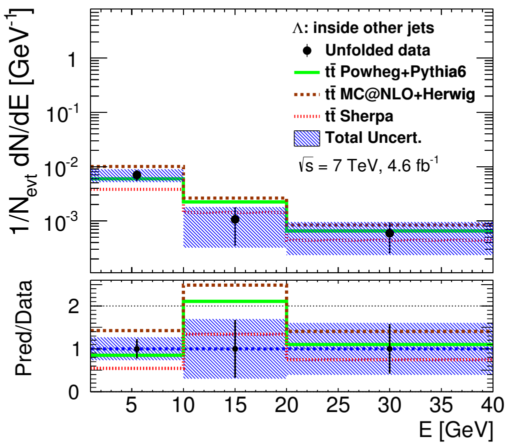
Figure 5.76: Kinematic characteristics for Λ production inside b-jets, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig and Sherpa generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



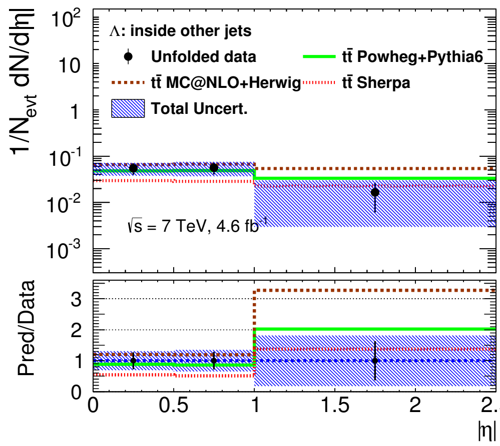
(a) p_{T} distribution



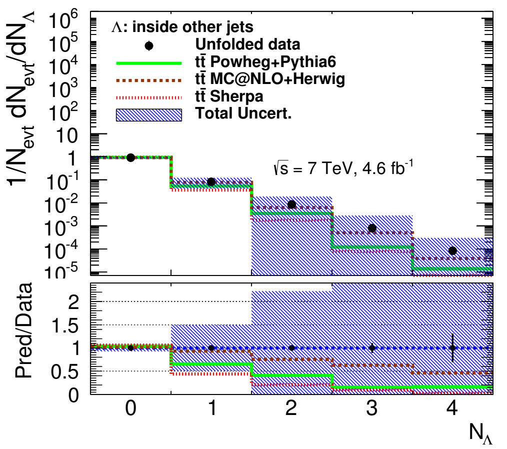
(b) Energy fraction distribution



(c) Energy distribution

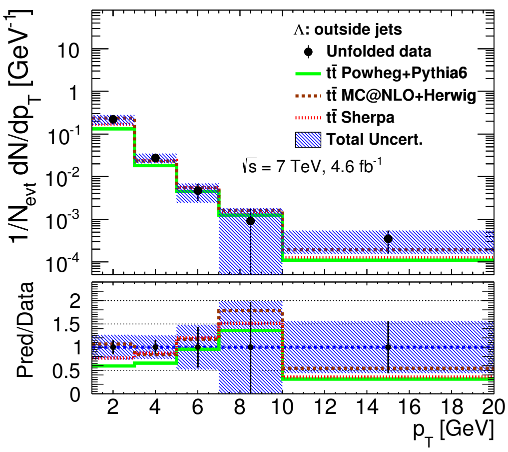


(d) $|\eta|$ distribution

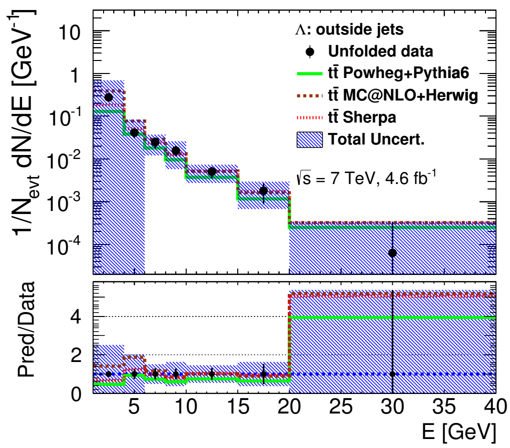


(e) Multiplicity distribution

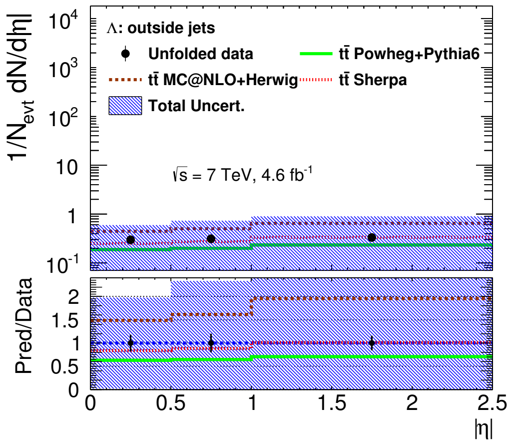
Figure 5.77: Kinematic characteristics for Λ production inside non b-jets, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig and Sherpa generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



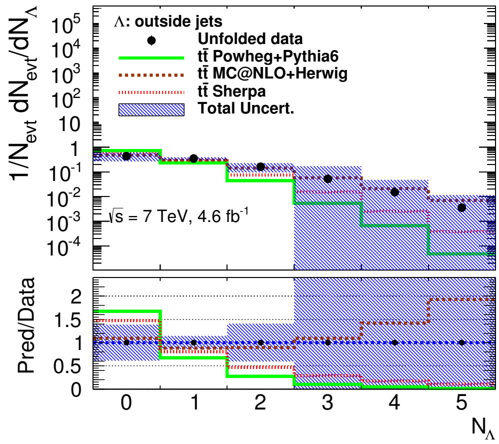
(a) p_{T} distribution



(b) Energy distribution

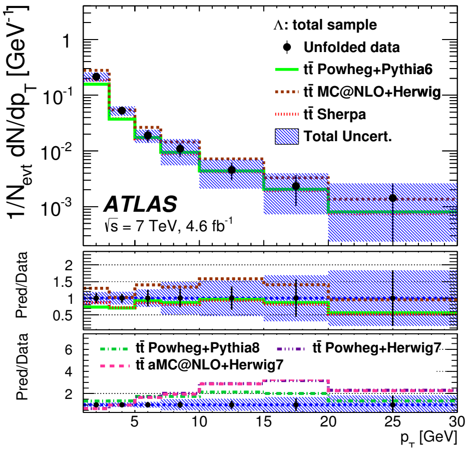


(c) $|\eta|$ distribution

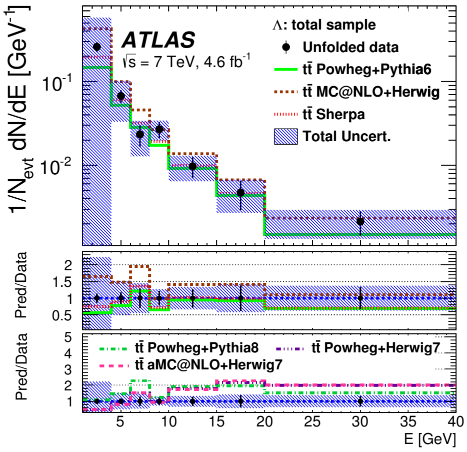


(d) Multiplicity distribution

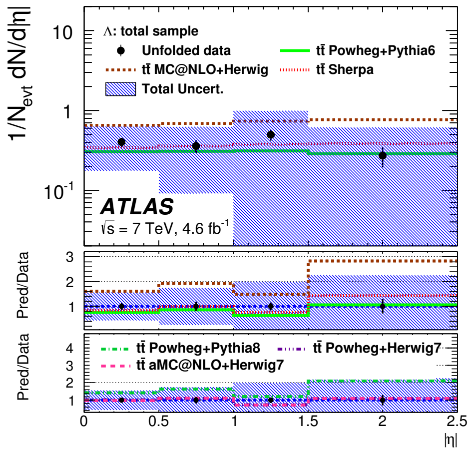
Figure 5.78: Kinematic characteristics for Λ production not associated with jets, for corrected data and particle level MC simulated with the POWHEG+PYTHIA6, MC@NLO+HERWIG and SHERPA generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



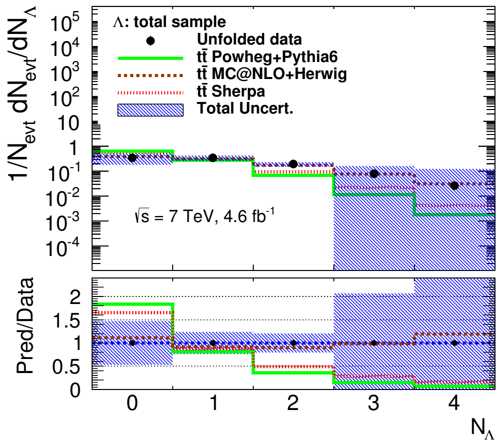
(a) p_{T} distribution



(b) Energy distribution



(c) $|\eta|$ distribution



(d) Multiplicity distribution

Figure 5.79: Kinematic characteristics for the total Λ production, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig, Sherpa, Powheg+Pythia8, Powheg+Herwig7 and aMC@NLO+Herwig7 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.

Table 5.25: Transverse momentum distribution unfolded to the particle level for the Λ total sample, and including invisible decays, along with the statistical and systematic uncertainties.

p_{T}[GeV]	$\frac{1}{N_{\mathrm{evt}}}\frac{\mathrm{d}N}{\mathrm{d}p_{\mathrm{T}}}$	Stat.	Model	Track	JES	JER	PU	Fid.
(1, 3)	0.215	0.033	0.114	0.011	$^{+0.003}_{-0.004}$	0.002	0.017	0.004
(3, 5)	0.053	0.007	0.004	0.003	$^{+0.001}_{-0.001}$	0.001	0.004	0.001
(5, 7)	0.019	0.005	0.002	0.001	$^{+<10^{-3}}_{-<10^{-3}}$	0.001	0.001	0.001
(7, 10)	0.011	0.003	0.001	0.001	$^{+<10^{-3}}_{-<10^{-3}}$	$< 10^{-3}$	0.001	$< 10^{-3}$
(10, 15)	0.0047	0.0016	0.0019	0.0002	$^{+0.0001}_{-0.0001}$	0.0001	0.0004	0.0001
(15, 20)	0.0025	0.0013	0.0010	0.0001	$^{+<10^{-4}}_{-<10^{-4}}$	0.0001	0.0002	$< 10^{-4}$
(20, 30)	0.0016	0.0035	0.0006	0.0001	$^{+<10^{-4}}_{-<10^{-4}}$	$< 10^{-4}$	0.0001	$< 10^{-4}$

Table 5.26: Pseudorapidity distribution unfolded to the particle level for the Λ total sample, and including invisible decays, along with the statistical and systematic uncertainties.

$\|\eta\|$	$\frac{1}{N_{\mathrm{evt}}}\frac{\mathrm{d}N}{\mathrm{d}\|\eta\|}$	Stat.	Model	Track	JES	JER	PU	Fid.
(0.0, 0.5)	0.404	0.050	0.220	0.020	$^{+0.009}_{-0.006}$	0.006	0.032	0.006
(0.5, 1.0)	0.361	0.070	0.260	0.018	$^{+0.006}_{-0.008}$	0.003	0.029	0.004
(1.0, 1.5)	0.498	0.076	0.492	0.024	$^{+0.011}_{-0.011}$	0.008	0.040	0.011
(1.5, 2.5)	0.273	0.075	0.334	0.013	$^{+0.003}_{-0.006}$	0.007	0.022	0.003

Table 5.27: Values of the χ^2 per degree of freedom and their corresponding p-values, for the total Λ sample, for POWHEG+PYTHIA6, MC@NLO+HERWIG and SHERPA predictions.

	$\chi^2/n.d.f.$ (p-value) Λ total sample		
	PYTHIA6	HERWIG	SHERPA
p_{T}	0.50 (0.83)	0.93 (0.48)	0.52 (0.82)
E	0.89 (0.51)	1.90 (0.06)	0.89 (0.51)
$\lvert\eta\rvert$	0.33 (0.86)	1.62 (0.17)	0.25 (0.91)

5.9 Comparisons with other MC samples

In previous Sections the observation has been made that neutral strange particle production outside jets is more sensitive to IFSR and UE effects than that inside jets. Therefore, the attempt is made in this Section to compare the particle level data with MC models where variations on

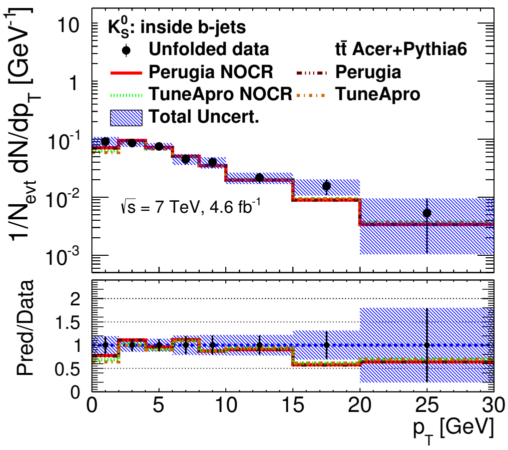
- UE tune
- fragmentation scheme
- colour reconnection effects

have been considered. Needless to say the the particle level MC distributions have been obtained after applying the same selection criteria as in Section 5.6.1.

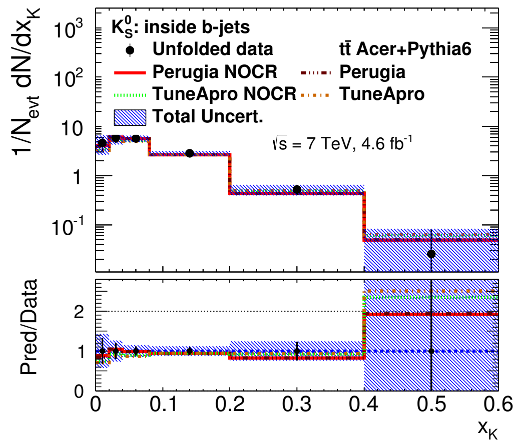
5.9.1 Acer+Pythia6: Perugia vs TuneAPro

In Figures 5.80 to 5.87, the results of a comparison to ACER+PYTHIA6 generators with PERUGIA2011C or TUNEAPRO tunes with and without colour reconnection effects, are presented. The results of a χ^2 test similar to that described in the previous subsection are summarised in Tables 5.28 to 5.31. The following conclusions can be drawn:

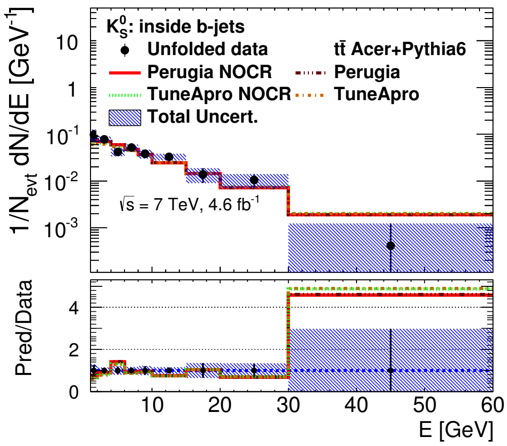
- no significant differences are found between these MCs and the reference POWHEG+PYTHIA6 as far as the description of neutral strange particle production inside jets is concerned
- Tune Apro is slightly disfavoured with respect to PERUGIA2011C in the description of the multiplicities for neutral strange particle outside jets.



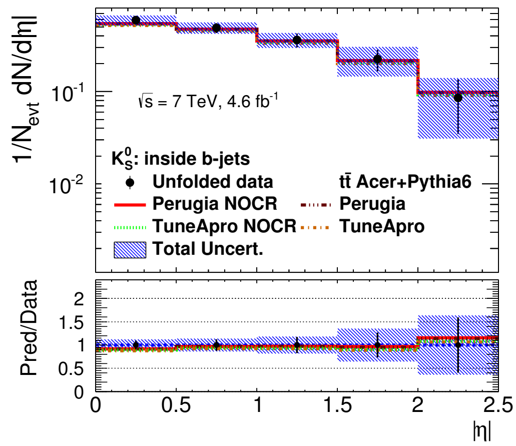
(a) p_{T} distribution



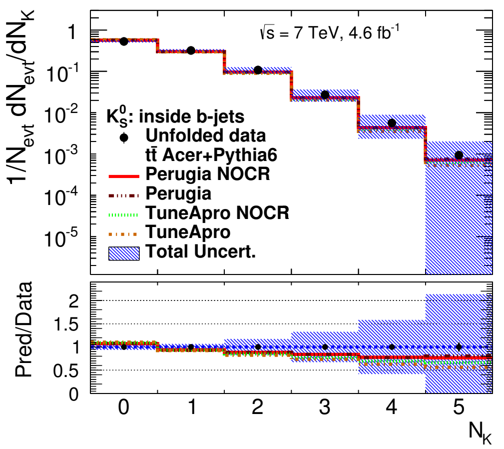
(b) Energy fraction distribution



(c) Energy distribution

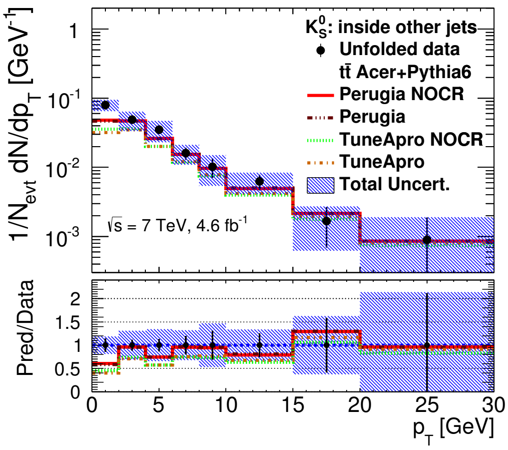


(d) $|\eta|$ distribution

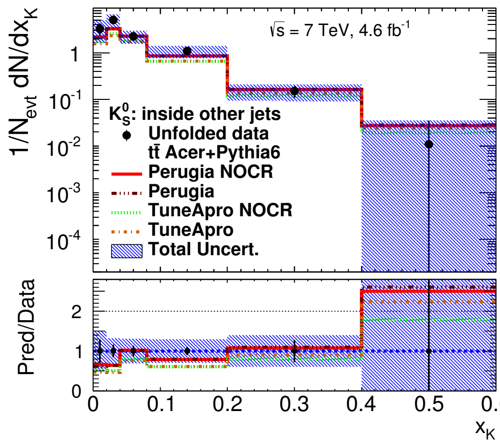


(e) Multiplicity distribution

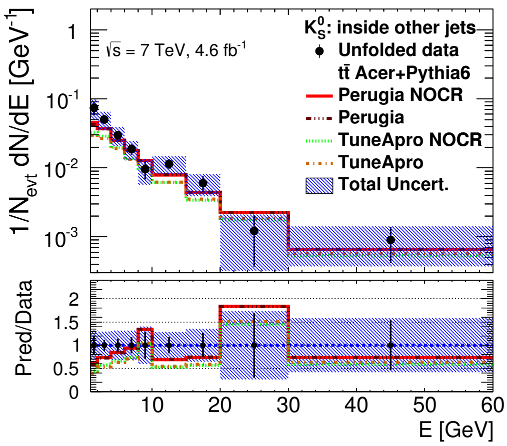
Figure 5.80: Kinematic characteristics for K^0_S production inside b-jets, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



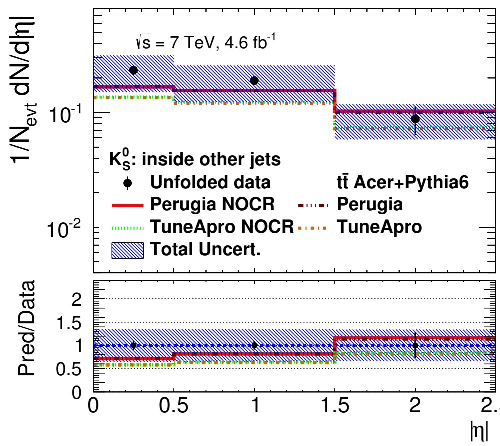
(a) p_{T} distribution



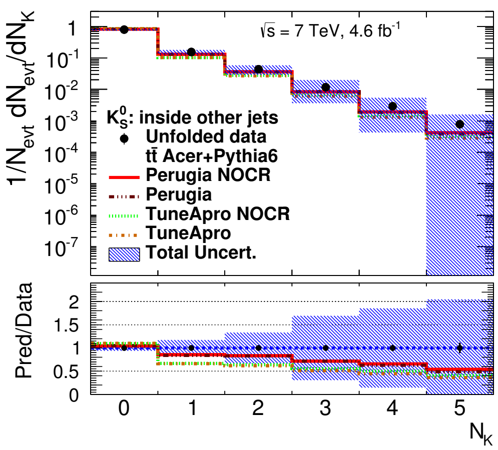
(b) Energy fraction distribution



(c) Energy distribution

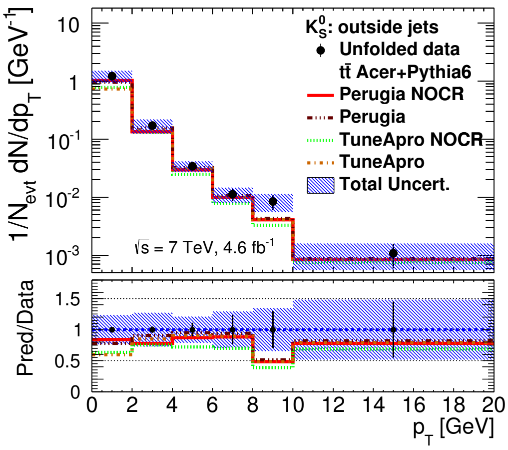


(d) $|\eta|$ distribution

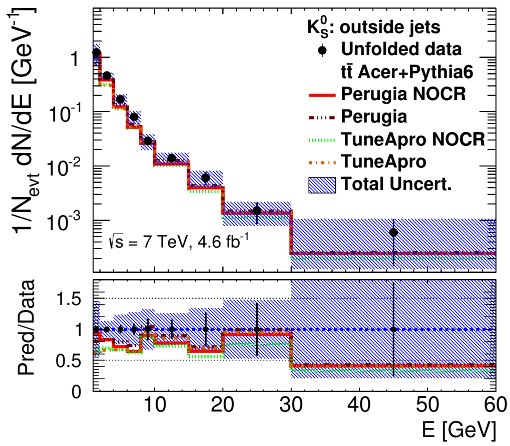


(e) Multiplicity distribution

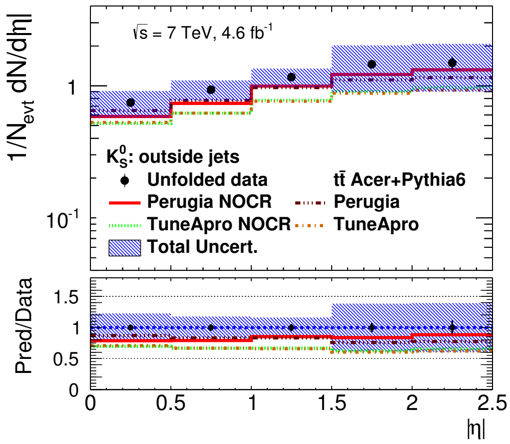
Figure 5.81: Kinematic characteristics for K^0_{S} production inside non b-jets, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



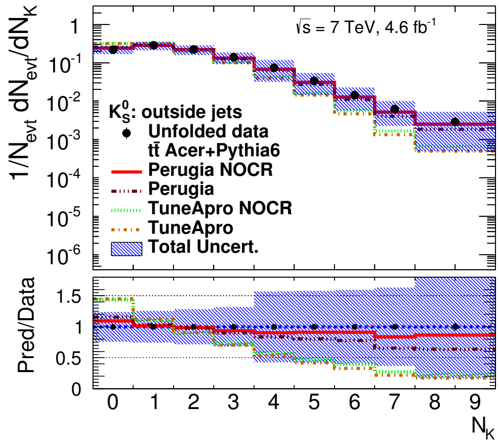
(a) p_{T} distribution



(b) Energy distribution

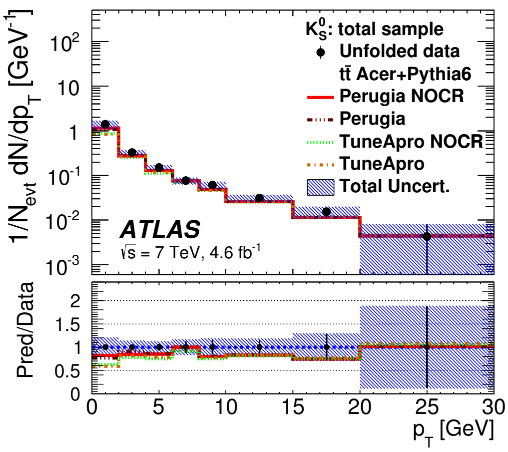


(c) $|\eta|$ distribution

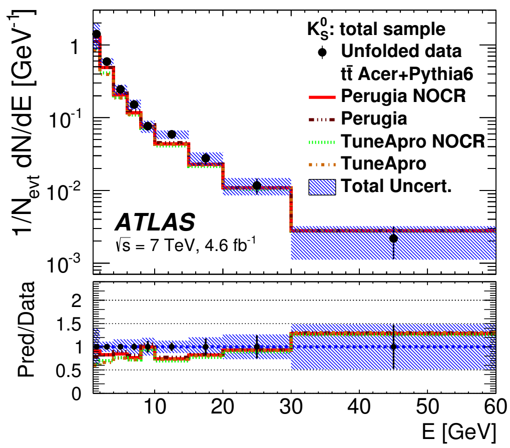


(d) Multiplicity distribution

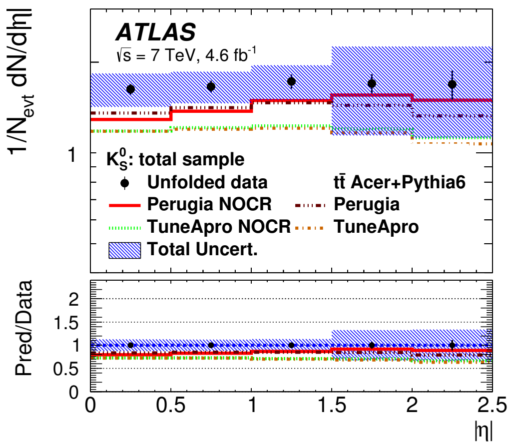
Figure 5.82: Kinematic characteristics for K_S^0 production not associated with jets, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



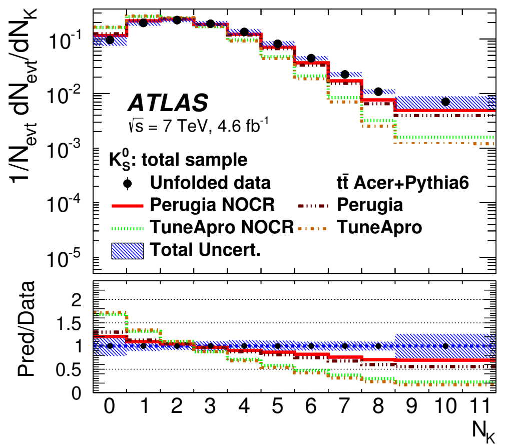
(a) p_{T} distribution



(b) Energy distribution

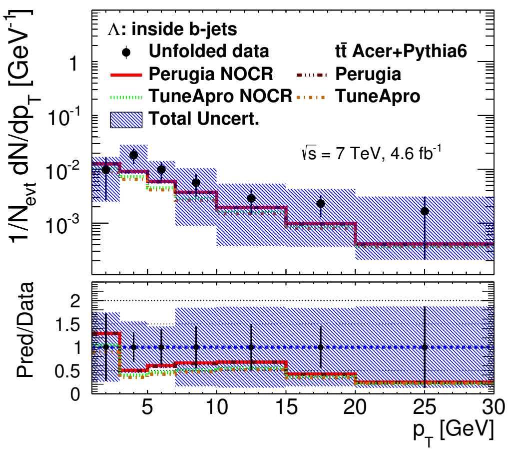


(c) $|\eta|$ distribution

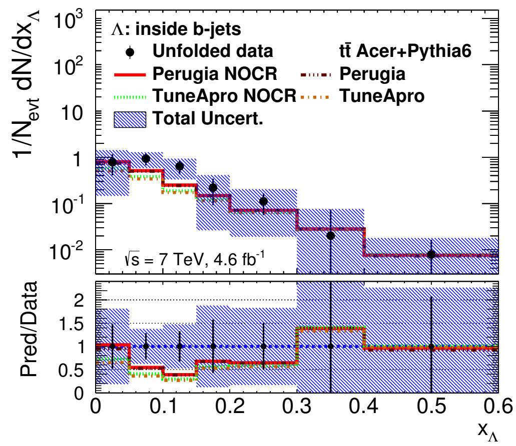


(d) Multiplicity distribution

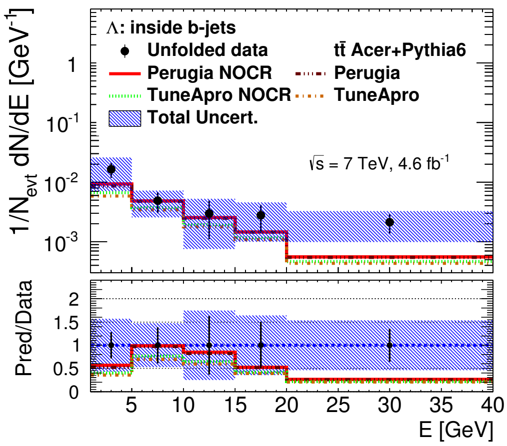
Figure 5.83: Kinematic characteristics for the total K_S^0 production, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



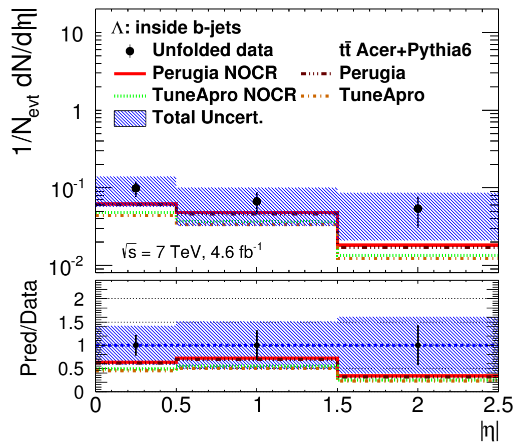
(a) p_{T} distribution



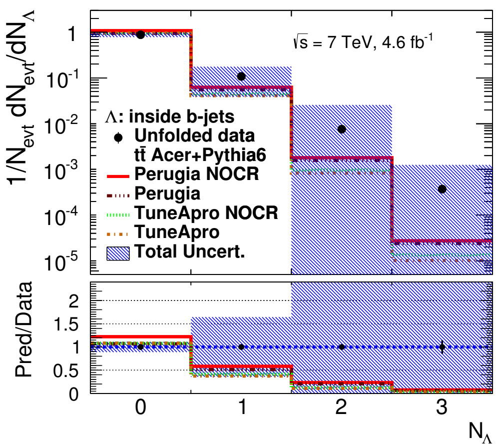
(b) Energy fraction distribution



(c) Energy distribution

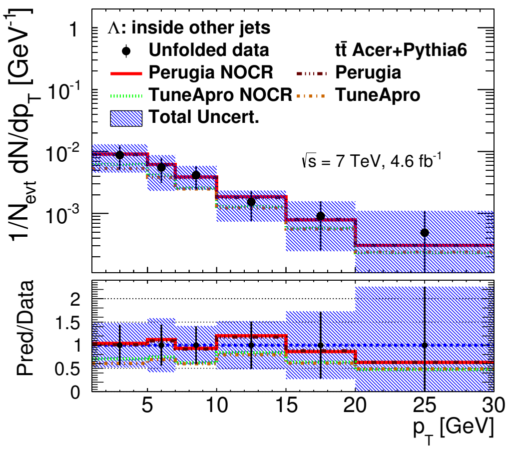


(d) $|\eta|$ distribution

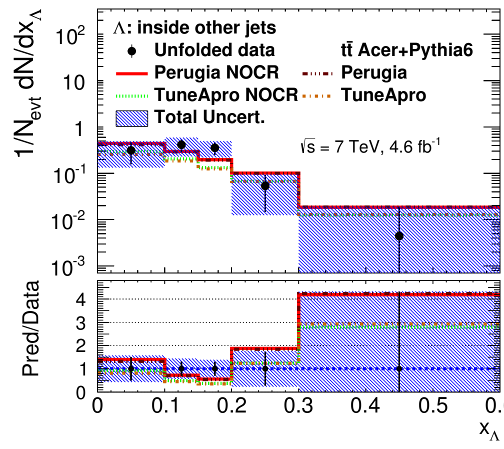


(e) Multiplicity distribution

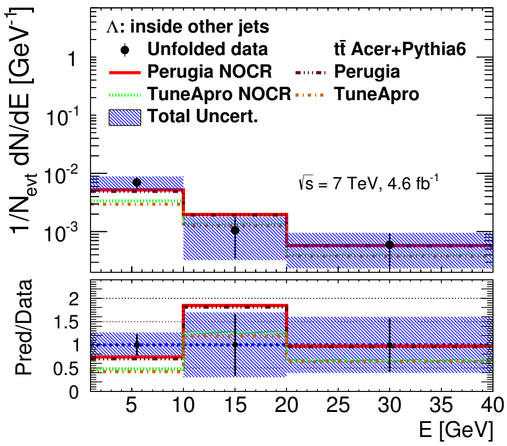
Figure 5.84: Kinematic characteristics for Λ production inside b-jets, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



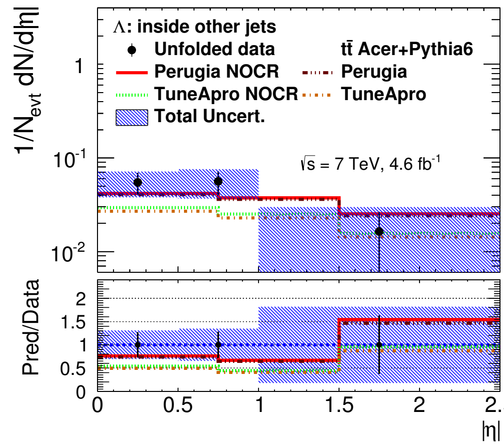
(a) p_{T} distribution



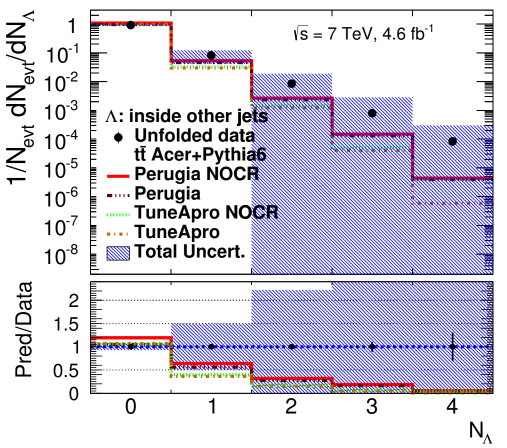
(b) Energy fraction distribution



(c) Energy distribution

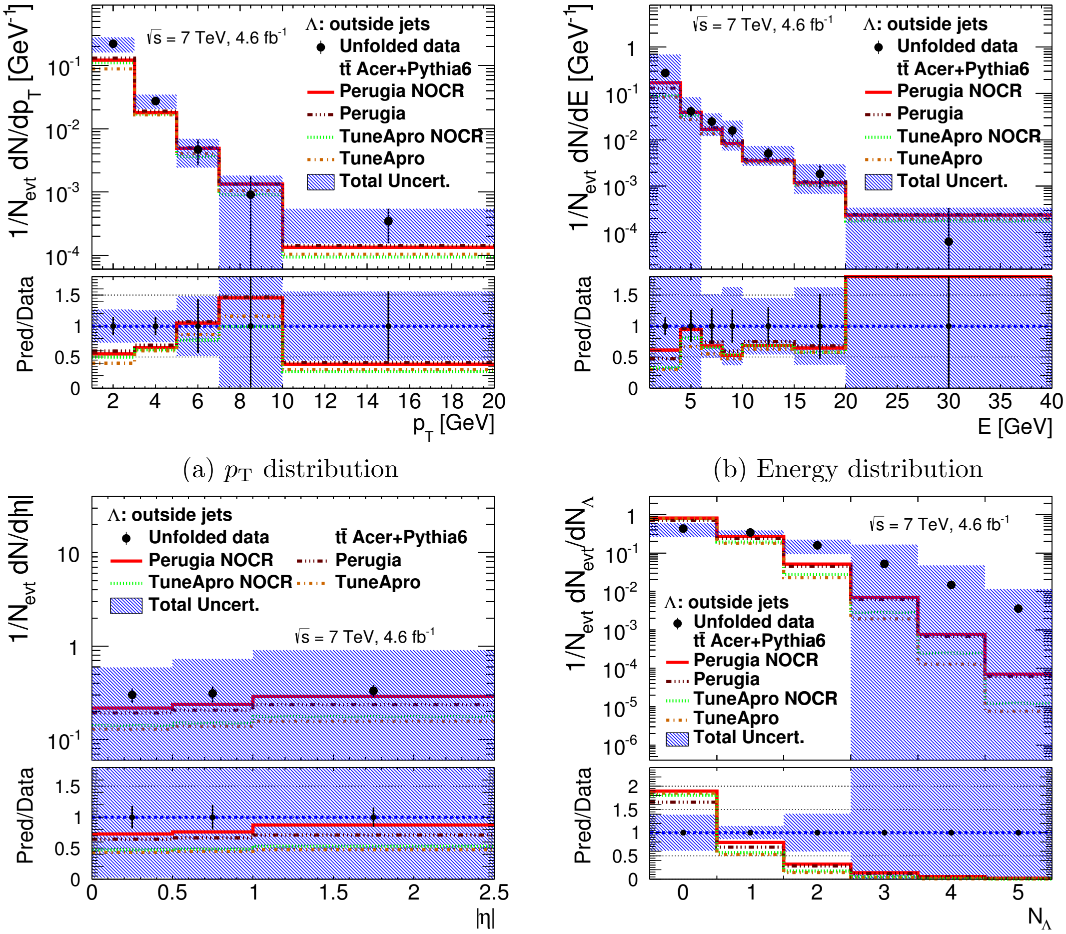


(d) $|\eta|$ distribution



(e) Multiplicity distribution

Figure 5.85: Kinematic characteristics for Λ production inside non b-jets, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



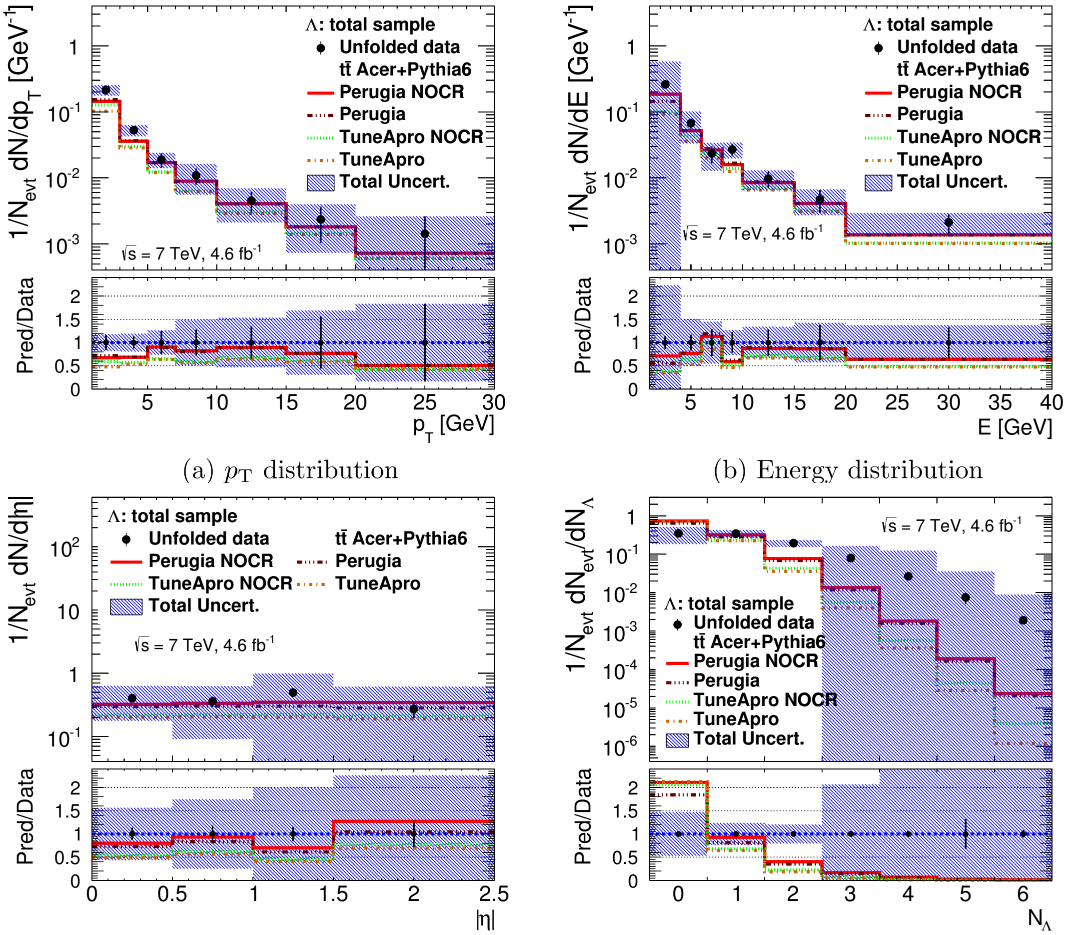
(a) p_{T} distribution

(b) Energy distribution

(c) $|\eta|$ distribution

(d) Multiplicity distribution

Figure 5.86: Kinematic characteristics for Λ production not associated with jets, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



(a) p_{T} distribution

(b) Energy distribution

(c) $|\eta|$ distribution

(d) Multiplicity distribution

Figure 5.87: Kinematic characteristics for the total Λ production, for corrected data and particle level MC simulation. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.

Table 5.28: Values of the χ^2, obtained from the covariance matrix of the unfolded data distributions for K^0_{S} production inside b-jets, for the ACERMC predictions with the following tunes: PERUGIA and TUNEAPRO (with and without colour reconnection).

	$\chi^2/n.d.f.$	K^0_{S} inside b-jets		
	PERUGIA	PERUGIA (no CR)	TUNEAPRO	TUNEAPRO (no CR)
p_{T}	0.55	0.53	0.59	0.52
E	1.48	1.49	1.62	1.54
$\lvert\eta\rvert$	0.24	0.23	0.38	0.31
x_K	0.36	0.36	0.33	0.28
N_K	3.94	3.61	5.21	4.24

Table 5.29: Values of the χ^2, obtained from the covariance matrix of the unfolded data distributions for K^0_{S} production inside non b-jets, for the ACERMC predictions with the following tunes: PERUGIA and TUNEAPRO (with and without colour reconnection).

	$\chi^2/n.d.f.$	K^0_{S} inside non b-jets		
	PERUGIA	PERUGIA (no CR)	TUNEAPRO	TUNEAPRO (no CR)
p_{T}	0.63	0.70	1.64	1.48
E	1.17	1.16	1.65	1.52
$\lvert\eta\rvert$	0.26	0.31	0.33	0.29
x_K	0.85	0.82	1.24	1.34
N_K	0.53	0.56	0.98	0.85

Table 5.30: Values of the χ^2, obtained from the covariance matrix of the unfolded data distributions for K^0_S production outside jets, for the ACERMC predictions with the following tunes: PERUGIA and TUNEAPRO (with and without colour reconnection).

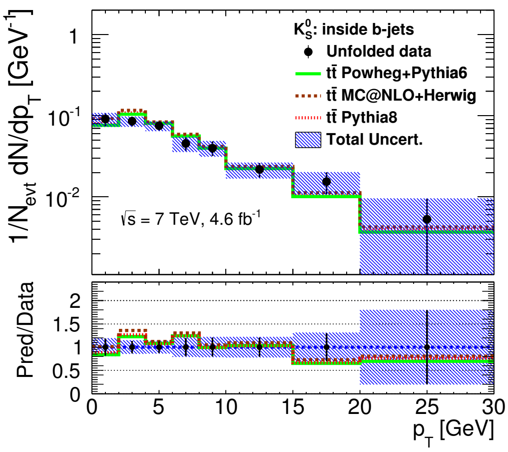
	$\chi^2/n.d.f.$ K^0_S outside jets			
	PERUGIA	PERUGIA (no CR)	TUNEAPRO	TUNEAPRO (no CR)
p_T	0.38	0.48	0.67	1.18
E	0.87	1.13	1.54	1.67
$\lvert\eta\rvert$	0.77	0.75	2.16	2.09
N_K	93.2	88.3	107.5	105.6

Table 5.31: Values of the χ^2 per degree of freedom and their corresponding p-values, for the total K^0_S production, along with the ACERMC+PYTHIA6 predictions with the following tunes: PERUGIA and TUNEAPRO (with and without colour reconnection).

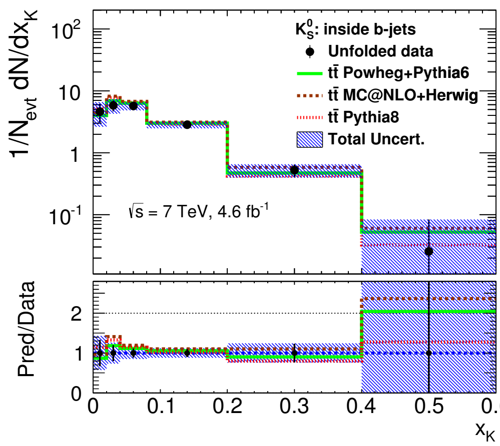
	$\chi^2/n.d.f.$ (p-value) K^0_S total sample			
	PERUGIA	PERUGIA (no CR)	TUNEAPRO	TUNEAPRO (no CR)
p_T	0.42 (0.91)	0.50 (0.85)	1.12 (0.35)	1.36 (0.21)
E	1.54 (0.13)	1.90 (0.05)	3.55 (0.0002)	3.37 (0.0004)
$\lvert\eta\rvert$	1.15 (0.33)	1.34 (0.24)	4.18 (0.0008)	3.90 (0.002)

5.9.2 Pythia8

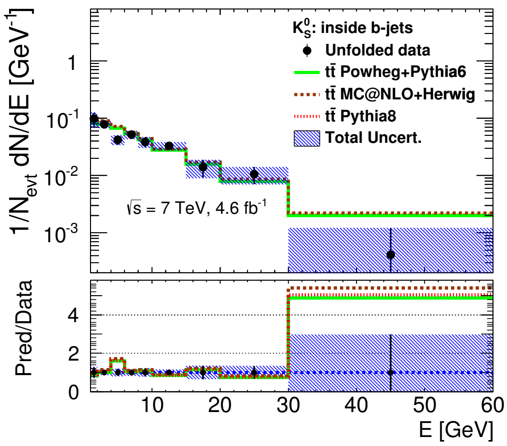
The unfolded distributions in the p_T, E, $|\eta|$ and x_K for K^0_S production inside b-jets, inside non b-jets, outside any jet and for the total sample are shown in Figures 5.88 to 5.91 including a comparison with the PYTHIA8 generator [71]. The description of the unfolded data spectra is similar to the PYTHIA6 prediction, which indicates that the new version of the generator treats neutral strange particle production in a similar way as the old one.



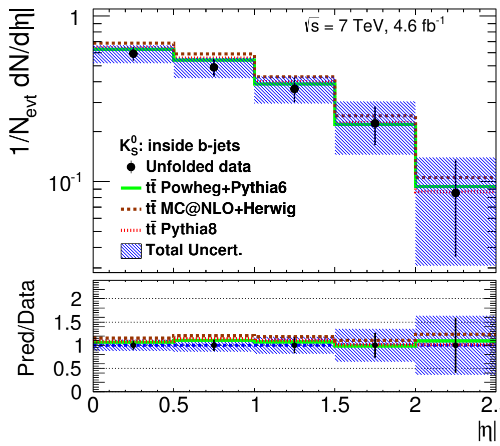
(a) p_{T} distribution



(b) Energy fraction distribution

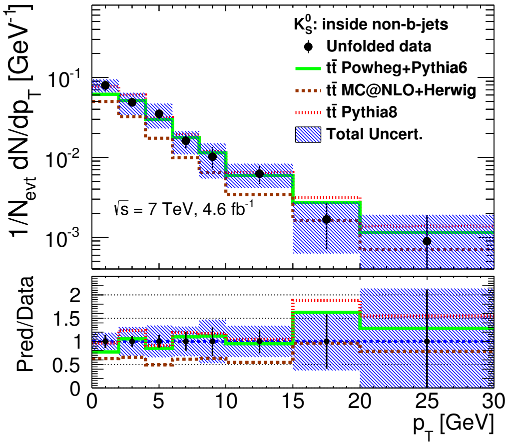


(c) Energy distribution

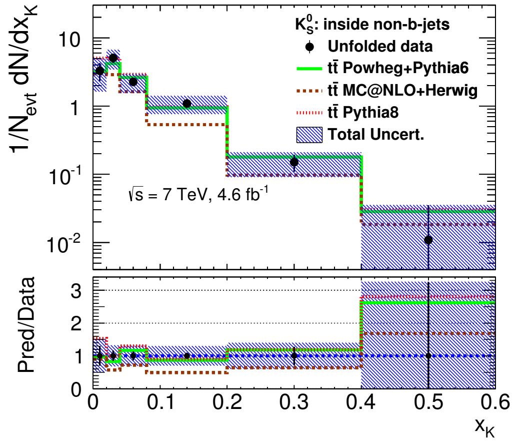


(d) $|\eta|$ distribution

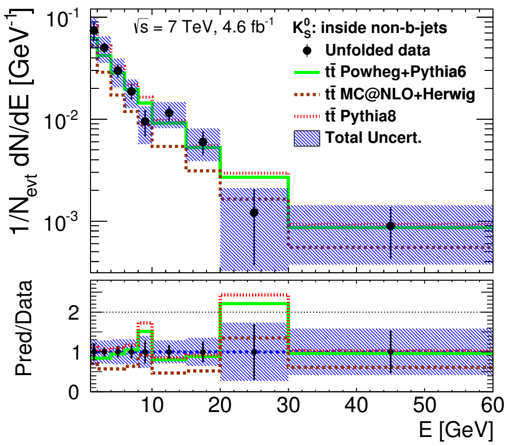
Figure 5.88: Kinematic characteristics for K^0_S production inside b-jets, for corrected data and particle level MC simulated with the POWHEG+PYTHIA6, MC@NLO+HERWIG and PYTHIA8 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



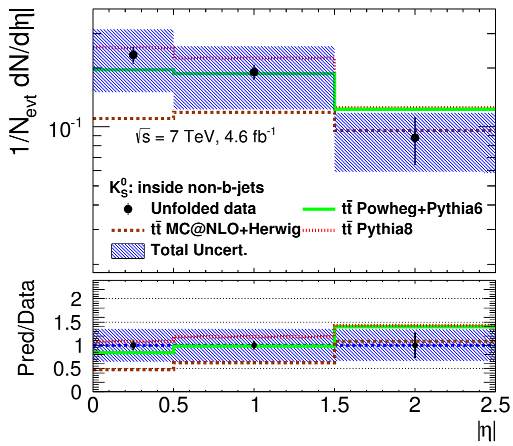
(a) p_{T} distribution



(b) Energy fraction distribution

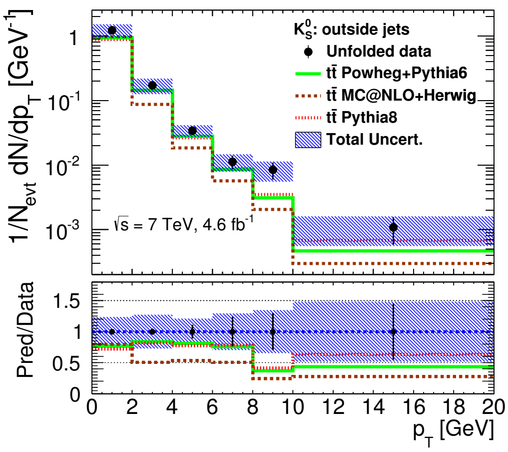


(c) Energy distribution

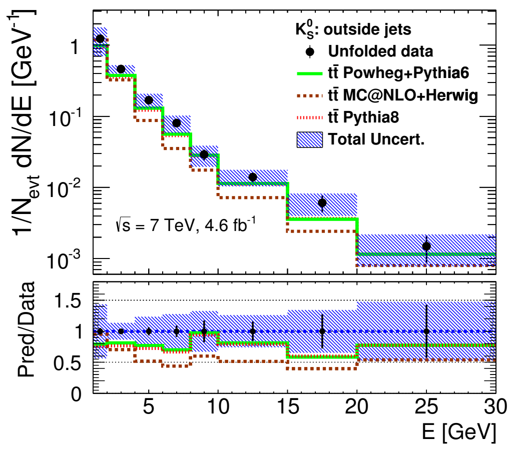


(d) $|\eta|$ distribution

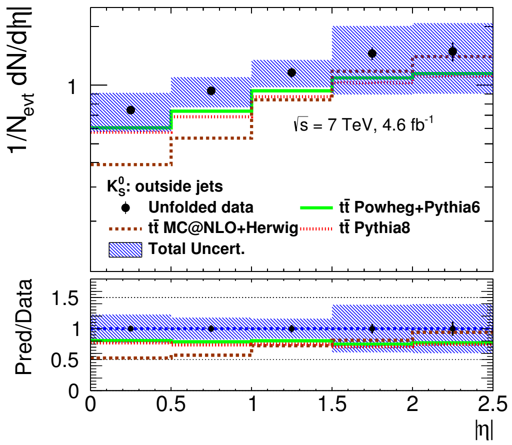
Figure 5.89: Kinematic characteristics for K_S^0 production inside non b-jets, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig and Pythia8 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



(a) p_{T} distribution

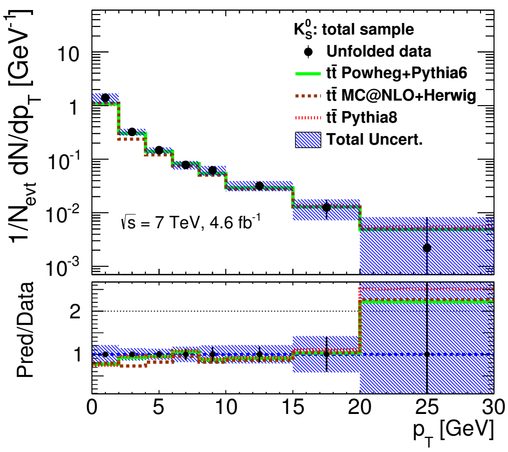


(b) Energy distribution

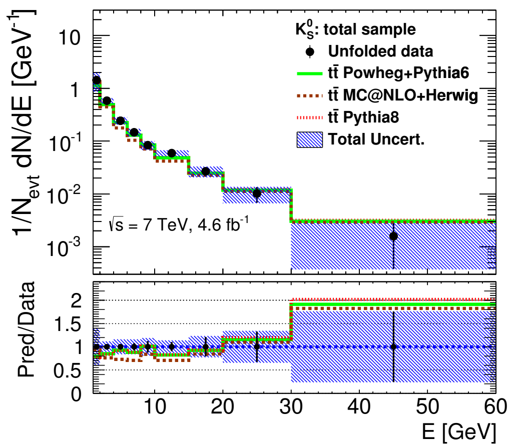


(c) $|\eta|$ distribution

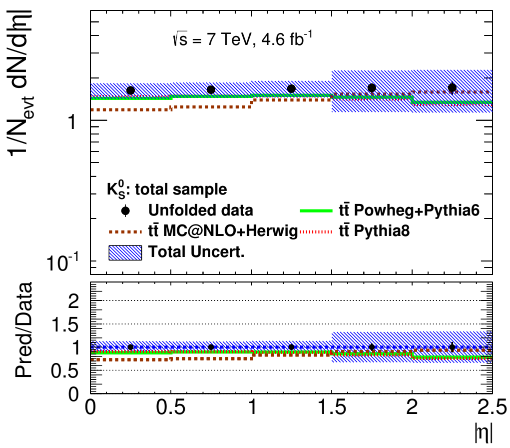
Figure 5.90: Kinematic characteristics for K^0_{S} production not associated with jets, for corrected data and particle level MC simulated with the POWHEG+PYTHIA6, MC@NLO+HERWIG and PYTHIA8 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.



(a) p_{T} distribution



(b) Energy distribution



(c) $|\eta|$ distribution

Figure 5.91: Kinematic characteristics for the total K_S^0 production, for corrected data and particle level MC simulated with the Powheg+Pythia6, MC@NLO+Herwig and Pythia8 generators. Total uncertainties are represented by the shaded area. Statistical uncertainties for MC samples are negligible in comparison with data.

5.10 Summary

First measurements of neutral strange particle, K^0_S and Λ, production in dileptonic $t\bar{t}$ final states are reported. The distributions in the energy, p_T and $|\eta|$ are presented for the complete sample as well as for three subsamples depending on whether the K^0_S or Λ are associated with jets, b-tagged or non b-tagged, or lie outside jets. The corresponding K^0_S and Λ multiplicities are also measured. The results are unfolded to the particle level using the K^0_S or Λ reconstruction efficencies in each distribution within the kinematical region given by $E > 1$ GeV and $|\eta| < 2.5$, respectively. The measurements are compared to current MC expectations where the $t\bar{t}$ matrix elements have been calculated up to NLO accuracy, according to Powheg, MC@NLO, aMC@NLO, AcerMC or Sherpa, and considering several variations on

- fragmentation scheme: Pythia6, Pythia8, Herwig, Herwig7 or Sherpa.
- UE effects: Perugia2011c, A14, TuneAPro, Jimmy or H7UE
- colour reconnection effects

The main conclusions to be drawn from the analysis are the following:

- strange baryon production is suppressed with respect to strange meson production, as expected
- neutral strange particle production outside jets is much softer than inside jets, as expected, and the pseudorapidity distributions flatter
- neutral strange particle multiplicities outside jets are larger than inside
- current MC models give a fair description of the gross features exhibited by K^0_S and Λ produced inside jets, while the yield for single neutral strange particles lie roughly 30% above the Pythia6+Perugia2011C and Herwig+Jimmy or Herwig7+H7UE MC predictions, with Pythia8+A14 falling short of the data by 15–20 %.

The fact that the current MC models describe properly the neutral strange particles spectra inside jets, gives confidence that these data can be used in the future for a direct determination of the $|V_{ts}|$ CKM matrix element.
A better description of the yields for K^0_S and Λ outside jets in $t\bar{t}$ final states would require further tuning of the current MC models, particularly the strangeness suppression mechanisms, and/or more elaborate models for MPI and colour reconnection schemes. For this purpose a Rivet analysis routine [145] and HEPData tables [142] are provided.
We are aware that the results for Λ production are hampered by poor statistics, they are included for the sake of completeness.

Resumen y Conclusiones

Hemos probado que los ‘jet shapes’ tienen un gran potencial para discriminar entre chorros adscritos a un quark b o a un quark ligero. Por tanto la inclusión de estas variables debe mejorar las eficiencias de ‘b-tagging’. Los algoritmos al uso de identificación de chorros iniciados por un quark b están basados en medidas de parámetros de impacto y en la reconstrucción de vértices secundarios. El poder de discriminación de los ‘jet shapes’ es casi tan alto como el de los ‘taggers’ normales.
En el caso concreto de la posible medida directa del elemento $|V_{ts}|$, la inclusión de los ‘jet shapes’ en el análisis multivariado basado en BDTs conduce a mejoras muy significativas. En cualquier caso este análisis multivariado ha sido realizado a nivel generador y por tanto está previsto repetirlo a nivel ‘reco’ a fin de tener en cuenta de manera apropiada los efectos de detector.
Para la separación del canal $pp \to Z(l^+l^-)H(b\bar{b})$ frente al $pp \to Z(l^+l^-) + b\bar{b}$ la inclusión de los ’jet shapes’ no conduce a mejoras significativas. Las variables cinemáticas utilizadas en el MVA son ya de por si poderosos discriminantes.

La contribución principal de esta tesis se centra en las primeras medidas de la producción de partículas neutras extrañas en los estados hadrónicos finales $t\bar{t}$ en los que los dos quarks top se desintegran leptónicamente. A este efecto se ha utilizado una muestra con una luminosidad integrada de 4.6 fb^{-1} tomada en 2011 a una energía en el centro de masas de 7 TeV con el detector ATLAS en el LHC. Hemos distinguido tres casos dependiendo de que las partículas extrañas estén embebidas en chorros etiquetados como b, en chorros no etiquetados como b o estén producidas fuera de un chorro.
En cada caso hemos estudiado las características de producción de las partículas extrañas en términos de variables cinemáticas como el momento transverso, la pseudorapidez, la energía o la multiplicidad.
Hemos corregido estas variables por efectos de detector y comparado con diversos generadores Monte Carlo, donde los elementos de matriz se han calculado a NLO como en Powheg, MC@NLO y Sherpa o a LO con Acermc. Hemos considerado distintas variaciones en:

- el esquema de fragmentación y 'underlying event': Pythia6+Perugia2011C o Pythia6+TuneAPro, Pythia8, Herwig+Jimmy, Herwig++ y Sherpa
- los efectos de reconexión de color

Las conclusiones principales son:

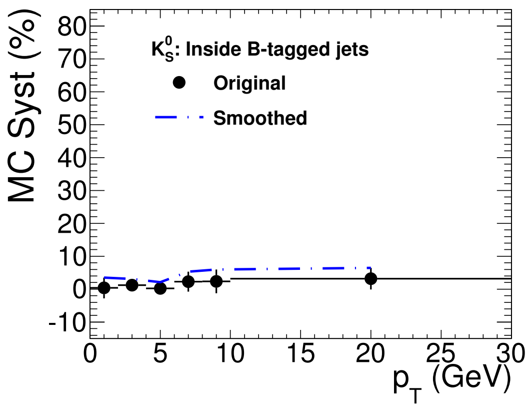
- la producción de Λ's está suprimida frente a la producción de K_S^0's como es de esperar
- el momento transverso de las partículas extrañas dentro de chorros, etiquetados como b o no, cae menos fuertemente que la de aquellas no asociadas a chorros como también era de esperar
- la pseudorapidez es más plana fuera de chorros
- la producción de partículas extrañas dentro de chorros está bien descrita por modelos MC al uso
- la producción de partículas extrañas no asociadas a chorros se ve infraestimada por los modelos al uso en un $30-40\%$.

Estos datos representan un 'input' necesario para el desarrollo de modelos más elaborados para las interaciones multipartónicas y distintos esquemas de conexión de color.

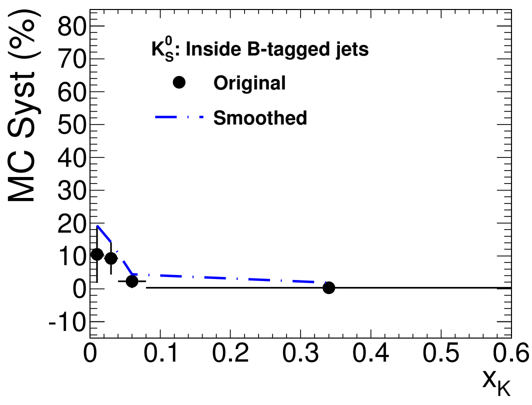
A Systematics Details

A.1 MC Generator uncertainties

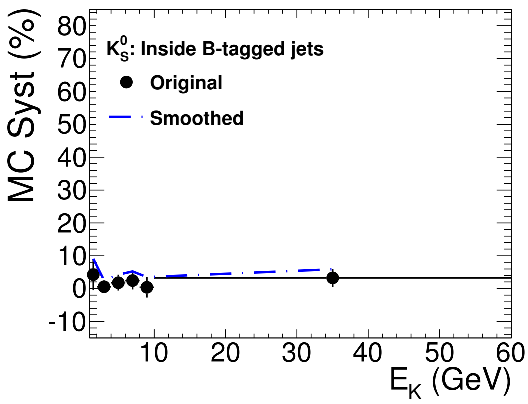
In this appendix details are shown for the calculation of the systematics uncertainties due to the choice of the MC generator for the unfolding, for K^0_S production inside b-jets, inside non b-tagged jets, outside any jets and for the total sample. As explained in Section 5.7.1 the distributions have been rebinned in order to bring down statistical fluctuations and the systematics uncertainties are conservatively taken as the sum of the central values plus the statistical uncertainty in the calculation.This is illustrated in Figures 92, 93, 94 and 95.



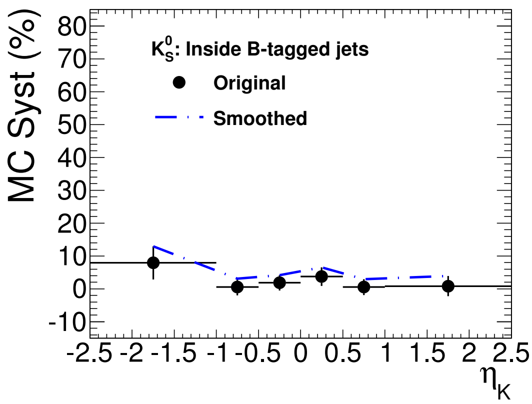
(a) p_{T} distribution



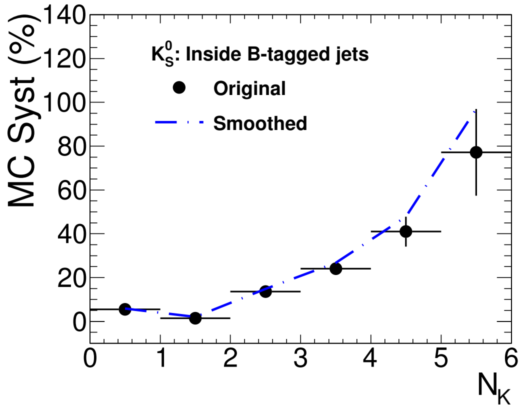
(b) Energy fraction distribution



(c) Energy distribution

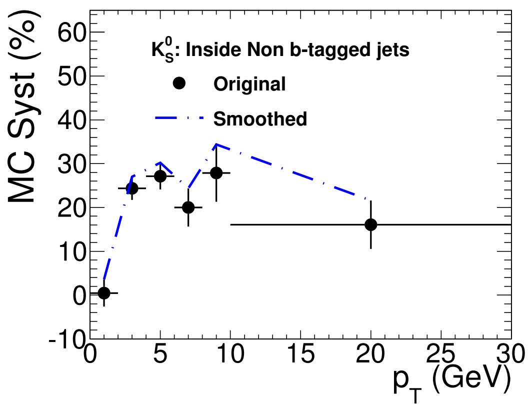


(d) $|\eta|$ distribution

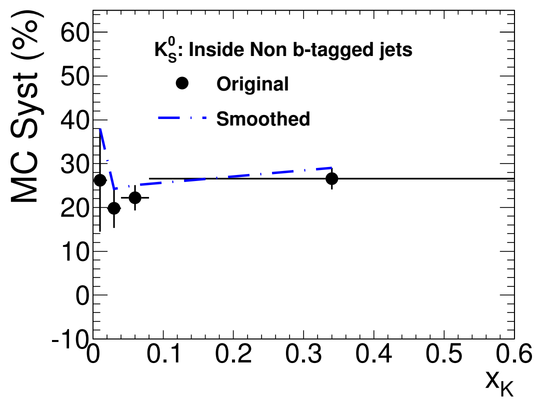


(e) Multiplicity distribution

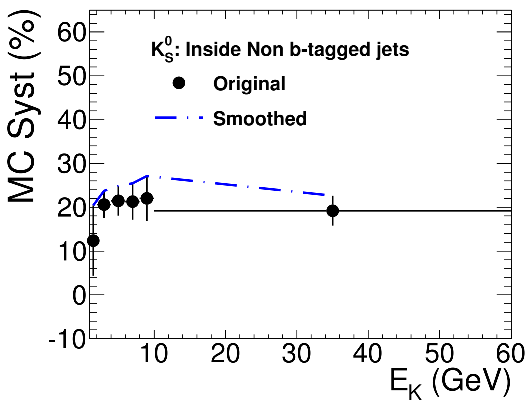
Figure 92: The K^0_{S} inside b-jets systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



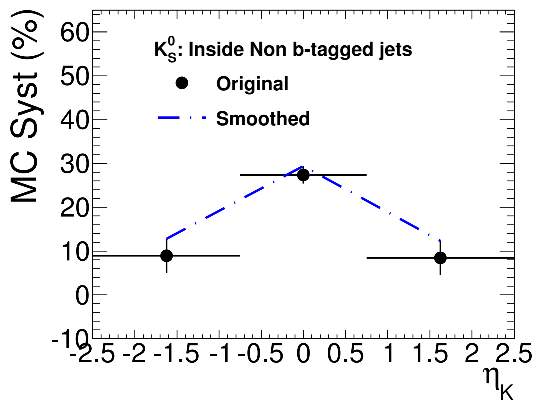
(a) p_T distribution



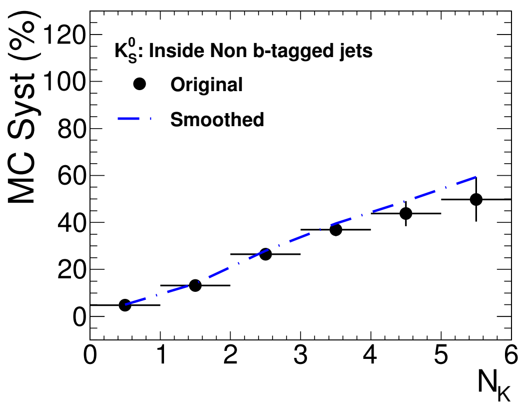
(b) Energy fraction distribution



(c) Energy distribution

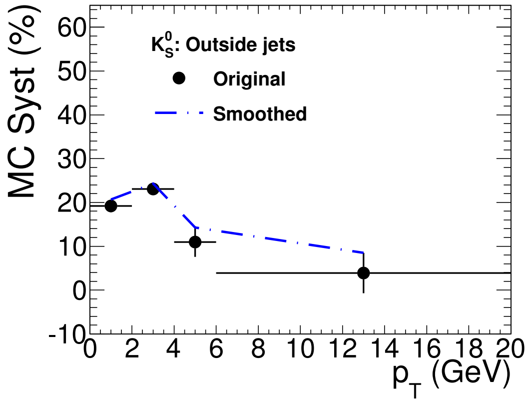


(d) $|\eta|$ distribution

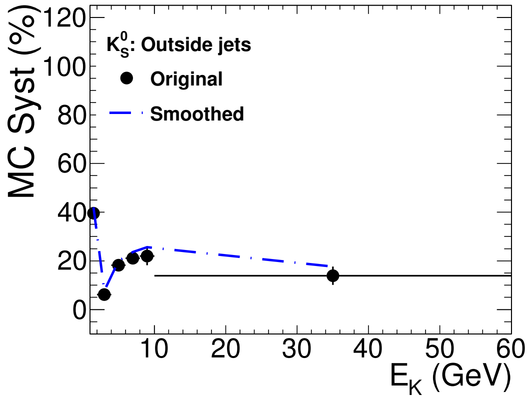


(e) Multiplicity distribution

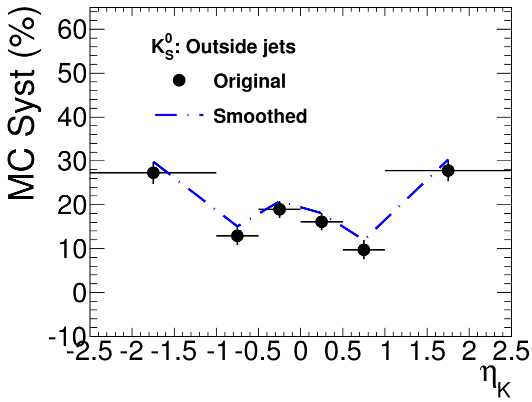
Figure 93: The K_S^0 inside not b-tagged jets systematic uncertainties due to the MC generator as a function of p_T, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



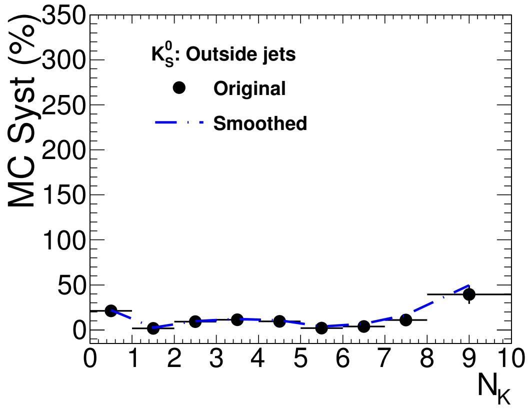
(a) p_{T} distribution



(b) Energy distribution

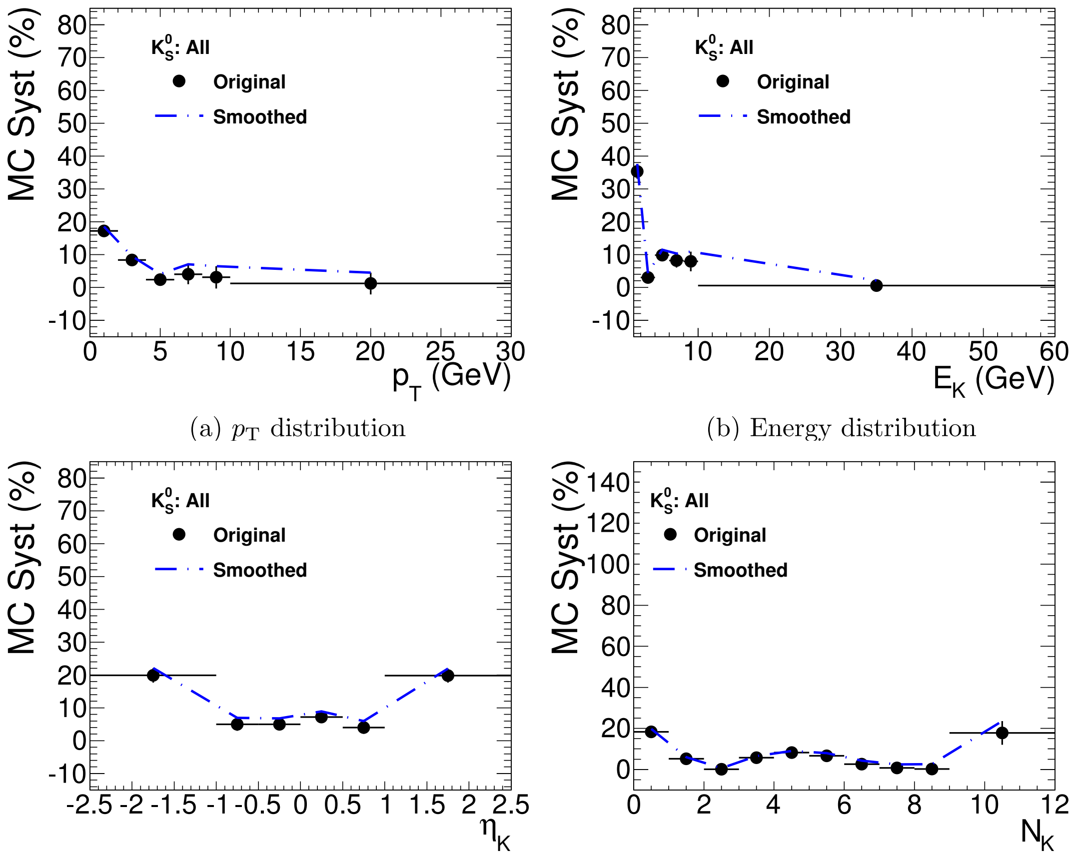


(c) $|\eta|$ distribution



(d) Multiplicity distribution

Figure 94: The K^0_{S} outside any jet systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



(a) p_{T} distribution

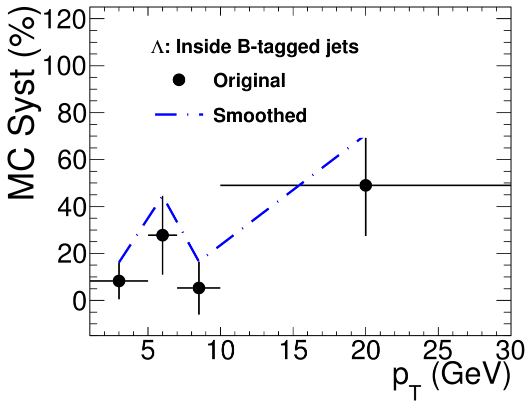
(b) Energy distribution

(c) $|\eta|$ distribution

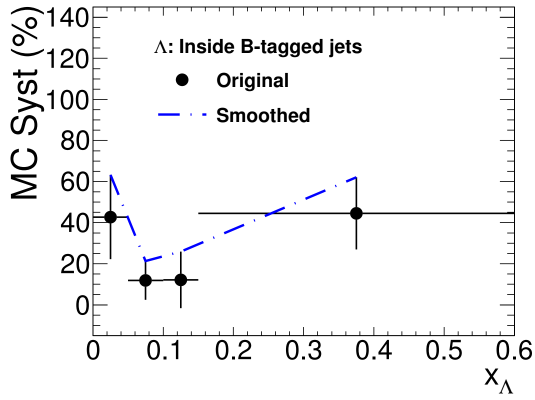
(d) Multiplicity distribution

Figure 95: The total sample K^0_{S} systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.

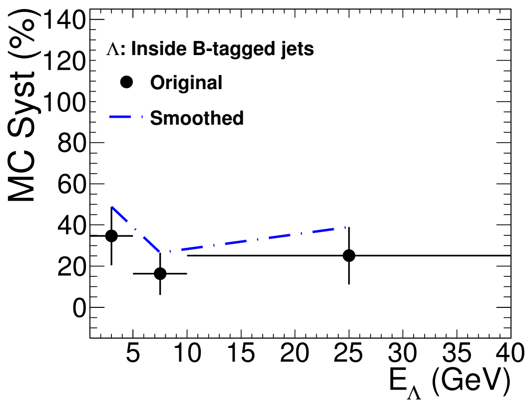
The Λ^0 results are shown in Figures 96, 97, 98 and 99.



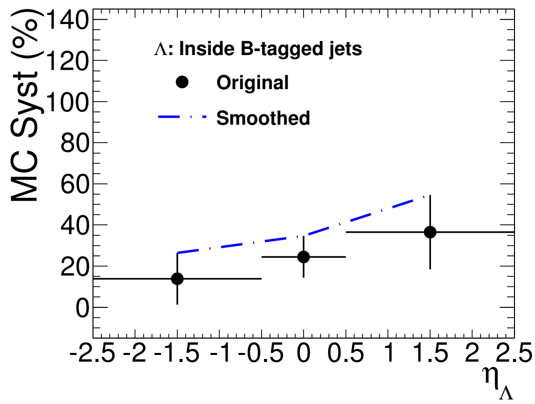
(a) p_{T} distribution



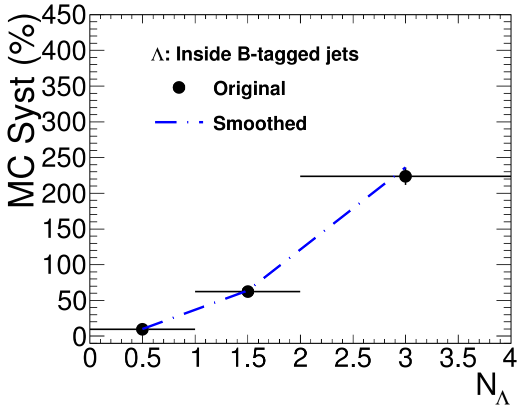
(b) Energy fraction distribution



(c) Energy distribution

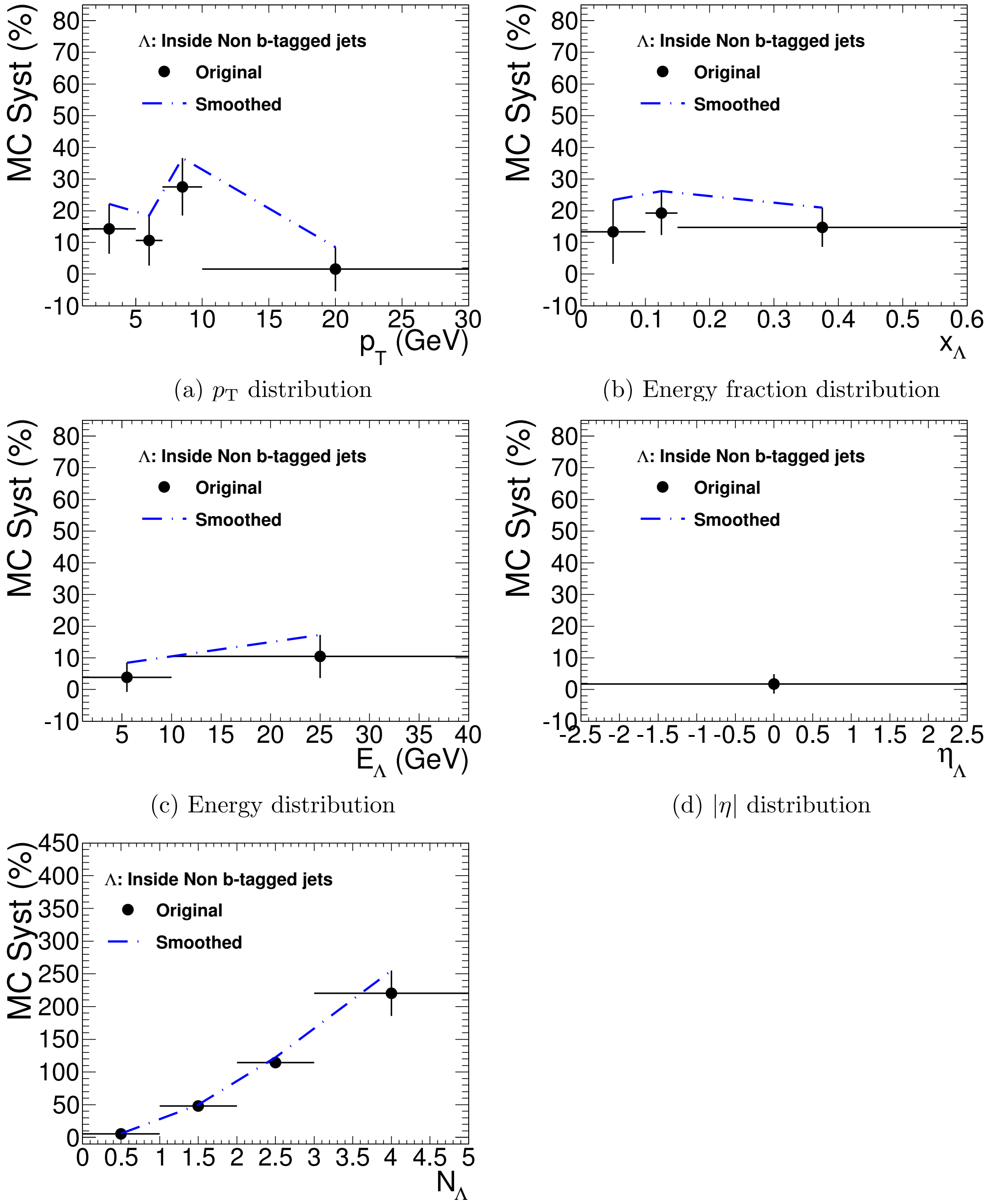


(d) $|\eta|$ distribution



(e) Multiplicity distribution

Figure 96: The Λ inside b-jets systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



(a) p_{T} distribution

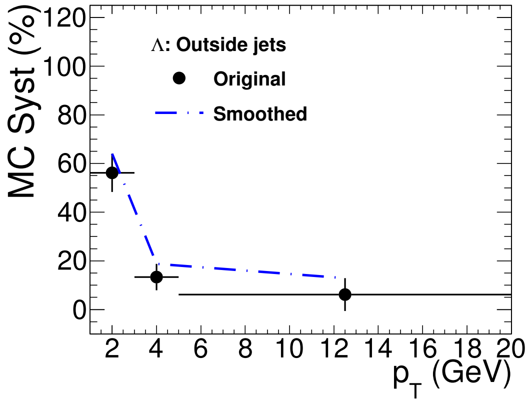
(b) Energy fraction distribution

(c) Energy distribution

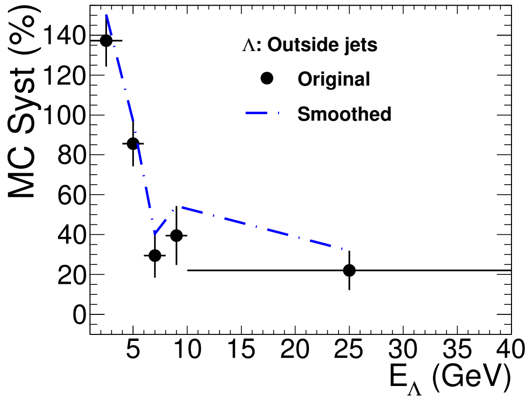
(d) $|\eta|$ distribution

(e) Multiplicity distribution

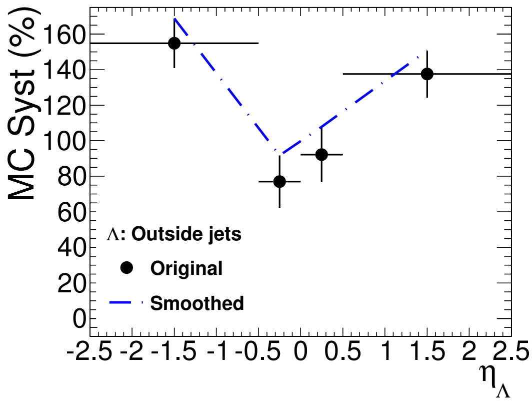
Figure 97: The Λ inside not b-tagged jets systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



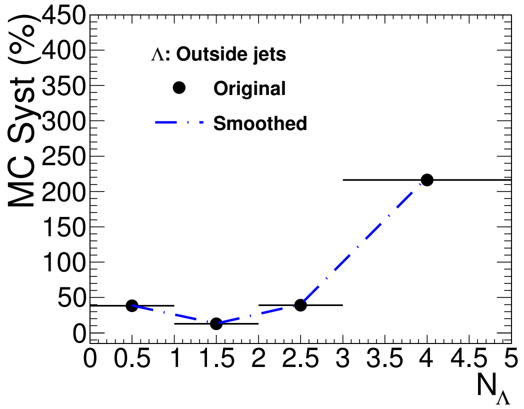
(a) p_{T} distribution



(b) Energy distribution

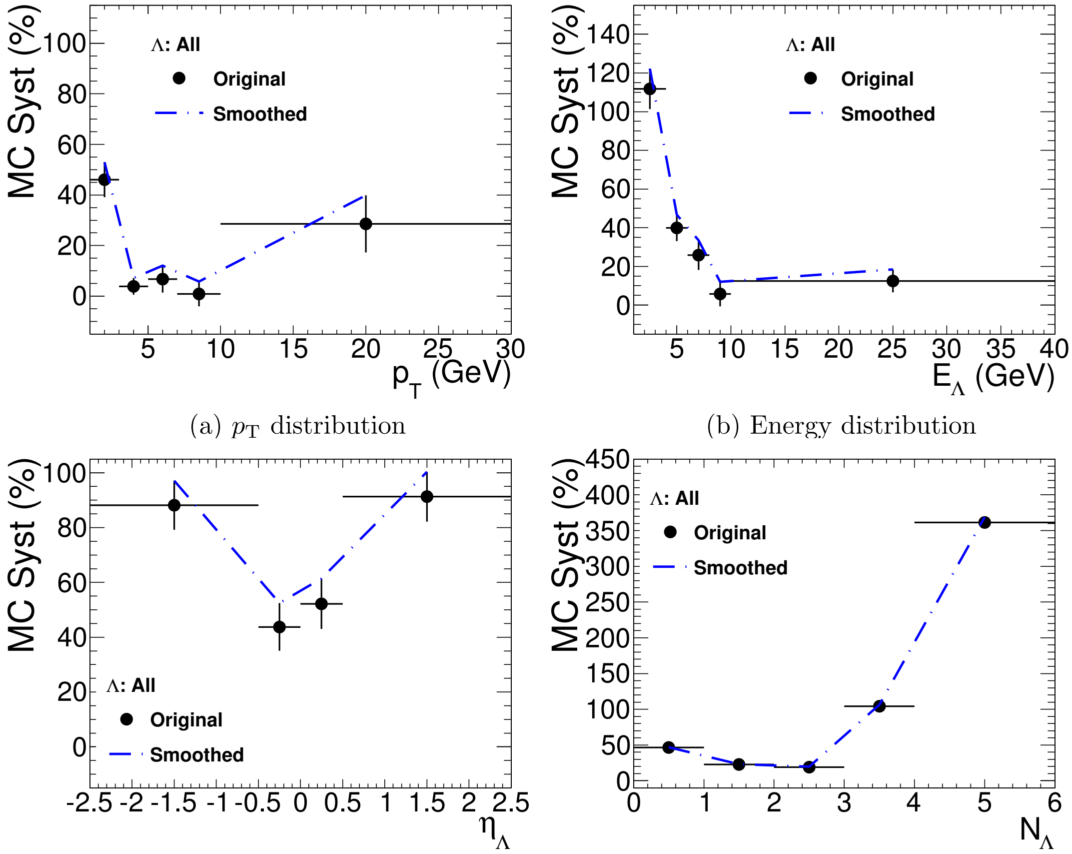


(c) $|\eta|$ distribution



(d) Multiplicity distribution

Figure 98: The Λ outside any jet systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



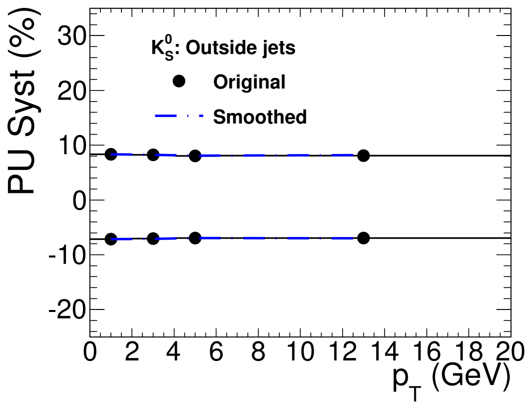
(a) p_{T} distribution (b) Energy distribution

(c) $|\eta|$ distribution (d) Multiplicity distribution

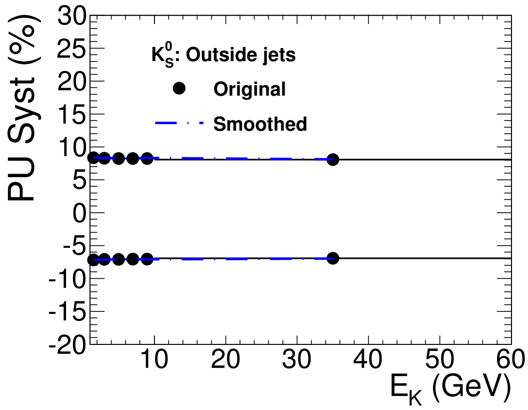
Figure 99: The total sample Λ systematic uncertainties due to the MC generator as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.

A.2 Pile-up uncertainties

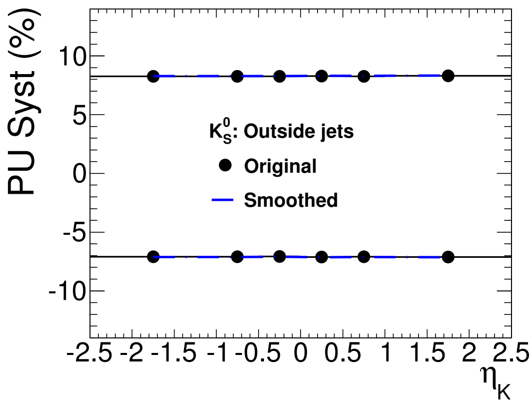
In Figure 100 details are shown of the pile-up systematics calculation following the data-driven procedure described in Section 5.7.2, for K^0_{S} production inside b-jets, inside non b-tagged jets, outside any jets and for the total sample.



(a) p_{T} distribution



(b) Energy distribution

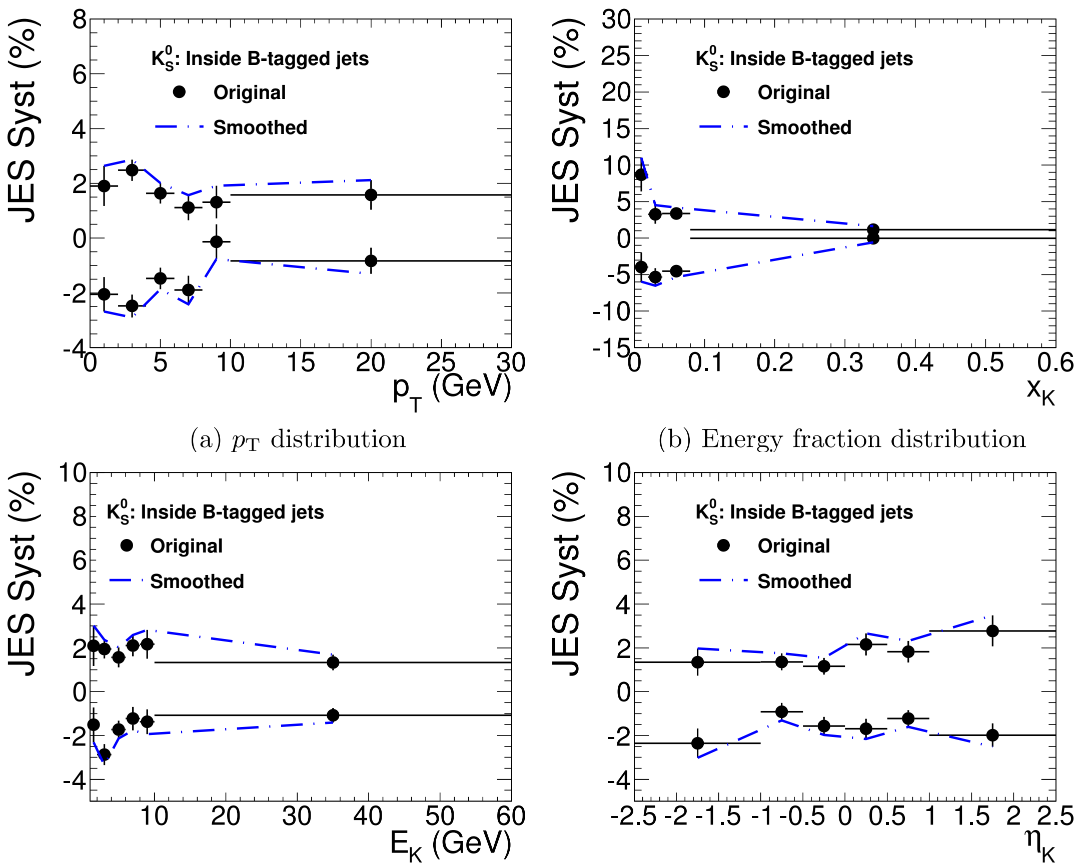


(c) $|\eta|$ distribution

Figure 100: The K^0_{S} outside any jets systematic uncertainties due to the pile-up modelling as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.

A.3 JES and JER uncertainties

In Figures 101, 102, 103 and 104, details for the calculation of JES systematics uncertainties are shown for K^0_{S} inside b-jets, inside non b-tagged jets, outside any jets and for the total sample.



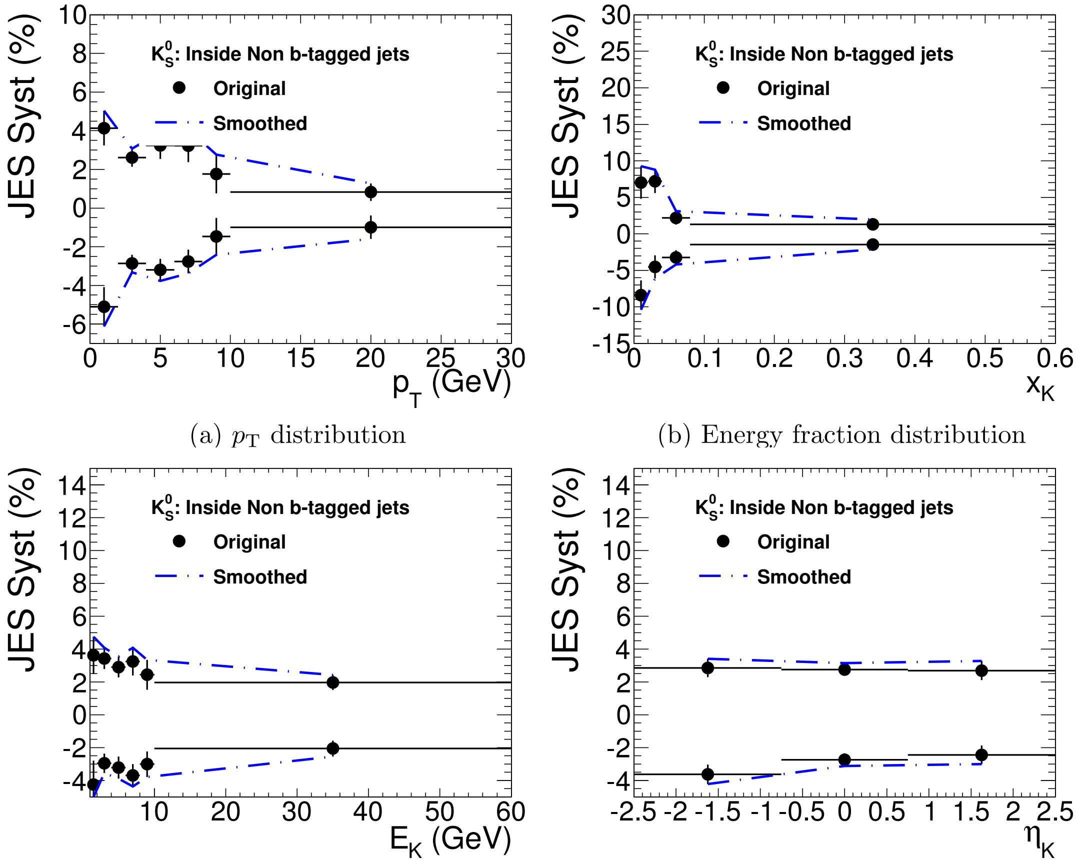
(a) p_{T} distribution

(b) Energy fraction distribution

(c) Energy distribution

(d) $|\eta|$ distribution

Figure 101: The K^0_{S} inside b-jets systematic uncertainties due to the JES uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



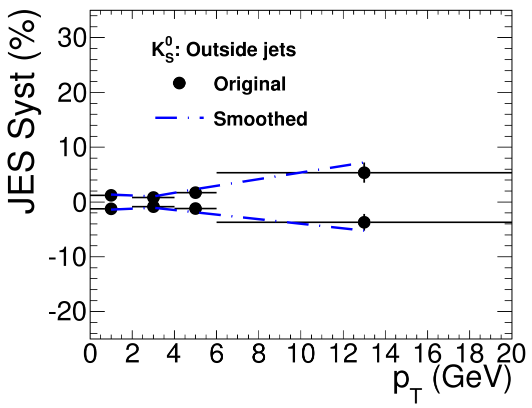
(a) $p_{\rm T}$ distribution

(b) Energy fraction distribution

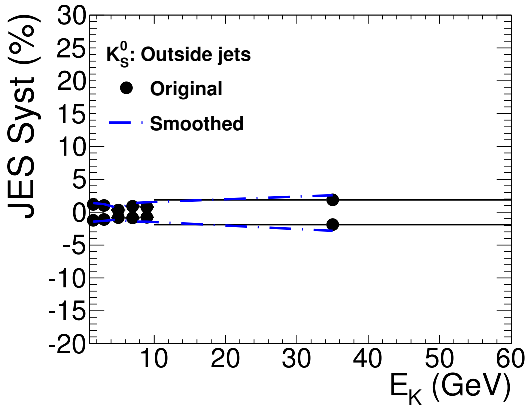
(c) Energy distribution

(d) $|\eta|$ distribution

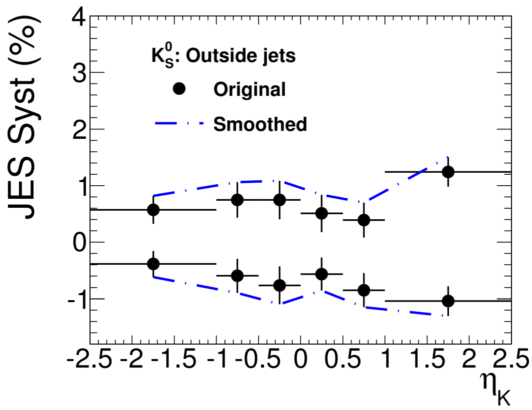
Figure 102: The $K^0_{\rm S}$ inside not b-tagged jets systematic uncertainties due to the JES uncertainty as a function of $p_{\rm T}$, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



(a) p_{T} distribution

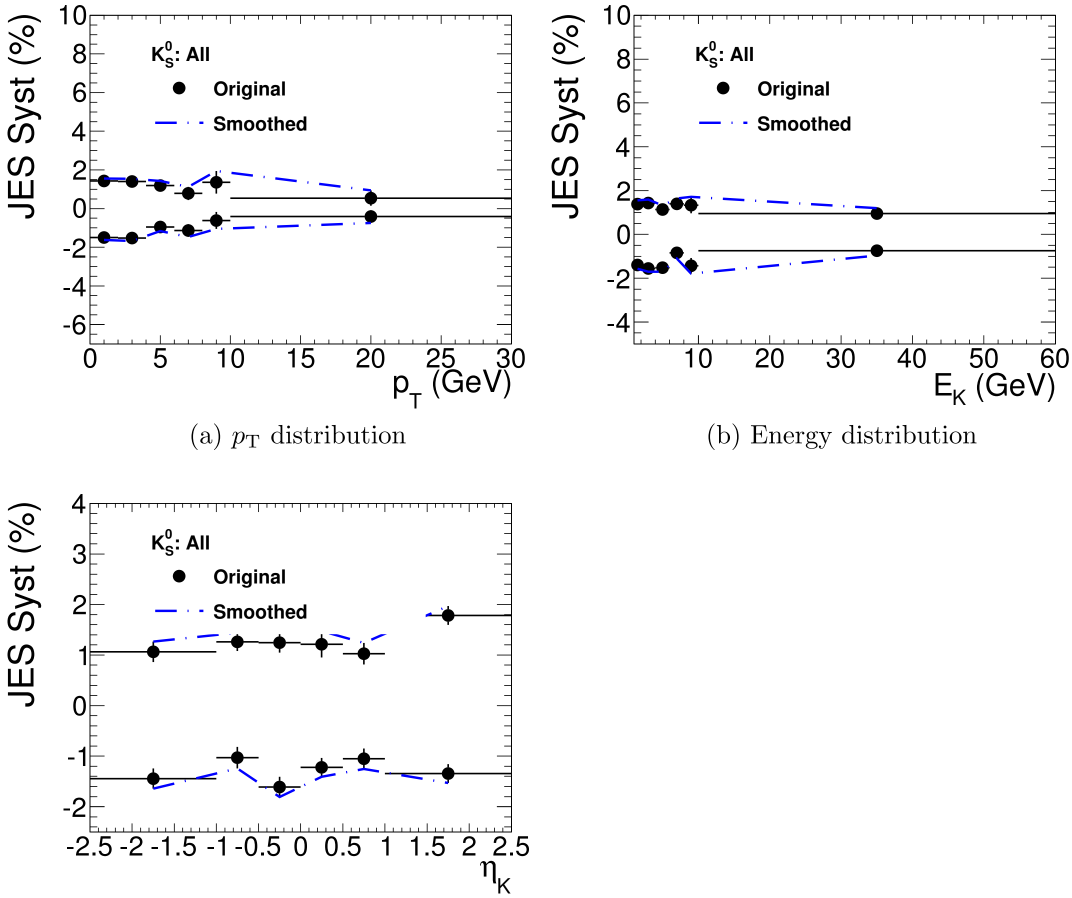


(b) Energy distribution



(c) $|\eta|$ distribution

Figure 103: The K^0_{S} outside any jet systematic uncertainties due to the JES uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



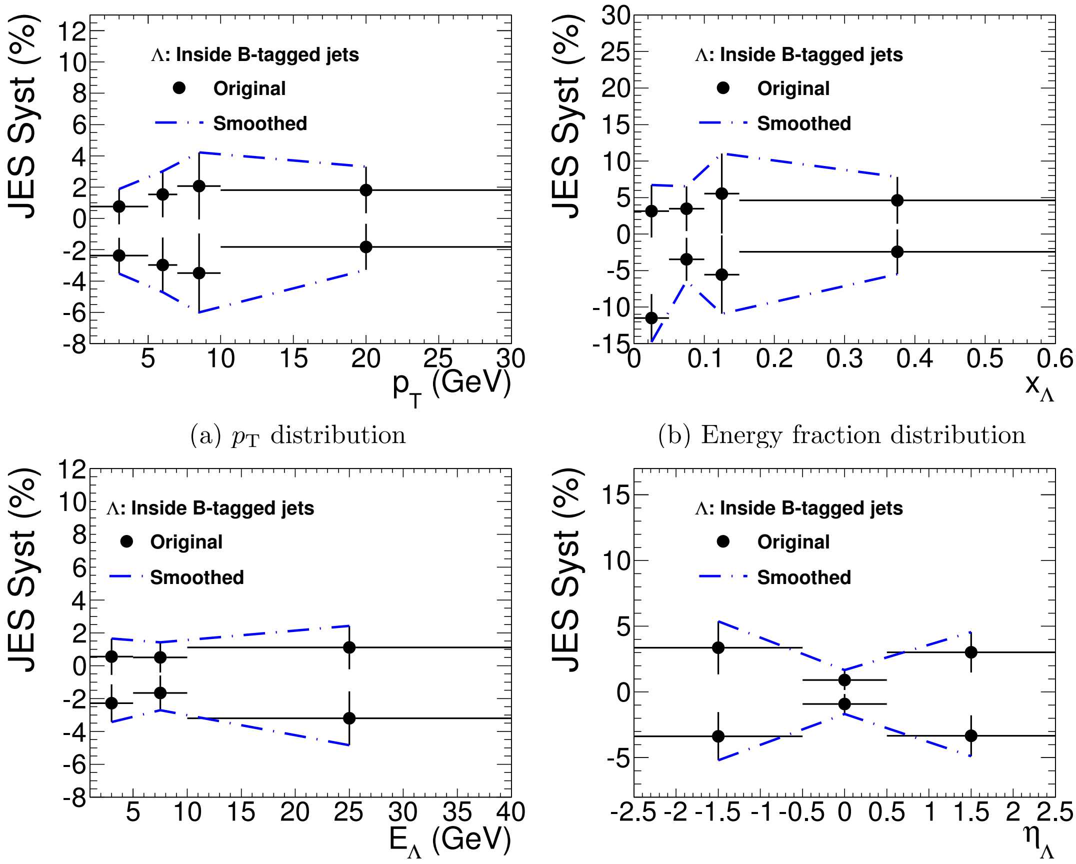
(a) p_{T} distribution

(b) Energy distribution

(c) $|\eta|$ distribution

Figure 104: The total sample K_S^0 systematic uncertainties due to the JES uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.

In Figures 105, 106, 107 and 108, details for the calculation of JES systematic uncertainties are shown for Λ inside b-jets, inside non b-tagged jets, outside any jets and for the total sample.



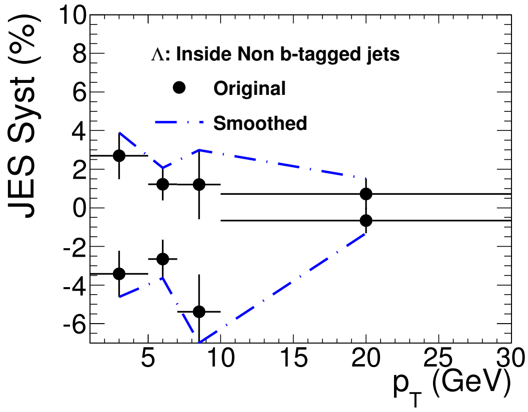
(a) p_{T} distribution

(b) Energy fraction distribution

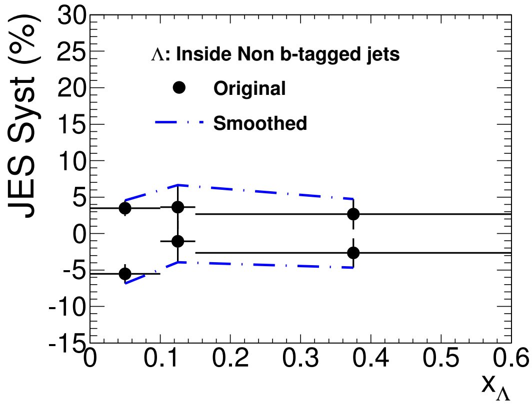
(c) Energy distribution

(d) $|\eta|$ distribution

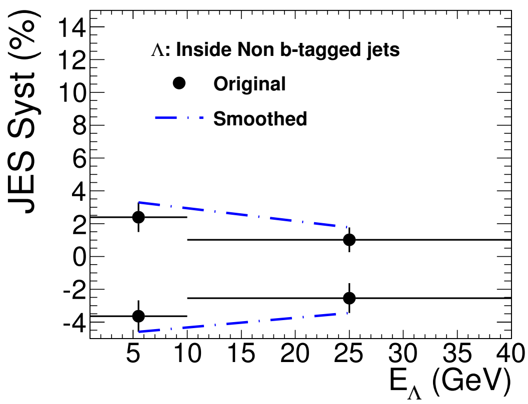
Figure 105: The Λ inside b-jets systematic uncertainties due to the JES uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_Λ), multiplicity and average multiplicity.



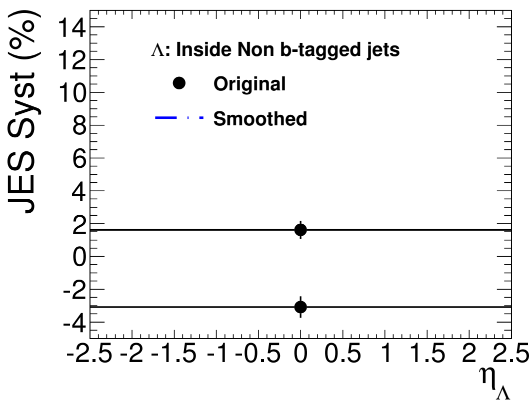
(a) p_T distribution



(b) Energy fraction distribution

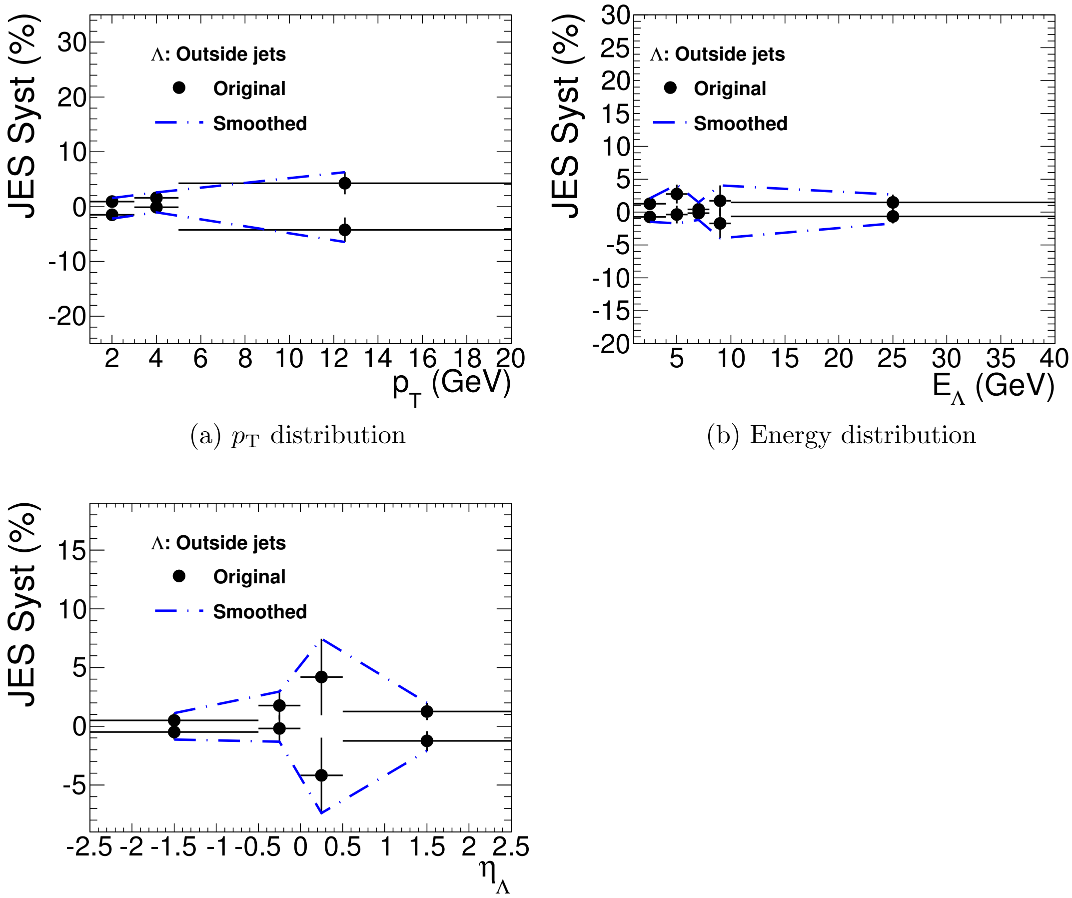


(c) Energy distribution



(d) $|\eta|$ distribution

Figure 106: The Λ inside not b-tagged jets systematic uncertainties due to the JES uncertainty as a function of p_T, $|\eta|$, energy, energy fraction (x_Λ), multiplicity and average multiplicity.

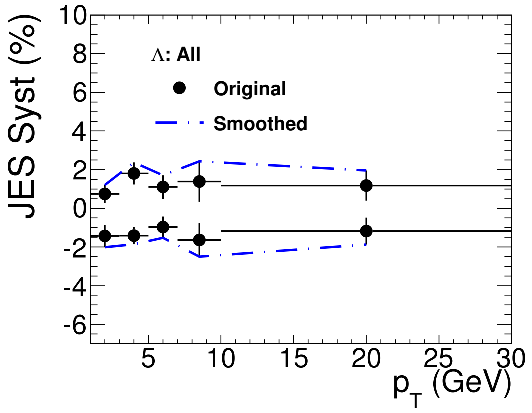


(a) p_{T} distribution

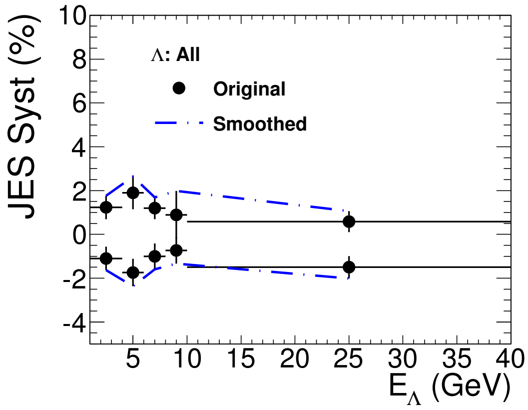
(b) Energy distribution

(c) $|\eta|$ distribution

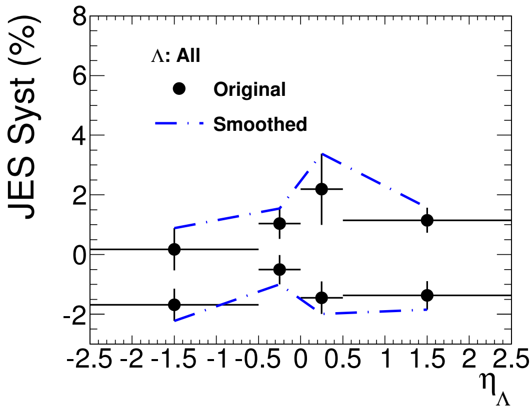
Figure 107: The Λ outside any jet systematic uncertainties due to the JES uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_Λ), multiplicity and average multiplicity.



(a) p_{T} distribution



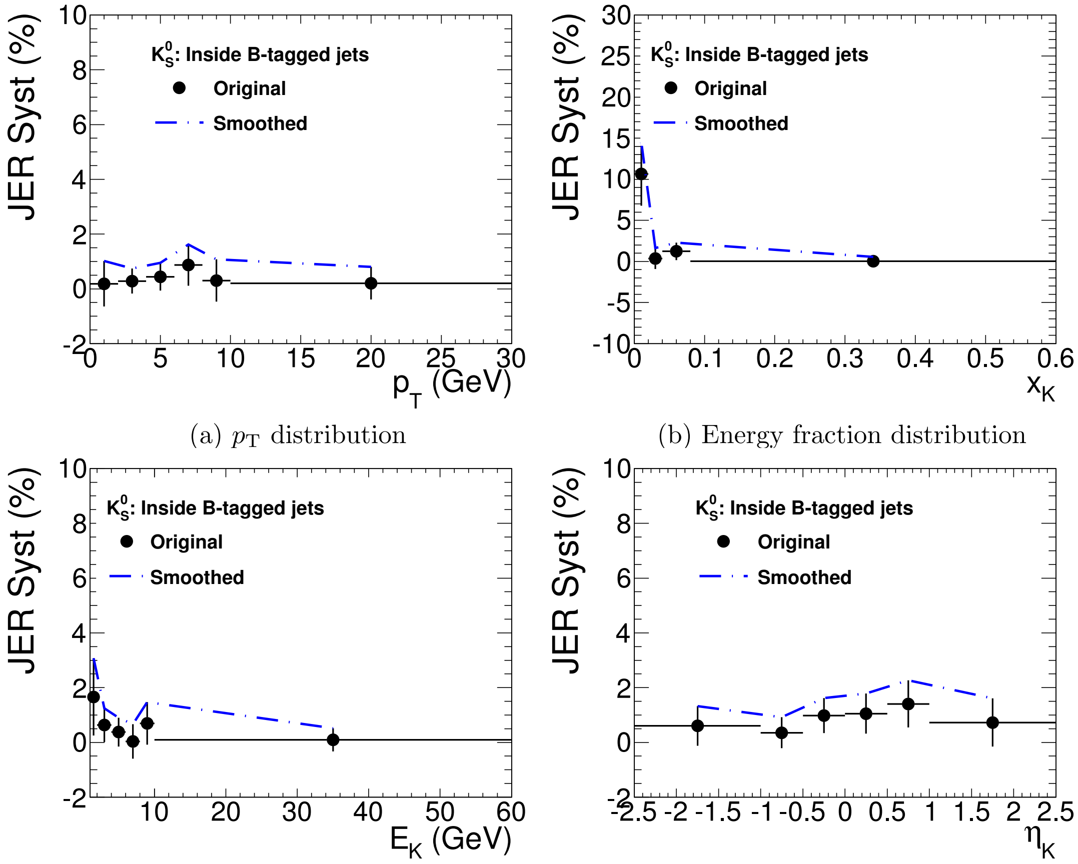
(b) Energy distribution



(c) $|\eta|$ distribution

Figure 108: The total sample Λ systematic uncertainties due to the JES uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_Λ), multiplicity and average multiplicity.

In Figures 109, 110, 111 and 112, details for the calculation of JER systematic uncertainties are shown for K^0_{S} inside b-jets, inside non b-tagged jets, outside any jets and for the total sample.



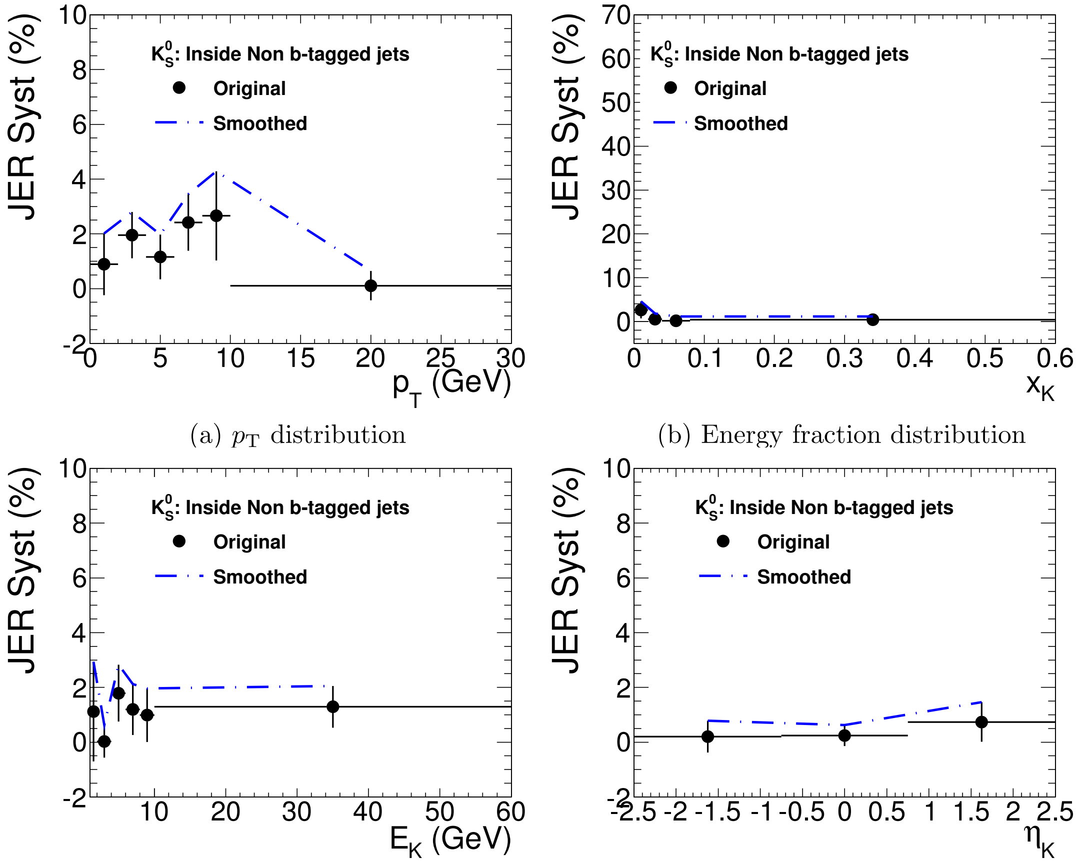
(a) p_{T} distribution

(b) Energy fraction distribution

(c) Energy distribution

(d) $|\eta|$ distribution

Figure 109: The K^0_{S} inside b-jets systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



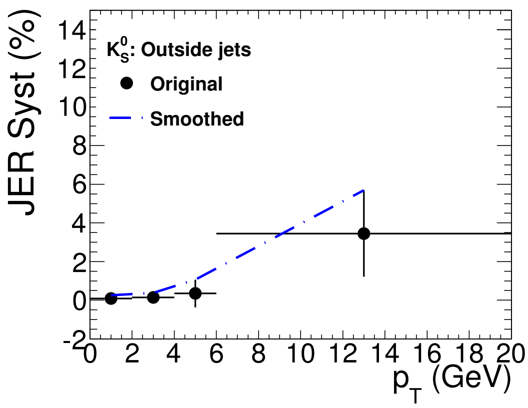
(a) p_{T} distribution

(b) Energy fraction distribution

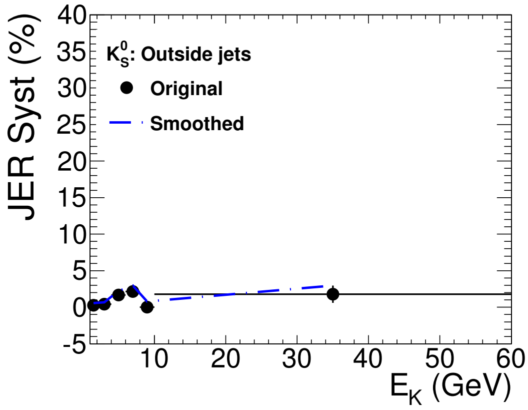
(c) Energy distribution

(d) $|\eta|$ distribution

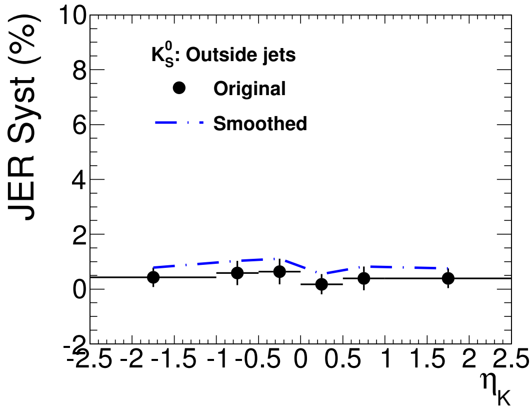
Figure 110: The K^0_{S} inside not b-tagged jets systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



(a) p_{T} distribution

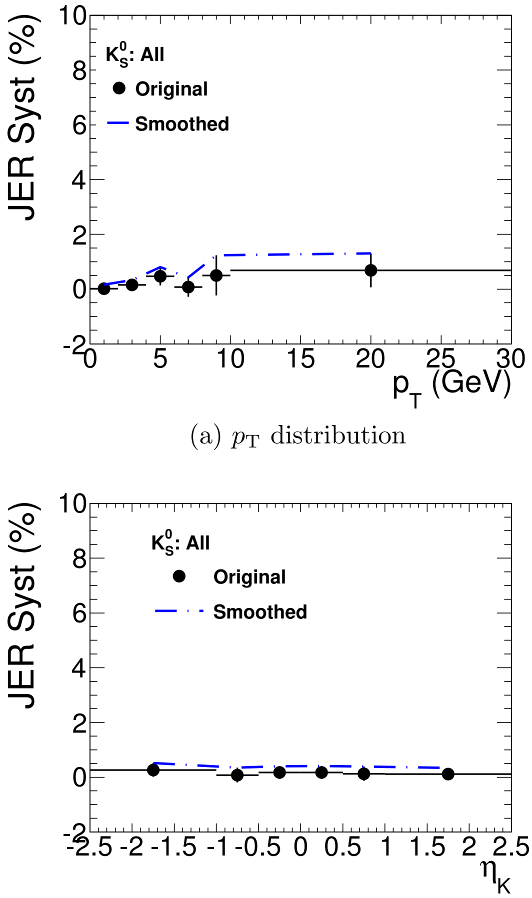


(b) Energy distribution



(c) $|\eta|$ distribution

Figure 111: The K_{S}^0 outside any jet systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.



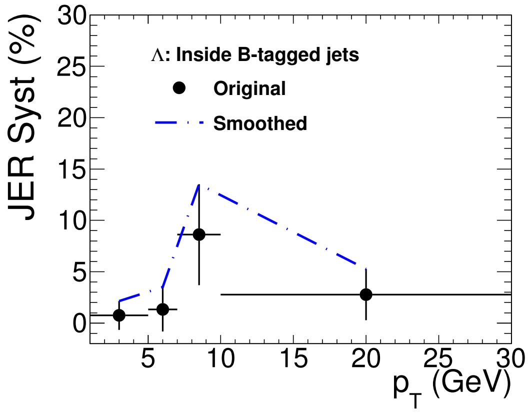
(a) p_{T} distribution

(b) Energy distribution

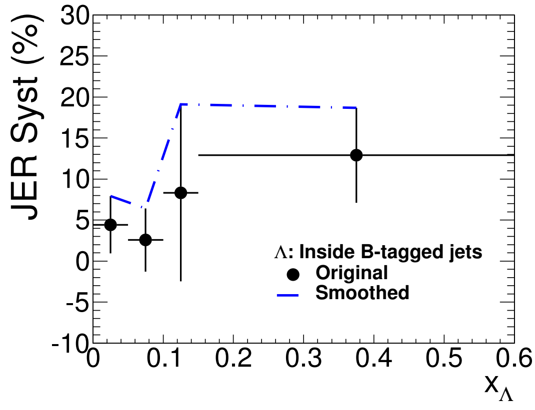
(c) $|\eta|$ distribution

Figure 112: The total sample K^0_{S} systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_K), multiplicity and average multiplicity.

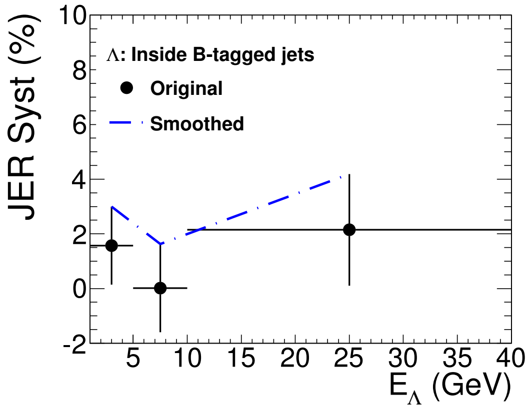
In Figures 113, 114, 115 and 116, details for the calculation of JER systematic uncertainties are shown for Λ inside b-jets, inside non b-tagged jets, outside any jets and for the total sample.



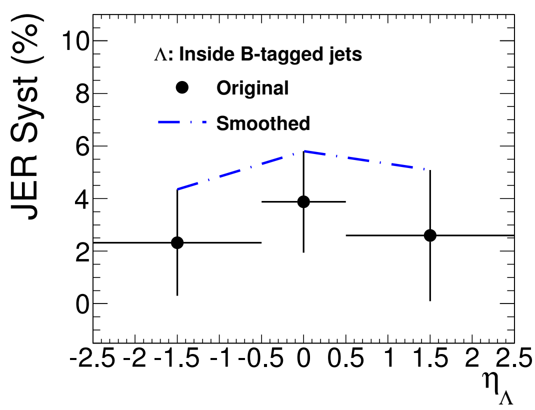
(a) p_{T} distribution



(b) Energy fraction distribution

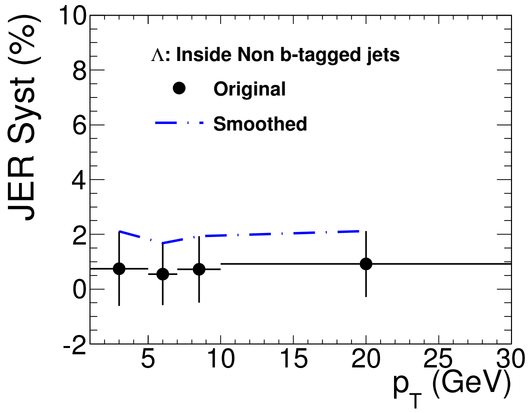


(c) Energy distribution

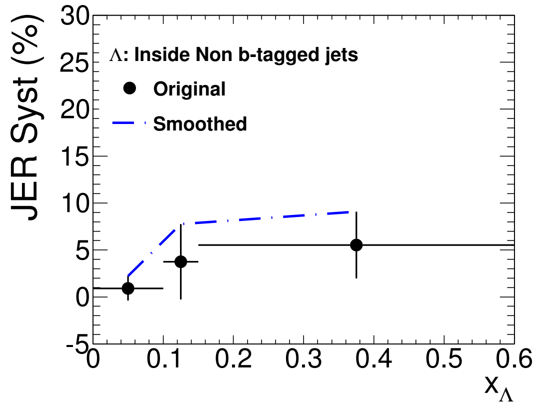


(d) $|\eta|$ distribution

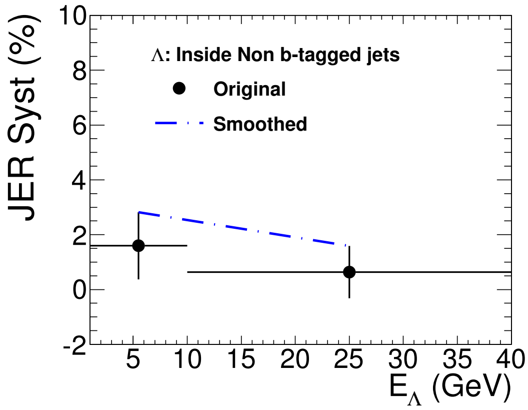
Figure 113: The Λ inside b-jets systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_{Λ}), multiplicity and average multiplicity.



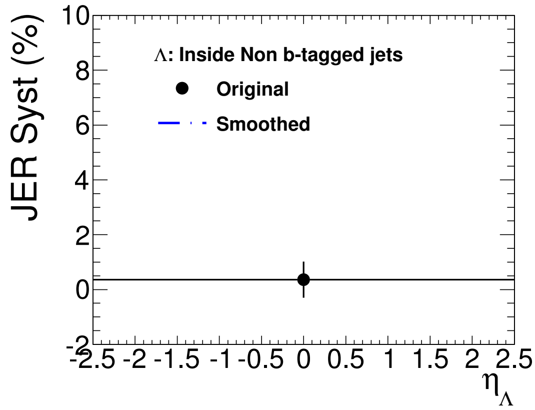
(a) p_{T} distribution



(b) Energy fraction distribution

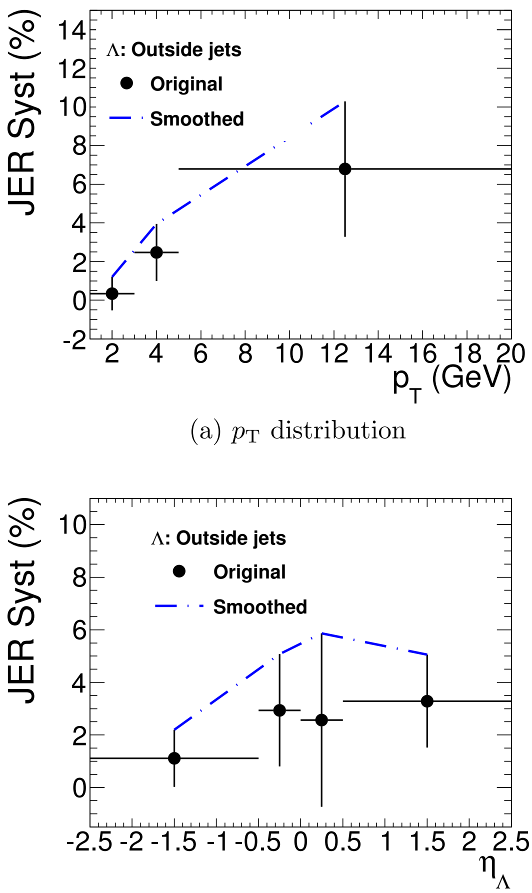


(c) Energy distribution



(d) $|\eta|$ distribution

Figure 114: The Λ inside not b-tagged jets systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_{Λ}), multiplicity and average multiplicity.

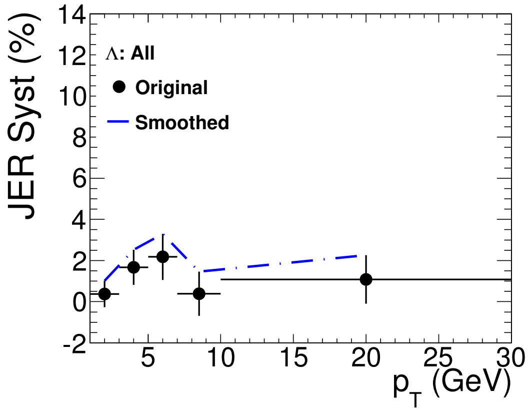


(a) p_{T} distribution

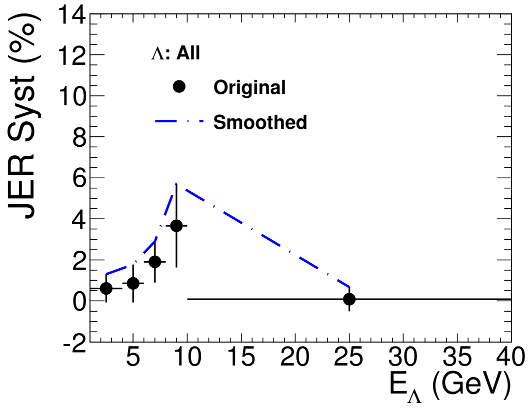
(b) Energy distribution

(c) $|\eta|$ distribution

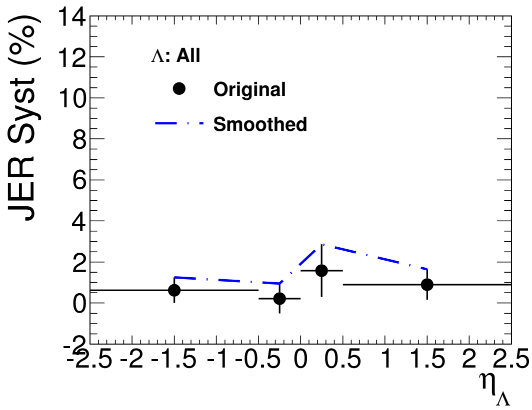
Figure 115: The Λ outside any jet systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_Λ), multiplicity and average multiplicity.



(a) p_{T} distribution



(b) Energy distribution



(c) $|\eta|$ distribution

Figure 116: The total sample Λ systematic uncertainties due to the JER uncertainty as a function of p_{T}, $|\eta|$, energy, energy fraction (x_Λ), multiplicity and average multiplicity.

References

[1] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. CRC Press, 2015.

[2] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley and Sons, 1984.

[3] S. Weinberg, *The quantum theory of fields*. Cambridge University Press, 2005.

[4] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*. Cambridge University Press, 1996.

[5] G. M. Prosperi, M. Raciti and C. Simolo, PLUTO Collaboration, *On the running coupling constant in QCD*, Prog. Part. Nucl. Phys. **58** (2007) 387–438, arXiv:hep-ph/0607209.

[6] M. Tanabashi et al., PDG Collaboration, *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001.

[7] ZEUS Collaboration, *Measurement of high-Q^2 charged current deep inelastic scattering cross sections with a longitudinally polarised positron beam at HERA*, Eur. Phys. J. C **70** no. 4, (2010) 945–963, arXiv:1008.3493 [hep-ex].

[8] R. P. Feynman, *Photon-Hadron Interactions*. W. A. Benjamin, 1971.

[9] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** no. 2, (1977) 298 – 318.

[10] H1 and ZEUS Collaborations, *Combined measurement and QCD analysis of the inclusive $e^{\pm}p$ scattering cross sections at HERA*, JHEP **01** (2010) 109, arXiv:0911.0884 [hep-ex].

[11] J. C. Collins and D. E. Soper and G. Sterman, *Factorization for short distance hadron-hadron scattering*, Nucl. Phys. B **261** (1985) 104 – 142.

[12] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10**, 531 (1963).

[13] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49**, 652 (1973).

[14] L. Wolfenstein, *Parametrization of the KM matrix*, Phys. Rev. Lett. **51**, 1945 (1983).

[15] A. Schaelicke et al., *Event generator for particle production in high-energy collisions*, Prog. Part. Nucl. Phys. **53** (2004) 329–338, arXiv:hep-ph/0311270 [hep-ph].

[16] V. V. Sudakov, *Vertex parts at very high-energies in quantum electrodynamics*, Sov. Phys. JETP **3** (1956) 65–71. [Zh. Eksp. Teor. Fiz.30,87(1956)].

[17] Y. I. Azimov and Y. L. Dokshitzer and V. A. Khoze and S. I. Troian, *Similarity of parton and hadron spectra in QCD jets*, Z. Phys. C **27** (1985) 65.

[18] B. Andersson., G. Gustafson, G. Ingelman and T. Sjöstrand , *Parton fragmentation and string dynamics*, Phys. Rep. **97** (1983) 31.

[19] R. D. Field, S. Wolfram, *A QCD model for e^+e^- annihilation*, Nucl. Phys. B **213** (1983) 65.

[20] B. R. Webber, *Fragmentation and Hadronization*, Z. Phys. C **A15S1** (2000) 577–606, arXiv:hep-ph/9912292 [hep-ph].

[21] ATLAS Collaboration, *Study of jet shapes in inclusive jet production in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, Phys. Rev. D **83** (2011) 052003, arXiv:1101.0070 [hep-ex].

[22] ATLAS Collaboration, *Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s}$ = 900 GeV and 7 TeV with the ATLAS detector*, Phys. Rev. D **83** (2011) 112001, arXiv:1012.0791 [hep-ex].

[23] T. Sjöstrand et al., *High-energy-physics event generation with PYTHIA 6.1*, Comput. Phys. Commun. **135** (2001) 238, arXiv:hep-ph/0010017.

[24] P. Z. Skands, *Tuning Monte Carlo generators: The Perugia tunes*, Phys. Rev. D **82** (1996) 074018, arXiv:1005.3457 [hep-ex].

[25] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv:hep-ph/0011363.

[26] J. Butterworth, J. Forshaw and M. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. C **72** (1996) 637, arXiv:hep-ph/9601371.

[27] S. Frixione, F. Stoeckli, P. Torrielli and B.R. Webber, *NLO QCD corrections in Herwig++ with MC@NLO*, JHEP **01** (2011) 053, arXiv:1010.0568 [hep-ex].

[28] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[29] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **0307** (2003) 001, arXiv:hep-ph/0206293v2 [hep-ph].

[30] G. P. Salam and G. Soyez, *A practical seedless infrared-safe cone jet algorithm*, JHEP **0705** (2007) 086, arXiv:0704.0292 [hep-ph].

[31] S. Catani and Y. L. Dokshitzerab and M. H. Seymour and B. R. Webber, *Longitudinally-invariant k_t-clustering algorithms for hadron-hadron collisions*, Nucl. Phys. B **406** (1993) 187.

[32] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (1993) 3160, arXiv:hep-ph/9305266 [hep-ph].

[33] Y. L. Dokshitzer and G. D. Leder and S. Moretti and B. R. Webber, *Better jet clustering algorithms*, JHEP **9708** (1997) 001, arXiv:hep-ph/9707323 [hep-ph].

[34] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv:1111.6097 [hep-ex].

[35] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ex].

[36] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[37] The ATLAS Inner Detector Community, *ATLAS inner detector: Technical design report*, CERN-LHCC-97-16, ATLAS-TDR-4 (1997). https://cds.cern.ch/record/331063.

[38] A. Negri on behalf of the ATLAS TDAQ Collaboration, *ATLAS TDAQ system integration and commissioning*, Nucl. Instr. Meth. A **623** (2010) 519.

[39] R. Fruehwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instr. Meth. A **262** (1987) 444.

[40] W. Lampl et al., *Calorimeter clustering algorithms: description and performance*, ATLAS-LARG-PUB-2008-002 (2008). http://cds.cern.ch/record/1099735.

[41] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. C **72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[42] S. Hassini et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, Nucl. Instr. Meth. A **572** (2007) 77.

[43] Th. Lagouri et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC at CERN*, IEEE Trans. Nucl. Sci. **51** (2004) 3030.

[44] D. Adams et al., *Track reconstruction in the ATLAS Muon Spectrometer with MOORE 007*, ATL-SOFT-2003-007 (2003). https://cds.cern.ch/record/685500.

[45] ATLAS Collaboration, *Properties of Jets and Inputs to Jet Reconstruction and Calibration withthe ATLAS Detector Using Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2010-053 (2010). https://cds.cern.ch/record/1281310.

[46] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. C **73** (2013) 2304, arXiv:1112.6426 [hep-ex].

[47] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 17, arXiv:1406.0076 [hep-ex].

[48] ATLAS Collaboration, *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*, Eur. Phys. J. C **73** (2013) 2305, arXiv:1203.1302 [hep-ex].

[49] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s}$=8 TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv:1510.03823 [hep-ex].

[50] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST **11** (2016) P04008, arXiv:1512.01094 [hep-ex].

[51] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102 (2011). https://cds.cern.ch/record/1369219.

[52] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039 [physics.data-an].

[53] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s}$ = 7 TeV with ATLAS*, Eur. Phys. J. C **72** (2012) 1844, arXiv:1108.5602 [hep-ex].

[54] S. Calvente, J. LLorente and F. Barreiro, *The importance of jet shapes for tagging purposes*, Nucl. Phys. B Proc. **273-275** (2016) 2761–2763.

[55] ATLAS Collaboration, *Measurement of jet shapes in top pair events at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2676, arXiv:1307.5749 [hep-ex].

[56] Javier LLorente, *QCD tests in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC.* PhD thesis, Universidad Autónoma de Madrid, 2015.

[57] CDF Collaboration, *Measurement of b-jet shapes in inclusive jet production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Rev. D **78** (2008) 072005, arXiv:0806.1699 [hep-ex].

[58] J. Gallicchio and M.D. Schwartz, *Seeing in Color: Jet Superstructure*, Phys. Rev. Lett. **105** (2010) 022001, arXiv:1001.5027 [hep-ph].

[59] CMS Collaboration, *Measurement of the t-channel single-top-quark production cross section and of the —Vtb— CKM matrix element in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **06**, 090 (2014), arXiv:1403.7366 [hep-ex].

[60] A. Ali, F. Barreiro and Th. Lagouri, *Prospects of measuring the CKM matrix element —V_{ts}— at the LHC*, Phys. Lett. B **693** (2010) 44, arXiv:1005.4647.

[61] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, M. Wiedermann, – *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034, arXiv:1007.1327 [hep-ex].

[62] A. D. Martin and W. J. Stirling and R. S. Thorne and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** no. 2, (2009) 189–285, arXiv:0901.0002 [hep-ph].

[63] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys. Lett. B **726** (2013) 88, arXiv:1307.1427 [hep-ex].

[64] CMS Collaboration, *Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to Z Boson Pairs*, Phys. Rev. Lett. **111** (2013) 081803, arXiv:1212.6639 [hep-ex].

[65] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173, arXiv:1806.00425 [hep-ex].

[66] CMS Collaboration, *Observation of $t\bar{t}H$ Production*, Phys. Rev. Lett. **120** (2018) 231801, arXiv:1804.02610 [hep-ex].

[67] ATLAS Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59, arXiv:1808.08238 [hep-ex].

[68] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, Phys. Rev. Lett. **121** (2018) 121801, arXiv:1808.08242 [hep-ex].

[69] ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072001, arXiv:1811.08856 [hep-ex].

[70] CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector*, Phys. Rev. Lett. **779** (2018) 283, arXiv:1708.00373 [hep-ex].

[71] T. Sjöstrand and S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[72] J. Gallicchio et al., *Multivariate discrimination and the Higgs+W/Z search*, JHEP **1104** (2011) 069, arXiv:1010.3698 [hep-ph].

[73] J. D. Bjorken and S. J. Brodsky, *Statistical Model for Electron-Positron Annihilation into Hadrons*, Phys. Rev. D **1** (1970) 1416.

[74] ATLAS Collaboration, *Measurement of K^0_S and Λ^0 production in $t\bar{t}$ dileptonic events in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 1017, arXiv:1907.10862 [hep-ex].

[75] PLUTO Collaboration, *Inclusive K^0 production in e^+e^- annihilation for* $9.3 < \sqrt{s} < 31.6$ *GeV*, Phys. Lett. B **104** (1981) 79.

[76] JADE Collaboration, *Charged particle and neutral kaon production in e^+e^- annihilation at PETRA*, Z. Phys. C **20** (1983) 1594.

[77] TPC Collaboration, *K^{*0} and K^0_S Meson Production in e^+e^- Annihilations at 29 GeV*, Phys. Rev. Lett. **53** (1984) 2378.

[78] MARK II Collaboration, *Measurement of K^* and K^0 inclusive rates in e^+e^- annihilation at 29 GeV*, Phys. Rev. D **31** (1985) 3013.

[79] HRS Collaboration, *Hadron production in e^+e^- annihilation at 29 GeV*, Phys. Rev. D **35** (1987) 2639.

[80] CELLO Collaboration, *Inclusive strange particle production in e^+e^- annihilation*, Z. Phys. C **46** (1990) 397.

[81] TASSO Collaboration, *Strange meson production in e^+e^- annihilation*, Z. Phys. C **47** (1990) 183.

[82] ALEPH Collaboration, *Production of K^0 and Λ in hadronic Z decays*, Z. Phys. C **64** (1994) 361.

[83] DELPHI Collaboration, *Production characteristics of K^0 and light meson resonances in hadronic decays of the Z^0*, Z. Phys. C **65** (1995) 587.

[84] OPAL Collaboration, *The production of neutral kaons in Z^0 decays and their Bose-Einstein correlations*, Z. Phys. C **67** (1995) 389.

[85] L3 Collaboration, *Measurement of inclusive production of neutral hadrons from Z decays*, Phys. Lett. B **328** (1994) 223.

[86] ATLAS Collaboration, *K^0_S and Λ production in pp interactions at $\sqrt{s} = 0.9$ and 7 TeV measured with the ATLAS detector at the LHC*, Phys. Rev. D **85** (2012) 012001, arXiv:1111.1297 [hep-ex].

[87] ATLAS Collaboration, *Measurement of the transverse polarization of Λ and $\bar{\Lambda}$ hyperons produced in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, Phys. Rev. D **91** (2015) 032004, arXiv:1412.1692 [hep-ex].

[88] CMS Collaboration, *Measurement of neutral strange particle production in the underlying event in proton − proton collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. D **88** (2013) 052001, arXiv:1305.6016 [hep-ex].

[89] ALICE Collaboration, *Strange particle production in proton-proton collisions at $\sqrt{s} = 0.9$ TeV with ALICE at the LHC*, Eur. Phys. J. C **71** (2011) 1594, arXiv:1012.3257 [hep-ex].

[90] LHCb Collaboration, *Prompt K^0_S production in pp collisions at $\sqrt{s} = 0.9$ TeV*, Phys. Lett. B **693** (2010) 69, arXiv:1008.3105 [hep-ex].

[91] UA5 Collaboration, *Kaon production in $\bar{p}p$ interactions at c.m. energies from 200 to 900 GeV*, Z. Phys. C **41** (1988) 179.

[92] G. Bocquet et al., *Inclusive production of strange particles in $p\bar{p}$ collisions at $\sqrt{s} = 630$ GeV with UA1*, Phys. Lett. B **366** (1996) 441.

[93] CDF Collaboration, *K^0_S and Λ^0 production studies in $p\bar{p}$ collisions at $\sqrt{s} = 1800$ and 630 GeV*, Phys. Rev. D **72** (2005) 052001, arXiv:hep-ex/0504048.

[94] ZEUS Collaboration, *Neutral strange particle production in deep inelastic scattering at HERA*, Z. Phys. C **68** (1995) 29, arXiv:hep-ex/9505011.

[95] H1 Collaboration, *Strangeness production in deep-inelastic positron-proton scattering at HERA*, Nucl. Phys. B **480** (1996) 3, arXiv:hep-ex/9607010.

[96] P. H. Stuntebeck et al., *Inclusive production of K^0_1, Λ^0 and $\bar{\Lambda}^0$ in 18.5-$\frac{GeV}{c}$ $\pi^{\pm}p$ interactions*, Phys. Rev. D **9** (1974) 608.

[97] R. Sugahara et al., *Inclusive strange particle production in $\pi^- p$ interactions at 6 GeV/c*, Nucl. Phys. B **156** (1979) 237.

[98] V. Blobel et al., Bonn-Hamburg-Munich Collaboration, *Multiplicities, topological cross sections, and single particle inclusive distributions from pp interactions at 12 and 24 GeV/c*, Nucl. Phys. B **69** (1974) 454.

[99] P. Bosetti et al., Aachen-Bonn-CERN-Cracow Collaboration, *Inclusive strange particle production in π^+p interactions at 16 GeV/c*, Nucl. Phys. B **94** (1975) 21.

[100] D. Bogert et al., *Inclusive production of neutral strange particles in 250 GeV/c π^-p interactions*, Phys. Rev. D **16** (1977) 2098.

[101] F. Barreiro et al., *Inclusive neutral-strange-particle production in π^-p interactions at 15 GeV/c*, Phys. Rev. D **17** (1978) 669.

[102] D. Brick et al., *Inclusive production of neutral strange particles by 147 GeV/c$\pi^+/K^+/p$ interactions in hydrogen*, Nucl. Phys. B **164** (1980) 1.

[103] H. Kichimi et al., *Inclusive study of strange-particle production in pp interactions at 405 GeV/c*, Phys. Rev. D **20** (1979) 37.

[104] F. Lopinto et al., *Inclusive K^0, Λ, $K^{*\pm}(890)$, and $\Sigma^{*\pm}(1385)$ production in pp collisions at 300 GeV/c*, Phys. Rev. D **22** (1980) 573.

[105] K. Abe et al., SLAC Hybrid Facility Photon Collaboration, *Inclusive photoproduction of neutral strange particles at 20 GeV*, Phys. Rev. D **29** (1984) 1877.

[106] I. Cohen et al., *Inclusive K_S^0 and Λ Electroproduction*, Phys. Rev. Lett. **40** (1978) 1614.

[107] R.G. Hicks et al., *Muoproduction of Neutral Strange Hadrons at 225 GeV*, Phys. Rev. Lett. **45** (1980) 765.

[108] H. Grassler et al., Aachen-Bonn-CERN-Munich-Oxford Collaboration, *Inclusive neutral strange particle production in νp interactions*, Nucl. Phys. B **194** (1992) 1.

[109] V.V. Ammosov et al., *Charged current events with neutral strange particles in high-energy antineutrino interactions*, Nucl. Phys. B **177** (1981) 365.

[110] J. Binnewies, B. A. Kniehl and G. Kramer, *Neutral kaon production in e^+e^-, ep and $p\bar{p}$ collisions at next-to-leading order*, Phys. Rev. D **53** (1996) 3573, `arXiv:hep-ph/9506437`.

[111] T.S. Biro, H.B. Nielsen and J. Knoll, *Color rope model for extreme relativistic heavy ion collisions*, Nucl. Phys. B **245** (1984) 449.

[112] Ch. Bierlich, G. Gustafson, L. Loennblad and A. Tarasov, *Effects of overlapping strings in pp collisions*, JHEP **03** (2015) 148, `arXiv:1412.6259 [hep-ex]`.

[113] S. Agostinelli et al., *GEANT4- a simulation toolkit*, Nucl. Instr. Meth. A **506** (2008) 250.

[114] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[115] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv:0707.3088 [hep-ph].

[116] J. Pumplin et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **07** (2002) 012.

[117] H.L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ex].

[118] T. Gleisberg et al., *Event generator with Sherpa 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ex].

[119] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[120] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021 (2014). http://cds.cern.ch/record/1966419.

[121] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv:1512.01178 [hep-ph].

[122] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[123] B. Kersevan and E. Richter-Was, *The Monte Carlo events generator Acer-MC version 1.0 with interfaces to Phythia 6.2 and Herwig 6.3*, Comput. Phys. Commun. **149** (2003) 142, arXiv:hep-ph/0201302.

[124] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. D **82** (2010) 054018, arXiv:1005.4451 [hep-ex].

[125] R. Gavin, Y. Li, F. Petriello, S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, Comput. Phys. Commun. **182** (2011) 2388, `arXiv:1011.3540 [hep-ex]`.

[126] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** (2013) 2518, `arXiv:1302.4393 [hep-ex]`.

[127] ATLAS Collaboration, *Measurement of the cross-section for top-quark pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using final states with two high-p_T leptons*, JHEP **05** (2012) 059, `arXiv:1202.4892 [hep-ex]`.

[128] M. Aleksa et al., *ATLAS combined test beam: Computation and validation of the electronic calibration constants for the electromagnetic calorimeter*, ATL-LARG-PUB-2006-003 (2006). `http://cds.cern.ch/record/942528`.

[129] ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, Eur. Phys. J. C **70** (2010) 1193, `arXiv:1007.5423 [hep-ex]`.

[130] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **73** (2013) 2304, `arXiv:1112.6426 [hep-ex]`.

[131] ATLAS Collaboration, *Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. D **86** (2012) 014022, `arXiv:1112.6297 [hep-ex]`.

[132] D.W. Miller, A. Schwartzman and D. Su, *Measurement of hadronic event shapes at high Q^2 with the ATLAS detector at the LHC*, ATL-COM-PHYS-2011-534 (2011). `https://cds.cern.ch/record/1352100`.

[133] ATLAS Collaboration, *Measurement of top quark pair differential cross-sections in the dilepton channel in pp collisions at $\sqrt{s}$ = 7 and 8 TeV with ATLAS*, Phys. Rev. D **94** (2016) 092003, `arXiv:1607.07281 [hep-ex]`.

[134] ATLAS Collaboration, *Charged particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New J. Phys. **13** (2011) 053033, `arXiv:1012.5104 [hep-ex]`.

[135] ATLAS Collaboration, *Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-072 (2010). https://cds.cern.ch/record/1281363.

[136] M. Bähr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639, arXiv:0803.0883 [hep-ph].

[137] G. D'Agostini, *A multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. A **362** (1995) 487.

[138] T. Adye, *Unfolding algorithms and tests using RooUnfold*, PHYSTAT 2011 Proc. (2011), arXiv:1105.1160 [hep-ex].

[139] ATLAS Collaboration, *Measurement of the inclusive jet cross-sections in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **09** (2017) 020, arXiv:1706.03192 [hep-ex].

[140] G. Aad et al., *Track Reconstruction Efficiency in $\sqrt{s} = 7$ TeV Data for Tracks with $p_T > 100$ MeV*, ATL-PHYS-INT-2011-001 (2011). https://cds.cern.ch/record/1319318.

[141] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv:1210.6210 [hep-ex].

[142] HepData. https://www.hepdata.net/record/ins1746286.

[143] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 tune*, Eur. Phys. J. C **74** (2014) 3024, arXiv:1404.5630 [hep-ph].

[144] S. Gieseke, P. Kirchgaeßer and S. Plätzer , *Baryon production from cluster hadronisation*, Eur. Phys. J. C **78** (2018) 99, arXiv:1710.10906 [hep-ex].

[145] Rivet. https://rivet.hepforge.org/analyses/ATLAS_2019_I1746286.html.