Physikalisches Institut
Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität
Freiburg Im Breisgau, Germany

A model-independent "General Search" for new physics with the ATLAS detector at LHC

von

Riccardo Maria BIANCHI

Betreut von
Prof. Dr. Gregor HERTEN

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Mathematik und Physik
der Albert-Ludwigs-Universität
Freiburg im Breisgau

Eingereicht im Januar 2014

Dekan:	Prof. Dr. Michael Růžička
Leiter der Arbeit:	Prof. Dr. Gregor Herten
Referent:	Prof. Dr. Gregor Herten
Koreferent:	Prof. Dr. Karl Jakobs
Prüfer (Experiment):	Prof. Dr. Markus Schuhmacher
Prüfer (Theorie):	JProf. Dr. Harald Ita

Tag der Verkündigung des Prüfungsergebnisses: 18.03.2014

Alla Professoressa Fernanda Pastore (1949–2010),
che per prima mi ha trasmesso il gusto per lo studio della Fisica delle Particelle
e che mi ha incoraggiato ad arricchirmi di esperienze sempre nuove,
con il suo intramontabile entusiasmo e la sua passione contagiosa,
sempre conditi da un gaio, inguaribile, buon umore.

In the memory of Prof. Fernanda Pastore (1949–2010),
who has been the first to instill in me a passionate interest for Particle Physics
and who always pushed me toward new work exciting and enriching experiences,
with her evergreen, contagious enthusiasm,
never without her own special smiley good humor.

❧ ❧

E al caro Giustino Peca (1936–2014),
per l'affettuosa amicizia, la simpatia e il genuino interesse
con cui ha sempre seguito i miei studi e il mio lavoro.
"Un grande abbraccio, Giustino!"

And in the memory of my dear friend Giustino Peca (1936–2014),
for his warm friendship and the sincere interest
he has always shown for my studies and for my work.
"A big hug for you, Giustino!"

Sê todo em cada coisa.
Põe quanto és No mínimo que
fazes.

Sii tutto in ogni cosa. Poni quanto sei nel minimo che fai.

Ricardo Reis *(Fernando Pessoa)*

Abstract

The LHC particle collider accelerates bunches of protons at energies never reached before, thus opening a completely new landscape of new physics. In this scenario the number of possible physics processes and signatures becomes virtually infinite, making the setup of dedicated analyses impossible. Moreover there are regions of the phase-space where signals of new physics are not very likely to be found, or where suitable theoretical models are missing, and it is important to be able to reveal new processes from such regions as well.

At the time this Thesis was started, no model-independent analysis had been set for the ATLAS experiment at LHC. The goal of this work was then to conceive and develop a new model-independent "General Search" for ATLAS, and to explore its possibilities.

The new analysis has been then implemented and run over the first data which have been collected by the ATLAS experiment at a centre-of-mass energy of 7 TeV. This work presents the motivation of the data analysis, describes its implementation and shows the results. The data have shown a very good agreement with the Standard Model expectation and no evidence of new physics has been observed. Nevertheless, an efficient strategy and methodology for a new model-independent "General Search" have been defined and they are ready to be used in a next version of this analysis, over a larger set of experimental data.

During this work, a new innovative software framework has been also conceived and developed to ease the implementation of physics analysis code using *Computer-Aided-Software-Engineering* (CASE) principles. The framework has been successfully used to analyze the very first LHC data, and then it has been transformed into an open-source modular framework for HEP data analysis, and presented at Physics and Computer Science international conferences.

Acknowledgments

With these few words I wish to thank all the nice people I met and who helped me, in very many different ways, during these many years of study. They all contributed to the success of this work.

Thanks!!

❧ ❧

First of all I would like to warmly thank *Fernanda Pastore*, my Professor at "Roma Tre" university, whose memory this Thesis is dedicated to. If I chose Particle Physics as my field of specialization, is also for the passion she injected into me with her cheerful and frank enthusiasm, during the many years I studied with her.

And I would like to warmly thank my supervisor, *Prof. Dr. Gregor Herten*: for the kind welcome in his group, his guidance and the warm friendship that he showed to me during these many — and sometimes difficult — years.

I also would like to thank *Sascha Caron* and *Renaud Brunelière* for their help, their ideas and the useful suggestions during these years of work; and to *Andrea Di Simone* for having reviewed and proofread this Thesis, contributing to it with his very useful suggestions. And thanks to *Prof. Dr. Ulrich Landgraph*, for his kind help during the years at Freiburg.

Thanks to *Simone Amoroso*, for the help and his precious suggestions.

I would like to lovingly thank *Cecilia Antonelli*, my wife, for her love, support friendship and happiness! Her bright, cheerful and contagious smile always guides me, every day of my life. "Thanks! And a big kiss for you, *Ceci*, my dear Love!"

Many thanks to *Francesco Santanastasio*, my dear friend; not only for our genuine and affectionate friendship during these many years together at CERN... but also for all the so beautiful and detailed explanations about physics and data analysis that he gave me during these years!! "A big hug for you, Francesco! And good luck with your new life in Rome!"

Ringrazio tanto, e affettuosamente, la mia famiglia: mio padre Mario*, mia madre* Giovanna *e mia sorella* Valeria *per tutto l'amore, l'affetto, il sostegno, l'incoraggiamento, la fiducia, la vicinanza, il supporto e l'amicizia durante questi miei 37 anni, nei momenti di grande difficoltà come nella vita di tutti i giorni. Un grande e affettuoso abbraccio per voi! E grazie! Di tutto. Sempre!*
E grazie al mio amato Rufone*, e a* Dik *e* Poldo*, per la vita che abbiamo condiviso.*

Thanks to my friends, *Stefano Colafranceschi* and *Renato Febbraro*, for their friendship and love, and for all the beautiful experiences that we lived together in the past years.

Thanks to my dear friend *Domenico Gelfusa*, for the great time that we lived together at the University, and for our warm friendship, which lasts since many years.

Thanks to my colleague and friend *Stephan Horner*, with whom I shared my Freiburg's office and a beautiful travel through the northern India: we shared a very nice time together!

Many thanks to *Kim Temming*, for her so kind and precious help with all that Freiburg bureaucracy: without her help, handing this Thesis in would have been much more difficult!!

Thanks to all my colleagues at Freiburg University, the old ones and the new ones, for the beautiful time we had together: *Tobias Rave, Song Xie, Florian Ahles, Michael Rammensee, Zuzana Rurikova, Adil Aktas, Ulrich Holland, Jan Erik, Martina Pagacova, Stephanie Zimmermann, Asen Christov, Debra Lumb, Janet Dietrich, Valerio Consorti, Manfredi Ronzani, Fabio Cardillo (who helped me a lot with the German translation of my Abstract), Kathrine Störig, Bernard Pfeifer, Jürgen Tobias, Christoph Kendel, Bernhard Meirose, Roman Karimi, Yusuf Erdogan, Yevgen Palaguta, Stefan König* and all the others who made my staying at Freiburg so enjoyable. And many thanks to *Susanne Ackermann*, for her precious and so kind help in organizing our work-life.

I also would like to thank my colleagues of the ATLAS TDAQ group, for the very good time we shared during my CERN fellowship, during the last past years: *Giovanna Lehmann Miotto, Benedetto Gorini, Alina Corso Radu, Raul Morillo Garcia, Luca Magnoni, Giuseppe Avolio, Wainer Vandelli, Andrei Kazarov, Igor Soloviev, Tommaso Colombo, Nicoletta Garelli, Serguei Kolos, Diana Scannicchio, Georgiana Lavinia Darlea* and all the others.

And many sincere thanks, for the wonderful time we are sharing, to my current supervisor, *Prof. Joseph Boudreau*, and all my colleagues at Pittsburgh University: *Prof. James A. Mueller*, *Carlos Escobar*, *Jun Su* and *Kevin Sapp*.

Thanks to all professors and researchers at "Roma Tre" university — where I first studied and where I obtained the *Bachelor* and the *Master* degrees — for their guidance, their help and their friendship. In particular, many sincere thanks to *Toni Baroncelli*, my mentor and friend, and to *Filippo Ceradini*, my professor and supervisor; and thanks to *Mauro Iodice, Fabrizio Petrucci, Alessandra*

Tonazzo, Domizia Orestano, Cristian Dan Stanescu, Stefano Mari, Paola Celio, Ada Farilla, Paolo Branchini, Enrico Graziani, Antonio Passeri and the whole ATLAS group at "Roma Tre".

I also would like to specially thank *Prof. Alessandro Pellegrinotti*, for his help, his trust and the faith he had on me, during a difficult period of my life.

Un grande grazie a tutta la mia famiglia, per tutto il genuino e caloroso affetto che mi ha sempre donato: grazie ai miei zii Lella *e* Stefano*,* Ezio *e* Fabiola*,* Armando *e* Susi*,* Alberto *e* Gigliola*; ai miei cugini* Silvia, Ele, Marco, Adriano, Lorenzo, Oscar*; a* Gloria *e* Tito*,* Manuela*,* Martina*,* Tommaso*,* Michele*,* Ginevra*,* Roberta *e a tutti gli altri!.*

Un grande grazie anche a tutta la famiglia Antonelli*, per il caloroso affetto con cui mi hanno da subito accolto, e per la loro contagiosa simpatia. Grazie* Alberto, Giovanna, Fiamma, Federico, Fabio, Chicco, Grazia, Paolo, Alessandra, Gabriele, Flavio, zia Silvina, Simone, Manuela, Simonetta, Roberto *e tutti gli altri!*

I would like to thank my old good friends *Fabrizio Ciaccia*, *Artiano Nasehatoen*, *Francesco Mori* and *Giuseppe De Marco* for the many beautiful years we have spent together, since the times of the high-school... when we were 14 years old! Thanks also to my old friend and "Civil Service" mate Andrea Gabrieli; "Good luck for everything, Andrea!".

Thanks a lot to all the very nice people I met in Bötzingen, the beautiful German village where I lived during my years at Freiburg University. In particular, many thanks to *Monika* and *Klaus Jakob* and *Beate* and *Klemens Rodemann*, for the very warm welcome, their friendship and for the help during those years.

Many thanks to *Dr. Sabina Renner*, *Dr. Marc Schnitzler* and *Dr. Frank Tietz* for the precious and kind help during these many years, and for their friendship.

Thanks to *Francesca Marroncini*, for her love and her precious help and support, sometimes during very difficult moments. Moreover, if I began this PhD, it's also for the positive encouragement she gave to me when was time to choose.

Grazie al compianto Giustino Peca*, per l'affetto, l'amicizia e per l'interesse e la simpatia che ha sempre dimostrato per il mio lavoro e i miei studi. "Un grande abbraccio, Giustino!"*

Grazie a Mario Rosati*, per l'amicizia che mi ha sempre mostrato in tanti anni che ci conosciamo.*

Un grande grazie e un abbraccio anche a Annamaria Rubino, Adriana Moretti, Mirella Valletta e Adriana Barbiellini, per il grande affetto che mi hanno sempre dimostrato in tutti questi anni.

Un grazie speciale va a Vittorio Giardiello*, per la simpatica amicizia e per aver sempre curato i nostri catorci a due e a quattro ruote!*

Thanks a lot to the fellows of the *"Keen8"*! With them I learnt how to sing and play together, and I lived the exciting experience of my first concert! Thanks to all of you Cecilia Antonelli (singer... and my wife!), Francesco "Il Maestro" Santanastasio (keyboards), Riccardo D'Aguanno (guitar), Nicolò Biancacci (guitar), Paul Harrison (bass guitar) and Alessandro Berti (drums)!

Vorrei ringraziare anche gli amici della Missione Cattolica Italiana a Ginevra: Giuseppe, Orlane, Cristina, Gianluca, Matilde, Alessandro, Chiara, Nicolò, Roberta, Michele, Giannarita, Adele, Annamaria, Lia, Serena, Massimiliano, Alice, Sr. Cleo, Sr. Marie Ange, Sr. Perle, Sr. Denise, Sr. Patricia, Sr. Jenny, Sr. Gentiane, Sr. Paulette, P. Jorge, P. Pierre, P. Francesco *e tutti gli altri, per avermi accolto con amicizia e genuino affetto e per tutti i bei momenti passati insieme durante questi anni ginevrini.*

Inoltre, un affettuoso grazie a P. Loris Piorar, P. Maurice Prestail, Don Luigi D'Errico, P. Luciano Cocco *e* P. Silvano Guglielmi *per avermi fatto scoprire, ognuno a suo modo e in tempi diversi, lo spessore e il valore della vita e l'importanza delle scelte.*

Infine, un grazie al Signore, per regalarmi ogni respiro della mia Vita.

Contents

Chapter 1

Introduction

> Absence of evidence is not evidence of absence.
>
> Carl Sagan

Nature has its rules: the Universe was born and evolves, and so they do all objects within. And, from the beginning of its Age, mankind has always looked for explanations: why objects in the Universe look like they appear, how their components behave, and how they interact among each other. From a physics point of view, the search for an answer to those questions has been translated in looking for the components of the Matter, and for the Forces that rule it out.

Giant steps have been taken in that direction, most of all in the last period of human history, when men have started to explore Nature with a rigorous scientific method. And in the very last century, physics succeeded in establishing an important milestone, depicting the Standard Model of particle physics: a collection of elementary particles and forces, which are the basic fundamental building blocks of our Universe.

And the ATLAS and CMS experiments at the LHC, have been recently set a very important result: they finally confirmed the existence of the Higgs boson and measured its mass, completing the picture of the Standard Model.

But we know that our understanding is not complete. Universe is composed not only by those objects and phenomena we can now directly see and measure, but also by other kinds of matter and forces, which we are currently unable to observe and study and whose origin is still unknown: for example, the Dark Matter and the Dark Energy, which together are supposed to fill the largest fraction of our Universe.

And thus, there's for sure much more to discover, compared to what we have revealed so far. The Standard Model of particle physics, in fact, is an incomplete model: it is not able to explain Gravity and does not contain particles that can be suitable candidate as Dark Matter components, for example; and it does not give explanations, either, for many physics evidences, like the Matter-Antimatter asymmetry we observe in our today's Universe. That's why it is important to look for new physics processes that can be able to fill the lacks in our current knowledge. The LHC accelerator can accelerate bunches of protons at an energy never reached before — up to 7 TeV per beam, when it will run at the full design energy — and a completely new landscape of possible observations has been opened, in the search for new physics.

Many theoretical models have been conceived, in order to complete the Standard Model. And many experimental dedicated analyses have been set up upon them, in order to search for evidences of such new physics processes. But, as already said before, Nature has its rules. And, hence, new processes could show themselves in ways that physicists don't think about. While looking at physics events with certain characteristics, new physics could hide behind other type of events, with different playing characters.

That's why being able to analyze experimental data also from point of views different from the standard approach is so important. And that's why we decided to start this work: to establish the strategy and the methodology for a new model-independent search for ATLAS. A model-independent analysis was missing in ATLAS, at the time this work started. But we thought that a General Search was a very important, valuable tool for a modern HEP experiment: to be ready to catch glimpses of new physics wherever they appear, no matter where; for example in very exotics processes, for which there's no space in analyses targeted at verifying specific theoretical models. It would be impossible, in fact, to analyze all kind of physics final states in a detailed and dedicated way; because the variables and the possible channels are too many. But with the General Search described in this work, a semi-automatic scan of data has been setup, where all events passing a very loose selection are classified and analyzed. In search for new physics.

1.1 Outline

A brief outline of this work is provided in the following.

In Chapter 2 will give a brief description of the Standard Model of particle physics. A short explanation of the particles landscape and of the basic fields and forces will be given.

In Chapter 3 a set of example theories beyond the Standard Model will be presented, in order to briefly describe the possible scenarios of new physics at the

LHC collider and of the physics processes and final states that can be used as their signatures.

Chapter 4 will give a short description of the LHC accelerator and of the ATLAS experiment. The different sub-detectors will be presented and analyzed.

An historical survey of model-independent searches in other HEP experiments will be found in Chapter 5. There, the different approaches to the problem will be presented, and a background will be set and analyzed for the approach taken in this work.

In Chapter 6 the new model-independent General Search for ATLAS will be presented. The strategy and the implementation will be analyzed, and the choices that have been taken will be motivated. Also the differences with other existent approaches to the problem will be outlined. Within this work an automatic algorithm has been conceived and implemented as well, starting from similar approaches in HEP, to scan data in search of discrepancies between the data and the Standard Model MonteCarlo expectation; this algorithm will be also described. Also, the comparison with the standard ATLAS 1-lepton Supersymmetry analysis will be shown, to validate the choices taken for this General Search. In the end of the Chapter, results from the analysis of 2fb^{-1} of experimental data from the ATLAS experiment will be presented, together with the statistical interpretation.

At the end, Chapter 7 will present an innovative software framework, which has been conceived and realized during this work, in order to ease the implementation of physics analysis code. The tool has been successfully used in ATLAS to analyze the very first data from LHC. The framework has been then presented in Physics and Computer Science international conferences and it has been transformed in an open-source project.

Chapter 2

The Standard Model of Particle Physics

> I think it is much more interesting to live with uncertainty than to live with answers that might be wrong.
>
> Richard Feynman

Particle physics phenomena, observed at the smallest scales and at the highest energy densities, are well described by the so-called *Standard Model of Particle Physics* (also called "SM" in the following). This chapter gives a brief overview on the Standard Model of particle physics and one of its major extension: Supersymmetry. This is not a complete description: its purpose is to introduce the main aspects of the theory, which are necessary to understand the other chapters in this Thesis. More complete information on the Standard Model of Particle Physics can be found in two classical reference books: [4][5], which also constitute the basis of the concepts and the explanations of this section.

The last part of this chapter introduces the phenomenology of *proton-proton* collisions.

Please notice the definitions used in this thesis: the speed of light c and the reduced Planck's constant $\hbar \equiv \frac{h}{2\pi}$ are considered as $\hbar \equiv c \equiv 1$; particle quantities like mass, energy and momentum will be given in multiples of *electronvolt*, eV.

2.1 The Standard Model of Particle Physics

In the SM, two fundamental entities rule the whole ordinary matter of our universe: elementary particles that make its structure, and fundamental forces that regulate the interactions among them.

To today's knowledge the matter basic elementary particles, called *fermions*, don't have an inner structure (they are point-like) and are divided into two types: *leptons* and *quarks*. The leptons — the *electron* e^-, the *muon* μ^- and the *tau* τ^- — carry a spin $S = 1/2$ and an electric charge $Q = -1$. For every lepton exists a corresponding neutral particle, called *neutrino* — ν_e, ν_μ and ν_τ respectively — with $Q = 0$. Quarks are the components of composite particles called *hadrons*. The six known quarks have different *flavour* — u, d, s, c, b, t — and they carry a fractional charge, Q, whose value is $\frac{2}{3}$, $-\frac{1}{3}$, $-\frac{1}{3}$, $\frac{2}{3}$, $-\frac{1}{3}$ and $\frac{2}{3}$ respectively. We summarize the elementary fermionic particles and their observable physical characteristics in tables 2.1 and 2.2.

The forces that exist between fermions, and which rule their interactions, are mediated by another class of particles, called *gauge bosons*, which have an integer spin $S = 1$. We summarize the bosonic particles and their observable physical quantities in table 2.3. The *photon* γ is the particle mediating the electromagnetic interactions; the eight gluon particles exchange the strong interactions among the quarks; and the three bosons Z^0, W^+ and the W^- mediate the weak interactions.

For every fermion, an anti-particle with same mass, but opposite quantum numbers, exists. Thus the *positron*, an electron-like particle with positive charge, is the corresponding anti-particle of the electron. Among the gauge bosons, the W^+ and the W^- are each other's anti-particle.

In the tables 2.1 and 2.2, as one can see, leptons and quarks can be divided into three categories, or "*generations*", which share the same quantum numbers. Actually, why these particles — which are all the only fundamental particles we know so far — are organized this way, is still an open question. All phenomena and forms of ordinary matter that one can commonly observe, are well described by the four fundamental forces (and their corresponding gauge bosons) and the particles belonging to the *first generation* only: the *up* and *down* quarks (which compose, for example, the neutron and the protons, forming the nucleus of any atom), the electron (which completes the atom structure, together with the nucleus) and the electron-neutrino ν_e, which is involved in the particle decays. All the other leptons and quarks only appear in phenomena observable in cosmic rays and in high-energy experiments. Thus that observation gives raise to several questions: the reason why there are three generations of particles; why they do differ in masses; and why they have such different masses, which span across more than three order of magnitude. These are still open questions, which the Standard Model failed to answer.

Standard Model Quarks				
Generation	*Particle*	*Mass* [MeV] [1]	*Charge* [e]	*Spin*
1	u	$2.3^{+0.7}_{-0.5}$	$+2/3$	$1/2$
	d	$4.8^{+0.7}_{-0.3}$	$-1/3$	$1/2$
2	c	$1.275 \pm 0.025 \times 10^3$	$+2/3$	$1/2$
	s	95 ± 5	$-1/3$	$1/2$
3	t	$173.2 \pm 0.6 \pm 0.8 \times 10^3$ [2]	$+2/3$	$1/2$
	b	$4.18 \pm 0.03 \times 10^3$	$-1/3$	$1/2$

[1] all values, except where stated, are taken from [66, pp. 21–22].
[2] 2011 CDF-D0 combined measurement [168] [120] (*the first error term is the statistical error, the second one is the systematic error*).

Table 2.1: Standard Model quarks

Standard Model Leptons					
Generation	*Particle*	*Symbol*	*Mass* [MeV] [1]	*Charge* [e]	*Spin*
1	electron	e	0.511	-1	$1/2$
	neutrino$_e$	ν_e	$< 2 \times 10^{-6}$	0	$1/2$
2	muon	m	105.7	-1	$1/2$
	neutrino$_m$	ν_e	< 0.19	0	$1/2$
3	tau	τ	1776.82 ± 0.16	-1	$1/2$
	neutrino$_\tau$	ν_τ	< 18.2	0	$1/2$

[1] all lepton values are taken from [66, pp. 13–14]; all neutrino values are from [65, pp. 623–624]

Table 2.2: Standard Model leptons

Standard Model Gauge Bosons					
Symbol	*Particle*	*Interaction*	*Mass* [GeV]	*Charge* [e]	*Spin*
γ	photon	Electromagnetic	0 [1]	0 [5]	1
Z	Z boson	Weak	91.1876 ± 0.0021 [2]	0	1
$W^\pm$	W boson	Weak	80.399 ± 0.023 [3]	± 1	1
g	gluon	Strong	0 [4]	0	1

[1] today's experimental limit: $< 1 \times 10^{-18}$ eV [66, p. 8].
[2] [66, p. 9].
[3] 2012 CDF-D0 combined measurement. See [119].
[4] theoretical value; but, as stated in [66, p. 12], we may not preclude a gluon mass as large as few MeV.
[5] today's experimental limit: $< 1 \times 10^{-35} e$ [66, p. 8].

Table 2.3: Standard Model gauge bosons. These particles are the "force carriers" of the three fundamental forces described by the Standard Model. In the table their symbols, masses and electric charges are listed[67]

2.1.1 The Standard Model

In the SM there are several quantities that are conserved. Some of them, like charge and energy, are conserved as consequence of symmetries in the model (more details in the following). While other quantities have been observed to be conserved, but there's no formal explanation or exact conservation law in the model: the conservation of those quantities is purely empirical, until now. Example of such quantities are the *leptonic number*, L, and the *baryonic number*, B. The number $L_{e,\mu,\tau}$ is defined as a quantity whose value is $+1$ for leptons and -1 for anti-leptons. The total leptonic number must be conserved in all processes involving leptons. The conservation of baryon number, B, is built into the SM to account for the fact that the decay of the proton has never been observed. The number has been defined as $B = 1$ for all baryons, like the proton p and the neutron n, and $B = 1$ for all anti-baryons. Mesons, like π, and leptons have $B = 0$.

Fermionic particles (quarks and leptons) and bosons, which compose the SM, are mathematically described by fields, and their interactions and dynamics are described by a relativistic Quantum Field Theory (**QFT**). The strength of the interactions are expressed by coupling constants. If the fields and the potentials of the field theory are described by symmetry groups — called *gauge groups* — the theory is called a *gauge theory*.

In the SM all the interactions between the elementary particles are formulated with gauge theories and they can be described by Lagrangian functions, which describe the dynamics of the system. The group of transformations of the field variables of the gauge theory, called *gauge transformations*, does not act on the basic physics of the quantum field, leaving it unaltered. Thus the Lagrangian function L is said to be invariant under the gauge transformations; one refers to this property with the expression *gauge invariance*. The gauge invariance condition gives the theory a defined symmetry. Noether's theorem [63] implies a conservation law for every symmetry causing gauge invariance, leading to a new gauge boson being introduced. In this picture the above mentioned photon becomes the gauge boson mediating the electromagnetic interaction. Quantum electrodynamics (**QED**), in fact, is a QFT describing the electromagnetic interaction. The theory features a symmetry group $U(1)_{em}$ and it describes, by means of the electromagnetic force, the interaction among particles carrying an electric charge. Being the physical observable quantities invariant with respect to the chosen electromagnetic potential, the theory leads to the electrical charge conservation and it introduces a new massless $S = 1$ gauge boson: the photon γ. The same reasoning can be done with the electroweak theory, leading to the other gauge bosons of the SM, which have been mentioned above.

The SM, in fact, is a non-Abelian Yang-Mills relativistic QFT based on the gauge invariance of the topological group shown in Equation 2.1.

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

The first topological group in Eq. 2.1, $SU(3)_C$, is the symmetry group of the gauge theory called *Quantum Chromodynamics* (**QCD**), which describes the interaction of the strong force. The group $SU(3)_C$ is based on the field represented by the quantum number named *color* [41]. Each quark has a *color* value $q_{i=1,2,3}$. The C flag in the topological group actually represents the *color* charge, and the number 3 represents its three possible values. In the framework of this theory, the gluon boson, which is a massless particle with spin $S = 1$, mediates the interaction of the strong force between the quarks. The gluon also carries a *color charge* (it is *colored*) and this causes its self-interaction. The self-interaction of the gauge boson makes the QCD a *non-Abelian* gauge theory. Gluons can have 8 different color states, which are mixed states of the fundamental color charges; they are presented in Table 2.4.

The color is a characteristic that is not experimentally observed. Thus the hadrons, which are the composite particles made of quarks, must be colorless states. Hadrons can be divided in two categories, according to the number of quarks they are composed of: *baryons*, which are made of three quarks of different colors ($q_i q_j q_z$) and which have a baryon number $B = 1$, and *mesons*, which are composed of a pair of quark/anti-quark of the same color ($q_i \bar{q}_i$) and with $B = 0$.

Quarks have been experimentally found only in 1968 by Friedman, Kendall and Taylor at SLAC, in Stanford, US. The theoretical interpretation of the results of the experiments, made by Feynman and Bjorken, ended up with two conclusions. The first one was that the electric charge inside the hadrons is concentrated in point-like particles; and this was the proof of the existence of the quarks. The second evidence was that those particles do not feel any strong reciprocal force inside the hadron, where they move like free particles. That last discovery collided with the evidence that no quark had been observed in isolation outside the nucleon. These were the first evidences of the *asymptotic freedom* of the strong force: quarks are confined in the nucleon, where they behave freely; but when they are pushed apart, the strong force tends to pull them back inside the nucleon. The strong nuclear force between colored quarks increases with distance, creating new quark-antiquark pairs from vacuum oscillations if the energy goes beyond a certain threshold. This behaviour, which is called *color confinement*, explains why it was impossible for the experiments to extract single colored quarks out of the nucleon and to observe them individually, and why the quarks and gluons inside a hadron behave like free particles and interact with a high-energy probe only at very short distances.

The second part of SM, the group $SU(2)_L \otimes U(1)_Y$ in Equation 2.1, represents the underlying topology of the Electroweak Theory (**EW**), which unifies the electromagnetic and the weak interactions. The EW theory has been first proposed

Gluon colour states	
$(r\bar{b}+b\bar{r})\sqrt{2}$	$-i(r\bar{b}+b\bar{r})\sqrt{2}$
$(r\bar{g}+g\bar{r})\sqrt{2}$	$-i(r\bar{g}+g\bar{r})\sqrt{2}$
$(b\bar{g}+g\bar{b})\sqrt{2}$	$-i(b\bar{g}+g\bar{b})\sqrt{2}$
$(r\bar{r}+b\bar{b})\sqrt{2}$	$(r\bar{r}+g\bar{g}-2b\bar{b})\sqrt{6}$

Table 2.4: The eight gluon colour states

by Glashow, Weinberg and Salam, and it includes the symmetry group $U(1)_{em}$ of the QED. The EW theory is gauge invariant with respect to two symmetries: the weak hypercharge Y symmetry and the weak isospin T symmetry. The relation between those two conserved quantities and the electric charge Q is described by the equation 2.2.

$$Q = T_3 + 1/2Y \tag{2.2}$$

The two symmetry groups act on different parts of the fields: while $SU(2)_L$ — where L stands for *left* — is the simmetry of only the left-handed component of the fermionic fields, the group $U(1)_Y$ represents the symmetry that acts both on the left- and the right-handed parts. The resulting unified symmetry features 4 massless gauge vector fields: one associated with $U(1)_Y$ (B_μ) and three linked to $SU(2)_L$ ($W^i_\mu, i = 1, 2, 3$). Those fields mix among themselves, giving 4 physical fields: W^1_μ and W^2_μ mix together to build the two fields W^+_μ and W^-_μ, while B_μ and A_μ build the fields A_μ and Z_μ, as shown in Equation 2.3, where Θ is the Weinberg's mixing angle. The fields $W^\pm_\mu$ gives the real vector bosons $W^\pm$, while the Z_μ defines the vector boson Z^0. The field A_μ represents, instead, the boson γ, which is the photon field of the SM.

$$\begin{aligned} W^\pm_\mu &= \frac{W^1_\mu \mp iW^2_\mu}{\sqrt{2}} \\ A_\mu &= B_\mu cos\theta_W + W^3_\mu sin\theta_W \\ Z^0_\mu &= -B_\mu sin\theta_W + W^3_\mu cos\theta_W \end{aligned} \tag{2.3}$$

In the theory the gauge bosons must be massless, because mass terms violate the chiral symmetry $SU(2)_L$. And so should be the corresponding vector bosons. But actually only the photon has been observed massless (and not self-interacting). The vector bosons $W^\pm$ and Z^0 have been observed as massive (and self-interacting) particles. To allow the vector bosons to get mass, the electro-weak symmetry, $SU(2)_L \otimes U(1)_Y$, must be broken. In the SM this happens by a spontaneous symmetry breaking, introduced with the Higgs mechanism, described in

Section 2.1.2. There is a *spontaneous symmetry breaking* (SSB) when a symmetric system — a system whose all vacuum states have the same probability — falls into a certain given state, which appears with a higher probability, namely 1. This means that one element of the symmetry group is different from the others, which makes the symmetry break. So the system is not symmetric anymore and it goes into a given vacuum state. A QFT system which spontaneously breaks a symmetry, features a Lagrangian equation that is invariant under the symmetry transformations; but the vacuum of the theory does not follow this invariance.

2.1.2 The Higgs mechanism

Proposed by Peter Higgs in 1964 to spontaneously break the gauge invariance of the SM [53, 52], the *Higgs mechanism* is the mechanism by which a vector boson gets mass without destroying the the gauge invariance of the Lagrangian density. Higgs suggested that the symmetry could break when adding at least one $SU(2)_L$ complex scalar doublet ϕ in Equation 2.4, named the Higgs field, and the corresponding Higgs potential $V(\phi)$, shown in Equation 2.5.

$$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^+ - i\phi_2^+ \\ \phi_1^0 - i\phi_2^0 \end{pmatrix} \tag{2.4}$$

$$V(\phi^\dagger\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \tag{2.5}$$

In the equation of the Higgs potential 2.5, μ^2 is the mass parameter. Depending on the sign of this parameter, two possible values of the vacuum expectation $< 0|\phi|0 >$ minimize the potential $V(\phi^\dagger\phi)$, being the *vacuum state* the state of lowest energy. The Higgs potential energy density $V(\phi)$, according to the sign of the mass parameter, is shown in Figure 2.1. The figure on the left shows the potential when $\mu^2 > 0$ (and $\lambda > 0$): the vacuum state is unique in this case and its expectation value is $< 0|\phi|0 >= 0$, such that the vacuum is symmetric with respect to $SU(2)_L \otimes U(1)_Y$ and, therefore, the symmetry is not broken. The figure on the right, instead, shows the potential $V(\phi)$ when $\mu^2 < 0$ (and $\lambda > 0$): in this configuration the vacuum state is not unique —because all values in the circular minimum minimize the potential —, the expectation value is different from 0, and the value of the scalar field ϕ is shown in Equation 2.6. The parameter ν corresponds to a choice for the vacuum ground state in Figure 2.1. The choice of the vacuum state corresponding to the value in 2.6 breaks the $SU(2)_L \otimes U(1)_Y$ symmetry.

$$\begin{gathered} \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} \\ |\phi^2| \equiv \frac{\nu^2}{2} = -\frac{-\mu^2}{2\lambda} \end{gathered} \tag{2.6}$$

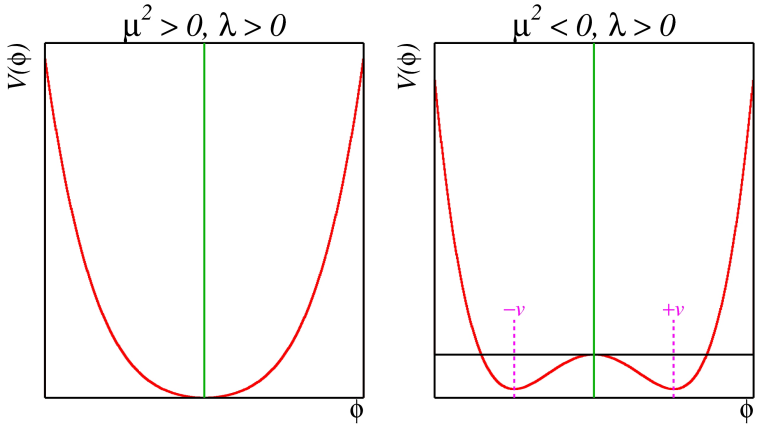


Figure 2.1: The Higgs potential energy density $V(\phi)$, depending on the sign of the mass term μ^2: on the left, with $\mu^2 > 0$, in the symmetric phase; on the right, with $\mu^2 < 0$, in the spontaneously broken phase [164].

The value of the ϕ scalar field in Equation 2.6 has to be expanded by small oscillation around the ground state, to let the calculations in perturbation theory converge. Thus an additional field $h(x)$ has to be added, which parametrizes the small deviations from the stable equilibrium configuration $\phi(x) = \phi_0$ [5] (see Equation 2.7 and Figure 2.2).

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix} \tag{2.7}$$

The complex scalar field has 4 degrees of freedom and the resulting particle spectrum features a massless gauge boson (the γ photon) and two massive intermediate bosons (corresponding to $W^{\pm}$ and Z^0). The remaining degree of freedom of the scalar field is a massive spin-0 boson known as the *Higgs boson*, whose value of its mass term is $m_H = \sqrt{-2\mu^2}$.

In conclusion, we can observe that the Spontaneous Symmetry Breaking (SSB) mechanism lets us obtain the right physical bosons we observe in the experiments, in particular [132]:

- Before SSB
  - 4 massless gauge bosons: $W^{\mu}_{1,2,3}$, B_{μ}
  - 4 massless scalars: The 4 real components of the field Φ: ϕ_1, ϕ_2, ϕ_3, ϕ_4
- After SSB
  - 3 massive gauge bosons: W^+, W^- and Z

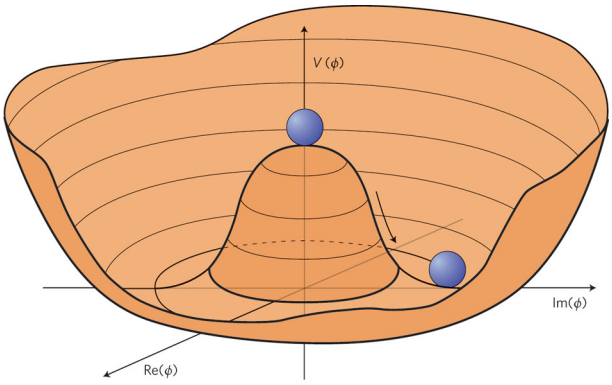


Figure 2.2: The Higgs potential energy density $V(\phi)$, when $\mu^2 < 0$. The *vacuum state* is the lowest-energy state. A potential like that one shown here features many states of minimum energy: all the points randomly chosen along the bottom of the potential curve. In this framework the Higgs boson is a spin-0 massive particle corresponding to quantum fluctuations in the radial direction, in the direction of the arrow [12].

- 1 massless gauge boson: γ
- 1 massive scalar: H

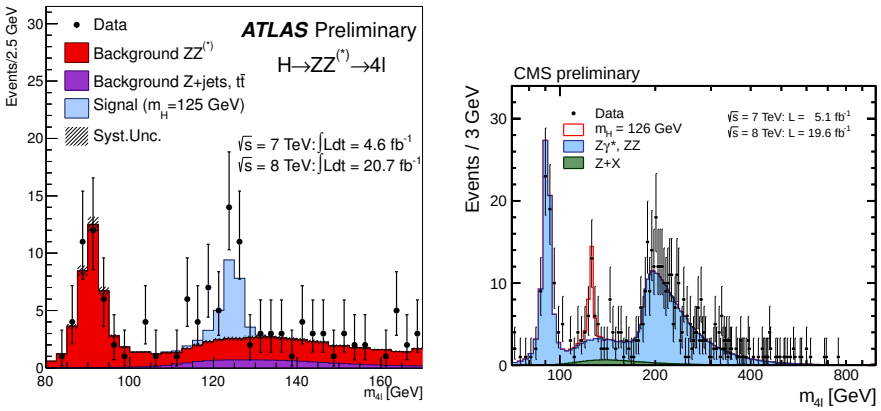
The additional Higgs field also generates the fermion mass terms, by the so called Yukawa interaction terms between the fermions and the Higgs scalar fields, also named *Yukawa couplings*. Hereby the strength of the coupling of the Higgs boson with a fermion is proportional to the mass of the fermion itself.

The Higgs boson has been first observed in 2012 by the ATLAS and CMS experiments and announced at a joint seminar on July the 4th 2012 [193, 160, 180, 15, 35], and its discovery confirmed in March 2013 by both experiments [99, 123] when they reported updated results that indicates spin 0 and positive parity for the new particle, in agreement with the predictions for the Standard Model Higgs boson. Figures 2.3 and 2.4 show the plots of the Higgs mass measurement, and in figures 2.5–2.6 on pages 18 and 19 two event displays of Higgs events are shown: a $Higgs \rightarrow \gamma\gamma$ in Figure 2.5 and a $Higgs \rightarrow 4mu$ in Figure 2.6.

2.1.3 The Standard Model Lagrangian

The SM is the gauge theory for electroweak interactions and has provided plenty of successful predictions with an impressive level of precision. As already said, it is based on the gauge symmetry of electroweak interactions, and it can be described by a Lagrangian function, which includes and summarizes the contributions of all forces and fields described until now. The Lagrangian is written in Equation 2.8 [132, 143]

$$L_{SM} = L_{EW} + L_{Strong} + L_{Yukawa} + L_{Higgs} \tag{2.8}$$



(a) **(b)**

Figure 2.3: The distribution of the invariant mass of the four leptons, measured by the ATLAS and CMS experiments, for the combined $\sqrt{s} = 8$ TeV and $\sqrt{s} = 7$ TeV data sets. The analyses — which combine the $4e$, 4μ and $2\mu 2e$ channels — have been presented at the Moriond Conference in March 2013. **(a)**: the 4-leptons invariant mass measured by ATLAS, compared to the background expectation in the mass range $80 - 170$ GeV. The signal expectation for the $m_H = 125$ GeV hypothesis is also shown. In the plot here, at low m_H values, the resolution of the reconstructed Higgs boson mass is dominated by the detector resolution [99]. **(b)**: the four-lepton invariant mass measured by CMS, in the full mass range [123].

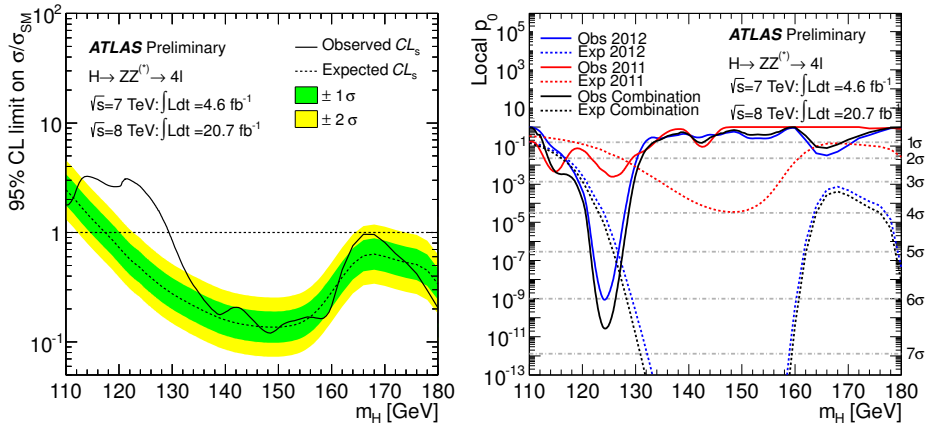
The first term in the Equation 2.8, L_{EW}, is the Lagrangian of the Electroweak interaction; as already described in the previous section, it is based on the symmetry group $SU(2)_L \otimes U(1)_Y$. The L_{EW} term can be expanded into the Equation 2.9 [38]:

$$L_{EW} = \overline{L} i\gamma^\mu D_\mu L + \overline{R} i\gamma^\mu D_\mu R - \frac{1}{4} W_i^{\mu\nu} W^i_{\mu\nu} - \frac{1}{4} B_i^{\mu\nu} B^i_{\mu\nu} \tag{2.9}$$

where:

- W^i_μ is the gauge vector field associated with $SU(2)_L$
- B^i_μ is the gauge vector field associated with $U(1)_Y$
- L indicates the left-handed weak isospin fermion doublets
- R denotes the right-handed fermion singlets
- D_μ is the covariant derivative defined as: $D_\mu = \partial_\mu + igT_i W^i_\mu + ig' \frac{Y}{2} B_\mu$, with g and g' the $SU(2)_L$ and $U(1)_Y$ coupling constants, respectively[132, p. 39]

The second term in the Equation 2.8, L_{Strong}, is the Lagrangian of the strong interaction, i.e. the QCD Lagrangian. It can be written as in Equation 2.10 [132, p. 35]:



(a) **(b)**

Figure 2.4: Higgs limits measured by ATLAS, for the combined $\sqrt{s} = 8$ TeV and $\sqrt{s} = 7$ TeV data sets (the data sets for the years 2011 and 2012). **(a)**: the plot shows the 95% CL upper limit on the Standard Model Higgs boson production cross section, as a function of mH in the low mass region and divided by the expected SM Higgs boson cross section. The black solid line represents the observed 95% CL upper limit, compared with the expected one, represented by the black dashed line. The green and yellow bands indicate the expected limits with 1σ and 2σ fluctuations, respectively. **(b)**: The combined observed local p-value is shown as the solid black line; while the horizontal dashed lines show the p-value, corresponding to local significances from 1 to 7σ. [99] .

$$L_{Strong} = L_{quark} + L_{gluon} + L_{interactions} \tag{2.10}$$

which can be expanded to Equation 2.11 [132, p. 14]:

$$L_{Strong} = \sum_i \overline{q_i}(x)(iD_\mu\gamma^\mu - m_i)q_i(x) - \frac{1}{4}F^\alpha_{\mu\nu}(x)F^{\mu\nu}_\alpha(x) \tag{2.11}$$

where:

- $i = 1, 2, 3$ is the quark index
- $\alpha = 1, 2, \ldots, 8$ is the gluon index
- the first term $\sum_i \overline{q_i}(x)(iD_\mu\gamma^\mu - m_i)q_i(x)$ contains the term $D_\mu = \partial_\mu - ig_s\frac{\lambda_\alpha}{2}A^i_\mu$, where:
  - λ_α are the $SU(3)$ generators;
  - A^i_μ are the gluon fields;
  - g_s is a dimension-less constant relative to the strong interaction, called the "strong coupling constant".

The whole term describes the gauge interactions among the quarks and gluons

- the term $F^{\alpha}_{\mu\nu}(x)$, which contains the gluon field A^{i}_{μ}, it is the gluon field strenght
- the whole term $\frac{1}{4}F^{\alpha}_{\mu\nu}(x)F^{\mu\nu}_{\alpha}(x)$ describes the 8 gluon fields, and their self-interaction.

The third and fourth terms in the Equation 2.8, L_{Higgs} and L_{Yukawa}, are the Lagrangian terms describing the Higgs mechanism and the Yukawa couplings of the fermions to the Higgs field. The terms are needed to provide masses to the SM: the M_W and the M_Z masses the first term; and the fermion masses the second one. These terms can be written as Equation 2.12 and 2.13 [132, p. 39]:

$$L_{Higgs} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\phi) = |D_\mu\Phi| - \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2 \tag{2.12}$$

$$\begin{aligned} L_{Yukawa} &= \lambda_e\bar{l}_L\Phi e_R + \lambda_u\overline{q}_L\tilde{\Phi}u_R \\ &\quad + \lambda_d\overline{q}_L\tilde{\Phi}d_R + \ldots 2^{nd} \textit{ and } 3^{rd} \textit{ generation} && \text{(2.13a)} \\ &= -g_f|\overline{L}\Phi R + \overline{R}\Phi_c L| && \text{(2.13b)} \end{aligned}$$

In Equation 2.12, Φ is the basic complex doublet, with hypercharge $Y = 1$; and $V(\Phi)$ is the simplest renormalizable potential. Both of them have already been introduced in Section 2.1.2.

The term D_μ in the Equation 2.12 is the same introduced in Equation 2.9, and it contains the W_μ and B_μ gauge fields of the $SU(2)_L$ and $U(1)_Y$, respectively.

The Equation 2.13 describes the Yukawa coupling of the fermions; and in Equation 2.13a l, q and Φ are such that:

$$\begin{aligned} l_L &= \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \\ q_L &= \begin{pmatrix} u_L \\ d_L \end{pmatrix} \\ \Phi &= \begin{pmatrix} \Phi^+ \\ \Phi_0 \end{pmatrix} \\ \tilde{\Phi} = i\sigma_2\Phi &= \begin{pmatrix} \Phi^\star_0 \\ -\Phi^- \end{pmatrix} \end{aligned} \tag{2.14}$$

where l_L and q_L are the left-handed lepton and quark respectively. The Yukawa coupling g_f to the Higgs field (in Equation 2.13b) generates the mass of a fermion

f. The term $g_f\overline{\psi_f}\psi_f\Phi$ (with Φ defined in Equation 2.7) gives a fermion mass, as in the Equation 2.15.

$$m_f = \frac{g_f\nu}{\sqrt{2}} \tag{2.15}$$

It is worth to be noticed that the physical Higgs field h couples to fermions through the term $g_f\overline{\psi_f}\psi_f h$, which is proportional to the mass of the fermion, m_f.

2.2 Limitations of the Standard Model

The Standard Model has been shown to be the best representation of fundamental physics up until now. During the last decades, many high-energy accelerators have contributed to test the Standard Model and the properties of the particles: the LEP at CERN, the Fermilab Tevatron (USA), ADA at Frascati (Italy), the SLC at SLAC (USA), HERA at DESY (Germany). In all experiments up to 100 GeV the theoretical calculations and predictions computed from the SM have been confirmed at a very high precision [11]. For example the top mass has been predicted, through SM calculations, several years before its discovery.

But the SM is also recognized as being an incomplete theory: for instance it does not explain the fermionic generations and it fails to include gravity. And, in general, it looks clear nowadays that the Standard Model has to be extended in some way, to make it able to describe and predict physics at high energy scale. In the following, some of the most problematic arguments will be listed and briefly discussed.

Origin of masses and number of generations

In the Standard Model, we find several quantities whose values cannot be predicted, because not provided by the theory. In its simplest version, the Standard Model features 19 parameters, which are unknown *a priori*: 9 fermion masses (charged leptons and quarks), 2 gauge boson masses (as the Higgs and Z bosons masses), 4 parameters related to the quark-mixing matrix (also known as CKM, from the name of its authors Cabibbo, Kobaywashi and Mascawa), 3 coupling constants, and 1 strong CP parameter. Moreover there is no explanation on the reason for the existence of the three generations of particles with two different flavors, as described in Section 2.1 (see also tables 2.2 and 2.1).

Gravity

So far, it has not been possible to include a quantum field theory of gravity, consistent with General Relativity, in the Standard Model. The problem is the inconsistency between the two theories: the Standard Model, which is a quantum field theory, and the General Relativity, which is a classical theory. And any

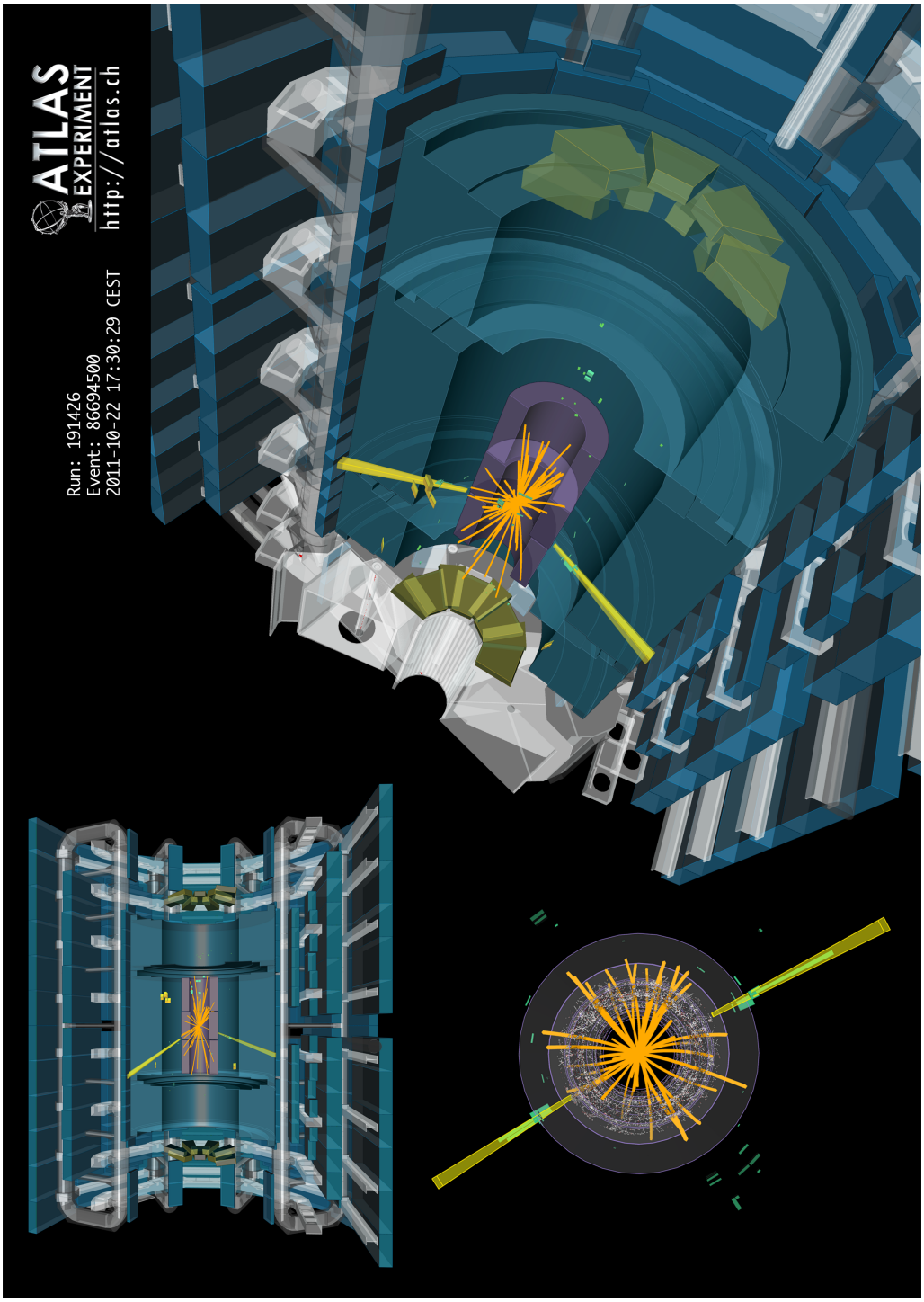


Figure 2.5: A $H \to \gamma\gamma$ event collected by the ATLAS experiment and visualized in the VP1 event display [81, 162]: the image shows a Higgs boson decaying into two photons [157].

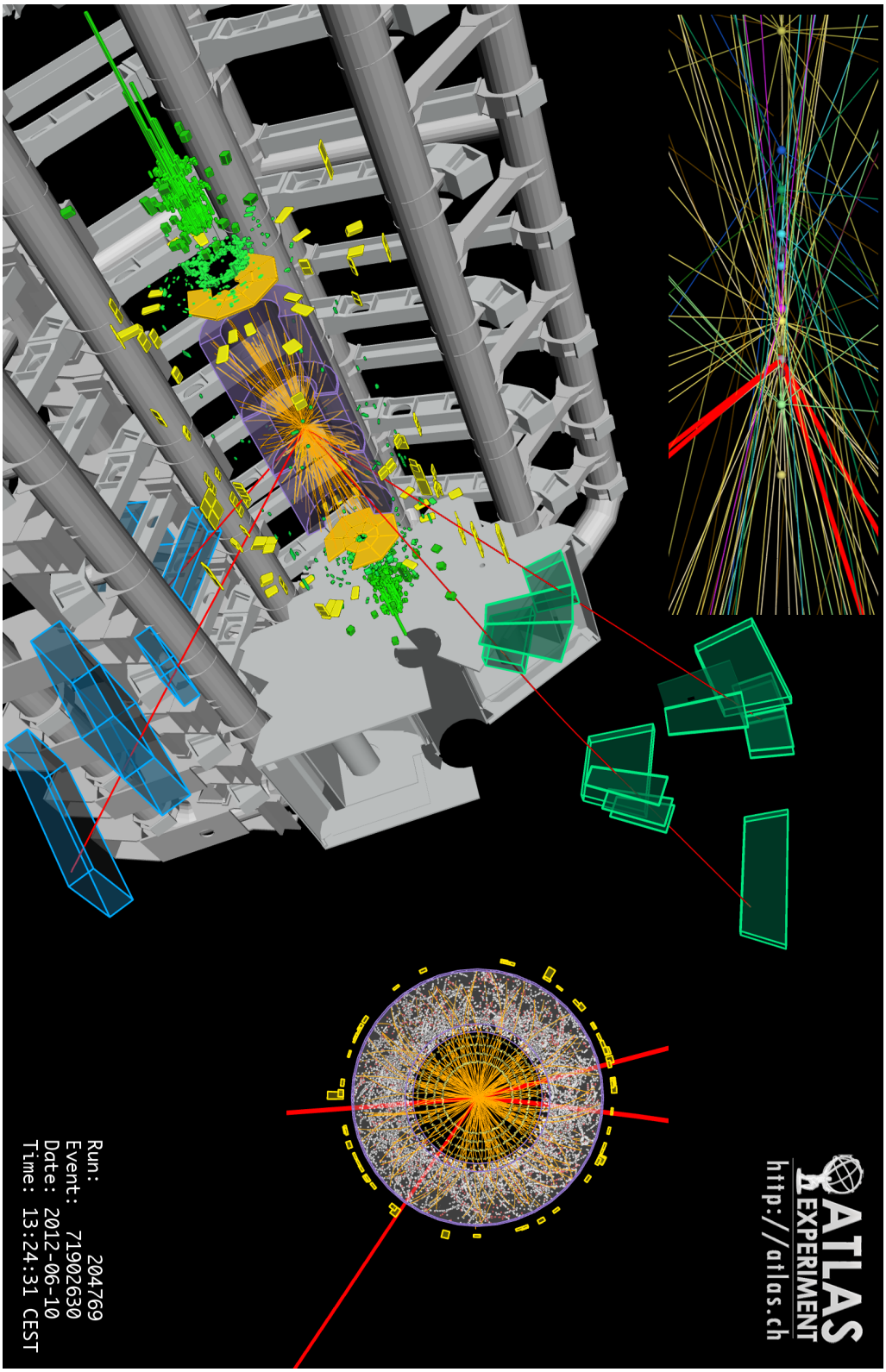


Figure 2.6: A $H \to \gamma\gamma$ event collected by the ATLAS experiment and visualized in the VP1 event display [81, 162]: the image shows a Higgs boson decaying into four muons [156].

attempt to quantize General Relativity failed. If an expansion in terms of Feynman diagrams is performed — the same technique used for the construction of the Standard Model — then infinities are found, which cannot be absorbed in a renormalization of the cosmological and Newton constants. It is a non-renormalizable theory and therefore it cannot be mixed with the Standard Model as it is known today [83]. And this is one of the strongest points, which makes appear the SM an incomplete theory: the fact that the SM cannot take into account gravity effects, is not a problem at the electro-weak energy scale (i.e. around 100 GeV), where the action of the gravity is negligeable; but it can cause problems if calculations and experiments at higher energies are involved.

Hierarchy problem

There are many orders of magnitudes between the strength of the Standard Model forces and the strength of gravity, with the electromagnetic force having a strength $\sim 10^{40}$ times greater than gravity [4]. By examining when the gravitational constant G_N approaches unity, we can have a very rough idea of the energy scale at which Standard Model forces — electroweak and strong forces — can be unified with gravity. Since the effects of the gravity are proportional to the mass or the energy of the particle, they grow at high energies. The energy scale where the strenght of gravity becomes comparable to the other Standard Model interactions is about the Planck's mass $M_P = \sqrt{\hbar c/G_N} \sim 10^{19}$ GeV[83][6].

The *Hierarchy problem* is about such a discrepancy: *why in Physics there is this hierarchy of energy scales?* According to the Standard Model, the mass of the Higgs boson has to be of the order of the electroweak scale $M_{EM} \sim 100$ GeV in order to be able to produce $W^{\pm}$ and Z^0 masses; but actually gravity becomes important only 16 orders of magnitude beyond this scale. Moreover the Standard Model, as known today, is heavily *fine-tuned*. The Higgs mass term is $M_H = -\mu^2 \sim -M_{EM}^2$, and therefore it is sensitive to the physics well beyond the Standard Model electro-weak scale.

These different scales result in very large quantum corrections to the bare masses of the Higgs boson and of the other particles. In order to obtain masses at the weak scale of around 1 TeV, the bare masses need to precisely cancel the renormalization corrections over many orders of magnitude.

Quantum loop fluctuations in renormalization lead to corrections of the mass of the Higgs boson with the a renormalization cutoff term, $f^2\Lambda^2$, determined by the scale Λ of the physics that can lies beyond the electroweak scale M_{EM} (the factor f merely represents other terms in the calculation):

$$-M_H^2 \sim -\mu_0^2 + f^2\Lambda^2 \sim -M_{EM}^2 \tag{2.16}$$

where μ_0 is the bare mass of the Higgs before renormalization, which has to be of the order of magnitude of the Planck's mass M_P. If there were no precise

cancelation between the quadratic correction term and the bare mass μ_0, the contribution of the quantum term to the square of the mass of the Higgs boson would be very large, making the value of the mass much larger than the mass observed by the first experimental evidences [15][35], which was the mass expected at the electroweak scale.

This fine-tuning of the bare masses is generally considered unnatural without a well-motivated physical mechanism to provide for it, opening possibilities in favor of new particles and fields existing between the electroweak and Planck scales; particles that could provide new opposite-signed corrections.

Grand unification

Following the electroweak unification, it is quite widely believed that the SM is a mere low-energy approximation (or effective theory) of a more fundamental theory. This belief is supported by the behaviour of the running couplings of the standard forces, which appear to almost converge in one point at very high energies, as shown by the dashed lines in Figure 2.7. This fundamental theory has been called GUT, or *Grand Unification Theory*, and combines, at a scale around 10^{16} GeV, the electroweak and the strong forces into a single interaction (featuring a larger gauge symmetry and one coupling constant), in a way similar to the unification of the electromagnetic and weak forces in the electroweak force at the scale of 100 GeV. The GUT has been first proposed in 1974 by Glashow and Georgi [42].

Matter-antimatter asymmetry

After the Big Bang, matter and anti-matter are assumed to have been produced in equal amounts. But wherever one looks at in the Universe, only matter can be found in large quantity. And in the Standard Model there is no mechanism that can justify this today's asymmetry that we observe in favor of the matter. And, unfortunately, the magnitude of CP violation [34] in the Standard Model is not sufficient to provide an explanation [71][49].

Dark matter

One of the biggest problem of the Standard Model is that it can describe and explain only a very small part of all the matter that composes the Universe, about 4%, while cosmological observations, like the measurements of rotational speed of spiral galaxies [24] or of the anisotropy of the cosmic microwave background [56], have established the existence of cold dark matter in the Universe, accounting for $\sim 25\%$ of the total mass. Many possible candidates for dark matter have been discussed in the past, but most of them were discarded, and some do only account for a small fraction of dark matter. One of the most recent and promising dark

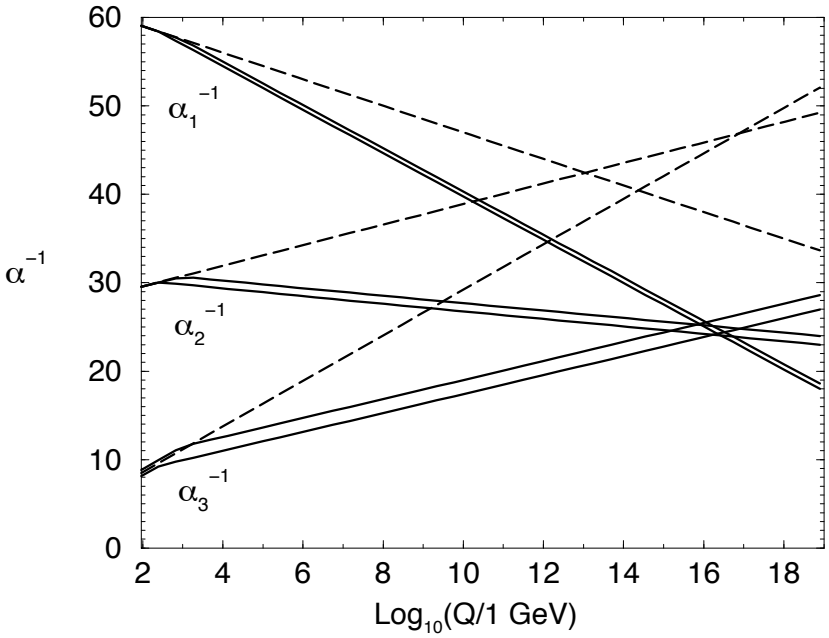


Figure 2.7: The evolution of the inverse gauge couplings α_i of the Standard Model (dashed line) and of the MSSM supersymmetric model. The indices corresponds to the electromagnetic (1), the weak (2) and the strong (3) coupling. In this figure one can see that the SM couplings *almost* converge at a scale around 10^{16}, while the MSSM supersymmetric extension enables a better unification [141].

matter candidates is the WIMP , or *Weak Interacting Massive Particle*, a particle that would be a relic of the Big Bang. In conclusion the problem is that, in the current form, the Standard Model does not contain a particle which can serve as a candidate for dark matter [25].

Chapter 3

Physics beyond the Standard Model at hadron colliders

In order to solve the problems which have been found in the Standard Model, many new models and extensions to it have been presented in the last decades: among them Supersymmetry, large extra dimensions, gravitons, long-lived particles, dark matter, microscopic black holes and leptoquarks. In general those models are referred to as Physics *beyond the Standard Model*, or simply BSM. Many of those models predict new physics at the TeV scale, which represents a strong motivation for searches at LHC, where the mass scale can reach ~ 3 TeV.

A selection of the most promising ones will be presented here in the following, together with a description of the experimental signatures and the characteristic final states of such processes, which later will be useful to understand the underlying rationale of the model-independent General Search presented in this work.

3.1 Supersymmetry (SUSY)

One of the principal motivations for the construction of the Large Hadron Collider — besides the hunt to the Higgs boson — is the search for low energy *Supersymmetry*, or SUSY.

Supersymmetry was first proposed by Miyazawa in 1966 [62] as a symmetry connecting mesons and baryons, and then extended by other groups to fundamental particles in the first half of 1970s. As an extension of the Standard Model, SUSY connects fermions with bosons, so that for each fermion (boson) there is a boson (fermion) with the same quantum number except the spin. The symmetry is presented as a transformation that turns a fermionic state into a bosonic

state by a spinor operator $\mathcal{Q}$, which rotates the spin of 1/2, leaving unalterated the other quantum numbers of the particle. The action of the transformation is expressed in the equation 3.1.

$$\begin{aligned} \mathcal{Q}|fermion> &= |boson> \\ \mathcal{Q}|boson> &= |fermion> \end{aligned} \tag{3.1}$$

The corresponding particles are called *superpartners* of each other, and more generally they can be referred to as *sparticles*. In more detail, let us notice that sparticles are usually labeled with an upper tilde (like $\tilde{q}$) and that the superpartners of fermions have a *s-* prefix in front of their names (like *squarks*), while the superpartners of bosons feature an *-ino* suffix in their names (like *gluino*).

The SM does not contain particles which can act as superpartners. In fact the symmetry within the theory implies the superpartners having the same mass of their counter-part. But no particles with the required quantum numbers have been observed so far in the mass range of the SM. Thus the SM has to be extended, and the symmetry must be broken in the ground state of the supersymmetric theory.

Many supersymmetric models have been proposed in the last years. But to date none of them has been confirmed by the experiments. In the following, only the MSSM model will be briefly introduced, as an example of a supersymmetric model. More complete information on the supersymmetric extension of the Standard Model can be found in many classical books and articles; among them [48][54][141][86][129][80], which also constitute the basis of the decription that follows.

MSSM – The Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model, or MSSM, is the minimal supersymmetric extension of the Standard Model, and it features a doubling of the SM particle content.

The fermions have bosonic superpartners, called scalar fermions or *sfermions*, with spin 0. The number of degrees of freedom here differs by a factor of two, thus each fermion is associated with two sfermions, labeled according to the handedness of the fermionic superpartner: e.g. $\tilde{e}_R$ represents the *selectron* whose superpartner is the right-handed electron. As for regular Standard Model fermions, the weak nuclear interaction only couples to left-handed sfermions.

The SM spin 1 gauge bosons have fermionic superpartners with spin 1/2, called *gauginos*. There are superpartners of the gluons, the *gluinos*, and of the electroweak gauge eigenstates ($W^{\pm}$, W^0 and B^0), called *winos* and *bino*: $\widetilde{W}^{\pm}$, $\widetilde{W}^0$ and $\widetilde{B}^0$.

In MSSM the Higgs boson has to be extended up to eight real degrees of freedom. That extension results in five Higgs bosons and in four sfermion partners called *Higgsinos*, two neutral and two electrically charged.

The mixed state of the neutral wino, the bino and the two neutral higgsinos produces four *neutralinos*, written as $\widetilde{\chi}^0$; while the charged winos and the charged Higgsinos mix together to give two pairs of so-called *charginos*, $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^{\pm}$.

The MSSM Lagrangian allows for the addition of terms violating the lepton- and baryon-number, whose associated processes have not been observed in experiments yet. But a new symmetry can be added as well, called *R-parity*, whose quantum number P_R is conserved and which forbids the violating terms. The R-parity quantum number can be written as:

$$P_R = (-1)^{3(B-L)+2S} \tag{3.2}$$

where s is the spin of the particle and L and B the lepton and baryon number respectively. In this way the SM particles have an R-parity value of $+1$, while their supersymmetric counter-parts have an R-parity number of -1. As a consequence, the *lightest supersymmetric particle*, or LSP, can not decay further — because there is no any lighter sparticle to decay to, while conserving R-parity — and therefore it has to be stable. Hence all the other sparticles in the end must decay in a final state with one or an odd number of LSPs. Another consequence of the R-parity, is that supersymmetric particles have to be produced in pairs.

The nature of the LSP depends on the nature and the formulation of the particular underlying supersymmetric model. But there are some constraints set by physical experimental observations. For example, it can not have electrical charge or color, otherwhise it could form bound states, which have not been observed;

The large mass of the LSP, and the fact that it is a stable particle, makes the LSP a candidate for Dark Matter. This is widely believed to be made of "cold" (i.e. non-relativistic) and massive non-baryonic particles, needed by the clustered structure of our universe [26][141][80]. The neutral electrical charge is another requirement for the LSP, to be consistent with the lack of astronomical observations.

3.1.1 Experimental signatures of MSSM supersymmetric particles at hadron colliders

In the following, the decay modes of the supersymmetric particles, in the framework of the MSSM, will be briefly listed. For the decay processes below, the R-parity is assumed to be conserved.

R-parity conservation implies, as already said, a pair production of SUSY particles, dominantly gluinos and squarks, through the strong interaction. The production process at leading order is depicted in Figure 3.1, which shows some

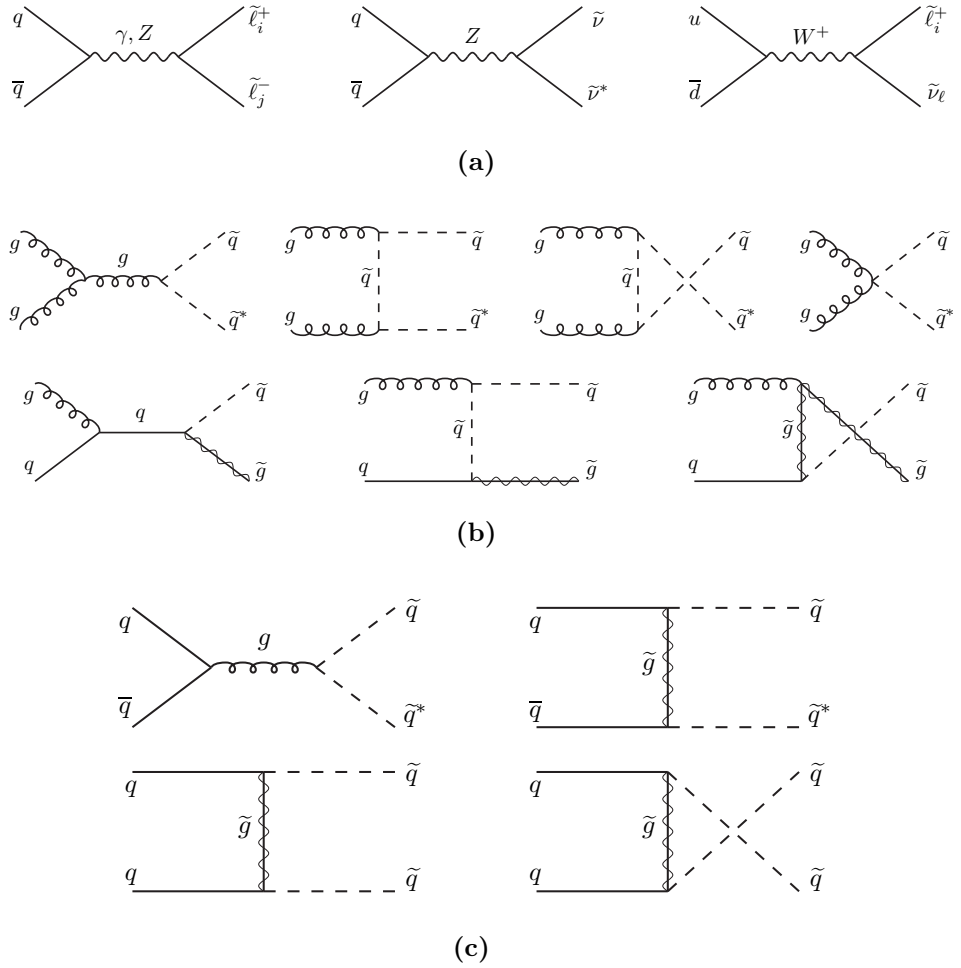


Figure 3.1: The production of supersymmetric particles at hadron colliders: **(a)**: an example of electroweak production of sparticles, at hadron colliders; **(b)**: squark and gluino production through gluon-gluon and gluon-quark fusion; **(c)**: squark production via strong quark-antiquark annihilation (above) and quark-quark scattering (below) [141].

of the production processes of supersymmetric particles at hadron colliders [141]. More in detail, Figure 3.1(a) shows an example of electroweak production of sleptons and $\widetilde{\nu}$. Figure 3.1(b) shows the production of squarks and gluinos through gluon-gluon and gluon-quark fusion. Figure 3.1(c) shows the squark production from strong quark-antiquark annihilation and scattering. Among SUSY models the masses of squarks and gluinos strongly vary, thus the dominant production process depends on the underlying theory. At the LHC, one could observe the production of gluinos and squarks by gluon-gluon and gluon-quark fusion dominating, unless the gluino and squarks are heavier than ≈ 1 TeV. The electroweak production, instead, should play a minor role at LHC [141].

The supersymmetric particles decay in several steps in what is called a *cascade-decay*, resulting in final states with two LSPs; those LSPs escape the detector undetected. The two LSPs should carry away at least the double of the neutralino mass, i.e. $2m_{\widetilde{N}_1}$, of missing energy. However at hadron colliders like LHC, the only observable component of the missing energy is the transverse to the colliding beams, denoted $E_{\mathrm{T}}^{\mathrm{miss}}$. Hence, in general, the observable experimental signature for supersymmetry at hadron colliders is:

$$n \cdot leptons + m \cdot jets + E_{\mathrm{T}}^{\mathrm{miss}} \tag{3.3}$$

In the expression above, n, m or both, can be 0. This results in a very wide range of possible combinations; giving a large number of possible experimental SUSY signatures.

From this large set of signatures, one of the classical SUSY channels is the

$$n \cdot jets + E_{\mathrm{T}}^{\mathrm{miss}} \tag{3.4}$$

without any isolated lepton. This could be very efficient in excluding Standard Model events from W decay while searching sparticles with sizeable branching ratios for channels where they decay with no leptons in the final states. At LHC also the 1-lepton channel could be observed:

$$1\ lepton + n \cdot jets + E_{\mathrm{T}}^{\mathrm{miss}} \tag{3.5}$$

despite the large SM background coming from W decays $W \to l\nu$, this channel could have a large branching ratio and therefore could be one of the SUSY golden channels at LHC.

Another promising SUSY channel could be the three-lepton one:

$$3\ leptons + n \cdot jets + E_{\mathrm{T}}^{\mathrm{miss}} \tag{3.6}$$

shown in Figure 3.2. At LHC the same final states are more likely coming from $\widetilde{g}\widetilde{g}$, $\widetilde{q}\widetilde{g}$ or $\widetilde{q}\widetilde{q}$ production, with one of the two squarks or gluinos decaying to a $\widetilde{N}_2$ and the other to a chargino $\widetilde{C}_1$. In that case one could also observe jets with very high p_{T} in addition to the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the three leptons. And this is the most likely three-leptons signature at LHC.

All the channels listed above, and many others, are currently under study at LHC, both at ATLAS and CMS. A full list of the channels under study at ATLAS, with hyperlinks to the public papers, can be found in [151]. So far they did not reveal any hint of the existence of Supersymmetry, but new data will be collected soon, and new results are expected.

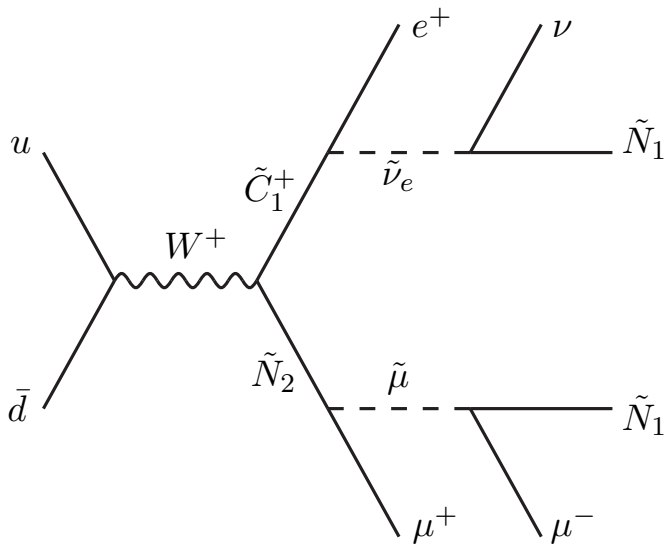


Figure 3.2: An example of SUSY signature: the three-leptons channel. In the final state there could be three leptons (here the two opposite-charge muons from the $\widetilde{N}_2$ decay chain and the electron from the chargino, $\widetilde{C}_1^+$) plus the transverse missing energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, coming from the two $\widetilde{N}_1$, which escape the detector without interacting, and so unrevealed [141].

3.2 Large Extra Dimensions

Models with large extra dimensions are of great theoretical interest, because they can explain the weakness of gravity, i.e. the hierarchy problem: the large difference between the electroweak symmetry-breaking scale, $M_{EW} \sim 1$ TeV, and the fundamental gravity scale $M_P \sim 10^{16}$ TeV. The model predicts new physics signatures appearing at the TeV-mass scale, such that they would accessible at the LHC.

The first proposal to solve the hierarchy problem by using large extra dimensions, was presented by Arkani-Hamed, Dimopoulos, and Dvali; and it is known as the ADD model [13]. In the ADD model, large extra dimensions (*large*, because the model predicts them as large as $\sim 1mm$) are added to the ordinary 3+1 space-time dimensions. In this multi-dimensional world, the Standard Model is constrained to the common 3+1 dimensions, while gravity is free to propagate through the whole N-dimensional space. The N-dimensional Gauss's Law reduces the gravitational flux in the ordinary 3+1 space-time dimensions. As a consequence of this, the fundamental Planck scale can be lowered to the electroweak scale, allowing the production of Gravitons at the TeVscale of the LHC [84]. Large extra dimensions can be revealed, for example, in direct graviton production or in physics processes involving micro black holes. In the following both of those

processes will be briefly described.

3.2.1 Large Extra Dimensions experimental signatures: micro black holes

A consequence of a TeV-scale quantum gravity, is the possible production of black holes in particle collisions at accelerators like the LHC [82]. It is expected that the production of micro black holes rapidly turns on, when the energy passes the threshold of the lowered Planck scale $M_P \sim 1$ TeV.

The production rate of micro black holes at colliders is expected to be quite large. If two partons with the center of mass energy $\sqrt{s} = M_{BH}$ collide head-on, and the collision impact parameter is lower than the corresponding *Schwarzschild radius*, then a black hole with the mass M_{BH} is formed. The total cross section of black hole production at hadron colliders can be then estimated from pure geometrical arguments, and it results of order of $\sim \pi R_S^2$, where R_S is the Schwarzschild radius. At LHC the total production cross section ranges between 15 nb and 1 pb for a lowered Planck scale in the range $[1 - 5]$ TeV [82].

Just after their production, the micro black holes quickly evaporate via Hawking radiation. Considering color and spin, $\sim 75\%$ of particles produced in micro black holes decays is expected to be composed by quarks and gluons, $\sim 10\%$ by charged leptons, another $\sim 5\%$ by neutrinos, and a last $\sim 5\%$ by W or Z bosons or photons; all of them high-p_{T}particles [82].

Thus, typical signals expected from micro black holes are high-multiplicity events, with large total transverse energy, featuring [84]:

$$n \cdot jets + m \cdot leptons + i \cdot photons + E_{\mathrm{T}}^{\mathrm{miss}} \tag{3.7}$$

3.2.2 Large Extra Dimensions experimental signatures: direct Graviton production and WIMPs

Gravitons are supposed to be free to propagate in the extra dimensions, and so they escape the detector undetected. Thus the production of a real graviton produced in association with an high-p_{T}hadronic jet or a photon, can be revealed from the amount of missing energy $E_{\mathrm{T}}^{\mathrm{miss}}$. Examples of processes involving gravitons are:

- $gg \to gG$; $\quad qg \to qG$
- $q\overline{q} \to Gg$; $\quad q\overline{q} \to G\gamma$

which can be translated into signatures like:

$$\begin{gathered} 1\ jet + E_{\mathrm{T}}^{\mathrm{miss}} \\ 1\ photon + E_{\mathrm{T}}^{\mathrm{miss}} \end{gathered} \tag{3.8}$$

which translate into channels with jets or photons and $E_{\mathrm{T}}^{\mathrm{miss}}$, i.e. $j\nu$ and $\gamma\nu$.

Analyses looking at these signatures are also sensitive to pair-production of WIMP (Weakly Interacting Dark Matter) particles χ, where the $\chi - \chi$ pair system recoils from an energetic photon or high-p_{T} hadronic jet from initial state radiation [84].

3.3 Warped Extra Dimensions

The warped extra dimension of the original Randall-Sundrum model (RS1), is another possible solution of the hierarchy problem of the Standard Model [68, 84].

The distinctive feature of this model is the existence of Kaluza-Klein gravitons (with spin 2), whose masses and couplings to the Standard Model are at TeVscale. These gravitons would show in experiments as resonances widely separated in mass, in contrast to the light gravitons predicted in large extra dimension models, whose masses are very close between them. The model features two parameters: the mass of the graviton and the coupling constant k/M_P, where M_P is the lowered Planck mass, and k is the curvature scale of the warped extra dimension.

An extension to the RS1 model also exists, called the *bulk graviton*, which addresses the flavour structure of the Standard Model [7], offering a geometric explanation of both the hierarchy and the flavour puzzles in the Standard Model. In this extension the production of graviton and its decay via light fermions is highly suppressed; and its decay into photons is negligible. But, depending on the model parameters, the production of bulk gravitons from gluon-gluon fusion can be significant. In this case the bulk graviton decays into W and Z longitudinally polarized bosons. Under certain choices of the model parameters, the production of bulk gravitons via vector boson fusion (VBF) can be significant as well.

The experimental signatures of the warped extra dimension depends on the model and on its parameters.

Warped extra dimension signatures: RS1 graviton

The experimental signatures of the RS1 model would be decays of gravitons to pairs of leptons, photons or light jets:

$$2n \cdot leptons + photons + jets \tag{3.9}$$

Both ATLAS and CMS started to look at these kind of signatures, in the $\mu\mu$, ee and $\gamma\gamma$ channels [19, 36, 18].

Warped extra dimension signatures: bulk graviton

As already stated above, in the bulk graviton extension of the RS1 model, the graviton decays into pairs of W and Z vector bosons, which decay into pairs of leptons or jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. In general the processes involve, for example:

$$\begin{gathered} G \to ZZ \to 2n \cdot leptons \\ G \to ZZ \to 2n \cdot leptons + 2m \cdot jets \\ G \to ZZ \to 2n \cdot jets + E_{\mathrm{T}}^{\mathrm{miss}} \\ G \to ZW \to leptons + E_{\mathrm{T}}^{\mathrm{miss}} \\ G \to ZZ \to 2n \cdot jets \\ G \to WW \to 2n \cdot jets \end{gathered} \tag{3.10}$$

3.4 New Heavy Resonances: W' and Z'

Different models predict the existance of new heavy resonances, like Z' and W'.

Many theoretical extensions to the Standard Model introduce additional symmetries, which involve the existence of new charged vector currents. Those currents have to be carried by new massive gauge boson, the W' . The W' boson appears in theoretical models that involve extra dimensions, from techni-colour models or from Little Higgs theories [107].

There are many models involving the hypothetic W', and many decay modes and experimental signatures. For example W' could decay semi-leptonically into a charged lepton (μ or e) and a neutrino [104], or through processes involving a top quark and b-tagged jets [107]:

$$\begin{gathered} W' \to 1 \text{ high-}p_{\mathrm{T}} \; lepton + E_{\mathrm{T}}^{\mathrm{miss}} \\ W' \to t\bar{b} \to electrons + muons + jets + E_{\mathrm{T}}^{\mathrm{miss}} \end{gathered} \tag{3.11}$$

Other extensions to the SM involve additional $U(1)'$ gauge symmetries, with its associated heavy gauge boson Z' [103]. The discovery of a new Z' boson would mean breaking the $U(1)'$ symmetry, which would require an extended Higgs boson, with important consequences for the problems of dark matter and the electroweak baryogenesis [58].

Leptonic and hadronic processes are involved in the Z' decay, like:

$$\begin{gathered} Z' \to 2 \text{ high-}p_{\mathrm{T}} \text{ opposite-sign same-flavour isolated } leptons \\ Z' \to 2 \text{ high-}p_{\mathrm{T}} \; jets \\ Z' \to 2 \text{ high-}p_{\mathrm{T}} \text{ different-flavour } leptons \\ Z' \to t\bar{t} \end{gathered} \tag{3.12}$$

3.5 Scalar Leptoquark

The Standard Model shows a symmetry between the properties of leptons and quarks. While other BSM theories, like e.g. the SU(5) grand unification and Pati-Salam SU(4), introduce a new symmetry relating leptons and quark and predict the existence of new boson particles, called *leptoquarks*. The leptoquark (or LQ) has fractional electric charge, is coloured, and carries both baryon and lepton number. At LHC, scalar leptoquarks can be produced either singly or in pairs. Pair production of scalar leptoquarks mostly occurs through *gluon-gluon* fusion, which is the dominant process for $m_L Q \sim 1$ TeV, and $q\overline{q}$ annihilation, dominant at larger masses. Single leptoquark production involves the unknown coupling γ_{LQ-l-q} [21].

Scalar leptoquark signatures

Leptoquarks can decay to leptons and quarks, violating the lepton and baryon numbers. Low mass leptoquarks could decay to a quark and a charged lepton with a branching fraction β, or to a quark and a neutrino with a branching fraction $(1-\beta)$ [84]. A summary of leptoquarks decay processes is shown here below:

$$\begin{gathered} 2 \text{ high-}p_{\mathrm{T}}\ leptons + jets \\ 1\ lepton + E_{\mathrm{T}}^{\mathrm{miss}} + jets \\ E_{\mathrm{T}}^{\mathrm{miss}} + jets \\ E_{\mathrm{T}}^{\mathrm{miss}} + \text{b-tagged-}jets \\ 2\ \tau + \text{b-tagged-}jets \end{gathered} \tag{3.13}$$

Those processes translate, for example, in the following channels:

- 1^{st} generation: $eejj$, $e\nu jj$
- 2^{nd} generation: $\mu\mu jj$, $\mu\nu jj$, $\nu\nu jj$
- 3^{rd} generation: $\nu\nu bb$, $\tau\tau bb$

3.6 Many possible signatures of new physics. Introducing model-independent searches

In this chapter only a very short selection of possible extensions to the Standard Model has been given. There are many other alternative theoretical models, which could explain phenomena beyond the SM. And yet, an already quite large number of physics channels has been listed so far. Table 3.1 shows a non-exaustive summary of possible discovery channels for the BSM models briefly described in this chapter.

A quite large fraction of those channels have been already addressed by HEP experiments in the past, and more recently by the ATLAS and CMS experiments at LHC; they are the object of dedicated searches.

But new physics could show in other channels, which are not listed here, behaving in ways that have not been described by any theoretical model so far. If that happened, the experiments could miss the signals of new physics, because the relevant signatures wouldn't have been considered and the corresponding channels wouldn't have been searched by any dedicated analysis.

That's why is important to have a model-independent search strategy as well, in addition to the dedicated analyses: model-independent searches are not bound to a particular theory and its experimental signatures, but they try to explore the largest number of physics channels, to being able to look at the largest number of possible processes. Already some experiments have performed model-independent analyses, not only at LHC, but also at Tevatron and HERA, from the recent past; those searches will be presented in Chapter 5.

But ATLAS lacked of this type of analyses until now. As already said, this is the motivation behind this Thesis: to establish a new model-independent search for ATLAS. The new ATLAS General Search will be presented and described in Chapter 6.

BSM Theory/Model	**selection of possible channels**			
Supersymmetry	νej	$\nu\mu j$	$\nu e2j$	$\nu\mu 2j$
	$\nu 2e2j$	$\nu 2e3j$	$\nu 2\mu 2j$	$\nu 2\mu 3j$
	νj	$\nu 2j$	$\nu 3j$	$\nu 4j$
	$\nu 3e2j$	$\nu 3e3j$	$\nu 3\mu 3j$	$\nu 3\mu 4j$
	...			
Large Extra Dimensions	νj	$\nu 2j2e2\gamma$	$\nu 2je\gamma$	$\nu 2j2\mu 2\gamma$
	$\nu\gamma$	$\nu j\mu\gamma$	$\nu 2j2e\gamma$	$\nu 2j\mu 2\gamma$
	$\nu je\gamma$	...		
Warped Extra Dimension	$2\mu 2e$	$2e\gamma j$	$2\mu 2j$	$2\mu\gamma 2j$
	2μ	$2\mu\gamma j$	$2e2j$	$4\mu\gamma 2j$
	$2e$	$4e\gamma j$	$\nu 2j$	$4\mu\gamma j$
	2γ	...		
Z'	$2e$	$t\bar{t}$	2μ	$e\mu$
	$2j$	...		
W'	$\nu e2j$	$\nu\mu 3j$	$\nu\mu 2j$	$\nu e3j$
	...			
Leptoquark	$2e2j$	νej	$2\mu 2j$	$\nu\mu 2j$
	$\nu e2j$	$\nu\mu j$	...	

Table 3.1: Short summary of possible channels of the BSM models described in this chapter. In this list ν indicates the presence of $E_{\mathrm{T}}^{\mathrm{miss}}$in the channel.

Chapter 4

The ATLAS detector at the Large Hadron Collider at CERN

> If your experiment needs statistics, you ought to have done a better experiment.
>
> Ernest Rutherford

ATLAS (acronym for A T*oroidal* L*HC* A*pparatu*S) is one of the four main detectors at the Large Hadron Collider (LHC). This chapter describes the structure of the detector and its main components. The description and the explanations that one can find in this chapter, follow to a great extent [125][130].

4.1 The Large Hadron Collider at CERN

The Large Hadron Collider (LHC), built at CERN (Geneva, Switzerland) between 2000 and 2008, is currently the world's most powerful largest hadron collider. The LHC is a two-ring-superconducting-hadron accelerator and collider, and it has been installed in the old 27 km tunnel constructed in 80's to host the earlier CERN accelerator, the electron-positron collider LEP [190]. The tunnel lies between 45 m and 170 m below the surface of the area between Geneva and the Jura mountains, on a plane inclined at 1.4%. The LHC "ring" is not a geometrical ring: it is composed, actually, of eight straight sections and eight arcs, the latter made up of bending magnets which guide the accelerated particles along the trajectory. 9,300 different types of super-conducting magnets are used in the LHC to steer and shape the particle beams as they travel around the 27 km loop of the collider.

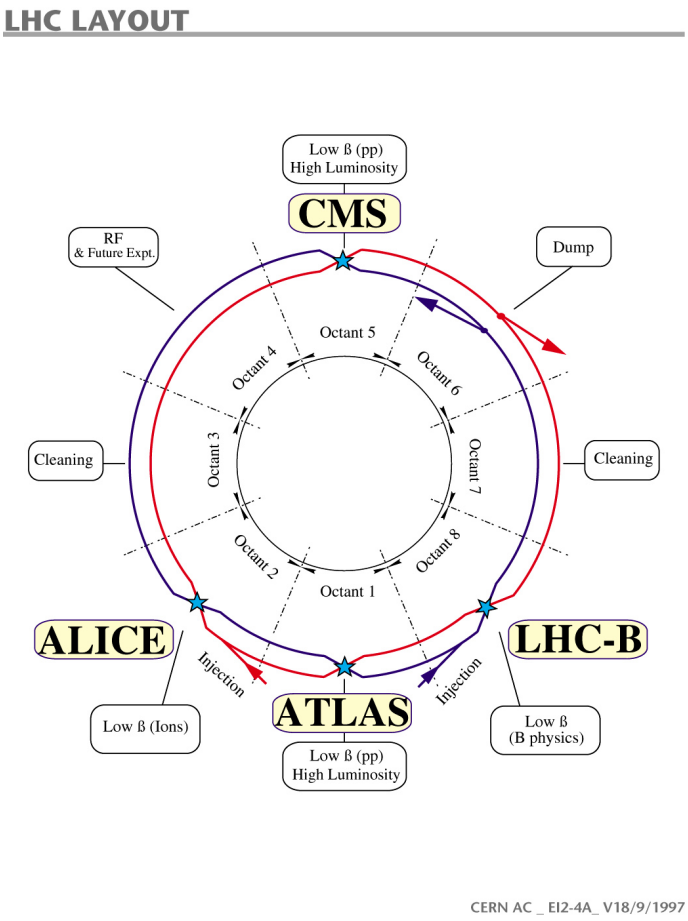


Figure 4.1: A schema of the Large Hadron Collider. The ring is divided into sections; the two counter-rotating particle-beams are shown, and the four colliding points where the main detectors (ATLAS, CMS, Alice and LHCb) have been built [177].

These include 1,232 di-pole magnets of 15 m length, which are used to bend the beams (shown here in Figure 4.3(b)), and 392 quadrupole magnets, each 5–7 m long, to focus the beams. Other kinds of N-pole super-conducting magnets are used to *squeeze* the beam near the collision points, in order to have the particles closer together and thus increase the chances of interaction among the colliding protons. A layout of the LHC ring is shown in Figure 4.1. In the figure, the two counter-rotating particle-beams are shown, and the four colliding points where the main detectors (ATLAS, [121], Alice and LHCb) have been built. There are also two transfer tunnels — at the bottom of the figure — each approximately 2.5 km in length, linking the LHC to the CERN accelerator complex, which acts as injector.

In fact other CERN machines act as pre-accelerators, accelerating the beams up to the injecting energy of the LHC. In Figure 4.5 one can clearly see how the

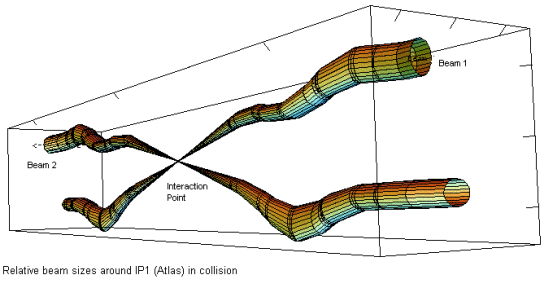


Figure 4.2: The two beams at the "P1" Interaction Point, in the ATLAS cavern. The 10^9 protons of a bunch are squeezed into a space of ~ 64 μm, to collide with a bunch of particles from the other beam. At the design luminosity the number of proton collision per bunch-crossing is around 20. With a rate of one crossing every 25 ns (at the design luminosity), that gives a number of 600,000,000 (600 million) collisions per second [178].

accelerating chain is based on almost all the CERN machines: Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), Super Proton Synchrotron (SPS). The LHC is the last ring (dark grey line) in this complex chain of particle accelerators, where the older and smaller machines are used in a chain to boost the particles up to the energy of 450 GeV, which is the injection energy of particles entering the LHC ring. Then LHC accelerates them up to the final energy: 3.5 TeV per beam up until now; and 7 TeV in the near future, after the upgrade of the machine, in 2015.

Coming back to the Large Hadron Collider itself, it is a proton-proton collider, and therefore it is made up of two rings with counter-rotating beams, as shown in Figure 4.1. The particles run in two separate beam pipes, as illustrated in Figure 4.3(a) and 4.3(b). In fact, unlike particle-antiparticle colliders, which can have both beams sharing the same phase space in a single ring because particles of different charge can be kept separated with electromagnetical fields, two beams of same-charge particles have to be kept in two physically different beam pipes. Only at the collision points the beams briefly share the same pipe as the magnets direct them to collide head-on, as shown in Figure 4.2. At the interaction points, about 20 collisions per crossing occur, with nominal beam currents at the design luminosity.

Particles circulating in the accelerator can collide with the gas molecules inside the beam pipes. These parasitic collisions affect the life-time of the beam, because the number of protons populating the beam decreases with the number of collisions; moreover these interactions, if they happen near the interaction points where the detectors are, can produce particles which can interfere with the experiments. In order to avoid those unwanted collisions with gas molecules, there is a ultra-high vacuum system establishing a pressure of $\sim 10^{-11}$ mBar in the beam pipes. Which gives the possibility to LHC physicists and engineers to

proudly call the LHC "*the emptiest space in the Solar System*"[191].

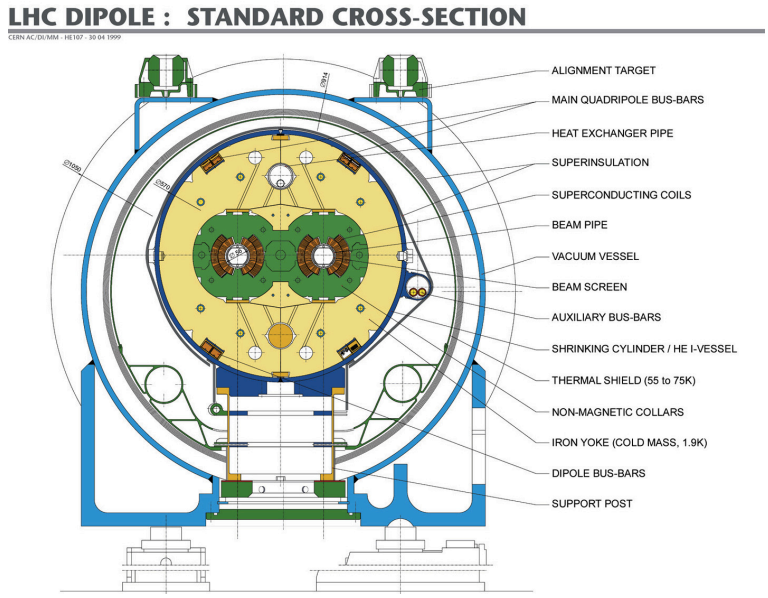
Protons are injected into the LHC in bunches, each containing $\sim 100 \cdot 10^9$ particles, which collide at discrete intervals 25 ns apart, at the design luminosity (50 ns during these first years of data taking). This interval is important because it lets the experiments distinguish data from subsequent collisions, recording the events as belonging to a particular *bunch crossing*. The acceleration of the protons is done at four locations where the particles pass through superconducting radio frequency (RF) cavities, as shown in Figure 4.1. The magnets must be cooled to 1.9 K (less than $-270.3°C$) so that the super-conducting coils can produce the required 8 T magnetic field strength. In Figures 4.3 and 4.4 different views of the LHC dipole modules are shown, together with the vacuum and cryogenic systems.

In principle, colliders can be designed and built for many different kinds of particles. The LEP collider, for example — the predecessor of the LHC —, was a matter-antimatter accelerator, which used leptons, electrons and positron, as colliding particles. Since leptons are elementary particles, the energy of the collision, in the centre-of-mass, is precisely defined: it is the energy of the colliding leptons; therefore such lepton accelerators are very well suited for experiments targeted to high-precision measures. On the other hand, in hadron colliders like LHC, the colliding objects are composite particles, like the protons in LHC; during the collisions, the actual interactions occur among the constituents of the proton, i.e. the quarks and the gluons. Each of such constituent carries a fraction of the proton energy, thus the actual collision energy can not be known *a-priori*, but it is a value between zero and the whole proton energy. Being uncertain the initial energy of the collisions, hadron colliders are not well suited for precise measures; but they offer a tremendous potential for the discovery of new Physics and new particles, because they allow collisions over a wide range of energy, which is otherwise not possible with lepton colliders.

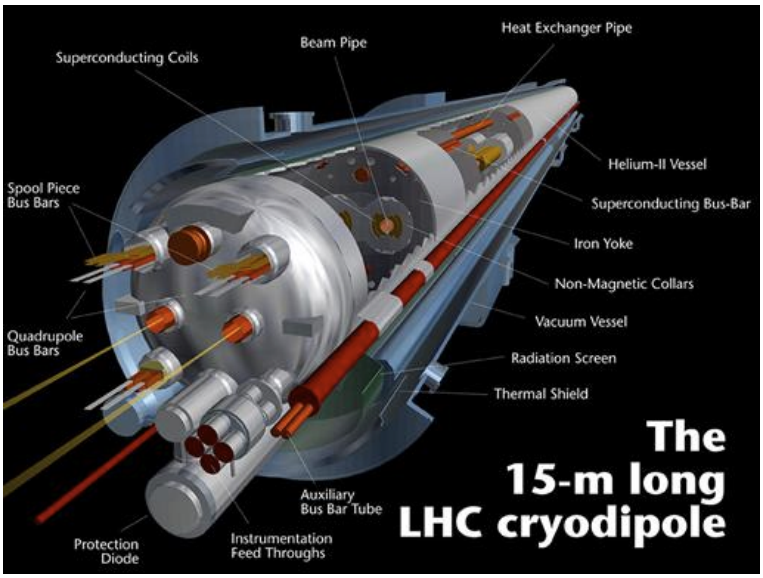
4.2 The ATLAS detector

ATLAS is one of the two major experiments at LHC, and it is a so-called *general-purpose* detector. Born to detect the Higgs particle and discover new physics beyond the Standard Model, its design has been conceived to be able to detect and measure a very broad range of different signals and particles, like leptons, hadrons, neutrinos and photons.

Like many other HEP detectors, ATLAS features an *onion-skin* structure, where each layer is composed by several sub-systems specialized in detecting a given class of particles and measuring their properties. The structure and the position of such layers — built as a series of ever-larger concentric cylinders around the nominal interaction point between the two proton beams — is moti-



(a)



(b)

Figure 4.3: The LHC. **(a)**: a section of a dipole of the LHC collider [175]; **(b)**: a computer-generated image of an LHC dipole magnet, showing the beam pipes, the magnets and their cooling system [169].

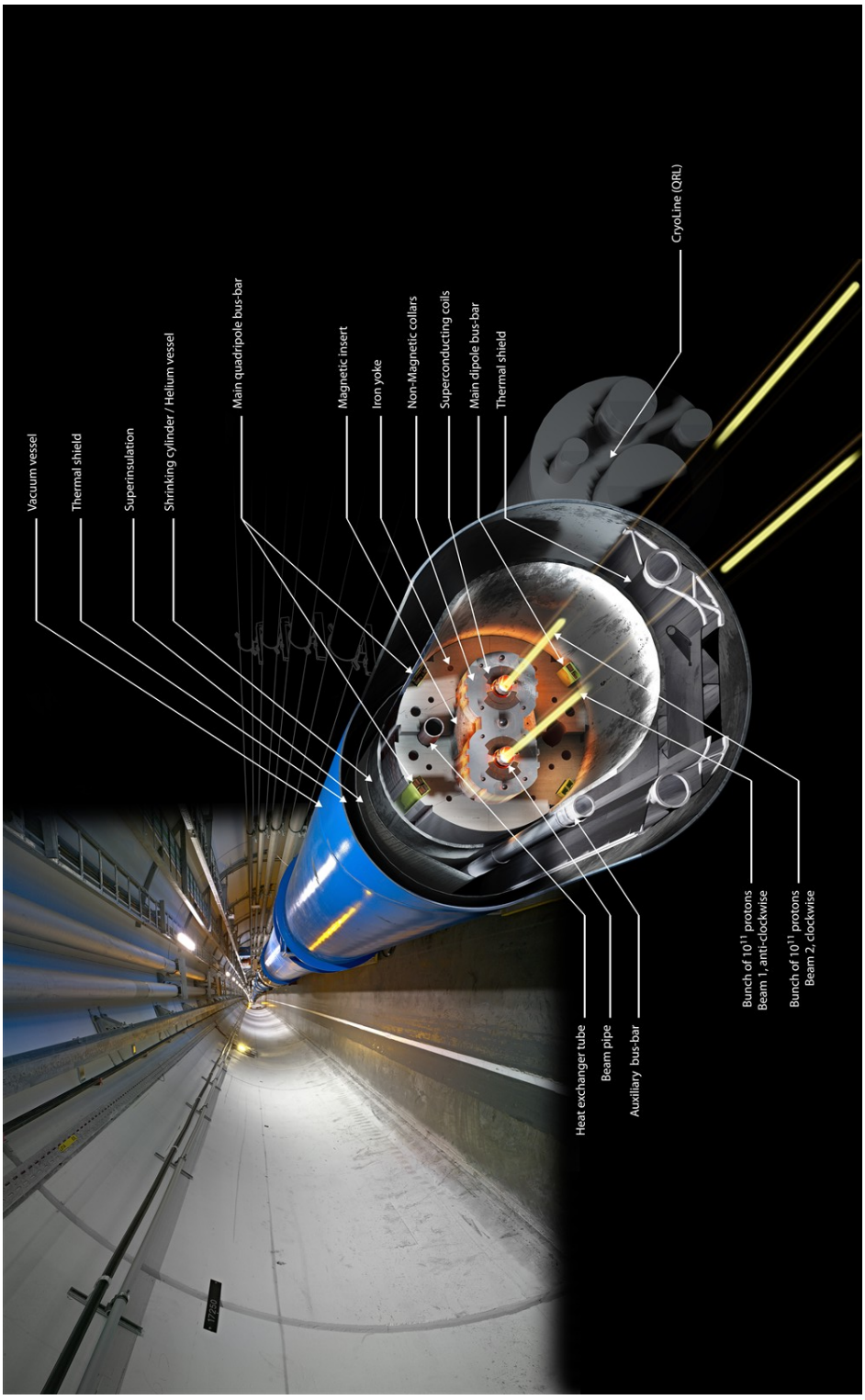


Figure 4.4: A computer-generated picture of a section of the LHC hadron collider in the tunnel, showing the beam pipes, the magnets and the other systems [174].

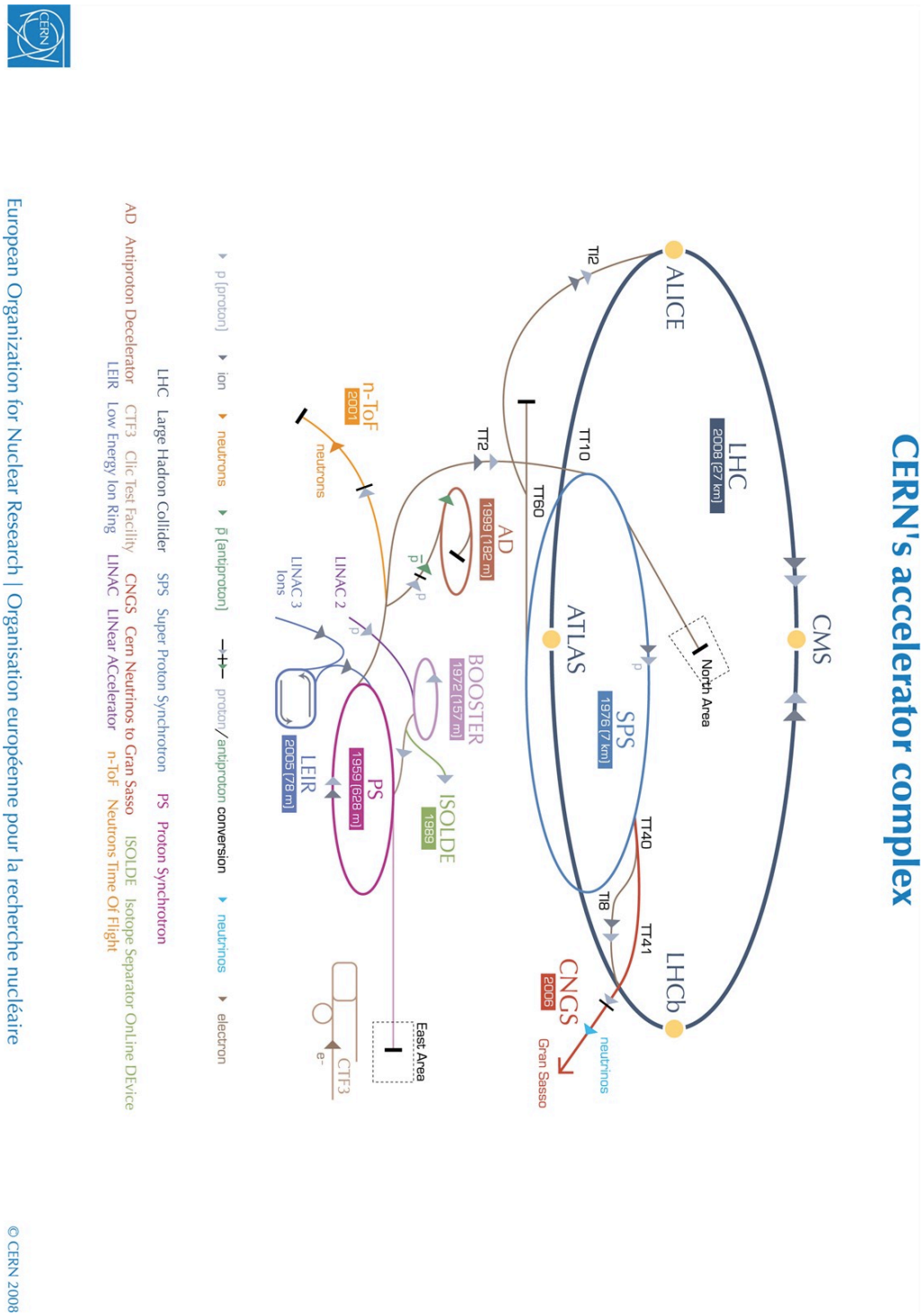


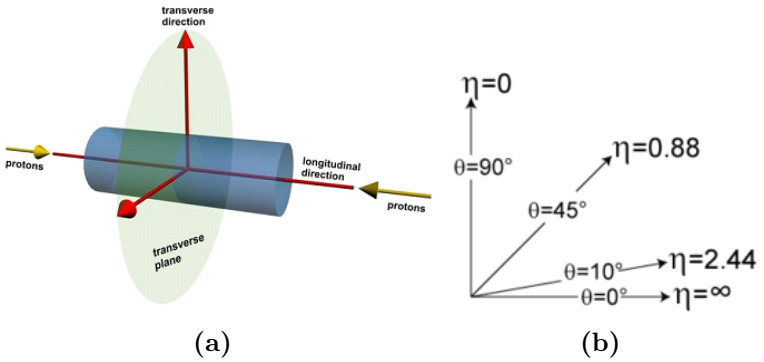
Figure 4.5: The layout of the accelerating-chain at CERN: the LHC is the last ring (dark grey line) in a complex chain of particle accelerators. The older and smaller machines are used one after the other to boost the particles to their final energies (besides providing beams to a whole set of smaller experiments) [179].

vated by the nature of the interactions of the different particles with the matter itself. Thus, as shown in Figure 4.7, we have sub-detectors targeted to strongly-interacting particles toward the collision point — the very centre of ATLAS — and those built to reveal particles that weakly interact with matter, positioned in the outer layer.

More precisely, the final-state products of a proton-proton collision are electrons, photons, jets, muons and neutrinos, from the more interacting to the lesser one. In the figure one can see a slice of the ATLAS detector. Leaving the collision point, the first layer is the *inner detector*, built to precisely reveal the tracks left by the charged particles (like electrons and muons). See section 4.3 for more details.

High-energy electrons and photons, interacting with matter, tipically produce showers of secondary particles due to bremsstrahlung and pair production. This energy is detected and measured by the second layer: the sub-system called *electromagnetic calorimeter*. Electrons and photons eventually loose all their energy in this calorimeter, and they stop. Also charged hadrons (like protons) leave a small amount of energy in this layer — due to the lower bremsstrahlung as a consequence of their larger mass — but not enough to get stopped. They move forward, toward the *hadronic calorimeter*, where they start to produce particle showers via strong interaction cascades, eventually loosing all their energy. Neutral hadrons, like neutrons, only loose energy, and get detected, in this layer. High-energy muons loose a very little amount of energy while travelling through all the sub-detectors, and they typically tend to escape the detector volume. Thus the outer layer of ATLAS is the *muon spectrometer*, where high precision chambers reconstruct the muon track; this is bended, because the muon travels through a strong magnetic field: from the curved trajectory is indeed possible to infer the energy of the travelling muon (see Section 4.5). The dashed line that escapes the detector, in the same Figure 4.7, represents the neutrino track. Neutrinos interact very weakly with matter, and so they practically leave the detector volume without any energy loss; this makes the neutrinos un-detectable, and their presence can only be inferred from conservation of momentum of all particles involved in a collision. The energy of neutrinos is part of the so-called *missing energy* (or $E_{\mathrm{T}}^{\mathrm{miss}}$), together with other un-detectable particles like the neutral exotics supersymmetric s-particles a part of the ATLAS collaboration is looking for.

The sub-detector layers are built all around the interaction point, in order to almost cover the whole solid angle around the spot where the collision takes place; this is to avoid missing potential interesting particles escaping from the proton-proton interaction, and to try to measure most of the energy produced in the collisions. A detector designed in such a way is called a "4π-*detector*", from the value of the complete, spherical, solid angle measured from the interaction



(a) (b)

Figure 4.6: The ATLAS coordinate system. **(a)**: a right-handed coordinate system where the z-axis points along the beam pipe, the x-axis points to the center of the LHC ring, and the y-axis points upwards [149]; **(b)**: values of the cylindrical θ coordinate, and the corresponding pseudo-rapidity η values commonly used in ATLAS: $\eta = 0$ for high-energetic particles leaving the interaction point at an angle of 90°; and $\eta = \text{inf}$ for a particle travelling along the beam [Figure: [199]].

point in sr, the 4π angle indeed.

In Figure 4.8 one can see the whole structure of the ATLAS detector. It is 44 m long and 25 m tall, and it weights around $\sim 7,000$ tons. Two magnetic fields are used in the detector, to bend particles trajectories: the solenoid, surrounding the inner detector, and the toroid magnet, situated outside the calorimeters and enclosing part of the muon chambers.

ATLAS uses a right-handed coordinate system with the z-axis along the beam pipe. The x-axis points to the center of the LHC ring, and the y axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle. The pseudo-rapidity η is defined in terms of the polar angle θ as:

$$\eta = -\ln(\tan(\theta/2)) \tag{4.1}$$

Thus $\eta = 0$ for high-energetic particles leaving the interaction point at an angle of 90°; and $\eta = \text{inf}$ for a particle travelling along the beam direction.

The distance in the $\eta\phi$ plane is defined as:

$$\Delta R = \sqrt{((\Delta\eta)^2 + (\Delta\phi)^2)} \tag{4.2}$$

4.3 The Inner Detector

The ATLAS Inner Detector is a composite tracking system consisting of silicon pixels, silicon strips and straw tubes in a 2 T magnetic field provided by

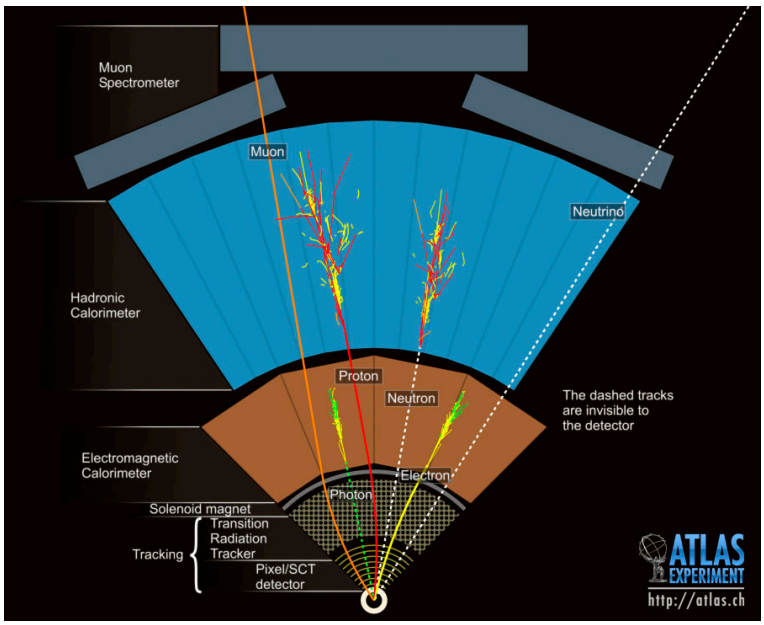


Figure 4.7: The interaction of particles with matter, in a slice of the ATLAS detector. On the bottom, represented by the white small circle, the collision point between the two proton beams [176].

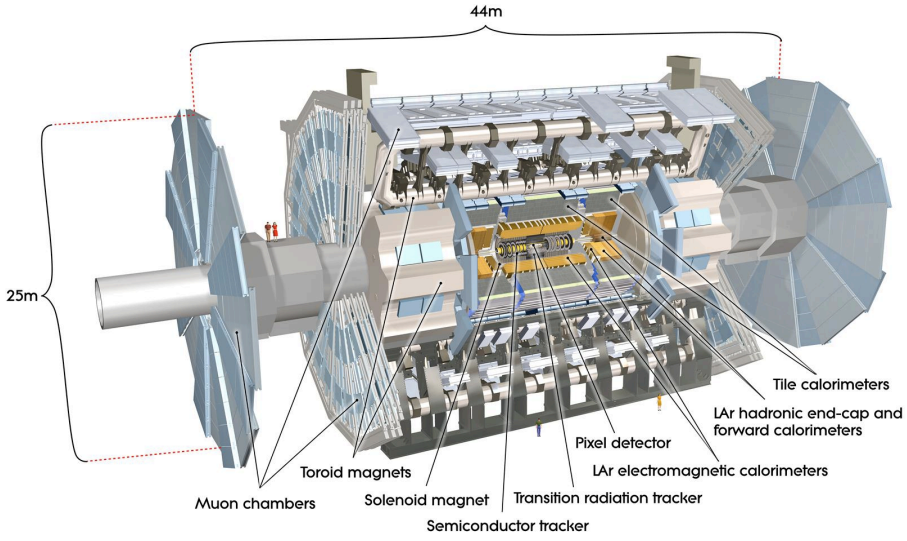


Figure 4.8: A computer-generated image of the whole ATLAS detector, with its sub-systems [173].

the solenoid magnet. The acceptance in pseudorapidity is $|\eta| < 2.5$ for particles coming from the LHC interaction point, with full coverage in the ϕ angle. The detector has been designed to provide a transverse momentum resolution of $\sigma p_{\mathrm{T}}/p_{\mathrm{T}} = 0.05\%$ in the plane perpendicular to the beam axis, and a transverse impact parameter resolution of 10 μm for high momentum particles in the central η region[125]. The Inner Detector comprises three complementary sub-detectors: the Pixel Detector, the Semi- Conductor Tracker and the Transition Radiation Tracker, shown in Figure 4.9. The *Pixel Detector* covers radial distances between 50.5 mm and 150 mm. It consists of 1744 silicon pixel modules arranged in three concentric barrel layers and two endcaps of three disks each, providing three measurement points for particles originating in the interaction region[108].

Each module, covering an area of 16.4 mm × 60.8 mm, contains 47 232 pixels, most of size 50 μm × 400 μm. The direction of the shorter pitch defines the local x-coordinate on the module and corresponds to the high-precision position measurement in the Rϕ plane. The longer pitch corresponds to the local y-coordinate and it is oriented approximately along the z direction in the barrel and along R in the endcaps. The total number of Pixel readout channels is 80 million, which is about half of the total read-out channels of ATLAS. Its very close proximity to the interaction point is essential for the measurement of secondary vertices caused by particle decays, like for example b quarks.

The SemiConductor Tracker (SCT) sensitive elements span radial distances from 299 mm to 560 mm. The detector consists of 4,088 modules of silicon-strip detectors arranged in four concentric barrels and two endcaps of nine disks each. the total number of readout channels is ≈ 6 million.

Both ID and SCT use silicon pn junctions operated at reverse bias and cover the region $|\eta| < 2.5$.

At larger radii, the Transition Radiation Tracker (TRT) sensitive volume covers distances from 563 mm to 1066 mm. The detector consists of 300,000 proportional drift tubes — the so-called *straws* — 4 mm in diameter, 144 cm long in the barrel and 37 cm in the end-caps, and are read out by 350,000 channels of electronics. The TRT assists in discriminating electrons from heavier charged particles (e.g. pions) by efficiently absorbing the transition radiation photons in the xenon- based gas mixture within the straw tubes.

The inner detector features a very fine global spatial resolution on track positions of about 200 μm.

4.4 The calorimeter

Moving further away form the interaction point, the particles escaping from the interaction point enter the calorimeter. This sub-detector, displayed in Figure 4.10, comprises two components: the *electromagnetic* and the *hadronic* calorime-

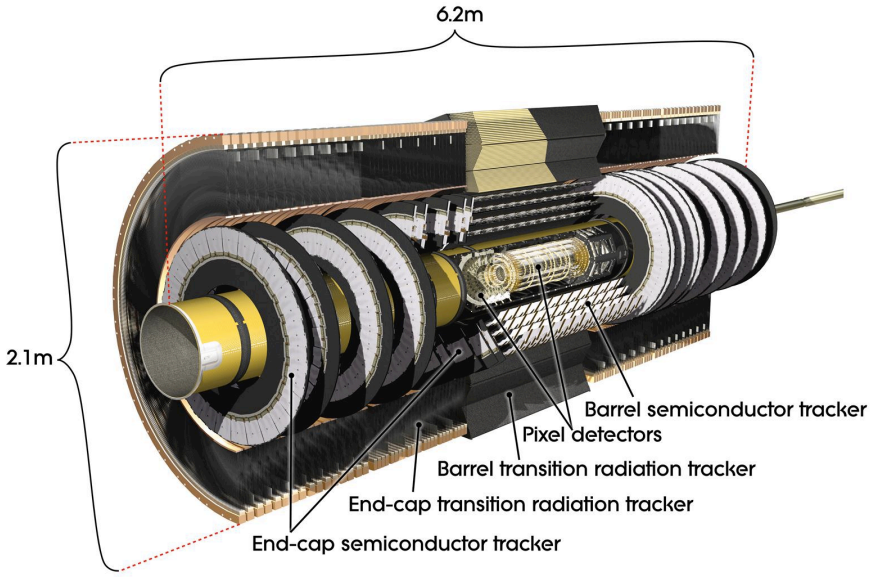


Figure 4.9: The ATLAS inner detector in a computer-generated image [171].

ter. Both are built as sampling calorimeters, alternating layers of high-density absorption material (metal) and sensitive layers (in ATLAS: liquid argon and scintillators) which measure the energy loss. This allows to measure the particle energy, inferring it from the sampled particle shower.

In the barrel region, the electromagnetic calorimeter, composed of lead absorption-layers and liquid argon sampling-layers, covers a pseudorapidity region up to $|\eta| < 3.2$. A liquid argon pre-sampler is also used in the region $|\eta| < 1.8$, in order to correct for the energy lost by electrons and photons upstream of the calorimeter. The hadronic calorimeter uses plastic scintillator tiles embedded in an iron absorber, and it covers the region $|\eta| < 1.7$.

In the end-cap region the hadronic calorimeters use again the liquid argon technology, due to intrinsic radiation tolerance needed in the high-particle-multiplicity forward region, together with copper plates as absorbers. This calorimeter covers the region $1.5 < |\eta| < 3.2$.

In the forward region there are three calorimeters: one electromagnetic layer using copper, in the region $3.1 < |\eta| < 4.9$, and two hadronic layers, using tungsten.

The ATLAS calorimeter system has been built to have a very fine energy resolution. With test beams, the combined energy resolution for the electromagnetic and the hadronic calorimeters together, has been measured, for pions, to be $(\sigma_E/E)^2 = (0.52/\sqrt{E(\text{ GeV})})^2 + (0.016/E(\text{ GeV}))^2 + 0.03^2$ in the barrel, and $(\sigma_E/E)^2 = (0.84/\sqrt{E(\text{ GeV})})^2$ in the end-cap regions [61].

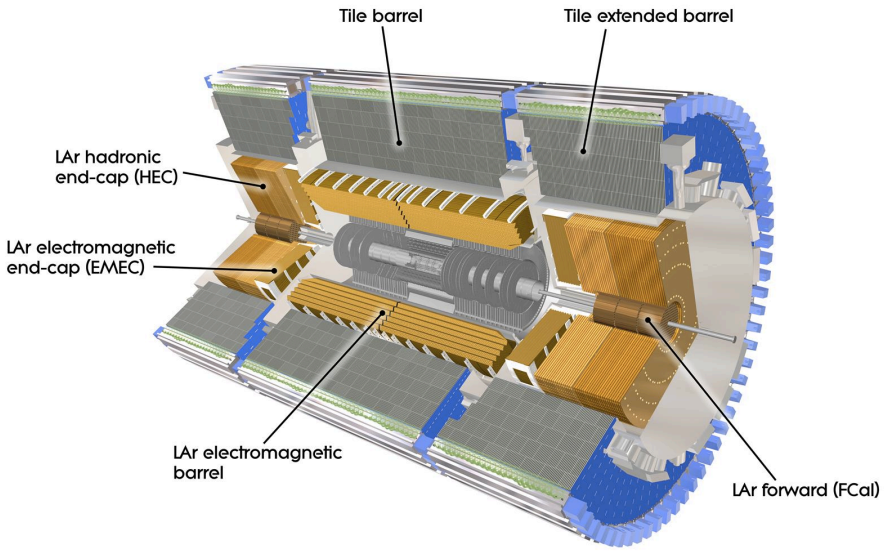


Figure 4.10: The ATLAS calorimeter system in a computer-generated picture [170].

4.5 The muon spectrometer

High-energetic minimum ionizing muons weakly interact with the detectors: carrying a charge they leave a track, but they loose a very little energy; in this situation it is impossible to measure the total energy of muons with calorimetric techniques. Thus, the energy of the high-p_{T} muons is inferred from the effect of a magnetic field on their trajectory. A charged particle travelling in a magnetic field, in fact, is deflected proportionally to its momentum.

Calling *sagitta* the divergence of the muon track due to the magnetic field and multiple scattering in the detector material, the track trajectory can be approximated by a segment of a circle, and assuming the angle of the segment is small enough one can write the sagitta as in Equation 4.3:

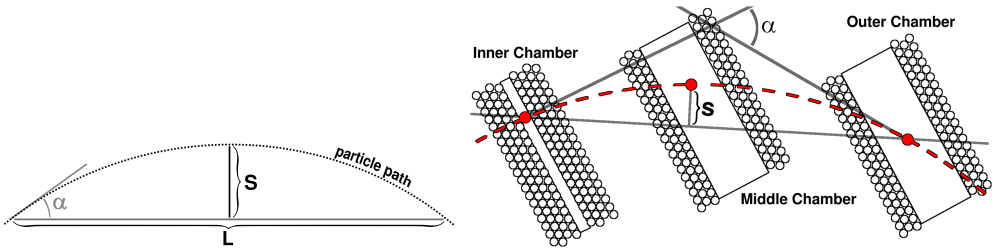
$$s = r(1 - \cos(\alpha/2)) \approx r(\alpha^2/8) \tag{4.3}$$

where α is the angle of the track segment, as shown in Figure 4.11(b).

After solving the angle, α, for the chord of the circle segment, and after having applied the equation of motion for a charged particle in a magnetic field 4.4, we can write the muon transverse momentum as in Equation 4.4.

$$p_{\mathrm{T}} \approx \frac{1}{8}\frac{L^2 \cdot B}{s} \tag{4.4}$$

Consequently, the measurement of the of the momentum of a track is dependent on the chord of the track, the magnetic field strength, and the sagitta[144].



(a) (b)

Figure 4.11: Muon trajectory in the muon spectrometer. **(a)**: definition of the sagitta s, the chord L and the angle α of an the actual muon track; **(b)**: in the ATLAS spectrometer the sagitta s is measured by measuring L and the track displacement. [186]

Since the magnetic field is known, the measurement of the chord and of the displacement of the particle track have to be measured, with great accuracy[136]. Being the goal of the muon spectrometer measuring the transverse momentum p_{T} of muons with $p_{\mathrm{T}} > 3$ GeV with a resolution of 4% up to p_{T} of 100 GeV, increasing to 10% at 1 TeV, the sagitta must be known to a resolution of 50 µm, the alignment of the muon spectrometer components be known to 30 µm and the uncertainty on the bending power of the magnetic field has to be around $\approx 10^{-4}$T [110][109].

In order to achieve these goals, the ATLAS muon spectrometer consists of a large air-core barrel and endcap toroid magnets with a $\int B \cdot dl$ between 2 and 6 T · m, and four types of trigger and precision tracking detectors, briefly described in the following. One can see a graphical representation of the many sub-systems in Figure 4.12. The spectrometer is composed by *Monitored Drift Tubes* (MDTs) chambers for precision tracking in the spectrometer bending plane, *Resistive Plate Chambers* (RPCs) and *Thin Gap Chambers* (TGCs) for triggering in barrel and endcap, respectively, and *Cathode Strip Chambers* (CSCs) for precision measurements in the high-rate endcap inner layer where MDTs would have occupancy problems. The spectrometer is designed so that muons cross three layers of MDT chambers for the sagitta measurement. The track coordinate in the bending plane of the spectrometer is measured by the precision chambers with a resolution of 40 µm. In comparison, the sagitta of a 1 TeV muon will be about 500 µm. The trigger chambers are placed on opposite sides of the middle MDT layer. The trigger chambers provide a trigger based on muon momentum in addition to identifying the bunch crossing time of the muon. They also provide the second coordinate measurement (non-bending plane) accurate to 5–10 cm. The resolution of the muon spectrometer is dominated by the energy loss of the muons in the calorimeter for low momenta, $p_{\mathrm{T}} < 30$ GeV; by multiple scattering effects for 30 GeV $< p_{\mathrm{T}} <$ 100 GeV; while, above 100 GeV, calibration and alignment of the

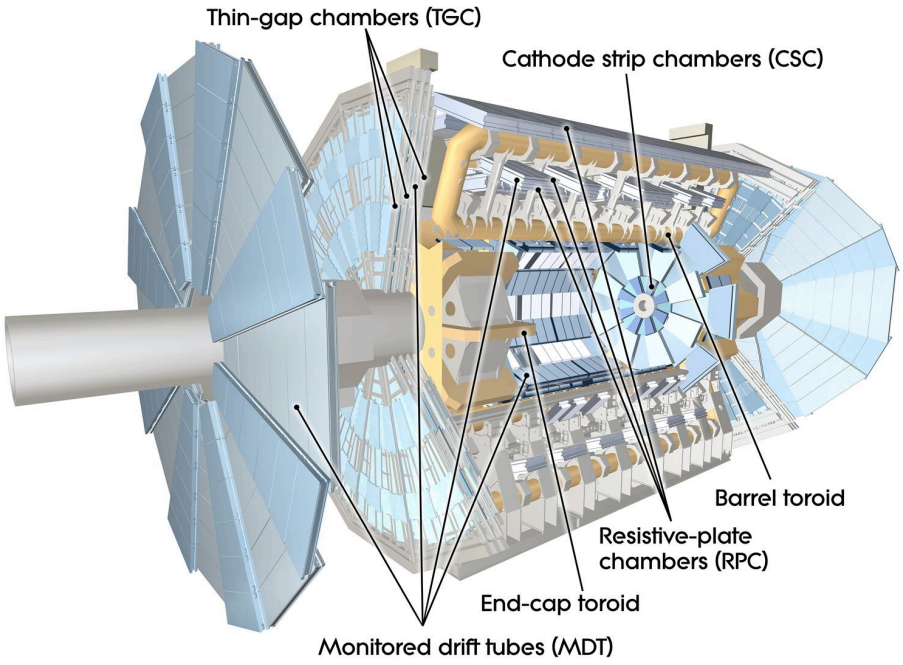


Figure 4.12: The ATLAS Muon sub-system in a computer generated image [172].

spectrometer itself become the most significant factors in momentum resolution.

4.6 The data acquisition system

4.6.1 Trigger

At the high luminosity at which LHC runs, the total *proton-proton* collision rate reaches ≈ 40 MHz (design value) and the average total size of one event is of the order of MByte. Hence, being the resulting amount of data too large to be all permanently written into physical storage, a pre-selection filter has been designed in order to reduce the total data flow without losing interesting physics events: the ATLAS trigger system.

In this brief description the design values will be used for presenting quantities like the trigger rates. In the end of the section the real values will be given, for comparison.

The ATLAS trigger system is organized in three levels as shown in Figure 4.13: the level-1, the level-2 and the level-3, also named *Event Filter*. Each step refines the decision taken at the previous step, by using a larger fraction of the event data and more advanced algorithms, analyzing more details of the event and filtering out those events which are not considered interesting. This reduces

the rate of data from the original 40 MHz to a much lower value of the order of few hundreds of Hz.

Level-1 trigger

The level-1 trigger, or LVL1, is a hardware-based trigger system which uses low granularity data from a given subset of detectors. The detectors are chosen to give a fast response and a rough but accurate enough information about the nature of the event being selected. The LVL1 has been designed to reduce the 40 MHz proton bunch-bunch crossing rate, to a rate of about 75 kHz. The latency time of LVL1 decision is ~ 2 µs. Two kinds of trigger have been implemented. The low-p_{T} trigger uses coincidence patterns that have been predefined in the $xy-$ and $rz-$plane. While in the barrel region, only the information coming from the RPC muon chambers in the middle stations is used. In the end-caps the two outer most TGC muon stations are combined to form a coincidence. The momentum cut-off of the low-p_{T} trigger is 6 GeV. Instead, the high-p_{T} trigger has a momentum cut-off of 20 GeVand requires coincidences in all three projective trigger stations [184]. If an event is accepted by the LVL1 trigger, then the full detector is read-out and the data is passed to the level-2 trigger.

Level-2 trigger

The level-2 trigger, or LVL2, is a software-based trigger which uses the output of the first level, the LVL1, and the full event information from all the subsystems, to select only the more interesting events and thus further reduce the data rate to 1 kHz. The specialized muon LVL2 trigger uses information from the precision chambers to improve the momentum estimate. In addition, a match with a track reconstructed in the inner detector can be required. The LVL2 trigger latency time depends on the complexity of the event, and it varies from 1 ms up to 10 ms for very busy events [184].

Level-3 trigger, or Event Filter

Events selected by the LVL2 trigger are passed on to the level-3 trigger, or LVL3, which uses the reconstruction algorithms also used during the offline event reconstruction, to fully reconstruct and understand the events. The LVL3 trigger runs on a computer farm located at CERN. The output rate of the LVL3 trigger is of the order of $\sim 100Hz$, which corresponds to a data rate of 100 MB/s [184]. All events selected by the LVL3 trigger are then permanently written to mass storage (disks or tapes) and are available for further analysis within the ATLAS offline software framework.

Actual Trigger rates

As already told, the values used in the description above are the design values of the Trigger system. For comparison, the actual rate values recorded during years 2010–2012 are given in Table 4.1. As one can notice, the collision rate is well below the design value: proton bunches, in fact, have injected at a lower rate during these first years of data taking: thus, e.g., in 2011 and 2012, collisions happened with an interval of 50 ns, instead of the nominal 25 ns. Moreover, one can also notice that while the L1 acceptance rate has been kept within the design range, the final output rate (the `EF` value, in the table) is much higher than the design value: a higher number of events have been selected by the *Event Filter* system to be saved to disc for the subsequent *offline* analysis.

Trigger Rates				
	Design value	*2010*	*2011*	*2012*
`Collision rate`	*40 MHz*	1 MHz	20 MHz	20 MHz
`L1`	*75 (100)*[1]*kHz*	20 kHz	50 kHz	70 kHz
`L2`	*3 kHz*	3.5 kHz	5.5 kHz	6.5 kHz
`EF`	*200 Hz*	350 Hz	400 Hz	1 kHz

[1] peak value.

Table 4.1: Trigger rates. In the second column the design values are given, while in the third, fourth and fifth columns the average values recorded during years 2010, 2011 and 2012 are shown, respectively.

4.6.2 Data Flow

The DataFlow system is responsible for moving data, which have passed the first level of selection to the High Level Triggers, and then for transferring the accepted data to mass storage [94]. See Figure 4.13 for a schematic view. The High Level Triggers have been designed such that their requirements in terms of the bandwidth needed for data movement are similar. The second level trigger operates only on a fraction of the data (≈2% of the full event), which has been tagged by the first level trigger as containing the relevant physics information (Regions Of Interest (ROI)). It has to be capable of handling events at up to 75 kHz and the average latency for the decision taking is of the order of 10 ms. The Event Filter on the other hand analyses the fully reconstructed events, but it operates at a rate of a few kHz (≈2 kHz). Here the latency for decision taking is in the order of a few seconds. The DataFlow is functionally decomposed in four building blocks: the ReadOut System (ROS), the ROI Collection, the Event Builder (EB) and the Event Filter I/O (EF I/O). The ROS is responsible for receiving data from the detector, forward them on request to the second level

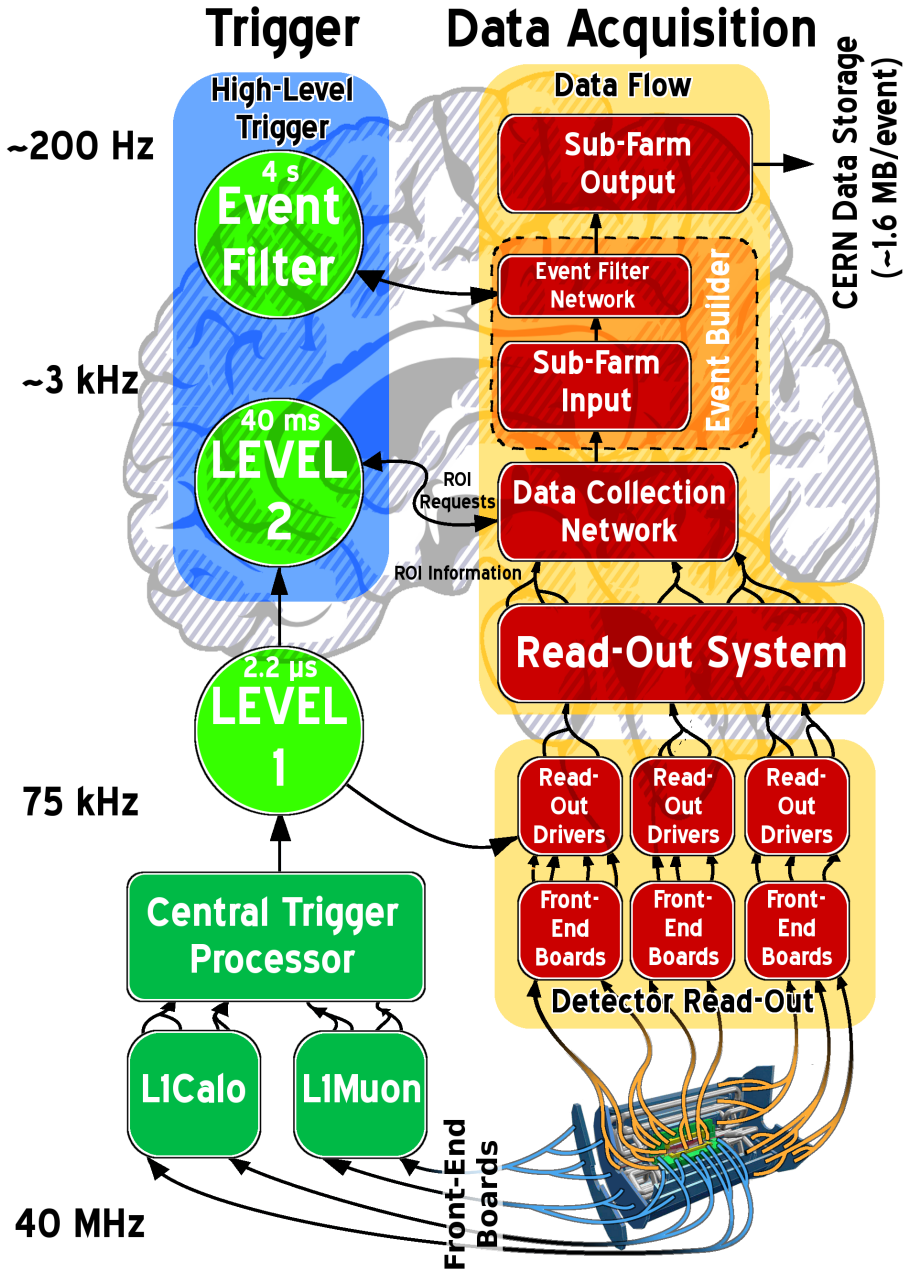


Figure 4.13: The structure of the ATLAS Trigger-DAQ (TDAQ) system [186][128]. In the picture, design values are shown; real values are given in Table 4.1.

trigger and Event Builder, and store the event data as long as it is explicitly told to delete them. The ROI Collection is responsible for gathering the data required by the second level trigger. The Event Builder is in charge of merging the event fragments coming from the ROS into a full event. The EF I/O forwards events to the last selection stage, retrieves the accepted events form the Event Filter and puts them on mass storage.

4.6.3 Control & Configuration

As already said, ATLAS relies on a complex and highly distributed Trigger and Data Acquisition (TDAQ) system [94] to gather and select particle collision data at unprecedented energy and rates. The Control and Configuration (CC) system is responsible for all the software required to configure and control the ATLAS data taking. This ranges from high level applications, such as the graphical user interfaces and the desktops used within the ATLAS control room, to low level packages, such as access and processes. Currently the CC system is required to supervise more than 30,000 processes running on more than 2,000 computers. At these scales, issues such as access, processes and resource management, distribution of configuration data and access to them, run control, diagnostic and especially error recovery become predominant to guarantee a high availability of the TDAQ system and minimize the dead time of the experiment. The CC system is made up of a central custom DB storing the whole ATLAS configuration (the so-called *OKS* database), a *run control* system handling the various sub-systems and sub-detectors of ATLAS in a sinchronized way along a common finite-state-machine, and various monitor and ancillary applications: from automatic error-recovery tools, to artificial intelligence and expert systems to effectively "store" experts' knowledge. [76][115]

4.7 Object reconstruction

Once data are stored on disk, the offline reconstruction phase starts. In this step all data are analyzed though specialized algorithms, running on computer clusters, that try to sort and combine all pieces of informations collected by the different sub-detectors in ATLAS, to reconstruct the final physics objects like jets, electrons, muons, $E_{\mathrm{T}}^{\mathrm{miss}}$. In the following a brief description of the object reconstruction will be given.

4.7.1 Jet reconstruction

Due to color confinement, final state gluons and quarks from QCD processes are not directly observable; they are observable only as colour-neutral particles. Hence the experimental signatures of hard QCD interactions are bursts of high-

energetic collimated hadrons. These bursts are called *hadronic-jets* , or simply *jets* . Jets are built by two processes, in two steps: *showering* and *hadronization*. Consequence of a proton-proton collision, a final state high-energetic parton first creates, through QCD radiation, a *shower* of gluons and quarks (*showering*), which then combine into colour-neutral particles (*hadronization*) [65] moving away from the interaction point, together in the same direction and within a cone of a certain radius.

Due to their nature, jets are not uniquely defined, but their definition depends on the *jet algorithm* chosen for their reconstruction. A jet algorithm is defined by different parameters and by the so-called *recombination scheme*, that describes the input objects and how their kinematic properties propagate to the jets [70]. Jet reconstruction algorithms are usually applied to *topological clusters* of calorimeter cells; but they can also be applied to simulated final state hadrons or charged-particle tracks.

Topological clusters, or *topo-clusters*, are three-dimensional groups of calorimeter cells, composed by cells with a signal-to-noise ratio above a certain threshold [137]. The signal-to-noise ratio is defined as the absolute value of the energy of the calorimeter cell, divided by the cell noise; where the noise of the cell is computed as the quadratic sum of the electronic noise and corrected to account for the effect of the proton-proton collision pile-up [100]. The topo-cluster is built adding adjacent cells together if their signal-to-noise ratio is above a given threshold. The calorimeter cells adjacent to the selected cells are also added to the cluster, without any constraint on their signal-to-noise ratio. In the end, a final four-momentum vector is defined for each cluster, considering the clusters massless, defining the energy of the clusters as the sum of the energy deposited in their constituent calorimeter cells, and their direction defined as the vector from the interaction point — which corresponds to the origin of the ATLAS coordinate system — to the energy-weighted barycentre of all the calorimeter cells which compose the cluster [112, 14].

The hadronic-jets used in this work are reconstructed with the "anti-k_t" algorithm with a distance parameter $R = 0.4$ [29]. The anti-k_t algorithm makes use of FastJet [166] to recursively and iteratively group input objects into jets; topo-clusters are taken as input objects for the reconstruction of the jets used in this analysis (see also Section 4.6.1).

4.7.2 Electron reconstruction

In ATLAS electrons are reconstructed combining the information from tracks in the inner detector with energy deposits in the calorimeter cells. Electron clusters are built in the LAr calorimeter using a sliding window of 3x5 cells to identify clusters whose energy is greater than 2.5 GeV. In the definition of the sliding window, the width in ϕ is larger than in η to account for the energy losses of

the *bremsstrahlung* contribution due to the action of the magnetic filed, which bends the tracks in the inner detector. Each cluster is then matched to a track in the inner detector, taking the track that is the closest in ΔR to the cluster barycenter. That track is taken as the electron track. Then the matching cluster is re-calculated into a larger window of cells (3x7 for barrel calorimeter cells and 5x5 in the end-caps) to be sure to get the correct whole energy deposit of the shower.

In the analysis presented in this Thesis, electrons are reconstructed using the "`medium`" level selections, as explained in [96, 95] and in Section 6.5. The "`medium`" algorithm requires that the angular distance of the matching track from the cluster barycentre must be $\Delta\eta < 0.01$, and it requires at least one hit in the pixel detector and at least 7 hits in the SCT chambers (introduced in Section 4.3), as well as a transverse impact parameter < 5mm ; moreover the algorithm includes additional requirements on the shape of the electromagnetic shower and on the hadronic leakage.

4.7.3 Muon reconstruction

The muons used in this work are reconstructed in ATLAS using the `STACO` reconstruction algorithm. `STACO` — or STAtistical COmbination — combines a muon track reconstructed in the muon spectrometer (see Section 4.5 for details) with a matching inner detector track; the outcome is called "combined muon". Then the algorithm also matches energy deposits in the calorimeter, a single hit or a segment of hits [97].

4.8 $E_{\mathrm{T}}^{\mathrm{miss}}$ definition and measurement

At hadronic colliders, the colliding objects are composite particles, made of subcomponents. At LHC the accelerated protons meet at the interaction points, and the actual collision occur among the proton partons. Due to the fact that each parton carries a variable fraction of the proton energy, it's impossible in experiments at hadron colliders, like ATLAS, to precisely know the centre-of-mass energy of a given single collision. However, in collisions at very high energy as like as at LHC, the longitudinal momentum of the particles is very small, and thus the sum of the transverse momentum of all particles in the event should sum to zero. When the sum is not zero, it means that a certain fraction of the initial energy has been brought away by weakly-interacting particles, which are not revealed by the detector, like neutrinos, or by yet unknown objects like neutral supersymmetric particles. The quantity of the initial energy of the event which is brought away by such particles is called "transverse missing energy", or $E_{\mathrm{T}}^{\mathrm{miss}}$, because it is missing from the final total energy sum. The measure of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is hardly dependent on the energy resolution of all the other reconstructed objects — muons, electrons,

jets, photons, taus — and any other additional energy revealed in the calorimeter cells that are not associated with any other object.

In this analysis the `MET_RefFinal` $E_{\mathrm{T}}^{\mathrm{miss}}$ definition has been used. It is calculated from muons and calorimeter cells. The calorimeter cells are calibrated according to objects that they are associated to: electrons, photons, taus, calibrated calorimeter "anti-k_T" topo-clusters (introduced in Section 4.7.1), isolated energy deposits in the calorimeter and muons. More details can be found in [161, 155] and [16].

Chapter 5

Model-independent searches in HEP

Strange events permit themselves the luxury of occurring. — Charlie Chan

Behind That Curtain
Earl Derr Biggers

The Standard Model (SM) of particle physics has been extensively tested with high precision in many of its predictions at a striking level of accuracy. But, as already described in Chapter 2, many theoretical hints and experimental observations suggest that the Standard Model cannot be a complete theory of fundamental interactions, and is instead just a low energy manifestation (or an effective theory) of the true underlying theory which would appear at higher scales. In particular, the current belief is that SM descriptions would break at energies of the order of the TeV scale. Therefore, around the energy scale of the LHC.

As seen in Chapter 3, various extensions of the SM have been proposed, involving new interactions, new spatial dimensions or new symmetries. The parameter space of possible new physics scenarios is so enormous, that it would be a real challenge (also in terms of manpower and time consumption) to cover all of the most promising proposed scenarios with dedicated searches; which makes the dedicated study of the full parameter space impossible *de facto*. Furthermore, even a complete study might miss a potential signal, since Nature may have chosen to not behave like the currently proposed models of new physics.

For these reasons most of the recent HEP experiments have set up model-independent analyses , aimed at searching for new physics phenomena in a large

number of channels, also in those ones lacking theoretical models describing them and which, otherwise, would be excluded, or simply not considered, by dedicated analyses.

In this chapter a selection of the model-independent searches from the most recent HEP experiments is presented and described.

In 2001 the D0 collaboration has set up a model-independent search in data samples of their top-quark analyses [37]. This analysis has used a multidimensional approach (called SLEUTH) which subdivided the data into Voronoi diagrams to search for deviations of the SM expectation. The use of a *search algorithm* allowed a calculation of the statistical trial factor, i.e. the fact that a deviation gets statistically more likely if more phase space regions are investigated. In 2004 the H1 collaboration at the HERA collider used an approach with a 1-d search algorithm, exploring the sum of transverse momenta distributions of their *Run 1* dataset. This was the first analysis to perform a search in every event class (i.e. final state) accessible to the experiment. A global model-independent search has also been successfully performed by the CDF collaboration at Tevatron [31] and, more recently, the CMS collaboration has studied event classes containing an electron or muon [122]. Those analyses will be presented in the next sections.

All these model-independent analyses aim at looking at the largest number of possible physics channels, to find discrepancies between the data and the background. Thus, covering a large phase space, they are sensitive to a large number of potential signals. Giving priority to a wide-angle view, however, affects the sensitivity: model-independent searches, in fact, are in most cases less sensitive to given specific signals than dedicated analyses. Their power is not in the sensitivity, but in their large coverage.

The interpretation of the results from a model-independent search is somehow delicate: any significant deviation from the background needs additional understanding and interpretation, to determine its origin. Possible causes could be detector problems, problematic event reconstruction, insufficient understanding of the detector simulation or the physics simulation or generation; or, indeed, genuine new physics processes in the experimental data. Thus the output of a model-independent analysis should always be considered as a "pointing tool": a first step in the potential discovery of new physics.

5.1 Model-independent search with the SLEUTH algorithm at D0

A model-independent search for new physics had been performed at Fermilab by the D0 collaboration in 2001, to analyze the first D0 data sample of 100 $\text{pb}^{-1}$of $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV, collected in the period 1992–1996 (Run I) [37]; and

an update of the analysis has been released in 2011 over the new D0 data sample of 1.07 $\text{pb}^{-1}$of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, collected in the period 2002–2007 (Run II) [127].

The event selection

The aim of the analysis is to spot evidences of new physics beyond the Standard Model; and, like the other approaches described in this chapter, and as well as in the analysis that is the subject of this Thesis, the only assumption made on the new physics is that it can show itself by discrepancies between the data and the SM expectation at high p_{T}. The D0 analysis looked for excesses of data events over the SM background.

In the 2001 analysis over Run I data [37], the events are selected requiring at least one or more isolated leptons (muons or electrons) whose $p_{\text{T}} > 15$ GeV, within the fiducial volume of the detector. Events are then classified into exclusive final states containing isolated jets, muons and electrons, $E_{\text{T}}^{\text{miss}}$and reconstructed W and Z electro-weak bosons. Jets are required to have $|\eta| < 2.5$, and $E_{\text{T}}^{\text{miss}}$ is considered as object if larger than 15 GeV. For each final state, and according to its content, the distribution of the sum of the p_{T} or of the $E_{\text{T}}^{\text{miss}}$, or simply the number of events, are then considered.

In the more recent update of the analysis [127], instead, only events with at least one electron or muon have been considered, and the events have been sorted into 7 inclusive final states, described in [127, pp. 8–9].

The search algorithm

The D0 analysis uses the Sleuth algorithm to find discrepancies in the distributions and to compute the probability of observing them. The algorithm basically divides the space of the variables into Voronoi diagrams, defining regions of N data events. The way of defining the interesting regions, helps to reduce the number of possible regions, from infinity to a value $\sim 2^N$, where N is the number of data events. For each region the value p_N is then computed: the probability of the background expectation to fluctuate up to the observed number of events, or above them. The p_N identifies the interesting regions: the smaller the p_N is, the more interesting the region is. Then, for each final state, the global probability of the Standard Model expectation to move up to or above the data points is computed, under the assumption that for enough large data sets there are always regions where the SM background can fluctuate and cover the discrepancies between the data points and the SM expectation. The computation is done using MonteCarlo techniques: a number of hypothetical pseudo-experiments are generated from the distribution of the SM expectation, and the fraction of the pseudo-experiments whose p_N^{sim} is smaller than the p_N from data is computed.

The region with the smallest p_N^{sim} is then the most interesting region of the considered final state. Then the global P value is computed. P takes into account the number of the regions considered within a given final state. The computation is performed using a second set of hypotetical pseudo-experiments, to determine the fraction of similar pseudo-experiments in which a deviation from the SM expectation is as interesting as the deviation found in the data. The probability P is a value between 0 and 1: the smaller values indicates interesting regions where new physics can show up.

Defined P_{min} as the smallest P, a $\widetilde{P}$ probability has been computed, which takes into account the number of the final states considered: the larger the number, the larger the probability that another final state would have fluctuated up to or above the excess found in data in the given final state. $\widetilde{P}$ is defined as the fraction of MonteCarlo-generated hypotethical *pseudo-runs* (made of one pseudo-experiment for each final state considered) whose P_{min}^{Run} is equal to or smaller than the P_{min} found in the data. $\widetilde{P}$ estimates the probability of having similar experiments and runs producing an excess as interesting as the excess found in the original data. The detailed description of the SLEUTH algorithm can be found in [37, pp. 3714–3715].

Results

The analysis performed by the D0 experiment on Run I data found the channel $ee4j$ as the most interesting final state, with $p = 0.04$. After taking into account the number of all the final states that have been considered, the corrected probability value resulted in $\widetilde{P} = 0.89$, which means that the 89% of a set of hypothetical MonteCarlo-generated pseudo-runs (see above section) would have produced more interesting results, i.e. a final state with a signal — or a deviation — more discrepant than the most interesting observed in the data analyzed by the D0 Run I analysis. The distribution of the p-value in units of standard deviation, P, is shown in Figure 5.1: as one can see from the plot there is no evidence of discrepancies from new high-p_{T} physics and a good agreement has been observed between D0 data and the Standard Model expectation.

5.2 General Search at H1

A model-independent search has also been perfomed by the H1 collaboration at HERA [45]. In that search configurations with high-p_{T} muons, electrons, jets, photons and missing energy have been systematically investigated. The analysis used the complete HERA I data sample, collected from 1994 to 2000, corresponding to an integrated luminosity of 117 pb^{-1}.

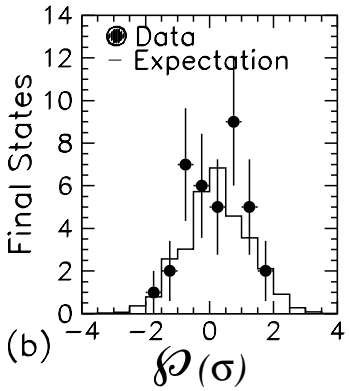


Figure 5.1: The final results of the SLEUTHanalysis performed by the D0 experiment. The distribution of the probability P (see Section 5.1) is shown. The plot shows good agreement, within uncertainties, between data and the Standard Model expectation background; no signal of new physics has been detected [37].

The search algorithm

The event selection of the H1 model-independent analysis requires at least one isolated jet, electron, photon or muon.

H1 developed a new algorithm to scan the events and quantify the agreement or the discrepancy between real data and Standard Model background; and the method has been proved to be able to spot single events, resonances and large regions of discrepancy [196].

All combinations with at least 2 objects from the set (μ, e, j, ν, γ) have been investigated; with objects in the selected final states such that $p_{\mathrm{T}} > 20$ GeV and $10^\circ < \theta < 140^\circ$. Then the selected events have been sorted and classified in exclusive classes, according to the type and number of selected objects in the event.

Once the events have been classified, two distributions have been investigated for discrepancies between the Standard Model background and the data: the scalar sum of transverse momenta $\sum p_{\mathrm{T}}$ and the invariant mass M_{all}; in these distributions new physics could show either as an excess or a deficit. Within these two distributions, regions of connected bins are built, with the requirement that the width of each region has to be at least the double of the resolution. For each region the number of data events and the number of SM background expected events are computed, together with the systematic uncertainty. The region of largest deviation is then found with a dedicated algorithm, and the probability of the deviation is computed, both for the single event class and globally (combining all classes together) to quantify the level of agreement between data and SM background. To find the most interesting region, the "p" statistical estimator

has been used. Here p is defined as the convolution of a Poisson and a Gaussian probability density function (PDF). The Poisson PDF takes into account the statistical fluctuations of data and background; while the Gaussian one takes into account the systematic errors. In Equation 5.1 the definition of the p-value is shown [45].

$$p = \begin{cases} A \cdot \int_0^\infty \mathrm{d}x\, G(x; N_{SM}; \Delta_{SM}) \sum_{i=N_{OBS}}^{\infty} \frac{e^{-x} x^i}{i!} & \text{if} N_{OBS} > N_{SM} \\ A \cdot \int_0^\infty \mathrm{d}x\, G(x; N_{SM}; \Delta_{SM}) \sum_{i=0}^{N_{OBS}} \frac{e^{-x} x^i}{i!} & \text{if} N_{OBS} < N_{SM} \end{cases} \tag{5.1}$$

In Equation 5.1 A is a unity normalization factor, defined in Equation 5.2. The definition of the p-value is such that if the Gaussian PDF would be replaced with a Dirac's delta-function $\delta(x - N_{SM})$, the p-value turns to the plain Poisson probability function.

$$A = \left[\int_0^\infty \mathrm{d}x\, G(x; N_{SM}; \Delta_{SM}) \sum_{i=0}^{\infty} \frac{e^{-x} x^i}{i!} \right]^{-1} \tag{5.2}$$

If an excess is observed — such that $N_{OBS} > N_{SM}$ — the p-value gives the probability of the SM expectation fluctuating upwards to at least the number of the observed data events, N_{OBS}, in the considered region. In the case of a deficit — i.e. $N_{OBS} < N_{SM}$ — then the p-value gives the probability of a fluctuation of the SM expectation downwards to at most N_{OBS}. When scanning the bins of the distributions for interesting regions, the smallest p-value indicates the region with the largest deviation. The smallest p-value per distribution is called p_{min}.

After computing the local greatest deviation, designated by the p_{min}, it is important to know the probability of such a fluctuation to show anywhere in the distribution. Thus an estimate of the probability of a fluctuation with a p-value at least small as p_{min} has to be evaluated, and that can be done with a MonteCarlo method, generating histograms following the SM expectation PDF. The approach is similar to the one used by D0 and described in Section 5.1: for every histogram the value p_{min}^{SM} is computed. The "corrected p-value" $\hat{P}$ probability is then a measure of the statistical significance of the discrepancy observed between data and the expected background. It is defined as the fraction of histograms featuring a p_{min}^{SM} equal to or smaller than the value of p_{min} obtained from the data. $\hat{P}$ can also be used to compare event classes, if they are exclusive. If the logarithm of $\hat{P}$ is used, then a value between 5 and 6 of the quantity $-\log_{10}\hat{P}$, corresponds to a region whose p_{min} value is $5.7 \cdot 10^{-7}$, which corresponds to a "5 σ" deviation [45][196].

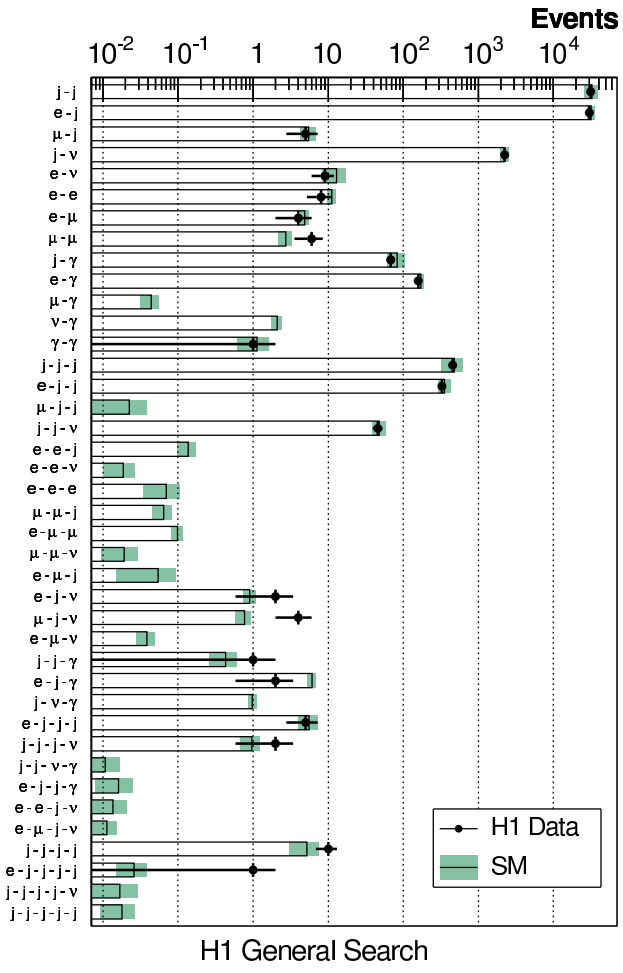


Figure 5.2: Final results of the model-independent search performed by the H1 experiment at HERA in 2004. The plot shows data events and the SM expectation for all the event classes with a Standard Model expectation greater than 0.01 events (integrated luminosity: 117 pb^{-1}) [45].

Results

The results of the model-independent search performed by H1 are shown in Figure 5.2. A general good agreement between data and the Standard Model expectation background has been found. The analysis spotted the largest deviation in the event topology $\mu j \nu$, but its significance has been found to be too small to indicate a physical interesting deviation. Figure 5.3 summarizes the outcome of the analysis: no channels showed significant deviations that can be related to new physics processes [45].

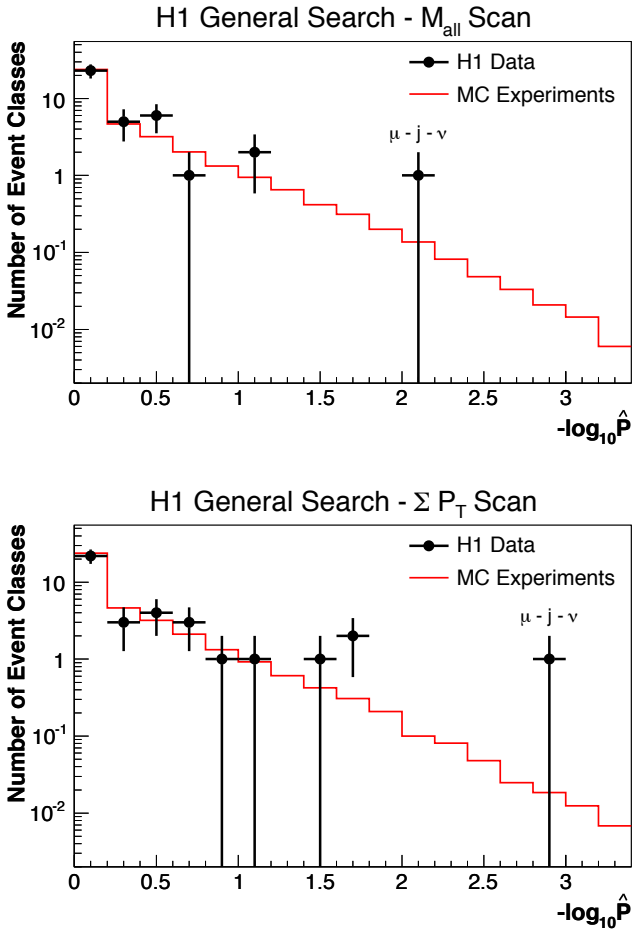


Figure 5.3: Distribution of the quantity $-\log_{10}\hat{P}$, summarizing the outcome of the model-independent search of the H1 experiment at HERA. A value between 5 and 6 (depending on the final states) corresponds to a region of a "5σ" deviation. The most interesting channel that has been found in the H1 analysis corresponds to the $\mu j\nu$, identified on the right of the plot with the corresponding label. The significance of that channel is however too small to indicate a signal from new physics [45].

5.3 Search with the VISTA+SLEUTH algorithms at CDF

The CDF collaboration has set up a model-independent search as well, to search for evidence of new physics at the electroweak energy scale [31]. The data set was composed by 927 $\text{pb}^{-1}$of $p\bar{p}$ collisions of the Run II data collected at Tevatron (Fermilab), with a center-of-mass energy of $\sqrt{s} = 1.96$ GeV, and the analysis has looked for discrepancies of data with respect to the SM expectation.

The search algorithm

The CDF analysis makes use of two algorithms simultaneously: VISTA, a model-independent algorithm which takes into account the gross features of the data, and SLEUTH, a quasi-model-independent algorithm also used by the D0 analysis (see Section 5.1), which is tuned for looking at events with large transverse momentum at the electro-weak scale (this tuning causes the presence of "quasi" in the definition of the algorithm). VISTA is sensitive to new physics featuring a large cross section, while SLEUTH is more adapted to find single final states showing discrepancies from the SM prediction and it has been built to identify new physics with small cross-section in the tails of the high-p_{T} distributions.

The whole CDF analysis has been set to find deviations between data and the SM background, in order to identify discrepancies which are worth of further and more detailed analyses.

For the analysis purpose three statistics have been used to spot and quantify the discrepancy: the number of events of exclusive final states (classified on the base of their particle content), the shape of the distribution of kinematic variables, and the identification of excesses in the tail of transverse high-p_{T} distributions.

VISTA uses "correction factors" to take into account source of errors and discrepancies, like the ratio of the not-available all order cross-section to the computed leading order cross-section (the k-factor) or the inefficiency of the object identification algorithms.

The SLEUTH algorithm used by CDF to analyze the Run II data sample, is basically the same developed by the D0 experiment (see Section 5.1), and then improved by the H1 experiment at HERA (see Section 5.2), with little changes [31, p. 15].

The event selection

The CDF VISTA analysis selects events with isolated high-p_{T} objects. For events containing single objects, it is requested, e.g., a muon with $p_{\text{T}} > 25$ GeV, an electron with $p_{\text{T}} > 25$ GeV, a photon of $p_{\text{T}} > 60$ GeV, a jet or b-jet with $p_{\text{T}} > 200$ GeV (or 40 GeV if pre-scaled). Additional cuts are requested for multi-objects events [31, p. 6]. The events are then sorted into orthogonal final state classes according to the particle content, and then subdivided into bins of p_{T} and η

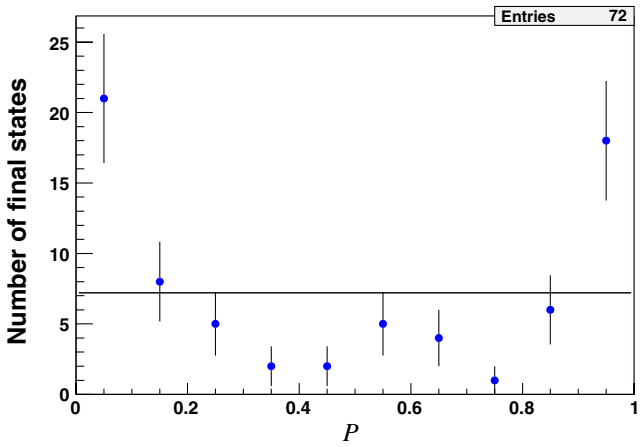


Figure 5.4: Final results of the model-independent search performed by the CDF experiment at Tevatron, on the Run II data sample, with the SLEUTH algorithm. The plot shows the distribution of the corrected P probability; no evidence of new physics has been found at high-p_{T} [31].

(pseudorapidity). MonteCarlo simulated events, after being corrected with the correction factors, provide the SM expectation for each bin.

The VISTA analysis looked at the population of the final states with at least ten data events, computing for each of them the Poisson probability p that the number of expected events of the background would fluctuate up to (or above) or down to (or down) the number of observed data events. Then, in order to take into account the *trials factor*, the probability p has been corrected for the large number of final states analyzed, with the formula shown in Equation 5.3, where N is the total number of final states.

$$p' = 1 - (1 - p)^N \tag{5.3}$$

The value of p', in Equation 5.3, is the probability of observing a more interesting discrepancy in the total sample, i.e. a deviation that corresponds to a probability lower than the value of p, which was found. The probability value p' has been then converted into units of σ solving the formula in Equation 5.4.

$$p = \int_{\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \mathrm{d}x \tag{5.4}$$

After correcting for the trial factor, a discrepancy $> 3\sigma$ is considered significant (it is analogous to the 5σ deviation in a standard analysis).

Results

After all the corrections applied in the VISTA analysis, none of the 344 populated final states considered in the analysis showed a significant discrepancy, thus no evidence of new physics has been claimed [31, p. 14].

The CDF VISTA analysis also looked at discrepancies in kinematic distributions, for each final state. A total number of $\sim$ 16500 distributions has been analyzed with a Kolmogorov-Smirnov test, in order to detect and quantify the discrepancy between data and the SM expectation, in each distribution. In the end none of all the distributions showed a deviation significant enough to spot an interesting feature in data, thus no evidence of new physics has been found with the VISTA algorithm.

The analysis performed using the SLEUTH algorithm analyzed a single variable in all the exclusive final states, i.e. the sum of the transverse momentum of all physical objects in the event: $\sum p_{\mathrm{T}}$. In Figure 5.4 the distribution of the final corrected P probabilities found by CDF with the SLEUTH algorithm is shown. In the end SLEUTH analyzed 72 final states and none of them showed a significant discrepancy from the Standard Model expectation. Thus no evidence of new physics has been found in the CDF Run II data sample.

5.4 Model-independent search at CMS: "MUSiC"

MUSiC (acronym for "Model Unspecific Search in CMS") is the search set up by the CMS experiment in 2010, in order to analyze the very first LHC data in a model-independent way [122]. Similarly to the other searches presented in this chapter, the aim of MUSiC is to scan the largest number of physics channels to find discrepancies between the data and the background, without assumptions on specific new physics models. MUSiC uses a modified H1 search algorithm (see Section 5.2 and [122]).

The event selection

The CMS analysis scanned the first 36 $\mathrm{pb}^{-1}$of data, selecting events that passed the single lepton triggers: a event should have at least an electron or a muon. To be sure to be in a region of large trigger efficiency, an high threshold is set on the p_{T} of the particles: 25 GeV for muons and 30 GeV for electrons. Additional cuts and vetoes are applied to suppress instrumental problems and backgrounds [122].

The search algorithm

As already said, MUSiC scans all the events to select those containing at least one charged lepton; events are then sorted into event classes, considering the following

physics objects: muons, electrons, photons, hadronic jets and $E_{\mathrm{T}}^{\mathrm{miss}}$($E_{\mathrm{T}}^{\mathrm{miss}}$is taken as an object if > 30 GeV). Events classes represent a single final state, depending on their content; and each event is sorted into precisely one event class. In the 2010 analysis [122] the scan led to about 250 event classes containing at least one real or simulated event.

After the assignment of events to classes, three kinematic distributions are built and analyzed, which are considered sensitive to new physics:

- the $\sum p_{\mathrm{T}}$ distribution: the scalar sum of the transverse momentum of all objects in the event; this distribution is the most sensitive to new physics models featuring new massive particles;
- the distribution of the invariant mass, M, of all objects (or of the transverse mass, M_T, if the class contains $E_{\mathrm{T}}^{\mathrm{miss}}$); this distribution is sensitive to new resonances;
- the $E_{\mathrm{T}}^{\mathrm{miss}}$distribution (if the event class features $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV), which allows the discovery of new physics involving boosted or heavy particles escaping the detector.

The kinematic distributions are built as binned histograms by a scanning algorithm, which compares the measured real data with the simulated SM prediction, looking for discrepancies. As like as in the H1 model-independent analysis (see Section 5.2), the width of the histogram bins is set to reflect the resolution of the considered variable. In addition, for the M (M_T) and $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, the scan is performed over a region of at least three bins, since discrepancies in those distributions very often affects large fractions of the spectrum. This requirement tends to suppress non-physical fluctuations, and to give importance to systematic effects. The scan is performed over regions of adjacent bins and, for each of them, the following quantities are computed or measured: the number of observed events N, the number of expected events B and its uncertainty σ. Then the algorithm calculates the p-value: the Poisson tail probability value that measures the probability of a random fluctuation to be as high as the observed one. The Poisson distribution is also convoluted with a Gaussian distribution, to take into account the systematic uncertainties. The p-value has already been introduced in Section 5.2 and in Equation 5.1 on page 68 and that can be expanded to Equation 5.5.

$$p = A \cdot \sum_i \int_0^\infty e^{\frac{-(\mu-B)^2}{2\sigma^2}} \cdot \frac{e^{-\mu}\mu^i}{i!} d\mu \tag{5.5}$$

In the equation 5.5 A is the normalization factor already introduced in the H1 formula 5.2, and the sum $\sum_i$ is performed for $i \in [0, N]$ if less data events

(N) than expected (B) are measured, and on $[N, \text{inf}]$ if more data events than expected are found.

For each distribution, the region with the smallest p-value — i.e. whose discrepancy is more likely to be interesting — is taken as "Region of Interest" (or "RoI").

After the calculation of the p-value, regions with no data events, or with the number of expected events smaller than 3σ from 0, are excluded. If all the regions of the whole distribution are excluded, then the event class is rejected. This selection lowers the number of event classes eventually considered in the analysis.

The p-value described in Equation 5.5 is a good estimate for local significances of a given distribution. But it has to be corrected for the "Look Elsewhere Effect" (LEE) when looking at many distributions [122]. The LEE effect arises when searching for discrepancies in a large number of distributions: a certain fraction of the examined distributions, in fact, is always expected to deviate, only due to statistical fluctuations. Thus a corrected p-value is computed, as like as in the H1 approach (see Section 5.2). For a given distribution, the corrected p-value used in the CMS analysis, the $\widetilde{p}$, is defined in Equation 5.6 as the ratio of the number of pseudo-experiments whose p-value is smaller than that one of the data region, and the total number of the pseudo-experiments.

$$\widetilde{p} = \frac{N_{pseudo}(p_{pseudo} < pdata)}{N_{pseudo}} \tag{5.6}$$

A MonteCarlo approach is needed, using pseudo-experiments generated from the SM expectation, because of the correlation of the RoIs, which does not allow an analytical solution of the corrected p-value.

An additional source of bias occurs from the fact of picking the most significant distribution. The same MonteCarlo approach has been used here, using pseudo-experiments: a $\widetilde{p}$ distribution has been generated from all the distributions, to show the confidence level of the choice by comparing this with the distribution of the measured $\widetilde{p}$ [122].

Results

The distributions of the $\widetilde{p}$ (i.e. the corrected p-value) of all analyzed event classes are shown in Figure 5.5. The plots show the results from 287 distributions, out of 118 event classes. The event classes analyzed by MUSiC feature up to 4 leptons (muons or electrons), 2 photons and 8 jets. The crosses are the data points (the $\widetilde{p}$ from the data, while comparing it to the SM), while the shaded area represents the MonteCarlo generated SM-to-SM $\widetilde{p}$.

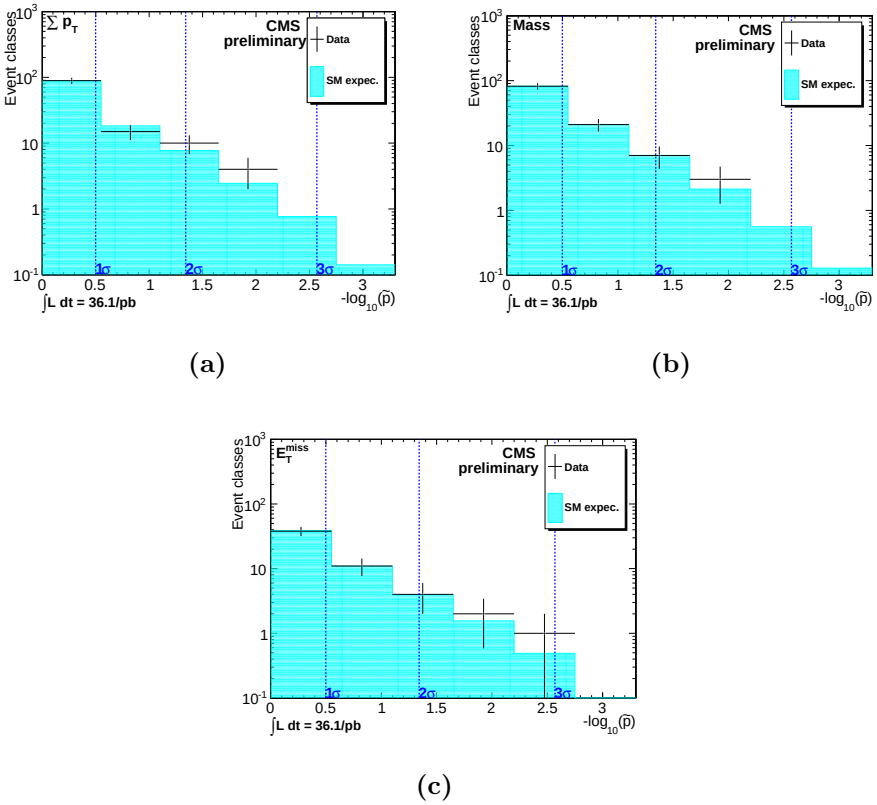


Figure 5.5: Final results of the model-independent MUSiC search performed at CMS with 2010 data. **(a)**: $\sum p_{\mathrm{T}}$ distribution of $\widetilde{p}$; **(b)**: invariant mass M distribution of $\widetilde{p}$; **(c)**: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of $\widetilde{p}$. Plots show good agreement between data and SM expectation within uncertainties; thus no signal of new physics has been revealed [122].

Within the uncertainties, the distribution of the data-to-SM $\widetilde{p}$ agrees very well with the SM-to-SM $\widetilde{p}$ distribution, and no significant unexpected discrepancies have been found in the CMS analysis [122].

Chapter 6

A new model-independent General Search for New Physics in ATLAS

All models are wrong, but some are useful.

George E.P. Box and Norman Draper [1]

Results of a model independent general search for new physics with the ATLAS detector at a centre-of-mass energy of $\sqrt{s}$=7 TeV will be presented in this chapter. Model-independent searches have been successfully set up in many HEP experiments, as already shown in chapter 5. In this search a similar global approach is used. Purpose of this work is to set the strategy and the methodology of a general search for ATLAS.

This model-independent general search performs a systematic comparison of data and simulated MonteCarlo (MC) expectations over the data collected in the year 2011 by the ATLAS detector. Data is analysed without bias: in contrast to common "model-driven" analyses, the only assumption made in this search is that new physics should appear in events with high p_{T}, i.e. events with particles with high transverse momentum.

Events containing leptons (μ, e), jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ are considered and subdivided into exclusive classes according to their final states. A basic idea of the approach presented here is that all events are sorted and classified into exclusive channels according to the type and the number of high p_{T} objects, in order to study all possible combinations — like, for example, muon-electron-jet, jet-jet or

neutrino-muon-jet. Channels are then merged with a custom *merging* algorithm developed for this study: its aim is to reduce the number of channels without loosing information on possible new physics. In each final event class, then, the number of events per class and the distribution of the Effective Mass M_{eff} are used to spot discrepancies between the data and the MC expectations, taking into account both statistical and systematic uncertainties. A search algorithm is also used to quantify the discrepancy in the M_{eff} distribution and to find the area of the greatest deviation.

Preliminary results over the first 2011 data, corresponding to an integrated luminosity of 2.052 fb^{-1}, are reported. A very good agreement with the Standard Model prediction is observed in most of the event classes, and no evidence of new physics has been found in the sensitivity limits of this analysis.

In the following sections the strategy of the analysis is described and the data and MC samples that have been used are outlined, together with their associated uncertainties. Then the description of the applied event selection and object reconstruction follows, and the mechanism of assigning events to channels and the merging algorithm are described. After the control distributions over the data and MC predictions, the search algorithm is presented. In the end the final results of this analysis are shown.

6.1 The analysis strategy

The approach pursued in this work is a model-independent general search for deviations from the SM predictions in all possible final states, and from all data streams collected by the ATLAS detector: `Muon`, `Egamma` and `JetTauEtmiss` data *streams*.

The data recorded by th experiment is subdivided into *streams* according to the objects that triggered the event. For example, in the ATLAS experiment, an event which contains at least an electron whose energy is enough high to pass the threshold of the `egamma` trigger , is pushed and stored in the "egamma" bin, that is the *egamma stream*; the same for events containing muons or jets. Of course events with many energetic particles of different classes are pushed into different streams: hence the same event can appear in different streams at the same time. Usually the analyses which combine more data streams together have to take this into account, to prevent double-counting of such events. In this analysis we analyze data from all the data streams, keeping them well separated.

In a model-independent analysis like this one, the usual notions of signal and control regions must be modified. Being unknown the exact region where new physics may appear — because we do not base the search on any theoretical model — all regions of the data are actually considered and treated as both signal and control. This analysis is not blind, but rather tries to spot, identify and

understand discrepancies between data and the SM MonteCarlo prediction. Good run criteria are imposed to the analysis, considering only the events from runs in the official ATLAS *good run lists*. Runs are flagged as good ones when all major sub-detectors are in good working conditions. SM backgrounds are estimated by generating samples of simulated Monte Carlo events, using the PYTHIA [145], Herwig [126], MC@NLO [185], JIMMY [165] and ALPGEN [139] generators, as explained in Section 6.2.2.

Event classification

Events are taken from all the three ATLAS data streams — `Muon`, `Egamma` and `JetTauEtmiss` — and from the simulated MC data samples reproducing the physics processes of the Standard Model (SM) background. Only very few, loose, event- and object-selection cuts are applied (see sections 6.4 and 6.5 for more details).

The events are then divided into exclusive channels according to their "topology" or "event class", i.e. according to the type and number of high energy objects identified in the event itself. In the following the terms "topology" and "event class" will be used as synonyms.

Channels with a very low expected number of events are merged to channels with the closest related features, which then become "inclusive". Channels are merged only if the significance of the lowest is not changed in the process. This "merging" is a new feature compared to previous model-independent searches, and it has been conceived and implemented to reduce the enormous amount of possible channels to a countable number without loosing discovery potential. The method is described in detail in section 6.6 and 6.7.

Number of channels

The total number of possible topologies N with n considered types of detector objects (e.g. jet, electron, muon, etc.) can be calculated as a sum over the number of combinations, including repetitions, assuming a total number of k objects in the final state [198]. This number represents the number of subsets of k elements from an initial set of n elements), and it can expressed with the equation 6.1.

$$N = \sum_k \binom{n+k-1}{k} = \sum_k \frac{(n+k-1)!}{k!(n-1)!} \tag{6.1}$$

So if we consider, for instance, up to 8 objects (or particles) in final states, the number of possible channels defined by a starting set of 3 objects like electron, muon and jets is 164 (with $n = 3$ and $k = 8$). Such a number of objects, 8, is quite large for SM processes when only leptons are considered, but even a bit too pessimistic if jets are involved; actually even higher jet multiplicities have already been studied in ATLAS at the same centre-of-mass energy.

Considering missing transverse momentum adds an object to the starting set — which is hence now composed by electrons, muons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$— and doubles the number of channels by 2; $E_{\mathrm{T}}^{\mathrm{miss}}$, in fact, can only be considered as a single object, which can be present in the event or not; we cannot consider more than one $E_{\mathrm{T}}^{\mathrm{miss}}$ object in the same event. In the end the number of possible "topologies" would be $2 \times 164 = 328$. At the present status of this model-independent search, only the 3 objects mentioned above plus $E_{\mathrm{T}}^{\mathrm{miss}}$ have been considered. But a logic extension of the analysis would consider also other particles like photon or b-jets. If we consider $n = 5$ objects (electron, muon, photon, jet, b-jet) and only final states with 6 objects, the number of topologies would already increase up to 2×461 channels; and 2×1286 if we consider up to 8 objects in final states. Considering also tau-objects and more object final states brings the number of channels above 10000. A rough count of the number of channels currently studied in the ATLAS SUSY and Exotics groups yields about 100–200 (considering the inclusive channels as well). Although these channels might cover a broad range of likely scenarios for new physics there is still the possibility that the "right" event topology is not studied. Topologies which are not currently studied in ATLAS are, e.g. $\gamma + E_{\mathrm{T}}^{\mathrm{miss}}$ or many-lepton channels. The work presented here aims to search in these more "exotic" channels with a systematic approach.

Search strategy

At the current state, this General Search only considers discrepancies between data and MC predictions in the number of events per topology and in the "Effective Mass" (or M_{eff}) distributions.

The distribution of the number of events per topology (as shown in Section 6.10) is a valuable tool to easily spot large discrepancies. Those discrepancies can arise from new physics, of course, but also from detection or reconstruction faults. In this way the General Search is also a tool to verify if the ATLAS experiment works as expected in a large variety of physics channels. For example if discrepancies were found in many jet channels, one could argue that there were problems in the reconstruction or the simulation of the particle jets. Up until now the ATLAS data showed a very good agreement with the MC expectation, as you can see for instance in figure 6.10, which is an evidence of the very good integration of the different parts of ATLAS in the whole chain from data-taking to the final analysis, through reconstruction and simulation.

The other variable that is searched for new physics in this work is the M_{eff}, which is the scalar sum of transverse momentum of all k particle-objects in the event, plus the missing energy $E_{\mathrm{T}}^{\mathrm{miss}}$; M_{eff} is defined in Equation 6.2.

$$M_{eff} = \sum_{k} |p_{\mathrm{T}k}| + E_{\mathrm{T}}^{\mathrm{miss}} \tag{6.2}$$

This variable is called "Effective Mass" in most ATLAS analyses, since it is sensitive to the mass-scale of new physics; the value of M_{eff} at which the signal exceeds the SM background, in fact, could provide a first estimate of the mass of particles coming from new physics, for example the mass of Supersymmetric particles [54] [146].

A search algorithm — derived from that one used in the H1 analysis (see Section 5.2 on page 66) — has been developed to automatically search the region of the largest discrepancy with respect to the SM expectation, for each event class. The algorithm also computes the significance of such deviation to occur, taking into account the trial-factor, for each channel. In the end the global significance is computed, for all event classes combined together, to give a global measure of the agreement between data and the MC expectation.

The search algorithm and its outcome are presented in more details in Section 6.10.2 on page 107.

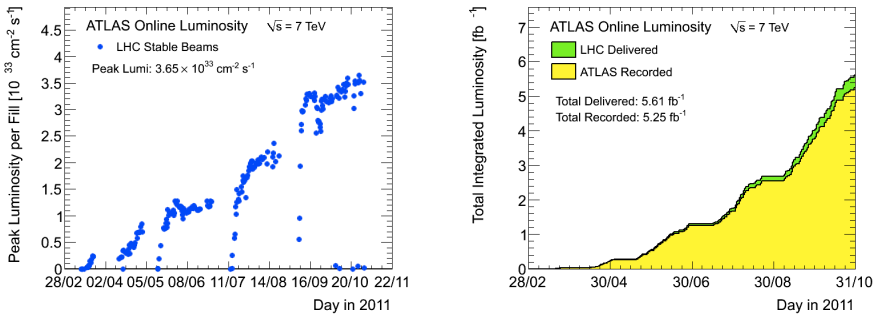
Reducing the trial factors

Looking at many different data streams, channels, distributions, selections and regions makes large statistical fluctuations more likely to occur. In fact the more places one looks at when searching new physics, the highest the probability that one can observe a discrepancy due only to random fluctuations governed by statistics: in those regions the background could fluctuate toward or away the signal, giving raise to a fake significant deviation, which might be misinterpreted as evidence of new phenomena. A dedicated analysis for every single channel would fix the issue, but a detailed analysis of so many channels is also technically difficult and impractical. Hence, broad searches like this one have to reduce their statistical significance by a factor determined by the number of *trials*, the "trial-factor".

To reduce the number of channels, and hence the trial-factor, we developed a "merging" algorithm to combine channels: channels with low probability — very low expected number of events from SM predictions — get merged with channels with similar features but higher probability, when that is possible without loosing information. More details on merging are provided in Section 6.7.

Future extensions

So far, only muons, electrons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ have been considered in this analysis; photons, b-jets and taus could be added in a future version of this General Search. And composite particles like Z and W could be added as well, in the future. These extensions would rise the number of channels a lot, giving a higher trial-factor to be taken into account, but they also open the search to a much broader landscape of possible signals in the search of new physics.



(a) **(b)**

Figure 6.1: **(a)**: Luminosity peak per fill in the whole 2011 dataset; **(b)**: Day-by-day integrated luminosity in the full 2011 data taking period [150].

6.2 Data and MC samples

In this section we outline the dataset and the different MC samples used in this analysis.

The analysis has been performed within the ATLAS framework, release 16, using all three data streams — `Muon`, `Egamma` and `JetTauEtmiss` — and the official MC background predictions from the ATLAS SUSY Group — literally SUSYD3PD filtered data files coming from the official ATLAS SUSY production, with tags `p601` and `p602`.

6.2.1 Data streams and equivalent luminosity

A dataset corresponding to the total integrated luminosity $\int \mathcal{L}\,\mathrm{d}t = 2.098 \pm 0.078$ fb^{-1} has been analyzed in this work. This corresponds to the data taking periods from `B2` to `K6`, produced with the release 16 reprocessing, with a luminosity uncertainty of 3.7% [98]. In Picture 6.1 the Luminosity values for the 2011 data taking period are shown.

A Good-Run-List (GRL) selection is applied to insure that all ATLAS subdetectors are fully functional and efficient at the time of taking the data used in this analysis. The latest SUSY GRL available for the release 16 reprocessing has been used, i.e. version `DetStatus-v28-pro08-07`. After the GRL requirement, the integrated luminosity available for the analysis is reduced to $\int \mathcal{L}\,\mathrm{d}t = 2.052 \pm 0.076$ fb^{-1}.

In this work all three ATLAS data stream are analyzed — `Muon`, `Egamma` and `JetTauEtmiss` — but they have been kept in separate analyses. This separation keeps the analysis robust and easier to be interpreted, and allows cross-checks and validations using channels that appear in more than one stream.

6.2.2 Background samples

The MC samples used in this analysis are coming from the ATLAS official `mc10b` production, with production tag `r2302_r2300`. The Pythia6 [145] and HERWIG [126] generators have been used for showering and GEANT4 [8] for the detector simulation. All the samples are normalized to the most precise theoretical calculations available, and corrected by the right "correction factors" provided by the ATLAS Physics groups. Correction factors can be of two types: theoretical and experimental. Theoretical correction factors reflect the actual difficulty of accurately computing SM physics processes. These factors — also known as "k-factors", or "knowledge factors" — represent the ratio between the not-available all order cross section and the currently computed leading order cross section. Experimental correction factors include the trigger and reconstruction efficiencies, and fake rates associated with a wrong identification of physics objects. All cross-sections and k-factors used in this analysis have been officially provided by the ATLAS SUSY Group.

Events have been generated with a mean number of *pile-up* interactions of $\langle\mu\rangle = 8$. In a proton-proton collider running at high energy, when a bunch crossing occurs the actual interactions take part among the colliding inner partons: quarks and gluons. This leads to events featuring multiple collisions *piling-up*: we observe then multiple vertices which superimpose within the recorded event, tending to hide the primary vertex of the highest-energy parton collision. The number of pile-up collisions grows with the beam energy. To take into account the difference between the MC simulated events and the real pile-up conditions observed in data, a standard ATLAS re-weighting procedure is applied to the MC samples, matching the average number of interactions per bunch crossing, averaged over bunch crossing ID, to the value in data [158]. A systematic uncertainty related to this pile-up effect can be determined by comparing nominal reweighting with a shift in the weight scale of 10%.

This General Search aims at detecting deviations in all kind of channels with high-p_{T} objects, so all relevant MC samples have been taken as background. The following backgrounds have been considered:

- **QCD**
  Multi-jet events are simulated with the PYTHIA event generator [145] with two-parton matrix elements. All additional jets are generated by parton showering. A muon-filtered PYTHIA production (requiring at least one muon) is used as background when analysing the Muon stream — instead of the common unfiltered `QCD` production — in order to increase the global statistics for muon events.

- **$t\bar{t}$ (semi-leptonic and full hadronic) and single top**
  Events are generated — assuming a top mass of 172.5 GeV— with the

MC@NLO generator [185], with a next-to-leading-order (NLO) precision correction to the matrix element for the hard processes. The generated sample is then interfaced to HERWIG [126] for parton showering and JIMMY [165] for the simulation of the underlying event.

– **Electroweak gauge bosons: Z/γ and $W + jets$**
Electroweak gauge boson events, in association with jets, are generated using the ALPGEN [139] generator. This includes leading order (LO) QCD and electroweak contributions for the multi-partons hard processes. The considered processes are $W(\rightarrow l\nu) + jets$, $Z(\rightarrow l^+l^-) + jets$ and $Z(\rightarrow \nu\nu)+jets$; the jets can be both light- and heavy-flavoured, and are generated with ALPGEN v2.13. Additional samples generated with PYTHIA[145] have been used to cover the low mass Drell-Yan production ($m_{ll} < 40GeV$). The generator is interfaced to HERWIG for showering and fragmentation, and to JIMMY[118][165] to simulate the underlying event[142][78].

– **Diboson**
WW, WZ and ZZ events are simulated using the HERWIG generator.

Some samples of BSM[1] signatures have been used as benchmark points to compare the sensitivity of the General Search to that of specific dedicated analyses (see Picture 6.14(a)).

A full list of all the background samples used in this work is given in Appendix 6.A.

6.3 Systematic uncertainties on the background

The theoretical and experimental systematic uncertainties considered for the various background processes are described in the following. All theoretical systematic uncertainties are added in quadrature to the statistical uncertainty.

Experimental uncertainties

For the luminosity measurements we assign an uncertainty of 3.7%; the value is taken from the official ATLAS measurement [98]. Out of the various systematic uncertainties affecting the shape of the different backgrounds only the Jet Energy Scale (JES) has been considered so far, being the largest uncertainty for most classes. The JES uncertainty — officially provided by the ATLAS Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$Combined Performance Group — is based on MonteCarlo studies and *in-situ* measurements and is evaluated by varying the scale in the MC samples up and down by one sigma [112][152].

[1] Acronym for *Physics beyond the Standard Model*.

Theoretical uncertainties

All MonteCarlo samples used in this work have been normalized to their most accurate inclusive cross sections. In this analysis only MonteCarlo expectations are taken into account to estimate the SM background, making a sensible choice for these errors critical. The following theoretical uncertainties have been assigned per type of process:

- **QCD**: multi-jet production is the most unreliable simulated process. However, due to the high-p_{T} threshold requested on particles in the object selection, it is of little importance for most of the event classes considered, except for the all hadronic classes. The multi-jet process is only available at the leading-order, and it has been assigned a systematic uncertainty of 50%.

- ***W/Z+jets***: the inclusive W/Z cross section is known at NNLO level, with an uncertainty of $\approx 5\%$; that uncertainty has been used for the samples with no jets, and an additional 20% uncertainty was added for each additional parton-jet in the final state;

- **ttbar and single top**: $t\bar{t}$ and single top cross sections are known to NNLO and NLO level respectively, and an uncertainty of 10% has been suggested by the ATLAS Top group;

- **Diboson**: di-boson processes are known at NLO level; an uncertainty of 10% has been used for such processes.

6.4 Triggers and pre-selection of events

The "Trigger", as already explained in Section 4.6.1, is the system which lets the physicists select the interesting events from proton-proton collisions. The first levels of the system — the so-called "L1" and "L2", see Picture 4.13 — are on-line filters, which means they are able to select the interesting events as they arrive from the detectors, before the events are written to the disks. The selection of the events is made on a "signature", that is on a set of physics objects with given features. Only the events that "pass the trigger" are retained, and written to disk. Once written to disk, the events are flagged with the triggers whose requests they satisfied.

For this analysis the choice was made to use events who passed single-object triggers, which are based on the characteristics of a single object in the event: a muon or an electron, for example. The single-object triggers very often are based on quite high thresholds on the particle energy, because of the high rate of events satisfying such a signature; using multi-objects triggers could have lowered

the threshold on the energy of the particles. But the choice to use only single-object triggers — when available — has been taken to impose as few constraints as possible to the event selection, in order to introduce the minimal amount of biases — or precondition choices — in this model-independent search.

Triggers can be *prescaled*, that means that, due to the very high rate of the given signature, only one event out of a given number is analyzed and retained. For this work, the choice to only use *unprescaled* triggers has been taken, to not lose statistics and to be able to interpret the results in a more robust way.

The list of the triggers used in this work follows. As one can see, pure single-object unprescaled triggers have been used for muon- and electron-events, for the whole data period; while more complex signatures had to be chosen for the jet-events: due to the very high rate of jet-events, in fact, those listed here below are the lowest-p_{T} unprescaled single-jet triggers available for the data-taking periods considered in this work.

The trigger signatures used in this analysis are listed below. As one can see, only triggers of the `EF` type have been used; that means that the trigger chains come from the last stage of event selection, the "Event Filter" (see Section 4.6.1 for more details). For events in the *Jet+Tau+*$E_{\mathrm{T}}^{\mathrm{miss}}$ data stream we use two combinations of triggers, each composed of an `OR` boolean relation between two jets triggers: the first one is a single-jet trigger, which selects events with at least one single jet whose $p_{\mathrm{T}} > 180$ GeV (`j180`); this `EF` trigger is seeded by the `L1_J75` and the `L2_j95` triggers at the `L1` and `L2` filtering steps, selecting jets with $p_{\mathrm{T}} > 75$ GeV and $p_{\mathrm{T}} > 95$ GeV respectively [153].

– **Jet+Tau+$E_{\mathrm{T}}^{\mathrm{miss}}$stream**:

  data period $<D$: `EF_j180_a4tc_EFFS || EF_j75_a4_EFFS_xe45_loose_noMu`

  data period $\geq D$: `EF_j180_a4tc_EFFS || EF_j75_a4tc_EFFS_xe45_loose_noMu`

– **Muon stream**: `EF_mu18_medium`

– **Electron+Photon stream**: `EF_e22_medium`

Due to the high rate of single-jet events (i.e. the probability this type of event occurs), the threshold on the jet p_{T} is quite high; for this reason another signature has been associated to the first one, with a lower p_{T} threshold on the leading jet, even if that means an additional constraint on other objects: the second signature selects events with at least one single jet whose $p_{\mathrm{T}} > 75$ GeV (`j75`), with a $E_{\mathrm{T}}^{\mathrm{miss}} > 45$ GeV (`xe45`) and with no muons (`noMu`). All the four inclusive jet triggers here above make use of the full-scan reconstruction algorithm (`EFFS` stands for "Event Filter Full Scan") from *topological clusters* (`tc`) in combination with the "anti-k_t" algorithm with a cone size $R = 0.4$ (`a4`) that is used to identify

the particle jets [113, 183]. More details on jet reconstruction can be found in Section 4.7.1.

For events in the *Muon* and *Electron+Photon* stream we use pure single-object triggers: that one selecting events with at least one isolated muon of $p_{\mathrm{T}} > 18$ GeV (`mu18`) in the first case; and a trigger selecting events with at least one isolated electron with $p_{\mathrm{T}} > 22$ GeV (`e22`) in the second case.

In order to speed up the analysis and to be sure to stay in the trigger *plateau* region — the region with constant, maximum, acceptance efficiency —, a pre-selection has been applied to the events, before any object- or event-selection. The requirements of this pre-selection, according to data streams, are in the list that follows. The threshold values have been chosen in accordance with the respective trigger *plateau* value, given by the ATLAS Trigger Performance group.

- **Jet+Tau+$E_{\mathrm{T}}^{\mathrm{miss}}$stream** : at least one leading jet with $p_{\mathrm{T}} \geq 300$ GeV
- **Muon stream**: at least one leading muon with $p_{\mathrm{T}} \geq 20$ GeV
- **Electron+Photon stream**: at least one leading electron with $p_{\mathrm{T}} \geq 25$ GeV

6.5 Event Selection and Object Definitions

Event selection

Several *cleaning cuts* have been applied to ensure to keep only true and well reconstructed events. Such cuts are listed here below:

1. selected events must have at least one primary vertex with more than 4 reconstructed tracks;
2. removal of events with (potentially) misreconstructed jets[2];
3. cosmic muons rejection;
4. *delta-phi cut* over the 4 leading jets, where *delta-phi* ($\Delta\phi$) is the distance in the angular coordinate ϕ, between a single jet and the $E_{\mathrm{T}}^{\mathrm{miss}}$. Being the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the jets the most difficult and critical objects to reconstruct, we cannot be enough confident in the existence of two such objects if their separation in ϕ angle is below a given threshold. In this analysis a cut of 0.1 is applied, following the suggestions of the ATLAS SUSY Group.

[2]The *Bad Jet* cut after *overlap removal* cut has been applied, and a *Lar Hole* veto of events; moreover, events flagged with *Lar error* (see page 92) have been removed. For a definition of *overlap removal*, see definition at page 92.

Jet Definition	
Cut	*Requirement*
Algorithm	AntiKtTopo
Acceptance	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.8$
JVF	$JVF > 0.75$
Overlap Removal	$\Delta R(jet, e) > 0.2$
Quality Cuts	rejects events with loose bad reconstructed jets

Table 6.1: Jet object definition.

The following corrections are also applied to take into account reconstruction discrepancies between real data and MonteCarlo simulations:

- *use of out-of-time pileup MC samples*: events from subsequent bunch crossings have been simulated in the MC samples, to match the data pile-up experimental conditions;
- *pileup re-weighting*: MC samples have been re-weighted to match the pile-up conditions of data, as explained in Section 6.2.2;
- *lepton reconstruction discrepancy re-weighting*: MC samples have been re-weighted to simulate the fake rate of lepton reconstruction in data.

Object definitions: Jets, Electrons and Muons

Object selection, or *object definition*, cuts assure the final particles are well reconstructed and reliable. These cuts have been chosen to extract the maximum of information from data while correcting detector deficiencies, following the recommendations from the ATLAS Detector and Performance groups.

These cuts are usually tailored for a particular analysis or model but, given the large number of topologies which this model-independent analysis aims to look at, a choice had to be done to choose a set of cuts which could properly perform on most of the channels, without the possibility of testing them in term of the performance on a particular channel. Hence, after considering that we focus mainly on high-p_{T} objects and large $E_{\mathrm{T}}^{\mathrm{miss}}$ — as explained in Section 6 — the same object definition cuts of the ATLAS 1-lepton SUSY search (See [90]) have been adopted.

The object definitions for jets, electrons and muons can be found in Tables 6.1, 6.2 and 6.3.

Jets are reconstructed using the `AntiKt4Topo` algorithm (more details can be found in sections 4.7.1 and 6.4) ; they must have a $p_{\mathrm{T}} > 20$ GeV and be in the central calorimeter fiducial volume, which means that their η angular coordinate

Electron Definition	
Cut	*Requirement*
Algorithm	AuthorElectron
Acceptance	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta^{cluster}\rvert < 2.47$
Quality Cuts	`medium` and `tight` flags
Overlap Removal	$0.2 < \Delta R(e, jet) < 0.4$
Track	with match

Table 6.2: Electron object definition.

Muon Definition	
Cut	*Requirement*
Algorithm	`STACO`, segment-tagged or combined
Acceptance	$p_{\mathrm{T}} > 10$ GeV, $\lvert\eta\rvert < 2.4$
Quality	`Loose`
Overlap Removal	$\Delta R(\mu, jet) > 0.4$
Quality Cuts	track quality

Table 6.3: Muon object definition.

should be in the range: $|\eta| < 2.8$. The requirement of a Jet Vertex Fraction (JVF) greater than 0.75 — where the JVF for a jet is defined as the p_{T} of the tracks coming from a given vertex over the p_{T} of all tracks — has been applied to all selected jets to reduce contamination from pile-up [159].

Electrons must have at least $p_{\mathrm{T}} > 20$ GeV and be within the acceptance region of $|\eta| < 2.47$; they also have to satisfy the `medium` isolation requirements of the electron reconstruction algorithm.

Muons must be either combined- or segment-tagged by the `STACO` reconstruction algorithm (already introduced in Section 4.7.3); their transverse momentum must be at least $p_{\mathrm{T}} > 10$ GeV and they have to be within the acceptance region of $|\eta| < 2.4$

Missing Transverse Energy

The missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is a very difficult quantity to measure, requiring the input and tuning of all other important physics objects: muons, electrons, jets, photons, taus and any other remaining energy in the calorimeter that is not associated with other physics objects.

In this analysis the "`SimplifiedRefFinal`" definition of $E_{\mathrm{T}}^{\mathrm{miss}}$ is used, which is equal to the modulus of the vector sum of the transverse momenta of all

the reconstructed objects, with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.9$, in the event: jets calibrated at the EM+JES scale, signal leptons, additional non-isolated muons and calorimeter clusters not associated with any object (the so-called `CellOut` term).

6.5.1 Overlap Removal

A set of *overlap removal* rules has also been applied after the object selection, to cope with the possible ambiguities in the object definition.

In this case, with *overlap removal*, one means the ability of separate electrons from jets. Since all electrons leave an energy deposit in the inner calorimeters, like the jets, at first all electrons are flagged as *jet-like* objects. Then electrons are identified in further details at later steps of the particle reconstruction and listed in the electron collection for that event. At the moment of analysing the event, all electrons are listed both in the electron collection and in the *jet-like* object collection. Thus, after selecting the good electrons with the *object selection* cuts, in order to separate real from fake electrons, one needs to loop over all electrons and compare their eta and phi angles with those of all jets in the jet collection: if the separation between the two objects in eta and phi is below a given threshold — which depends on the reconstruction accuracy and which is given by the Performance Groups of the experiment —, the jet is rejected because it is actually an electron.

In this analysis a cut of $dR(jet, electron) < 0.2$ between electrons and jets is applied: that means that any jet within a cone of $\Delta R = 0.2$ of a good electron candidate is discarded. After that, a similar cut is also applied on a second pass, between the remaining leptons — which means muons in this analysis — and jets: leptons lying within $\Delta R = 0.4$ from a jet are discarded, keeping the jet.

6.5.2 The LAr hole problem and its treatment

The label "LAr hole problem" refers to an accident occurred on April 30th 2011, on a crate of the front-end electronics, which leaded to a loss of the 0.4% of the Liquid Argon (LAr) calorimeter cells [90, 106]. The problematic region is located at $0.0 < \eta < 1.4$ and at $-0.74 < \phi < -0.64$, which defines a *jet dead region* as defined here below, which affects only jets falling at a distance $\Delta R < 0.1$ from the dead region, i.e.:

$$-0.1 < \eta < 1.5$$
$$-0.9 < \phi < -0.5$$

The jet dead region is defined larger than the actual dead region, to account for the jet resolution.

The hole in the LAr calorimeter produces events with fake or wrong missing transverse energy, because the loss of energy in the dead material affects the jet reconstruction, and so it affects the jet energy response and its resolution. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is a very important quantity for many new physics models — it represents, in fact, the amount of energy which is potentially taken away by neutral particles like neutrinos, neutralinos, and so forth — a detailed study on the impact of the LAr hole has been performed by several Physics and Performance groups within the experiment.

To avoid simply applying a veto which rejects events if an object hits the dead region — which would lead to a too large event rejection — a *smart* event veto was developed in the context of the 0-lepton SUSY ATLAS analysis [89]. This rejects bad events, assuring the good agreement between data and the MonteCarlo simulation, which does not include the LAr hole problem. Studies [90] have been proved that the LAr hole has a small impact on the 1-lepton analysis, which the object selection of this work is based upon, as stated above. But nevertheless, as this general search looks at many topologies, with and without leptons, the same smart event veto is applied in this analysis.

In brief, events are rejected if these conditions are simultaneously satisfied (more details can be found in [89]):

- any jet in it points in the direction of the LAr Hole dead region and if it has $p_{\mathrm{T}} > 20$ GeV
- the jet contribution to the global $E_{\mathrm{T}}^{\mathrm{miss}}$ is large enough; which means:

$$\Delta E_{\mathrm{T}}^{\mathrm{miss}}(jet) > 10 \text{ GeV} \quad \texttt{OR} \quad \frac{\Delta E_{\mathrm{T}}^{\mathrm{miss}}(jet)}{E_{\mathrm{T}}^{\mathrm{miss}}} > 0.1$$

where $\Delta E_{\mathrm{T}}^{\mathrm{miss}}(jet)$ is the impact of the energy contribution of the jet in the dead region, on the global $E_{\mathrm{T}}^{\mathrm{miss}}$ of the whole event.

6.6 Topological classification through Dynamic Topology

Once the events are selected, their content is analyzed. At this step, all events are gathered in different classes, or *topologies*.

A *topology* label is dynamically assigned to events, according to the final states. For example, starting from the electron collection, an "`e`" flag is given to the event if the first electron in the collection satisfy some given criteria; an additional "`e`" flag is given for a second "good" electron and so on. The same for the other particles in the event. All the flags and thresholds used in this analysis so far, are listed here below:

- "`v`": if there is an amount of $E_{\mathrm{T}}^{\mathrm{miss}} \geq 125$ GeV

- "`m`": for each selected *muon* with $p_{\mathrm{T}} \geq 20$ GeV
- "`e`": for each selected *electron* with $p_{\mathrm{T}} \geq 25$ GeV
- "`j`": for each selected *jet* with:
  - $p_{\mathrm{T}} \geq [300, 40, 40, \dots]$, if we run on `JetTauEtmiss` data stream
  - $p_{\mathrm{T}} \geq [60, 25, 25, \dots]$, if we run on `Muon` and `Egamma` (Electron+Photon) data streams, or if leptons are present in the topology

As detailed in the above list, two different jet-energy thresholds are applied, according to the nature of the data stream. When the `JetTauEtmiss` stream is analyzed, the "`j`" flag is given to the leading jet if its $p_{\mathrm{T}} \geq 300$ GeV, and to the second and all subsequent jets if their $p_{\mathrm{T}} \geq 40$ GeV. When `Egamma` and `Muon` streams are analyzed, thresholds are lowered to 60 and 25 for the leading and the subsequent jets respectively, in order to take into accounts the lower acceptance cuts of those data streams.

As a first step, leptons are also distinguished according to the charge: a "`m`" is given to a muon with negative charge and "`M`" to a muon with positive charge; and the same for the electrons: "`e`" for the electron and "`E`" for the positron. Charge information is not used in the final analysis at the moment, as it would lead to a factorial growth of the event classes; but it is used for control, debugging and cross-checks during the analysis process; other ways to exploit it will be studied in the future development of the analysis.

The large value of the E_{T} cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ (125 GeV) is imposed by the trigger choice: as explained in Section 6.4 the threshold of 125 GeV is necessary to stay in the *plateau* region, where the trigger efficiency is ≈ 1.

Below, two examples can be found of this dynamic flagging of events based on their particle content. As a first example, one could consider an event featuring the physics objects listed below (the objects are assumed to have passed the object selection, already):

- missingEt $= 140 GeV$
- 2 muons: $p_{\mathrm{T}1} = 32$ GeV, $p_{\mathrm{T}2} = 23$ GeV
- 1 electron: $p_{\mathrm{T}} = 27$ GeV
- 3 jets: $p_{\mathrm{T}1} = 62$ GeV, $p_{\mathrm{T}2} = 45$ GeV, $p_{\mathrm{T}3} = 22$ GeV

According to the cuts presented above, all particles but the last jet would pass the *topology* cuts; the last jet, whose $p_{\mathrm{T}} = 22$ GeV, is not flagged as a good jet candidate for this topology study. The event hence would be given a *dynamic topology* flag of "`vmmejj`".

A second example event. The particle content, after object selection, is listed here below:

– 3 muons: $p_{\mathrm{T}1} = 32$ GeV, $p_{\mathrm{T}2} = 23$ GeV, $p_{\mathrm{T}3} = 6$ GeV

– 3 electron: $p_{\mathrm{T}1} = 30$ GeV, $p_{\mathrm{T}2} = 26$ GeV, $p_{\mathrm{T}3} = 8$ GeV

– 2 jets: $p_{\mathrm{T}1} = 70$ GeV, $pt_2 = 22$ GeV

In this case all the particles with the lowest energy do not pass the topology cuts, and hence the event would be flagged as "`mmeej`".

6.7 The "Merging Channels" algorithm

As already explained in the previous sections, this model-independent General Search aims to look for discrepancies in the largest possible number of topologies. If that assures the largest coverage on the new physics phase space on one hand, on the other hand it also leaves us with a potentially very large number of channels, which makes the statistical interpretation of the analysis more difficult. Moreover the largest the number of topologies, the largest the statistical "Look-Elsewhere" effect affecting the analysis, as explained in more details in Section 6.10.2 on page 107.

So one had to consider this double aspect. The largest number of covered topologies is needed, to be able to spot new physics looking at topologies which were not already considered; but one also needs to have the smallest possible number of channels, in order to reduce the statistical uncertainty and its effects. Hence there was a need for an algorithm to reduce the number of topologies, without loosing the information they carry.

Considering this, a merging algorithm with a double threshold has been developed for this analysis; the algorithm's goal is merging topologies together, without burying interesting and rare signatures under common, and so not-interesting, topologies. With this merging action, the potential infinite number of channels is lowered to a smaller definite set of channels, without loosing potential interesting information.

The merging makes use of the background expectation of topologies, i.e. the number of events of a certain class observed in the simulated samples. This approach has several advantages: the number of starting topologies is the smallest possible; and the interpretation is easier and more robust due to a smaller "Look-Elsewhere" effect.

When a new event topology is found, the algorithm tries to merge it based on the background topologies that are present. If the merging fails, the new topology is kept *as-is*, in order to not loose potential new physics evidences. Details follow.

Let one say that only the channels listed here below have been found in the background (they have been chosen without any special reason, only for the purpose of explanation, and the values are not real):

– `m`: SM expectation: 10^3

– `2mj`: SM expectation: 10

– `2m2j`: SM expectation: 0.002

– `2m3j`: SM expectation: 10^{-5}

– `2m4j`: SM expectation: 10^{-9}

As stated above, the algorithm tries to merge topologies, to reduce the number of channels to analyze; but also avoiding to loose information by hiding interesting channels under common ones. Hence, obviously, in this example, the algorithm cannot merge all channels under the, let's say, "minimalistic" `m` channel. This would be very bad, because all the information about the other channels would be missed.

For that reason, the "Merging Channels" algorithm is based on two thresholds:

– `threshold`: 0.001

– `threshold_up`: 0.05

First of all a threshold of 0.001 has been set, which is the minimum SM expectation value for a topology to not be merged; all topologies whose SM expectation is under that threshold will be considered for merging.

To try to merge the topologies, a particle is eliminated, following a certain order, and the SM expectation for that new topology, if any, is considered. The particles are taken out from the topology in the following order: first jets, then electrons, then muons and then $E_{\mathrm{T}}^{\mathrm{miss}}$, following an order based on the importance of a particle in a topology: the reconstruction of particle-jets is less accurate than electrons, and, moreover, the number of jets is usually larger in comparison with the other particles; for these reasons jets are the first objects considered to be taken out. In the same way an electron is considered "less important" than a muon in a topology. $E_{\mathrm{T}}^{\mathrm{miss}}$ is the last object to be removed since it represents the amount of energy taken away by neutrinos and by potential new particles (like, e.g., supersymmetric neutralinos).

Back to the above example, the "`2m4j`" has a SM expectation of 10^{-9}, which could be the background expectation of an interesting deviation from data $> 5\sigma$: the topology could be an interesting one, carrying a signal from new physics. But its SM expectation is under threshold, because $10^{-9} < 0.001$, hence the algorithm would try to merge it, removing one objects after the other — always starting with jets — to look for a good "host" channel accepting the channel to be merged.

A first jet from the `2m4j` is removed, ending up with the topology `2me3j`. The `2m3j` channel is already present in the example channel list, and due to its SM

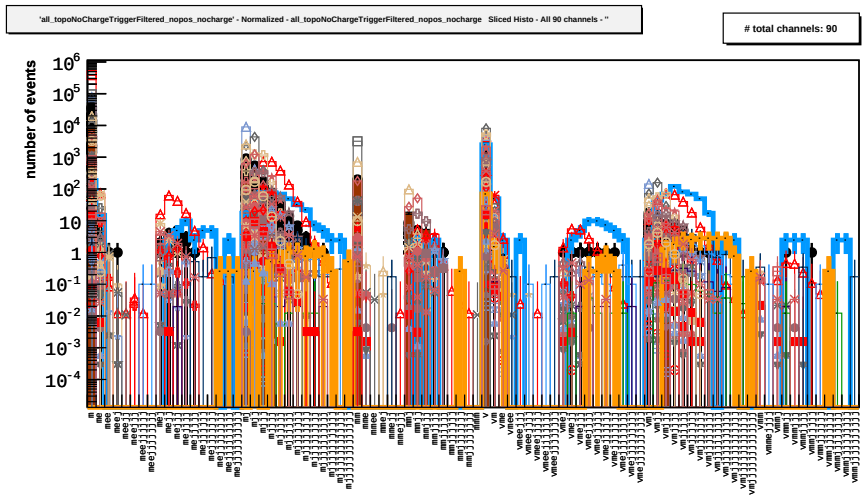
expectation of 10^{-5}, is also an interesting channel, but its SM expectation is lower than the threshold; the algorithm would try to remove one more jet, ending up with the `2m2j`. The `2m2j` channel is still interesting, because its SM expectation is 0.002, and it is a good "host channel" because its expectation is larger than `threshold`. Before merging, in order to try to reduce the number of channels at the maximum, the algorithm would try to remove another jet, to see if another topology could host the channel as well; and it would end up with the `2mj`. `2mj` channel's SM expectation is 10, which is greater than the value of `threshold`, and hence it could be a good host channel for merging; but its expectation is larger than the value of `threshold_up` as well, and that prevents the channel `2mj` to be used as hosting channel: the `2mj` topology is too "crowded" to accept a rare, interesting event; if merged into the `2mj`, the starting channel `2m4j` would be hidden under it, loosing the information it carries. At this step the algorithm would stop the iteration and the `2m4j` topology would be merged with the class `2m2j`.

Channels get merged into channels with larger expectation, but which are not too crowded; in this way the information brought by them is not lost and the potential deviation can still be flagged as "interesting".

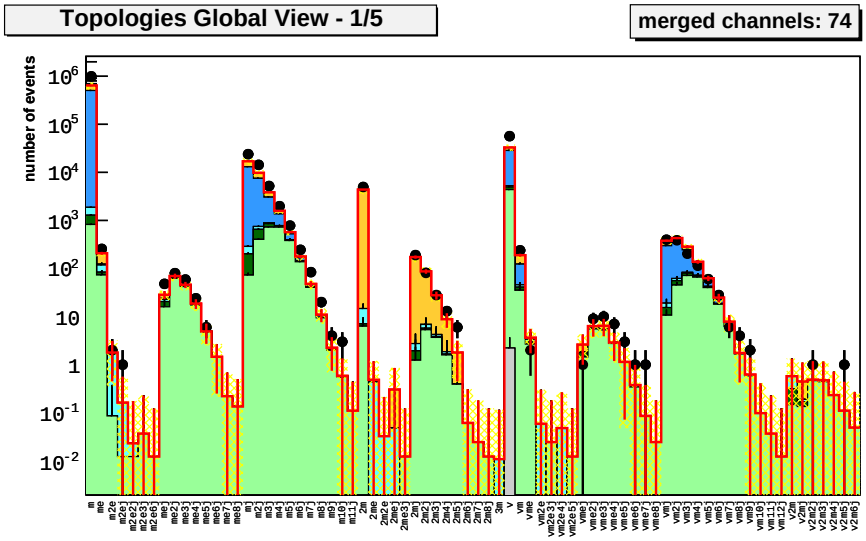
In case the Standard Model expectation of the `2m2j` channel were larger than the `threshold_Up` value as well, or no suitable topologies were found in the SM background, then the starting channel `2m4j` would have been kept *as-is*, to not loose the information it carries.

After removing all "objects" from the topology string, if no topology is found able to accept the channel or to stop the merge operation, the starting channel goes to the "empty channel". This is to be avoided, because that causes the loss of information about the starting channel. For that reason in this analysis the "empty channel" is monitored and cross-checked, to be sure it does not get populated.

After merging two channels the number of events of the starting channel is added to the number of events of the host channel. In the end a new histogram is built with the distribution of the merged channels. In Figure 6.2 on page 98, one can see an example of the raw plot of the distribution of all the channels, before and after the merging. The number of channels lowered of about $\approx 20\%$, from 90 to 74. This is the channel reduction which has been observed in this analysis with the amount of data analyzed up until now ($\approx 2\text{fb}^{-1}$). Of course the amount of reduction could be much larger when the amount of analyzed data will increase, due to the larger number — potentially *infinite* — of not-conventional events one could find in data.



(a) dynamic topology *distribution before merging*



(b) dynamic topology *distribution after merging*

Figure 6.2: Data and simulated events of the `Muon` stream divided according to their *dynamic topology* content: **(a)** before merging channels; **(b)** after merging channels. In this example the merging channels algorithm lowers the number of channels of about $\approx 20\%$, from 90 to 74.

6.8 Comparison with the official ATLAS 1-lepton analysis

In order to validate the event selection, the object selection and the *dynamic topology* flagging mechanism, a comparison has been performed between the cut-flow of channels flagged as "1-lepton" in this analysis, and the official results of the ATLAS approved SUSY standard 1-lepton analysis, over the first 662.16

$fb^{-1}$of the year 2011 ATLAS data.

The selection cuts are summarized here below:

- ATLAS standard SUSY 1-lepton analysis in 2 jet signal region:
  - at least 2 selected jets with $p_{\mathrm{T}} \geq 60, 25$ GeV;
  - $E_{\mathrm{T}}^{\mathrm{miss}} \geq 125$ GeV;
  - exactly one selected `tight` isolated lepton: electron with $p_{\mathrm{T}} \geq 25$ GeV XOR muon with $p_{\mathrm{T}} \geq 20$ GeV.
- General Search:
  - we asked for topologies of the type MET + lep + Njets where $N \geq 2$,
  - $E_{\mathrm{T}}^{\mathrm{miss}} \geq 125$ GeV;
  - $electron > 25$ GeV;
  - $muon > 20$ GeV;
  - $jet > 60, 25, 25, ...$ GeV.

The two analyses lead to the selection of exactly the same particles and the same events: a sign that the strategy and implementation of this analysis is compatible with the ATLAS standards.

6.9 Control distributions

In order to validate the analysis, control distributions have been defined from channels used as probe. These distributions allow to check if the MonteCarlo samples used in the analysis correctly describe the Standard Model background. And they also allow to check the correctness and the efficiency of the event and object selections used in this work.

In the figures of this section, some control plots from some of the candle processes will be shown, mainly from Z and W boson production.

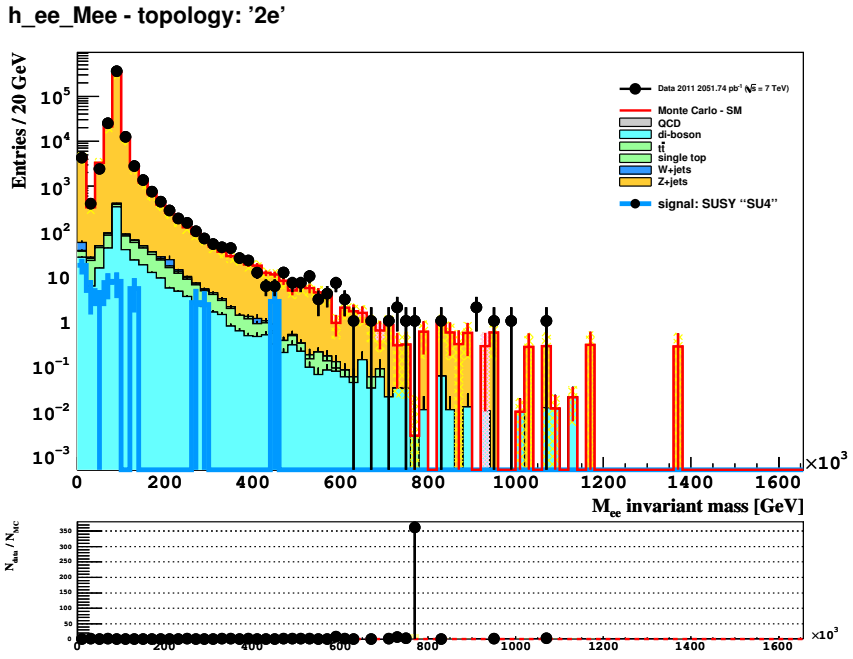
Z+jets production In Figures 6.3 and 6.4 the distributions of the invariant mass of the di-lepton pairs are shown. More in details, the Figure 6.3 shows the invariant mass distribution of channels from `2e` up to `2e4j`. One can see that the distribution of the background well describes the physics process, and the very good agreement with data, within uncertainties, up to 4 jets. The Figure 6.4, instead, shows the very good agreement, within uncertainties, between data and the SM background for the muon channels, from `2m` up to `2m2j`. In addition, Figures 6.5 and 6.6 show the distribution of the M_{eff} (see Equation 6.2 for a definition) of the di-lepton pairs, up to

3 jets; also in these distributions one can see the good agreement, within uncertainties, of data and the SM background.

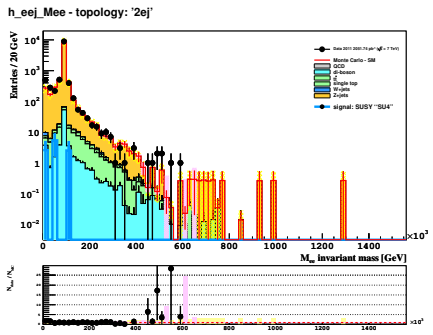
In figures 6.5 and 6.6 the distribution of M_{eff} is shown for the electron and muon channels, respectively. In this case too, data are in good agreement with the SM background, within uncertainties.

W+jets production In Figures 6.7 and 6.8 the transverse mass of the lepton-neutrino system is shown, in channels with a lepton, $E_{\mathrm{T}}^{\mathrm{miss}}$ and up to 4 jets. These distributions are used to control the $W + jets$ samples. One can see the good agreement, within uncertainties, in all distributions.

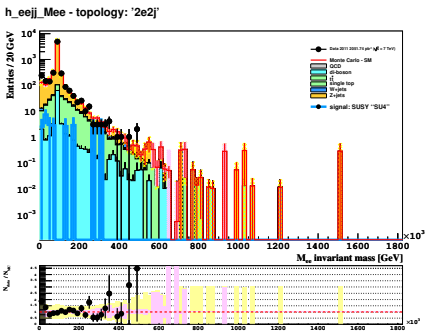
ZZ production Figure 6.9 shows the M_{eff} distribution for the **4e** event class; the plot shows the good description of the background of the di-boson MC sample used in this work, and checks the sensitivity of this analysis to di-boson events.



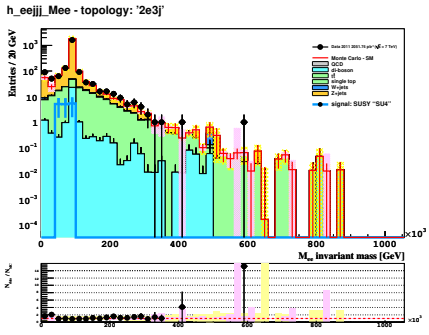
(a) *Di-electron invariant mass for the channel* **2e**



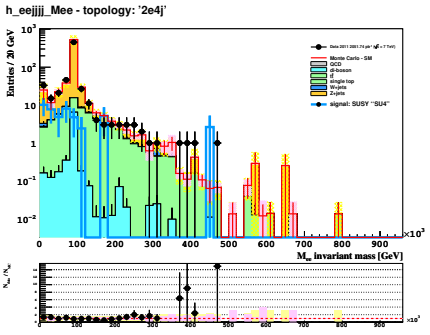
(b) *2ej*



(c) *2e2j*

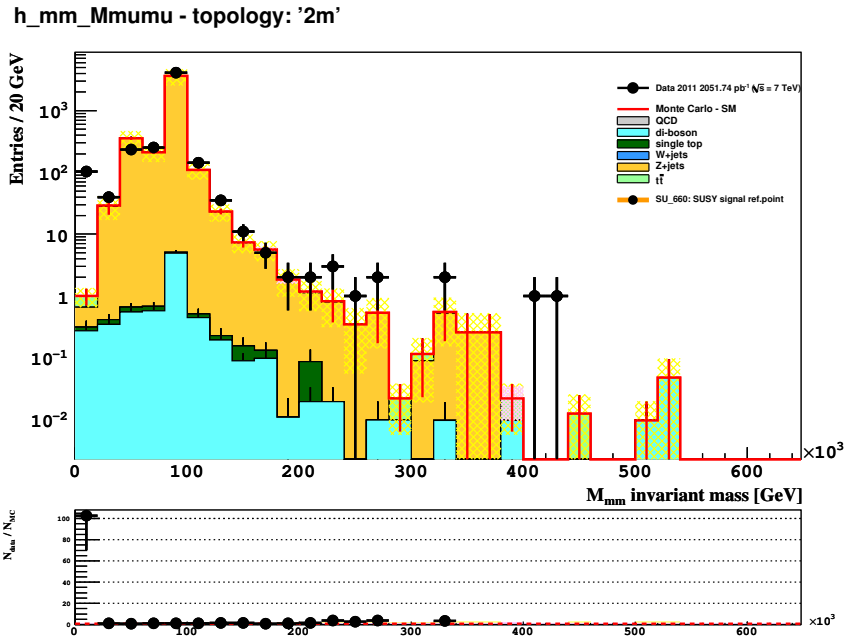


(d) *2e3j*

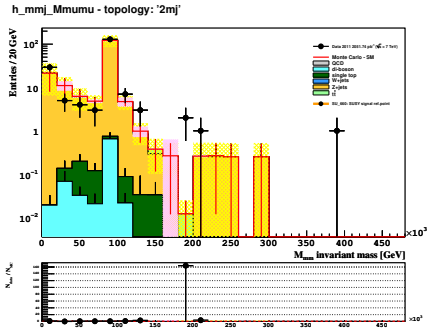


(e) *2e4j*

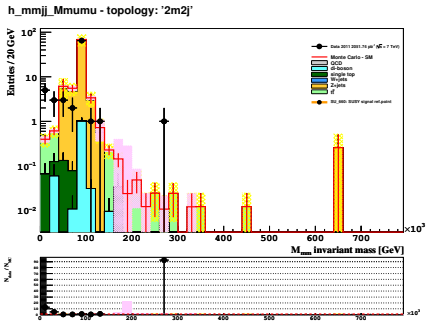
Figure 6.3: Di-electron invariant mass. The invariant mass is computed with only the first two electrons in the event. (a): channel **2e**; (b): channel **2ej**; (c): channel **2e2j**; (d): channel **2e3j**; (e): channel **2e4j**. The plots show a very good agreement, within uncertainties, between data and the SM background.



(a) *Di-muon invariant mass for the channel* ***2m***

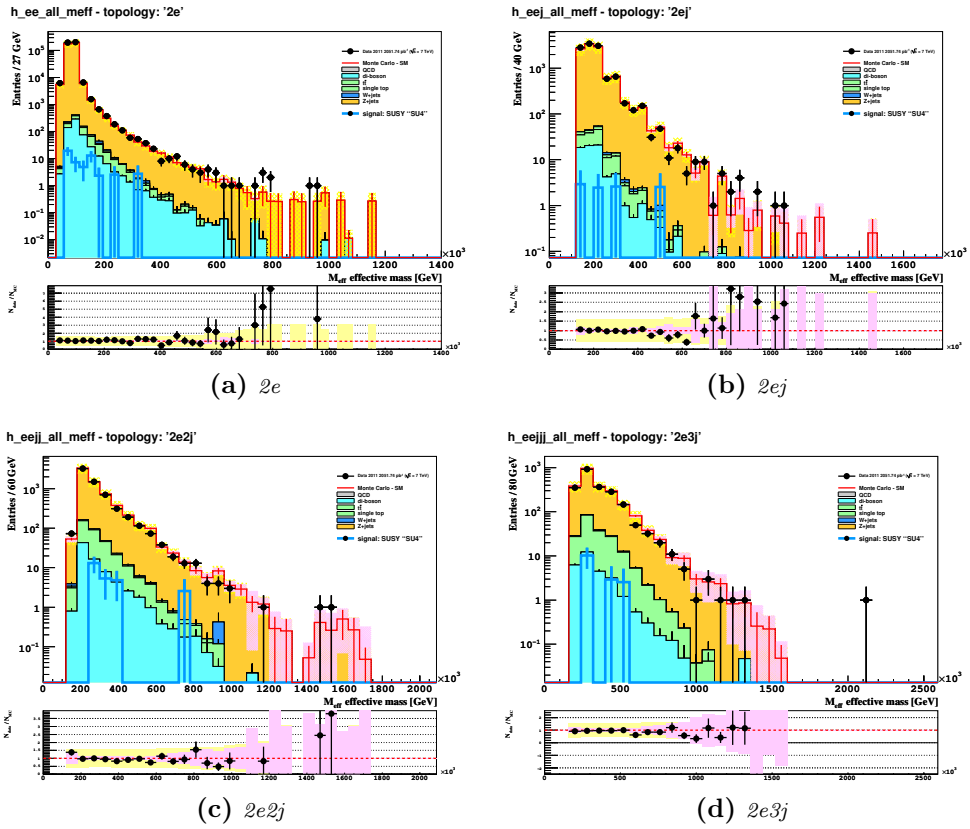


(b) *2mj*



(c) *2m2j*

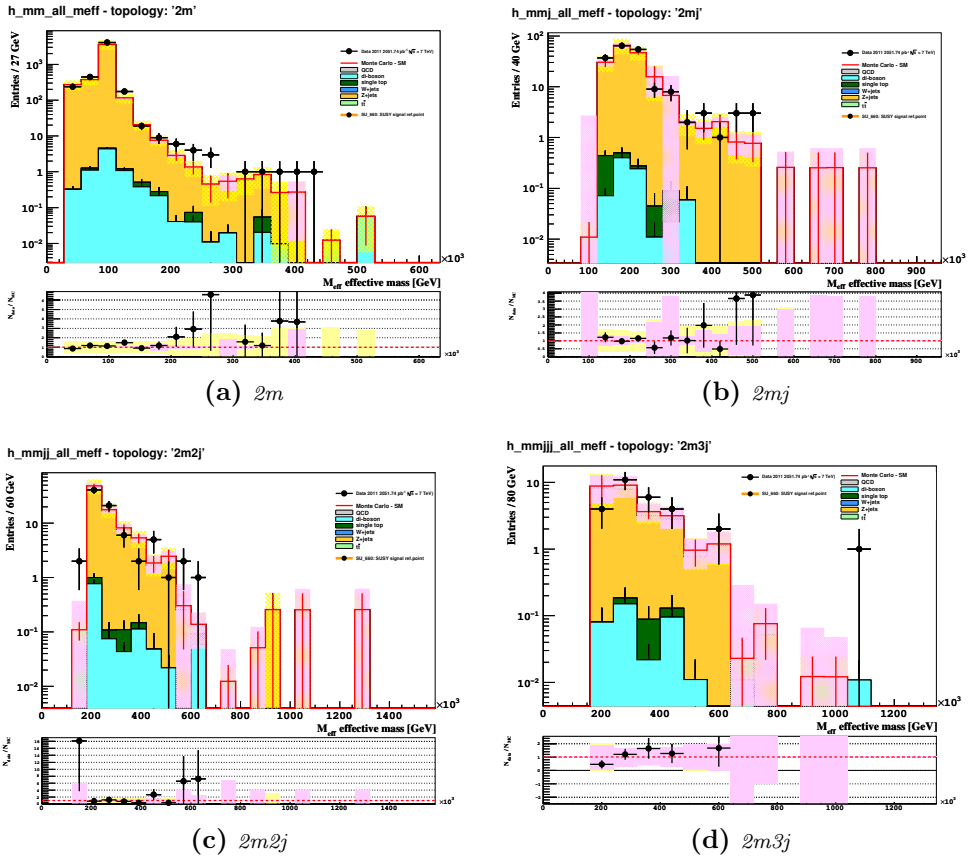
Figure 6.4: Di-muon invariant mass. The invariant mass is computed with only the first two muons in the event. **(a)**: channel **2m**; **(b)**: channel **2mj**; **(c)**: channel **2m2j**. The plots show a good agreement, within uncertainties, between data and the SM background.



(a) *2e* **(b)** *2ej*

(c) *2e2j* **(d)** *2e3j*

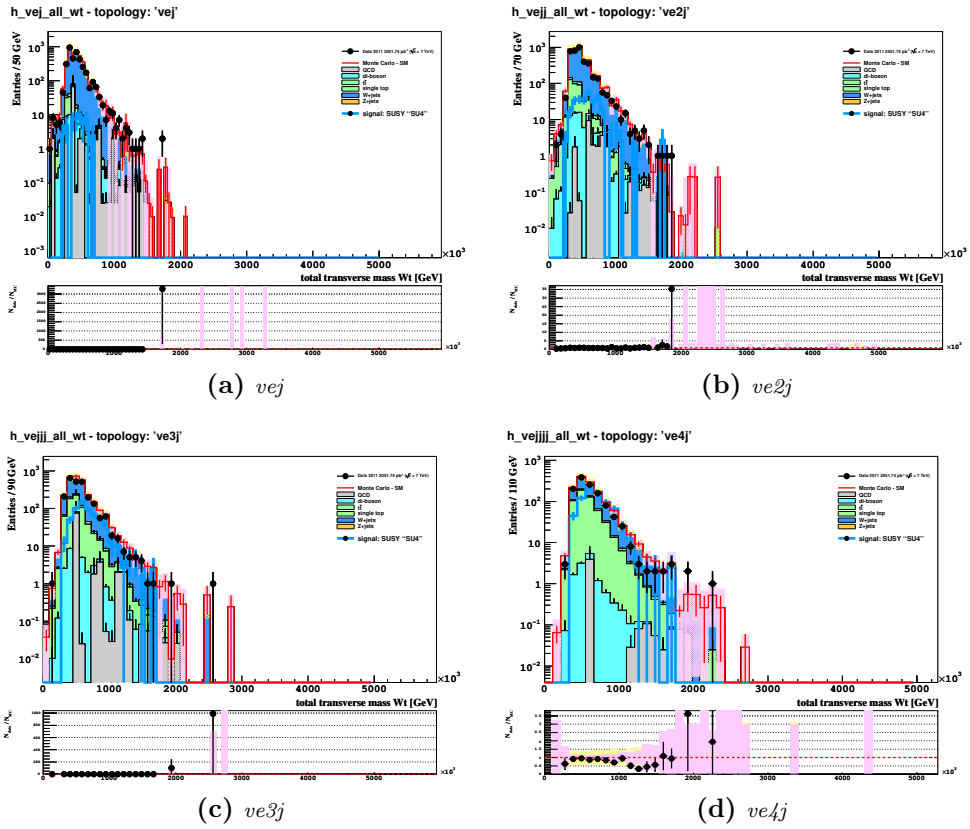
Figure 6.5: Effective mass M_{eff} of the di-electron channels. (a): channel **2e**; (b): channel **2ej**; (c): channel **2e2j**; (d): channel **2e3j**. The plots show a good agreement, within uncertainties, between data and the SM background.



(a) *2m* **(b)** *2mj*

(c) *2m2j* **(d)** *2m3j*

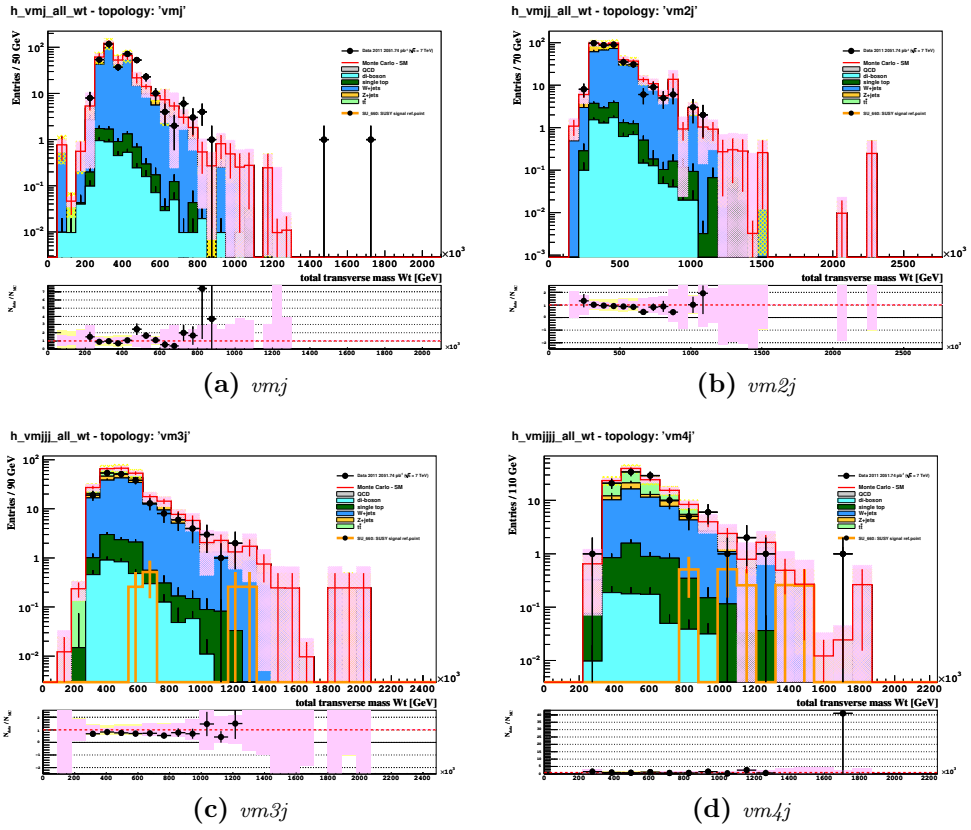
Figure 6.6: Effective mass M_{eff} of the di-muon channels. **(a)**: channel **2m**; **(b)**: channel **2mj**; **(c)**: channel **2m2j**; **(d)**: channel **2m3j**. The plots show a good agreement, within uncertainties, between data and the SM background.



(a) *vej* **(b)** *ve2j*

(c) *ve3j* **(d)** *ve4j*

Figure 6.7: Electron-neutrino transverse mass. (a): channel **vej**; (b): channel **ve2j**; (c): channel **ve3j**; (d): channel **ve4j**. The plots show a good agreement, within uncertainties, between data and the SM background.



(a) *vmj* **(b)** *vm2j*

(c) *vm3j* **(d)** *vm4j*

Figure 6.8: Muon-neutrino transverse mass. **(a)**: channel **vmj**; **(b)**: channel **vm2j**; **(c)**: channel **vm3j**; **(d)**: channel **vm4j**. The plots show a good agreement, within uncertainties, between data and the SM background.

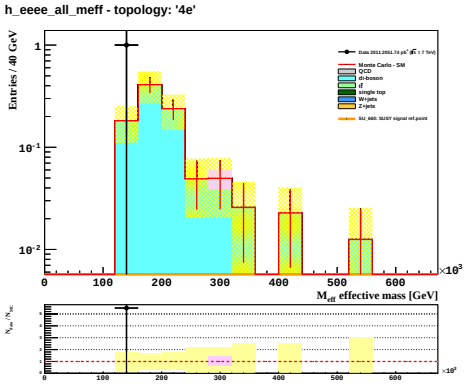


Figure 6.9: Effective mass of the **4e** event class. The plot shows a good description of the MC for the di-boson background and the good sensitivity of the General Search to di-boson events.

6.10 General Search results

In this section the results of this model-independent General Search over the first 2.1 fb^{-1} of the ATLAS data will be shown.

6.10.1 Global event yields

In figures 6.10–6.22, on pages 110–122, the global event yields for all event classes are shown. In Figure 6.10 the global result from the analysis of the `Egamma` stream is shown; in Figure 6.11–6.13 the same information is presented "sliced" in different plots, for a better view. The figures 6.15 and 6.16–6.17 show the results from the `Muon` stream, in global and sliced versions. The figures 6.19 and 6.20–6.22 show the results from the `JetTauEtmiss` stream.

A good agreement, within the uncertainties, of the total event yields between data and background has been observed, except from the event class `e15j` in the `Egamma` analysis — in Figure 6.11(b) on page 111 — featuring 1 electron and 15 jets, for which no background expectation has been obtained from the MonteCarlo simulation, due to limited statistics of the sample. Also the event class `v`, featuring only $E_{\text{T}}^{\text{miss}}$, shows discrepancies in the `Muon` and `JetTauEtmiss` analyses — in figures 6.17(a) on page 117 and 6.21(a) on page 121 — probably due to trigger issues.

6.10.2 Scan of the M_{eff} distribution

In order to quantify the agreement between the data and the Standard Model background expectation, and to identify regions where discrepancies occur, a search algorithm to scan over the M_{eff} distribution has been used in this analysis. The search algorithm has been developed for this analysis, along the general lines of the algorithm used by the H1 collaboration (already described in Section 5.2).

Significance and the "Look-Elsewhere" effect

In searches like this one, where as many channels and topologies as possible are analyzed, it is crucial to account for the *trials factor* and the *significance* of a measure.

When we compare data points to any background, there is a probability that the background fluctuates toward the data, simply due to statistics. The probability for a fluctuation of the background to be at least as large as the observed maximum data excess is represented by the *p-value*; while the probability for a background excess anywhere in a specified region is represented by the global *P-value*. More details on the *p-value* can be found in Section 5.2.

This background fluctuation probability can be evaluated by generating sets of MonteCarlo simulated data, starting from the background distribution.

In fact fluctuations at the level of three, or more, standard deviations[3] are expected to appear in the analysis, simply by the fact of the large number of data points and regions that are considered. The larger the number of regions, the larger the probability of observing an "interesting deviation" anywhere in the whole region only due to random fluctuations, the larger the global *P-value* for the specified region compared to the local *p-value*. This is called the "Look-Elsewhere" effect, which has to be taken into account in order to correct the results.

The algorithm

The search algorithm discussed here has been applied only to one distribution: the M_{eff} distribution (see Eq. 6.2). This choice has been taken in order to get a good balance between the sensitivity of the analysis — the M_{eff} is sensible to new physics, because of its dependency on the masses of all objects and on the $E_{\mathrm{T}}^{\mathrm{miss}}$ — and the "Look-Elsewhere" effect, keeping the number of places where the analysis looks for deviation as low as possible.

The algorithm consists in several steps, described here below.

1. for each topology, the M_{eff} distribution is taken and the most significant discrepancy between data and MC background is found, analyzing regions of adjacent bins. For each region of the histogram, the *p-value* is computed, to quantify the significance of the discrepancy. The lower the *p-value*, the smaller the probability that the discrepancy is due to a random fluctuation, the more significant the discrepancy. From all the regions of the histogram, the lowest *p-value* is kept. In the end one obtains the $p\text{-value}_{max}$ for each topology, as shown in Figure 6.23 on page 123;

2. for each topology, the MC background of the same M_{eff} distribution is used to randomly generate an high number of toy-MC "pseudo-experiments";

3. each pseudo-experiment — considered as "fake-data" — is compared with the original MC background histogram and the most significant discrepancy is found as in step 1. In the end one gets the $p\text{-value}^{toy}_{max}$ for each topology, describing the probability of interesting deviations being found and originated from background statistical fluctuations only.

[3]Both the local *p-value* and the global *P-value* can be expressed as a corresponding number of standard deviations using the one-sided Gaussian tail convention. A significance *p-value* of $3 \cdot 10^{-7}$ corresponds to a probability of about 1 in 3.5 million, which corresponds to a 5σ deviation. In this context a 5σ significance tells that the probability of the background alone locally fluctuating up by the observed amount or more is about 1 in 3.5 million, which is very small, and so very unlikely.

Histogram binning

The binning of the histograms — i.e. the choice of the width of the histograms' bins — plays an important role in the statistics interpretation of the analysis. In fact, the smaller the width of the bins, the larger the number of bins, the larger the number of regions considered in the M_{eff} scan and the larger the statistical "Look-Elsewhere" uncertainty affecting the analysis.

Thus, in order to avoid to unnecessarily increase the trial factors, the binning of the M_{eff} histograms has been chosen to be comparable to the expected experimental resolution of each event class.

In this analysis, a value of 10 GeV has been used for the resolution of muons and electrons; while a value of 20 GeV has been considered for jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. So, for example, the bin width of a M_{eff} distribution plot of an event class like `vm5j` would be 130 GeV.

In addition, since discrepancies between data and background are expected to be wider than one single bin in the M_{eff} distribution, the algorithms uses regions of at least 2 consecutive bins.

6.10.3 M_{eff} scan results

The $-\log_{10}$ value of the $p\text{-}value_{max}$ and the $p\text{-}value^{toy}_{max}$ is then plotted, as shown in figures 6.24 on page 124, for the analyses on all three data streams. A value between 5 and 6 of the quantity $-\log_{10}(p\text{-}value)$ has been proved to correspond to a region whose value of $p\text{-}value_{max}$ — the smallest $p\text{-}value$ — is $5.7 \cdot 10^{-7}$, which corresponds to a "5σ" deviation [45, 196].

In this analysis, in all three data streams, no data, neither background, show discrepancies featuring large deviations.

6.11 Conclusions

In this model-independent General Search all the three ATLAS data streams have been analyzed, in the search of new physics. Events have been selected without any base theoretical model, with the idea of looking for discrepancies in the largest possible number of channels. The results have been statistically treated in order to take into account the trials factor and the "Look-Elsewhere" effect. In the end the global event yields and the scan on the M_{eff} distribution of all event classes showed no interesting discrepancies.

No evidence and no sign of new physics have been observed in this analysis.

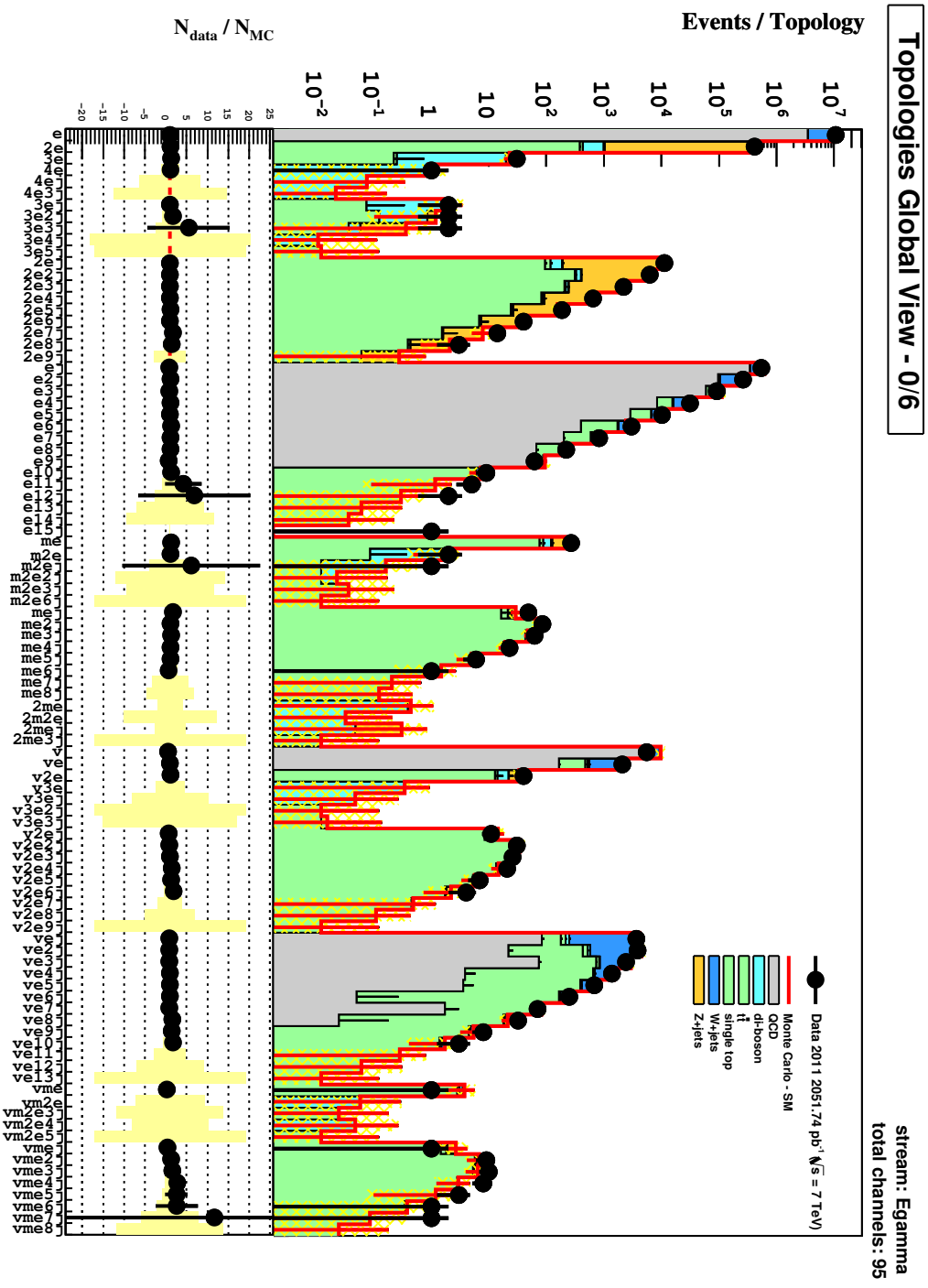
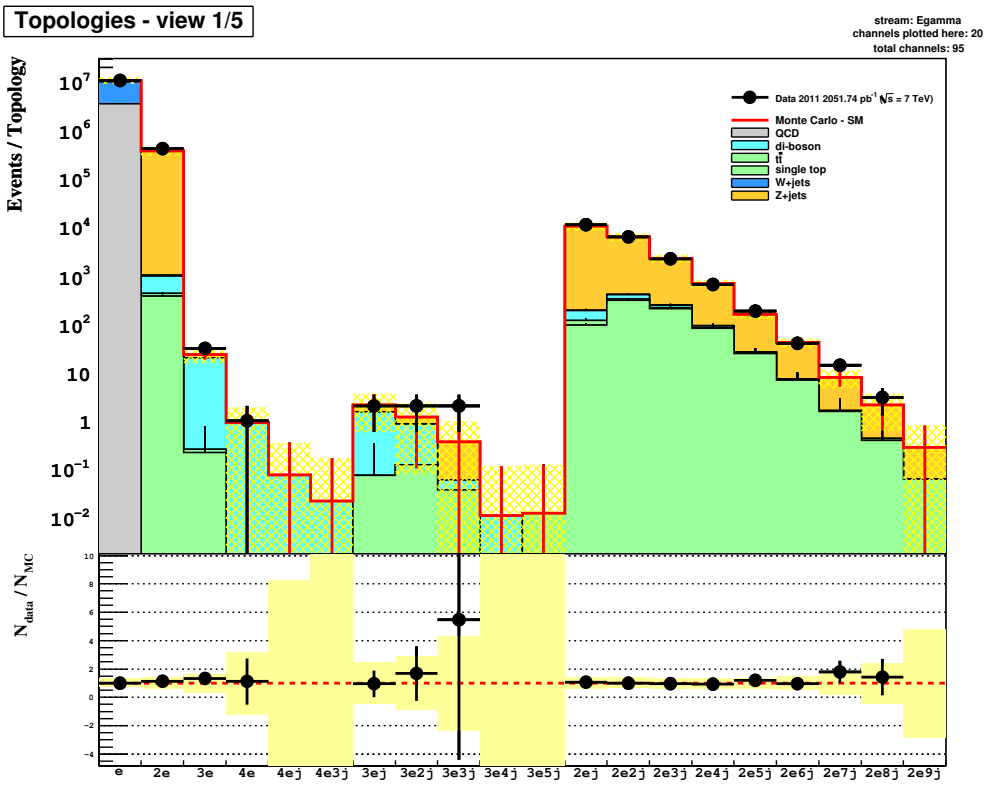
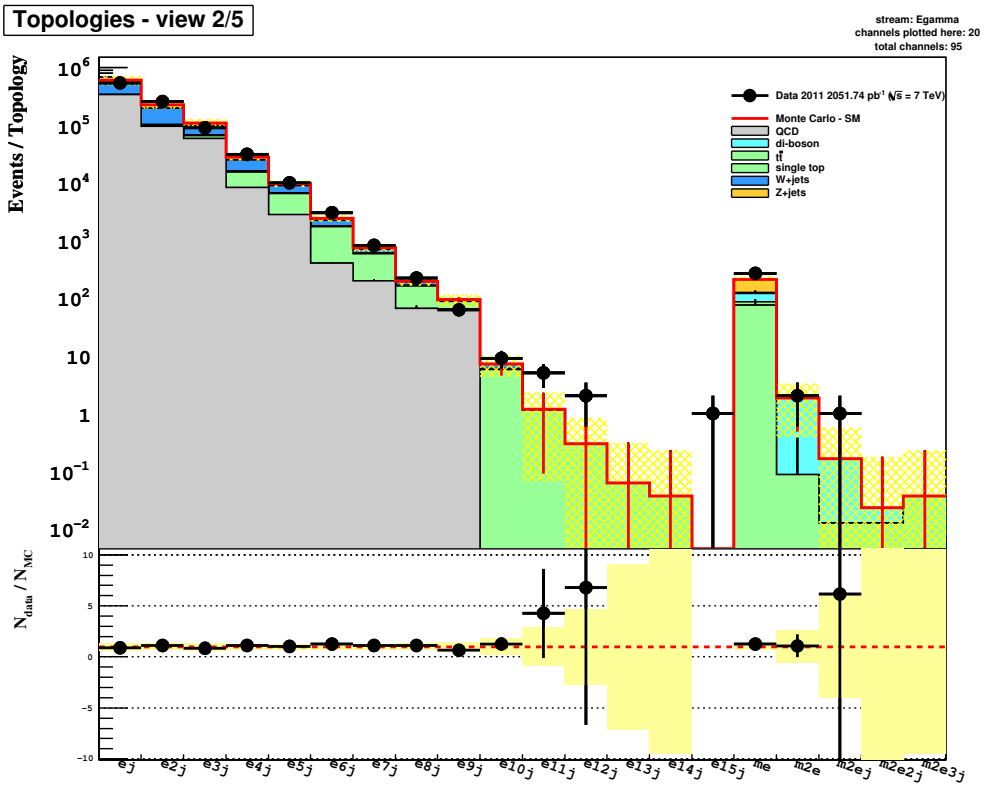


Figure 6.10: All events of `Egamma` stream divided according to their *dynamic topology* content.

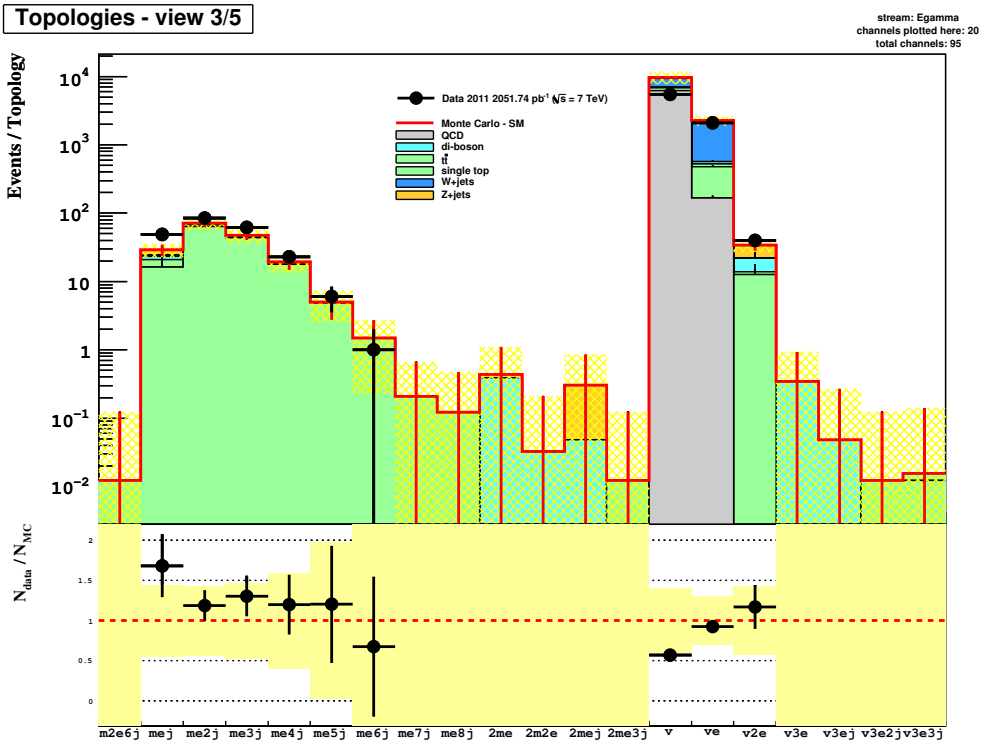


(a)

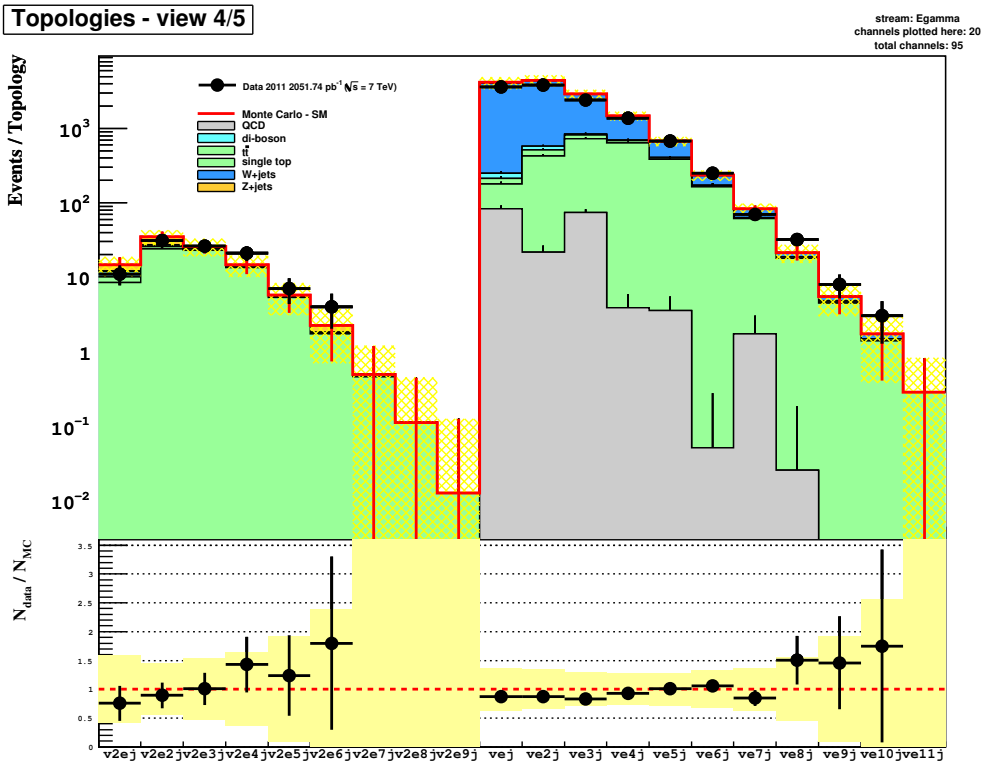


(b)

Figure 6.11: The `Egamma` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig. 6.10 is shown, sliced into sub-plots. [*Part 1/3 – First two plots*]

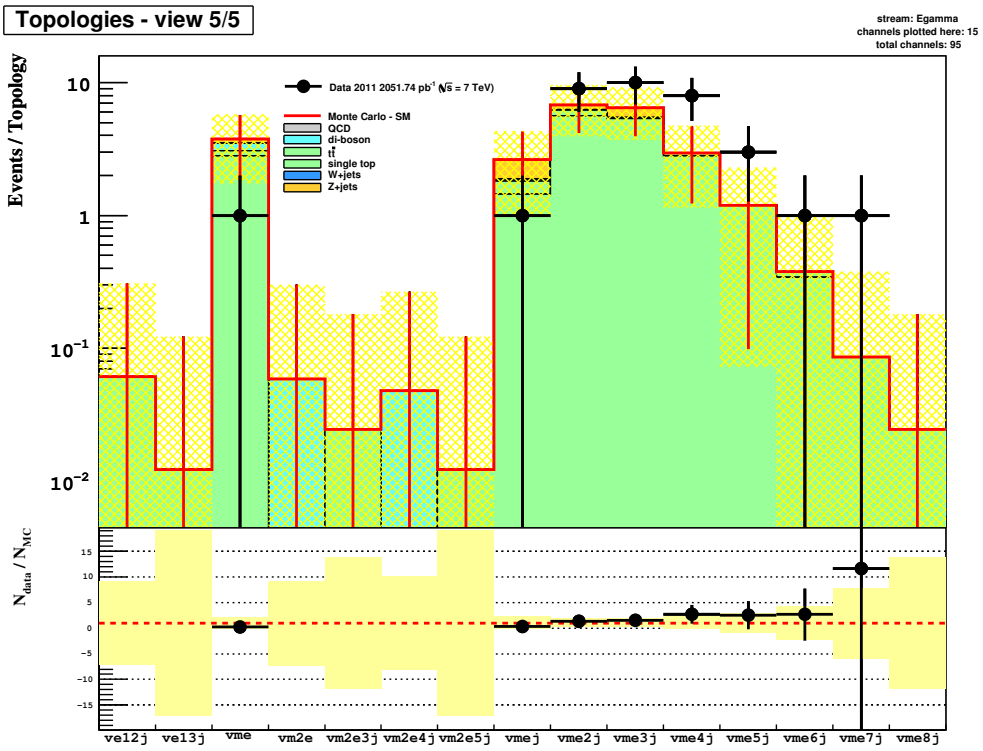


(a)



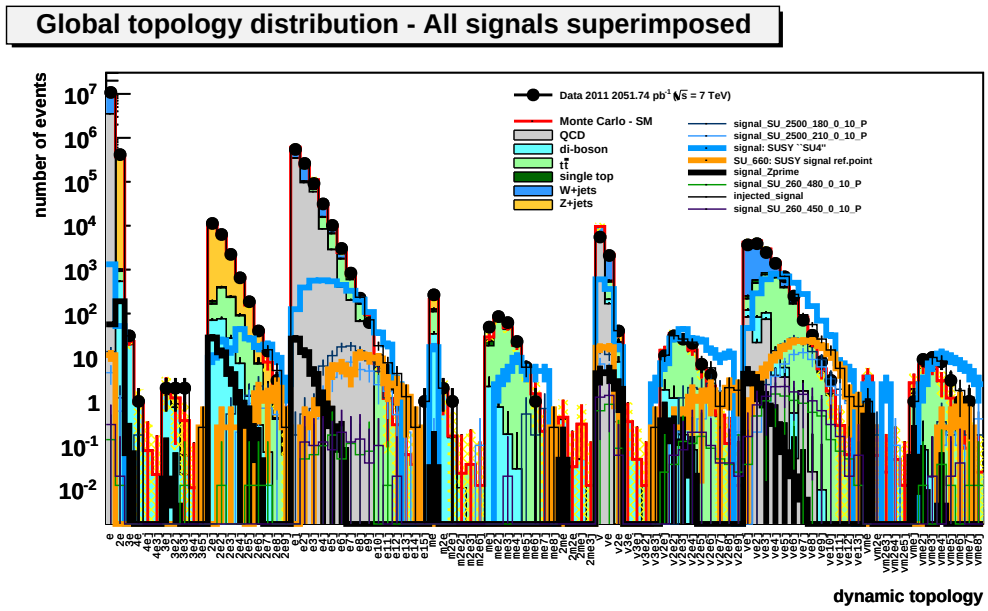
(b)

Figure 6.12: The `Egamma` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig. 6.10 is shown, sliced into sub-plots. [*Part 2/3 – Second two plots*]

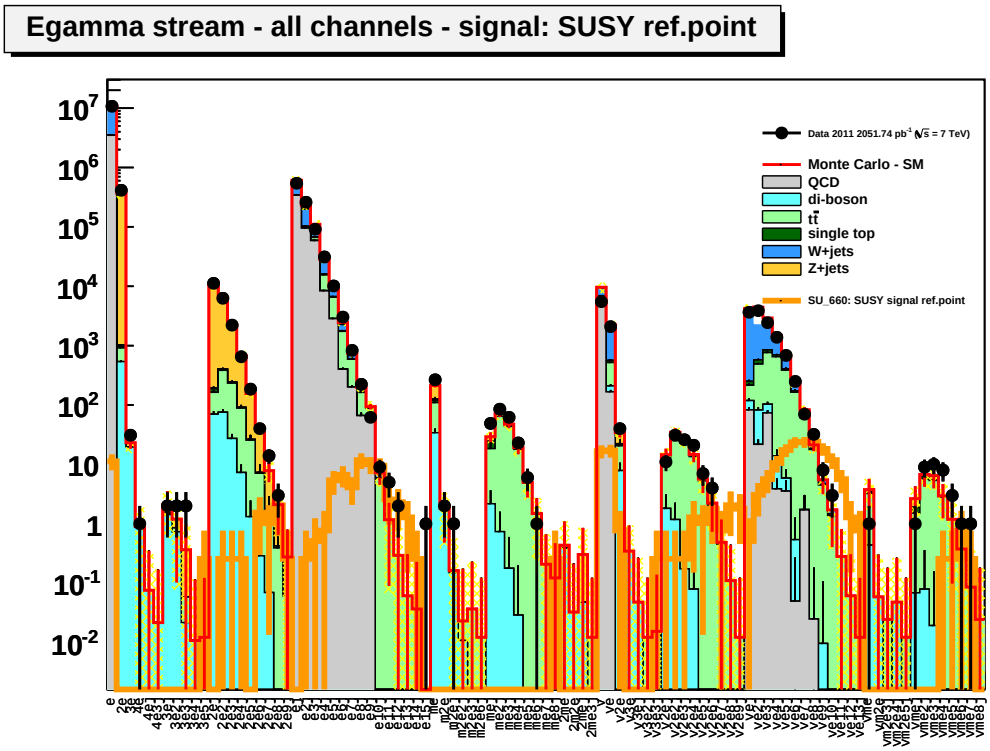


(a)

Figure 6.13: The `Egamma` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig. 6.10 is shown, sliced into sub-plots. [*Part 3/3 – Last plot*]



(a) *Data, MC background and all signals*



(b) *Data, MC background and the SUSY reference-point signal*

Figure 6.14: All the events of the `Egamma` stream divided according to their *dynamic topology* content: data events, MC simulated background events and signals. (a) in this plots some example signals of exotics processes are shown, superimposed to the main plot; (b) here only the SUSY reference point "660" is shown.

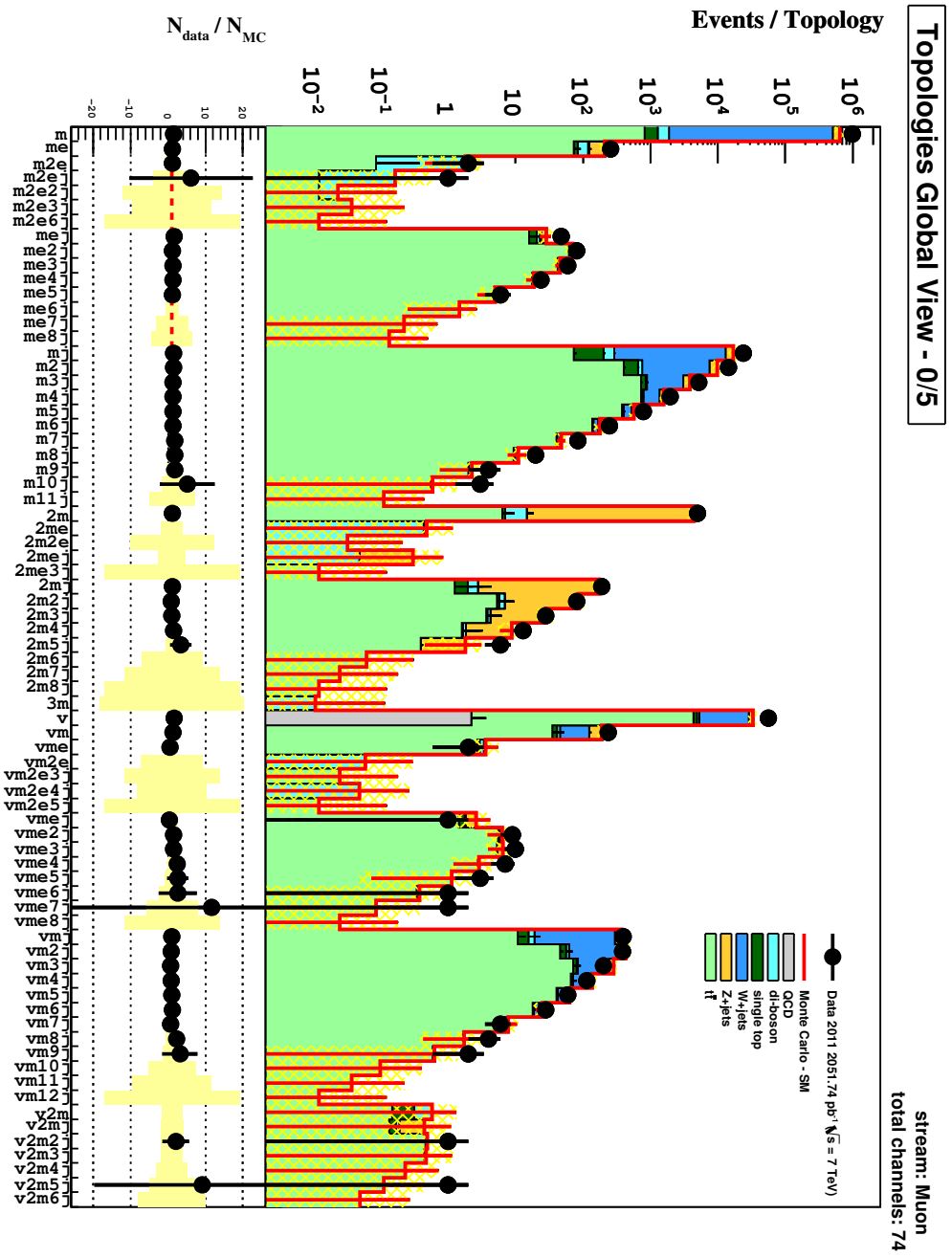
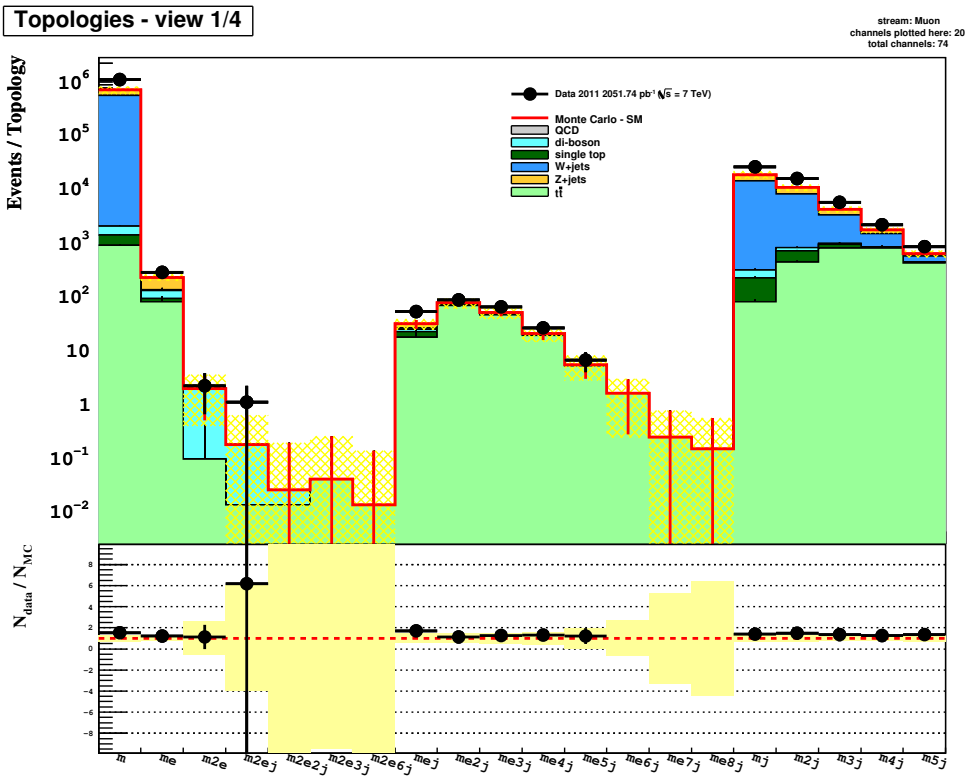
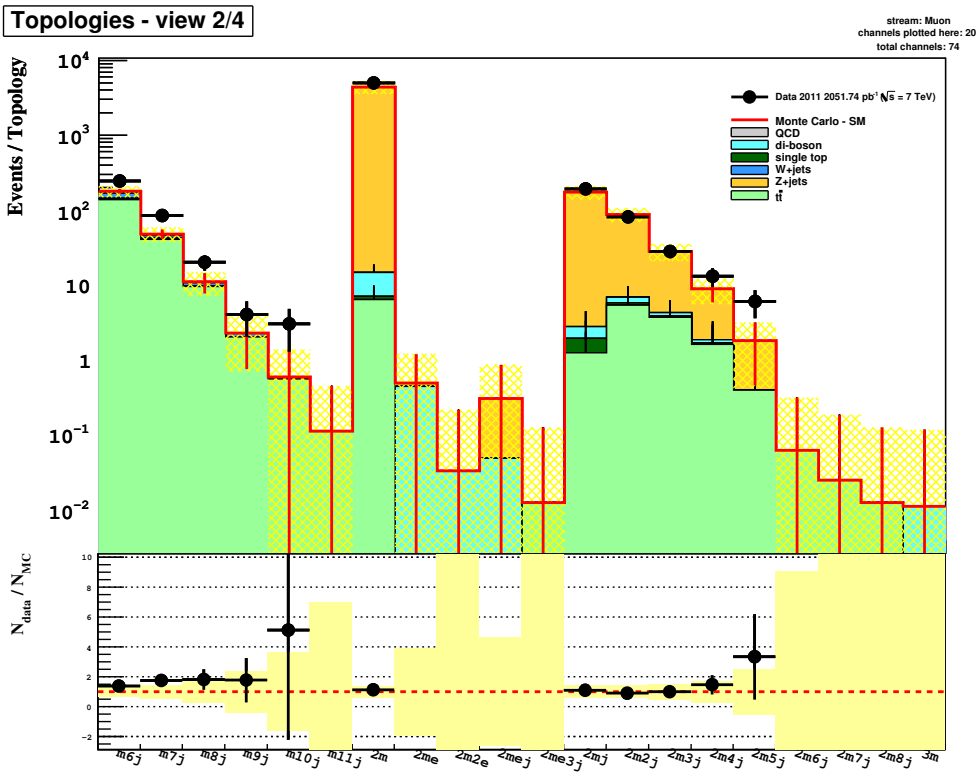


Figure 6.15: All events of the `Muon` stream are here divided according to their *dynamic topology* content.

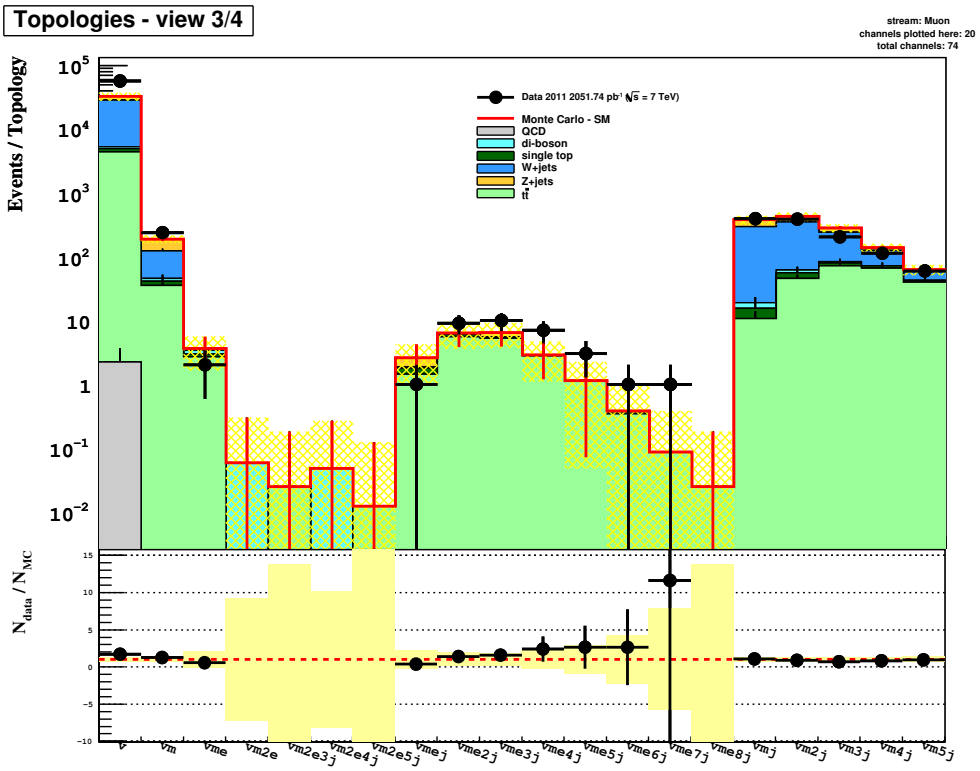


(a)

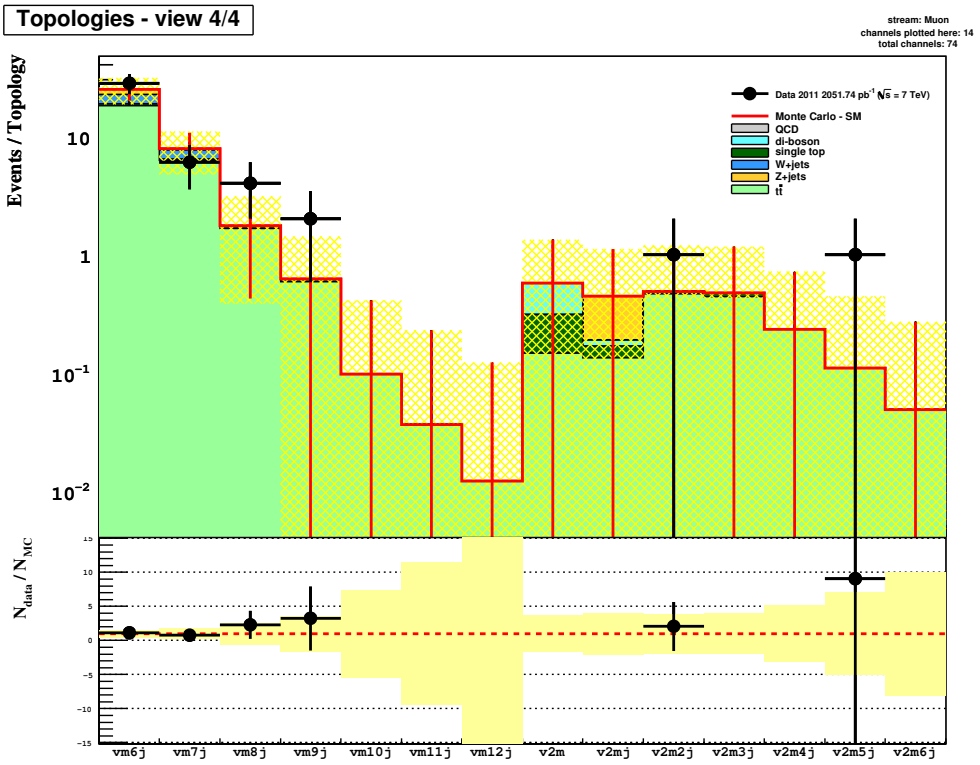


(b)

Figure 6.16: The `Muon` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig.6.15 is shown, sliced into sub-plots. [*Part 1/2 – First two plots*]

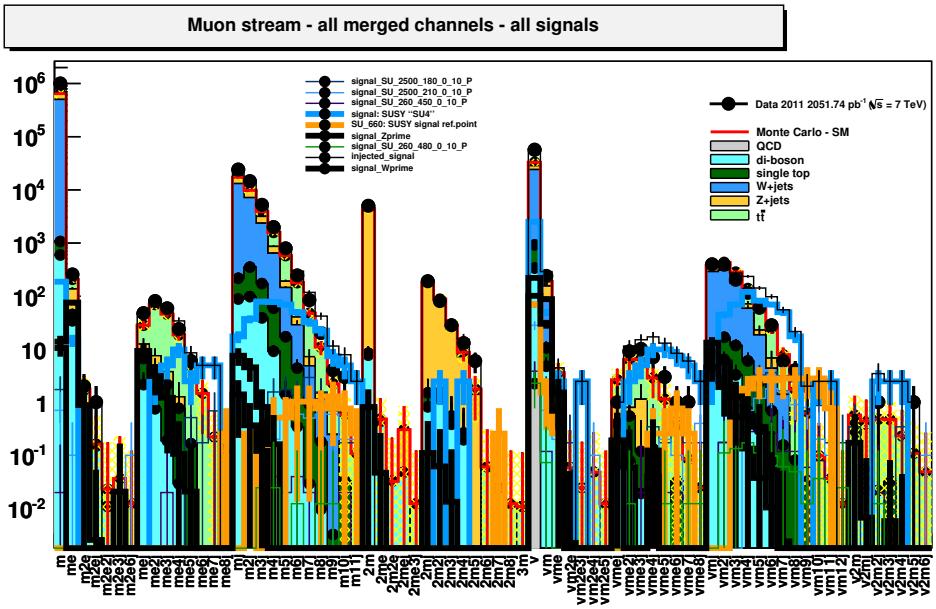


(a)

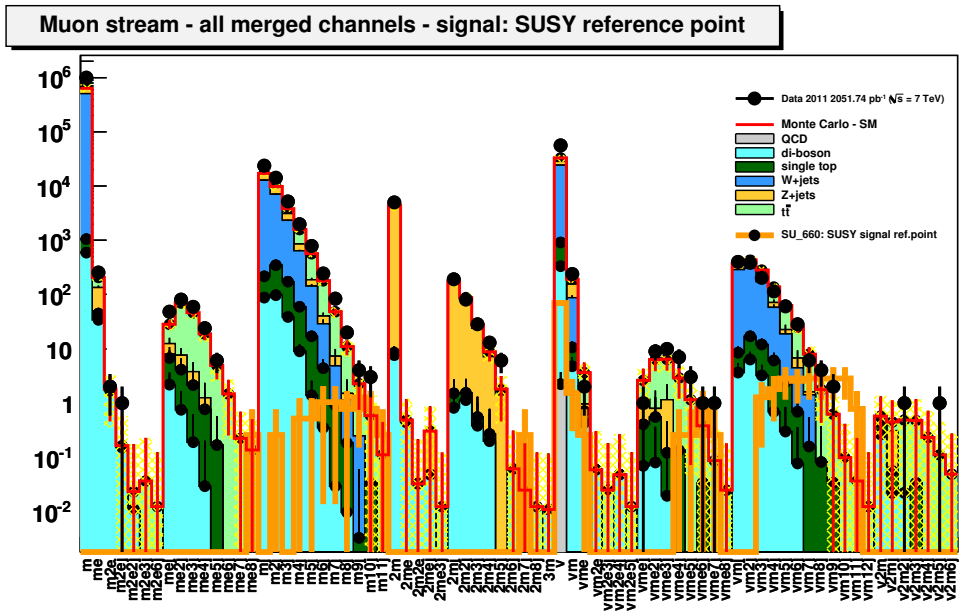


(b)

Figure 6.17: The `Muon` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig.6.15 is shown, sliced into sub-plots. [*Part 2/2 – Last two plots*]



(a) *Data, MC background and all signals*



(b) *Data, MC background and the SUSY reference-point signal*

Figure 6.18: Events of the `Muon` stream divided according to their *dynamic topology* content: data events, MC simulated background events and signals. **(a)** in this plots some example signals of exotics processes are shown, superimposed to the main plot; **(b)** here only the SUSY reference point "660" is shown.

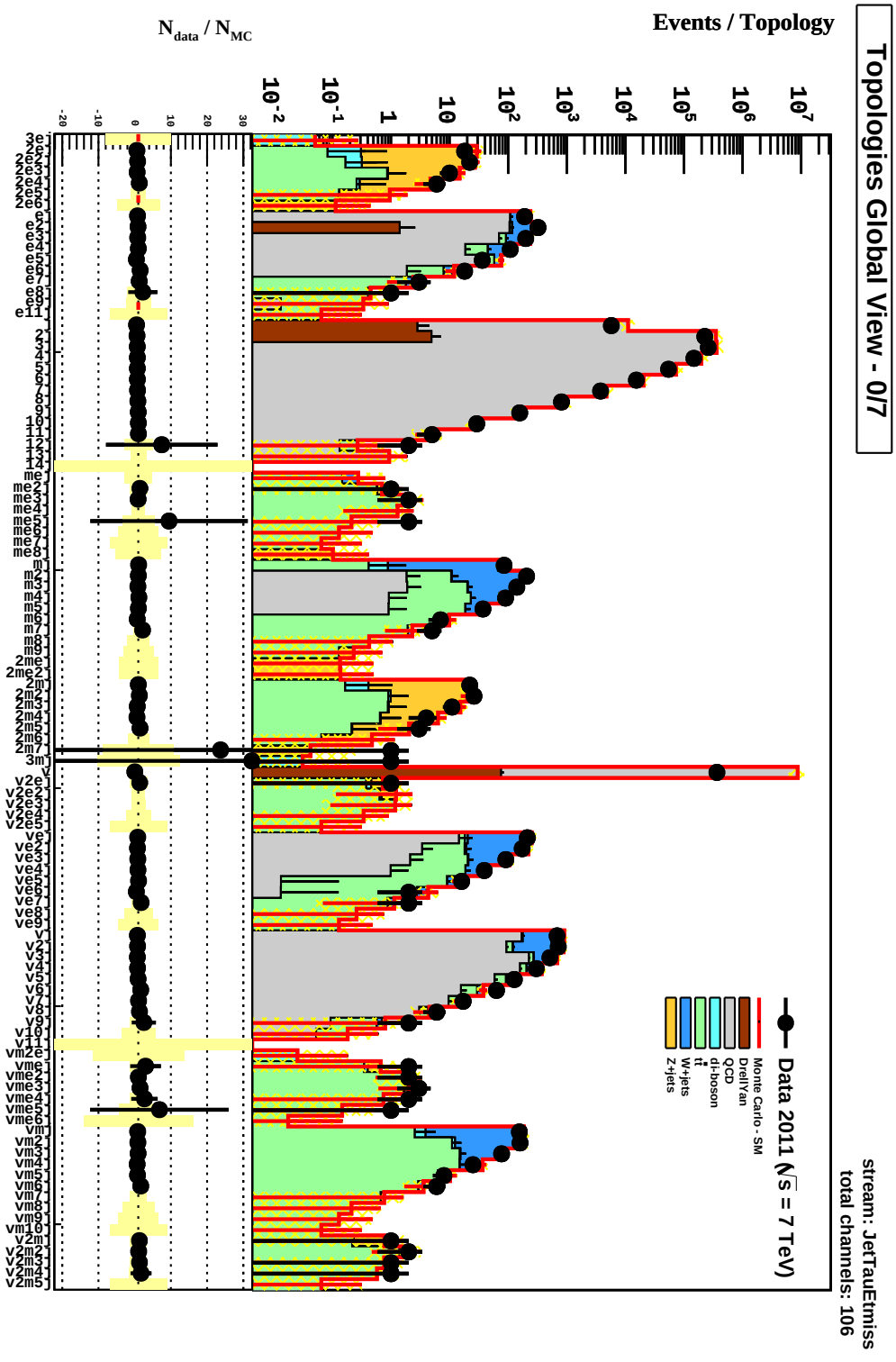
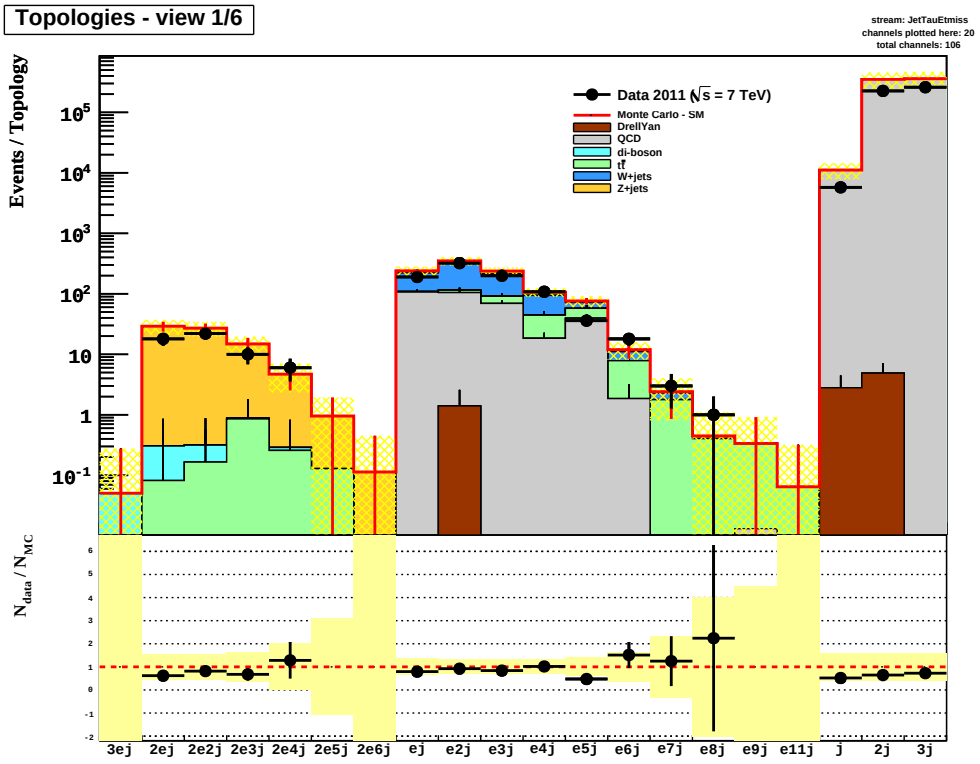
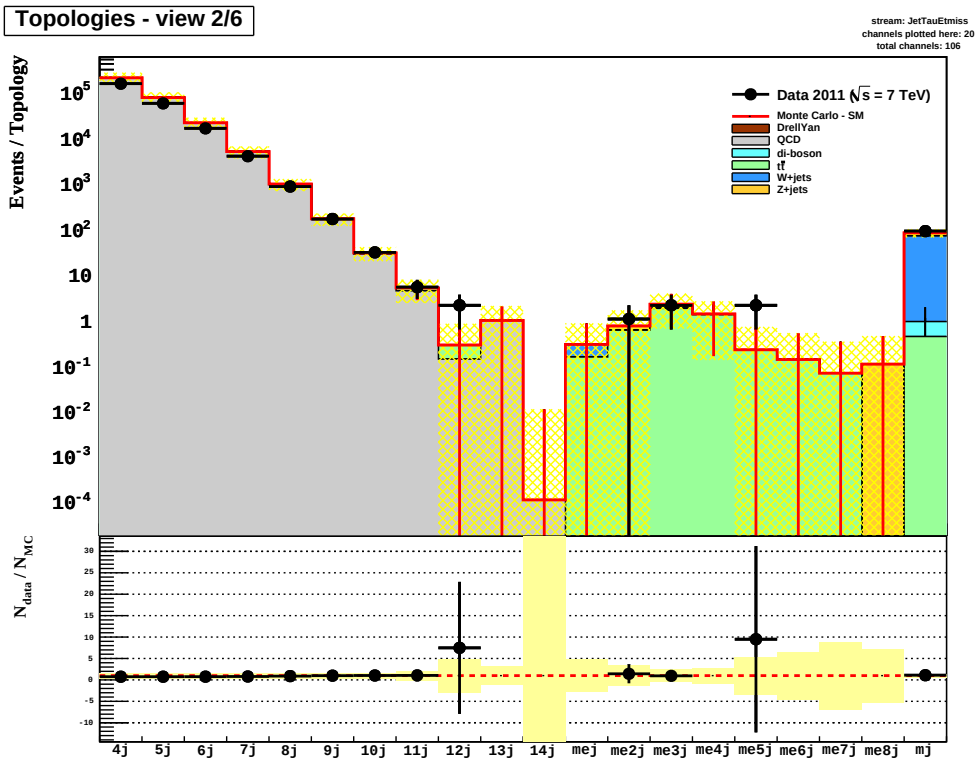


Figure 6.19: All events of the `JetTauEtmiss` stream are here divided according to their *dynamic topology* content.

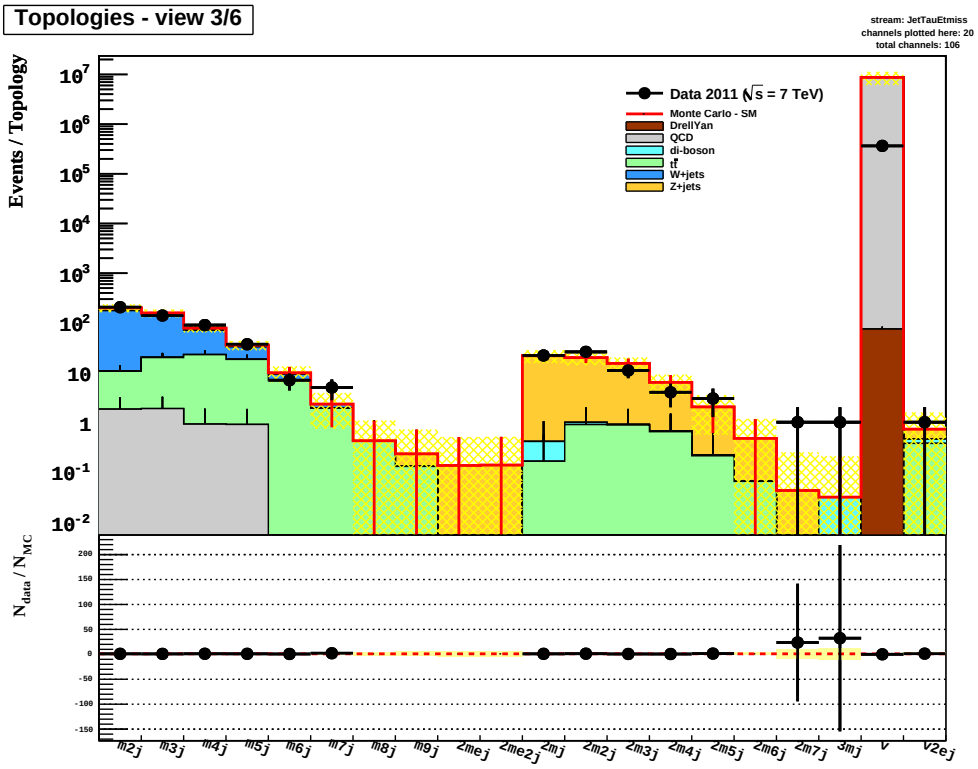


(a)

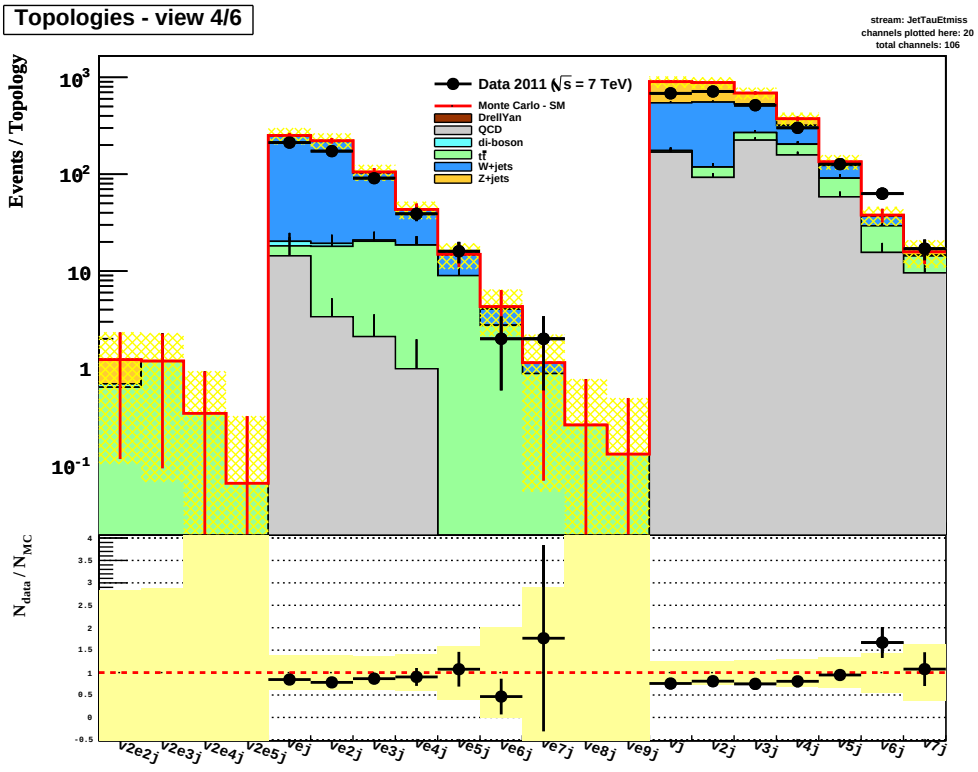


(b)

Figure 6.20: The `JetTauEtmiss` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig.6.19 is shown, sliced into sub-plots. [*Part 1/3 – First two plots*]

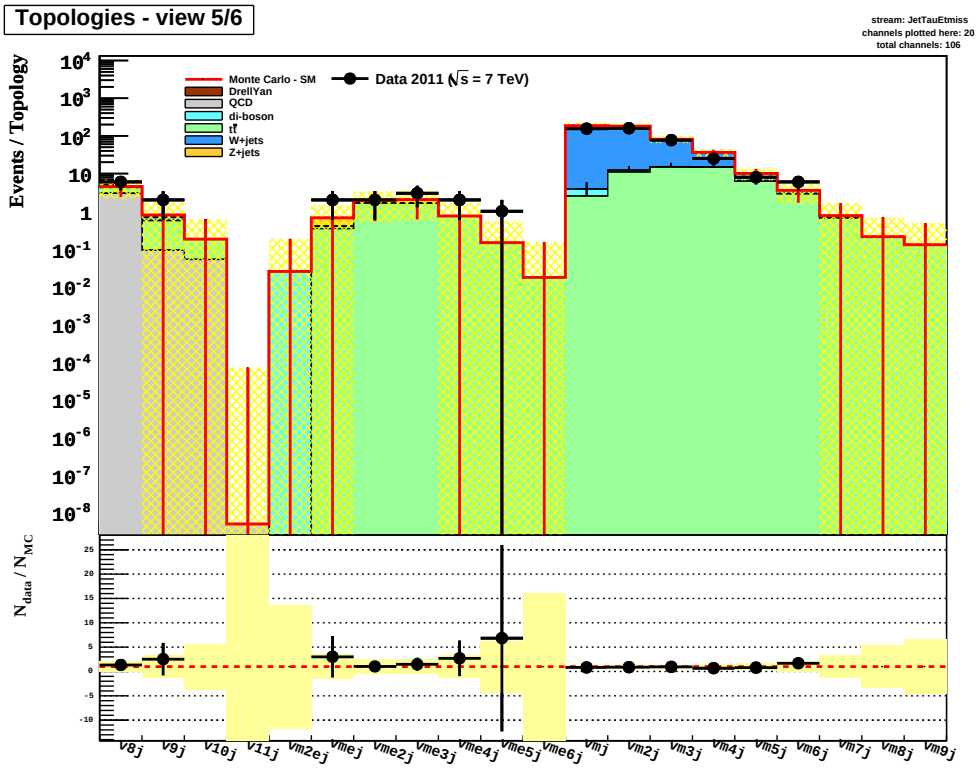


(a)

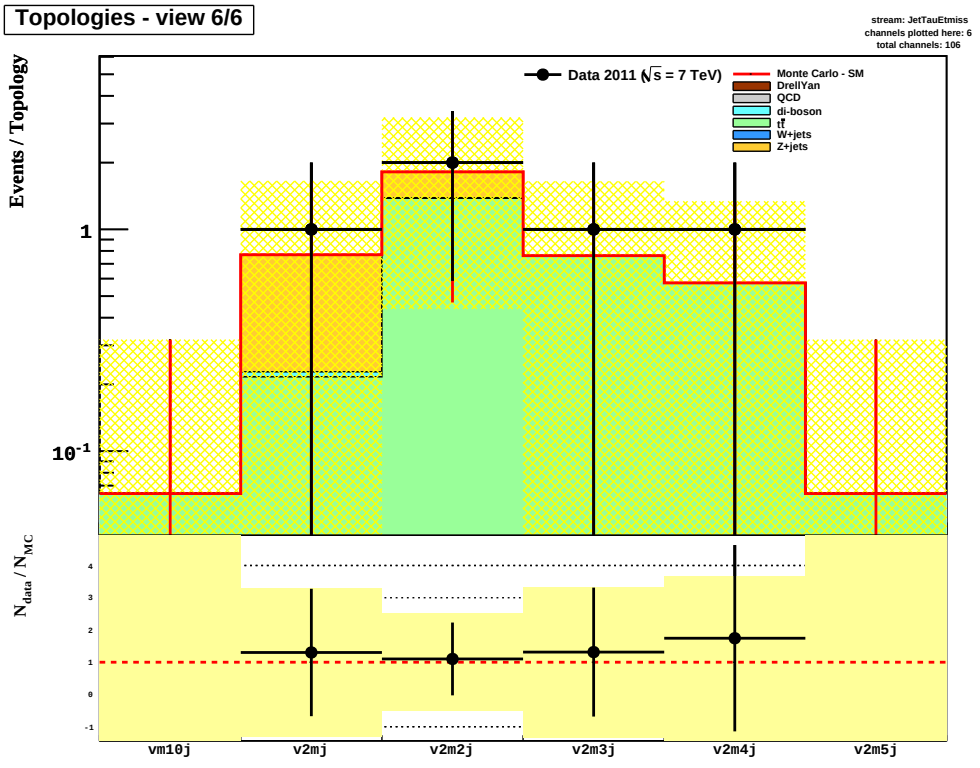


(b)

Figure 6.21: The `JetTauEtmiss` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig.6.19 is shown, sliced into sub-plots. [*Part 2/3 – Second two plots*



(a)



(b)

Figure 6.22: The `JetTauEtmiss` stream events divided in topologies according to their *dynamic topology* content. Here the same plot of Fig.6.19 is shown, sliced into sub-plots. [*Part 3/3 – Last two plots*]

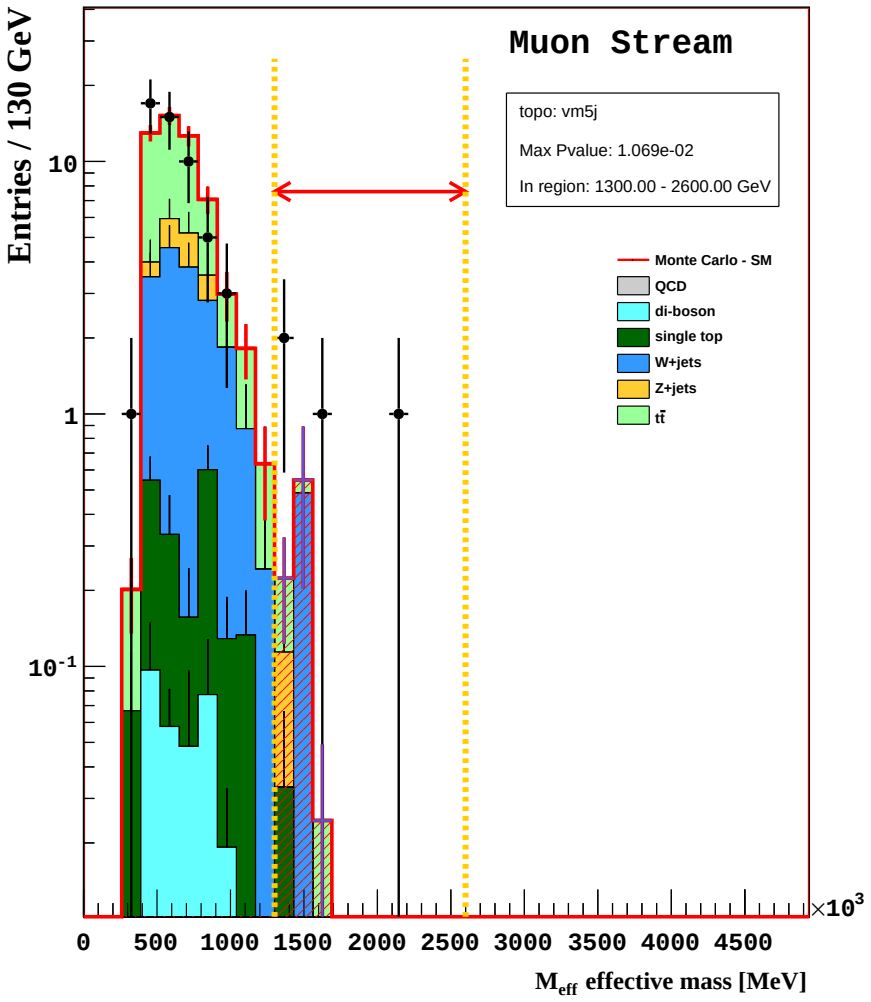
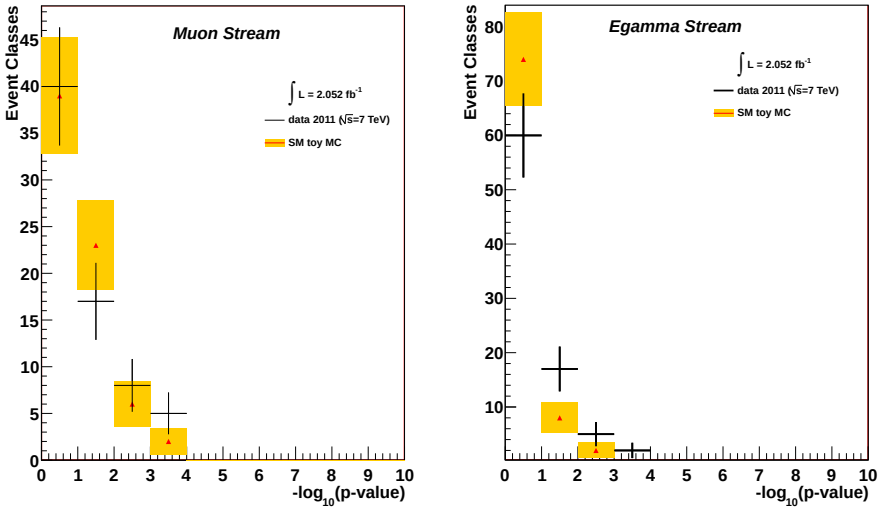
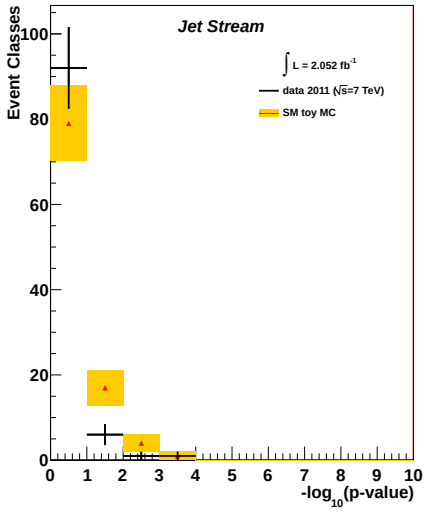


Figure 6.23: An example of the M_{eff} scan in the `Muon` stream analysis. M_{eff} distributions from all event classes are automatically scanned by the algorithm in order to find the region with the smallest *p-value*, the $p\text{-}value_{max}$. In this example a $p\text{-}value_{max}$ of $1.069 \cdot 10^{-2}$ has been found for the topology `vm5j` in the region 1300–2600 GeV.



(a)

(b)



(c)

Figure 6.24: The global distribution of the $-\log_{10}$(p-value) of the M_{eff} scan is shown, both for data and for *toy MC* "pseudo-experiments". Pseudo-experiments are generated from the M_{eff} distributions of the event classes of the SM background, as explained in Section 6.10.2. No interesting deviations have been observed in data.

Appendix 6.A List of background samples

Table 6.4: Simulated processes used as Standard Model background in this General Search. *(the table continues on the next 2 pages...).*
All samples come from the official ATLAS SUSY Production Group.

Background: simulated processes and samples			
Background group	process name	process	sample
QCD *(Pythia)*	J0_pythia	105009	mc10_7TeV.105009.J0_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J1_pythia	105010	mc10_7TeV.105010.J1_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J2_pythia	105011	mc10_7TeV.105011.J2_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J3_pythia	105012	mc10_7TeV.105012.J3_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J4_pythia	105013	mc10_7TeV.105013.J4_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J5_pythia	105014	mc10_7TeV.105014.J5_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J6_pythia	105015	mc10_7TeV.105015.J6_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J7_pythia	105016	mc10_7TeV.105016.J7_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
	J8_pythia	105017	mc10_7TeV.105017.J8_pythia_jetjet.merge.NTUP_SUSY.e574_s934_s946_r2299_r2300_p601/
QCD_1muon *(Pythia)*	J4_pythia_1muon	109280	mc10_7TeV.109280.J4_pythia_jetjet_1muon.merge.NTUP_SUSY.e574_s933_s946_r2301_r2300_p601/
	J5_pythia_1muon	109281	mc10_7TeV.109281.J5_pythia_jetjet_1muon.merge.NTUP_SUSY.e574_s933_s946_r2301_r2300_p601/
	J6_pythia_1muon	109435	mc10_7TeV.109435.J6_pythia_jetjet_1muon.merge.NTUP_SUSY.e574_s933_s946_r2301_r2300_p601/
	J0_pythia_1muon	109276	mc10_7TeV.109276.J0_pythia_jetjet_1muon.merge.NTUP_SUSY.e574_s933_s946_r2301_r2300_p601/
	J1_pythia_1muon	109277	mc10_7TeV.109277.J1_pythia_jetjet_1muon.merge.NTUP_SUSY.e574_s933_s946_r2301_r2300_p601/
	J2_pythia_1muon	109278	mc10_7TeV.109278.J2_pythia_jetjet_1muon.merge.NTUP_SUSY.e574_s933_s946_r2301_r2300_p601/
	J3_pythia_1muon	109279	mc10_7TeV.109279.J3_pythia_jetjet_1muon.merge.NTUP_SUSY.e574_s933_s946_r2301_r2300_p601/
top *(McAtNLO+Jimmy)*	st_schan_enu	108343	mc10_7TeV.108343.st_schan_enu_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	T1	105200	mc10_7TeV.105200.T1_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	TTbar_FullHad	105204	mc10_7TeV.105204.TTbar_FullHad_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	st_tchan_munu	108341	mc10_7TeV.108341.st_tchan_munu_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	st_tchan_taunu	108342	mc10_7TeV.108342.st_tchan_taunu_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	st_schan_taunu	108345	mc10_7TeV.108345.st_schan_taunu_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	st_schan_munu	108344	mc10_7TeV.108344.st_schan_munu_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	st_tchan_enu	108340	mc10_7TeV.108340.st_tchan_enu_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	st_Wt	108346	mc10_7TeV.108346.st_Wt_McAtNlo_Jimmy.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/

Table 6.4: (. . . continued) Simulated processes used as Standard Model background in this model-independent General Search.

Background group	process name	process	sample
Zee *(Alpgen+Jimmy)*	ZeeNp0	107650	mc10_7TeV.107650.AlpgenJimmyZeeNp0_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZeeNp1	107651	mc10_7TeV.107651.AlpgenJimmyZeeNp1_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZeeNp2	107652	mc10_7TeV.107652.AlpgenJimmyZeeNp2_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZeeNp3	107653	mc10_7TeV.107653.AlpgenJimmyZeeNp3_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZeeNp4	107654	mc10_7TeV.107654.AlpgenJimmyZeeNp4_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZeeNp5	107655	mc10_7TeV.107655.AlpgenJimmyZeeNp5_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
Zmumu *(Alpgen+Jimmy)*	ZmumuNp0	107660	mc10_7TeV.107660.AlpgenJimmyZmumuNp0_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZmumuNp1	107661	mc10_7TeV.107661.AlpgenJimmyZmumuNp1_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZmumuNp2	107662	mc10_7TeV.107662.AlpgenJimmyZmumuNp2_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZmumuNp3	107663	mc10_7TeV.107663.AlpgenJimmyZmumuNp3_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZmumuNp4	107664	mc10_7TeV.107664.AlpgenJimmyZmumuNp4_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZmumuNp5	107665	mc10_7TeV.107665.AlpgenJimmyZmumuNp5_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
Znunu *(Alpgen+Jimmy)*	ZnunuNp2	107712	mc10_7TeV.107712.AlpgenJimmyZnunuNp2_pt20_filt1jet.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	ZnunuNp3	107713	mc10_7TeV.107713.AlpgenJimmyZnunuNp3_pt20_filt1jet.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	ZnunuNp4	107714	mc10_7TeV.107714.AlpgenJimmyZnunuNp4_pt20_filt1jet.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	ZnunuNp5	107715	mc10_7TeV.107715.AlpgenJimmyZnunuNp5_pt20_filt1jet.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	ZnunuNp0	107710	mc10_7TeV.107710.AlpgenJimmyZnunuNp0_pt20_filt1jet.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	ZnunuNp1	107711	mc10_7TeV.107711.AlpgenJimmyZnunuNp1_pt20_filt1jet.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
Ztautau *(Alpgen+Jimmy)*	ZtautauNp0	107670	mc10_7TeV.107670.AlpgenJimmyZtautauNp0_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZtautauNp1	107671	mc10_7TeV.107671.AlpgenJimmyZtautauNp1_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZtautauNp2	107672	mc10_7TeV.107672.AlpgenJimmyZtautauNp2_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZtautauNp3	107673	mc10_7TeV.107673.AlpgenJimmyZtautauNp3_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZtautauNp4	107674	mc10_7TeV.107674.AlpgenJimmyZtautauNp4_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
	ZtautauNp5	107675	mc10_7TeV.107675.AlpgenJimmyZtautauNp5_pt20.merge.NTUP_SUSY.e737_s933_s946_r2302_r2300_p601/
DiBoson *(Herwig)*	WW	105985	mc10_7TeV.105985.WW_Herwig.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	ZZ	105986	mc10_7TeV.105986.ZZ_Herwig.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/
	WZ	105987	mc10_7TeV.105987.WZ_Herwig.merge.NTUP_SUSY.e598_s933_s946_r2302_r2300_p601/

Table 6.4: (. . . continued) Simulated processes used as Standard Model background in this model-independent General Search.

Background group	process name	process	sample
Wbb *(Alpgen+Jimmy)*	WbbNp0	106280	mc10_7TeV.106280.AlpgenJimmyWbbNp0_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	WbbNp1	106281	mc10_7TeV.106281.AlpgenJimmyWbbNp1_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	WbbNp2	106282	mc10_7TeV.106282.AlpgenJimmyWbbNp2_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	WbbNp3	106283	mc10_7TeV.106283.AlpgenJimmyWbbNp3_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
Wenu *(Alpgen+Jimmy)*	WenuNp0	107680	mc10_7TeV.107680.AlpgenJimmyWenuNp0_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	WenuNp1	107681	mc10_7TeV.107681.AlpgenJimmyWenuNp1_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	WenuNp2	107682	mc10_7TeV.107682.AlpgenJimmyWenuNp2_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/
	WenuNp3	107683	mc10_7TeV.107683.AlpgenJimmyWenuNp3_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/
	WenuNp4	107684	mc10_7TeV.107684.AlpgenJimmyWenuNp4_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/
	WenuNp5	107685	mc10_7TeV.107685.AlpgenJimmyWenuNp5_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/
Wmunu *(Alpgen+Jimmy)*	WmunuNp0	107690	mc10_7TeV.107690.AlpgenJimmyWmunuNp0_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	WmunuNp1	107691	mc10_7TeV.107691.AlpgenJimmyWmunuNp1_pt20.merge.NTUP_SUSY.e600_s933_s946_r2302_r2300_p601/
	WmunuNp2	107692	mc10_7TeV.107692.AlpgenJimmyWmunuNp2_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/
	WmunuNp3	107693	mc10_7TeV.107693.AlpgenJimmyWmunuNp3_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/
	WmunuNp4	107694	mc10_7TeV.107694.AlpgenJimmyWmunuNp4_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/
	WmunuNp5	107695	mc10_7TeV.107695.AlpgenJimmyWmunuNp5_pt20.merge.NTUP_SUSY.e760_s933_s946_r2302_r2300_p601/

Chapter 7

WatchMan - Computer Aided Software Engineering for High Energy Phisics

> A big computer, a complex algorithm and a long time does not equal science.
>
> Robert Gentleman

LHC experiments, as said before, provide a large amount of data to be looked at, of the order of several petabytes each year. And all modern HEP experiments rely on custom dedicated software frameworks to handle, store, reconstruct and analyze such a very large amount of data. Thus, nowadays, analysing such data means, not only knowing how to write an analysis piece of software, but also being able to do it within these custom software frameworks, which are often quite complex. Moreover, a new piece of software implementing the analysis has to be written for each physics channel one wants to analyze.

In order to ease the writing of data analysis code, a software-generator has been conceived and built within this work: WatchMan [163]. The main idea is that the users insert the settings and the parameters of their physics analyses in an easy way, and the final analysis code is dynamically and automatically generated, ready to be run on data.

In the following sections more details about the motivation behind the development of WatchMan and about its implementation will be provided, starting with the main idea, then its architecture, continuing with the modular interfaces

and the user front-end; and ending with a simple example of analysis implementation in WatchMan.

The work presented in this chapter has been presented to the Physics and Computing communities during two talks: the first one at the *"13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research"*, on February 22–27 2010, in Jaipur, India; and the second one at the *"EuroSciPy 2010 - 3rd European meeting on Python in Science"*, on July 8–11, 2010, in Paris, France. From those talks, two papers followed: a conference proceedings paper [77] and a peer-reviewed journal article [27].

7.1 The problem: analyzing billions of HEP data in complex software frameworks

To analyze data physicists have to know, or learn, how to access and select data, interacting with the experiment software framework through dedicated tools. The frameworks which run today's experiments like ATLAS, are very large and complex; thus writing the analysis code is not always an easy task, and the learning curve is usually quite steep. So physicists not only have to know the physics and how to test it, writing the right physics algorithms, but they also have to learn very many technicalities about the custom software framework of the experiment, related to common operations like configuring and accessing the framework packages, importing necessary modules, loading data, accessing data, accessing containers in data files, looping over collections of particles, booking and filling histograms, handling files and so forth. So, beside the physics-related lines of code, physicists usually have to write a large amount of framework-related code, which has nothing to do with the physics one wants to explore, but they are necessary in order to run the code to analyze data; and for some tasks the amount of such extra-code is much larger than the amount of code that is directly related to the physics analysis.

Moreover, usually, for each physics channel, a dedicated analysis is set up, in order to have a better control on it and being able to cross-check the results among all the groups of researchers involved in such an analysis, validating in this way the final results. But very often the analyses that have been set up for different channels, they actually share a large part of common code. And very often the analysis code for a new channel is made by copy-and-paste, starting from another, similar, piece of analysis code and changing only the parts that are relevant to the given analysis. And when one starts to work on different analyses, the writing of so many pieces of analysis code can become tedious, time consuming and, most important, error prone.

A certain number of software frameworks and tools used in HEP experiments already tried to simplify the task, offering helper classes and functions,

but the right usage of them requires from the user, nevertheless, a certain knowledge about the framework beneath. Moreover, many of such frameworks are customized for a specific experiment, or bound to a specific data format, and only few are publicly available, i.e. not restricted to collaborators of a given experiment.

Thus, in order to ease the writing of data analysis code, a software-generator has been conceived and built within this work: the main idea is to let the users insert the settings and the parameters of their physics analyses, and the final analysis code is dynamically and automatically generated, ready to be run on data. Moreover, the software-generator accepts different data format and framework interfaces as plug-in modules, in order to make it usable by physicists working on different experiments or with different data files.

This software-generator tries to imply, as much as possible, the separation of concerns: a first layer is the physics analysis itself, with its rationale and its algorithms; while the second layer is the coding part, involving the framework of the experiment and the specific data format. In this view the user only takes care of the physics part, directly and easily translating an idea into analysis code, leaving all the technical subtleties and details of communicating to the experiment framework and handling data to the machinery beneath.

As implementation of this idea, WatchMan, a "data analysis construction kit" and a highly automated analysis code generator, has been conceived and developed. WatchMan takes as input the parameters and the user's settings from a simple text-like "steering file"; then, after having parsed the input file, it dynamically generates the complete analysis code, ready to be run over data, both locally or on the GRID. WatchMan has been implemented in `Python` and `C++`, using *Computer Aided Software Engineering* [167] (or, briefly, CASE in the following) principles. And it can be interfaced to different data formats or experiments via a modular interface mechanism. WatchMan implements a new idea in the HEP field, the usage of CASE to build reliable, easy to maintain and easy to validate data-analysis code, mainly aimed at analyzing new data from the experiments at the LHC collider.

WatchMan has been used with success at the beginning of the LHC data taking, in order to analyze the first data collected by the ATLAS experiment. It has been used as the main data-reading tool within the Freiburg group, and partially by the ATLAS SUSY Group, to implement different Supersymmetry analyses, and for the first version of the model-independent General Search presented in Chapter 6 of this work. Then, as the experiment evolved, new centralized tools for data reading have been developed and maintained by the ATLAS Physics Analysis Group (PAT). At that point, the ATLAS layers have been moved to a plug-in module, and WatchMan has been then converted to an open-source general public tool for analyzing data. A new interface to the publicly available Delphes format [182] was added, and the code has been made public on the web.

7.2 A CASE package to automatically generate HEP data analysis software

As said in the previous section, other software packages and frameworks have been developed, to help physicists to write analysis code. But the problem with all those tools is that the analysis code has to be written by the user anyway — usually using the `C++` or `Python` programming languages — and a certain knowledge about the framework beneath is required.

The approach of WatchMan is different from that one adopted by other tools. What this project does, is providing to the user a tool to automatically generate the analysis code, instead of helping him to write it.

The main idea is an ideal separation between the analysis strategy, related to the physics one wants to explore, and the actual and practical implementation of such an analysis.

In such a view the process of analyzing physics data — from the point of view of a physicist — should rely mainly on the design of the analysis and on hypothesis tests: with the ideal possibility of testing many hypotheses in an easy and straightforward way. But nowadays this is burdened by the need of writing complex code.

As said before, in order to write analysis code, physicists are now forced to learn a large amount of technicalities about the software framework of the experiment, beside of the analysis techniques and the physics they want to explore. Then a lot of lines of extra-code has to be written, to configure the analysis. And with some software frameworks, or for certain data formats, the amount of such extra-code can be much larger than the code directly related to the physics analysis.

Moreover, when one starts to work on different analyses, the number of such pieces of software grows, and those files become very difficult to be written, maintained, updated and validated. For example, let us imagine what would happen if the name of several containers changed in the data file format; or if a function to access a piece of data changed in the experimental framework; things that can happen, especially in experiments which rapidly evolve. In that case one would need to open and edit all files to get rid of the changes, and that can be an important source of errors.

Another source of errors, while writing repetitive code, is recurring formulas. Very often, in physics analysis code, the same formula is used many times and places — for example in cut-based analyses, where one selects particles and events through boolean expressions on thresholds values of specific quantities, like the particle energy or the direction of its trajectory. Such formulas are often quite simple and small, but they are used many times and in many different places, in order to implement the analysis strategy. In such a scenario, changing a formula

is a tedious task, forcing the programmer to change it in all places, going through all the pieces of software where it is used. And that is even more error prone if the analysis code is based on large and complex formulas.

WatchMan addresses all those problems. As said, WatchMan has been build as a CASE package, which means that the final code is automatically written, based on the user's settings. Easing, in such way, the development and the maintenance of the HEP analysis code.

WatchMan is the first "HEP analysis code generator" on the market, and its main goal is to relieve the user, as much as possible, of the burden of writing the analysis software code, letting her or him concentrate over the strategy of their physics analyses. In Figure 7.1 the main idea of WatchMan is presented: trying to provide to the user a tool to implement a modern HEP analysis as one would do with a pencil on a sheet of paper, just directly translating the ideas into actual analysis code, without wasting time with framework-related technicalities and, at the same time, avoiding introducing errors.

WatchMan takes care of the common code necessary to run the analysis, besides helping with the design and the implementation of the analysis, providing the most common algorithms used in HEP as building-blocks.

After having parsed user's settings, WatchMan adds the right code to implement the common or recurring tasks and the calls to the functions of the underlying framework, and generating the analysis code.

And a library of recurring common formulas has been also included in WatchMan, to be used by users in their analysis. In this way, changing a formula is easy: it only has to be changed once and in one place only, inside the collection of formulas. The presence of the formula library and the fact that the code is automatically generated, makes the analysis code very easy to validate: once the formulas in the collection are validated, all the generated analysis code is automatically validated as well, with the only exception of custom code added by the single user.

Another valuable feature of WatchMan is the possibility to handle many analyses at the same time: the user can define as many analyses as needed, and WatchMan generates the code for all of them together. Once the generated code is then run over the data, the output file contains the results for all the analyses, while remaining quite small and lightweight. This is made possible by a flagging mechanism: all objects inside the output file are flagged according to the analysis they belong. In this way, while sharing objects and disk space, the different analyses are still well separated. More on this in Section 7.7.

The frameworks used in HEP experiments, like Athena [93] in ATLAS, generally provide low- to medium-level functions and methods to access and analyze data. WatchMan, on the contrary, offers a very high-level interface, providing tools to access and handle data that are independent of the data format or of the

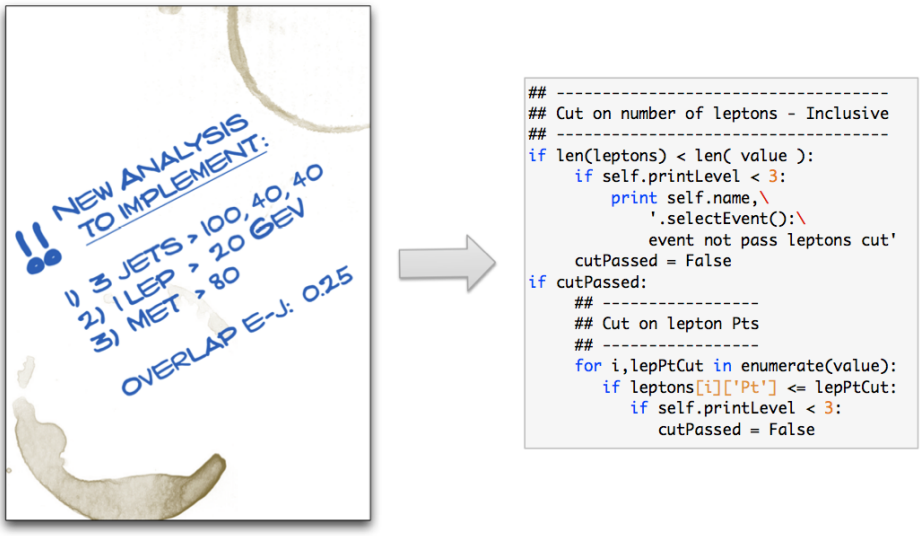


Figure 7.1: Main idea of WatchMan: letting the user easily implement analysis ideas.

underlying software framework. In WatchMan, which is also a framework itself, interfaces for different experiments and data formats can be easily added, and the user can expand it adding custom formulas to be used in the data analysis. Three interfaces for three data formats have been provided with the framework, so far, as detailed in Section 7.4.

As said before, the user's settings are provided in a text-like form in a "steering file", without taking care of the given interface. Thus, the same steering file can be used to define the same physics analysis using different frameworks or different data formats, simply by changing the corresponding option.

7.3 The WatchMan framework

WatchMan is a framework featuring a core containing common algorithms, which are presented in this section, and modular interfaces for specific experiments or data formats, which will be described in Section 7.4. A parser which takes user's input, and which handles all the modules, completes the layout.

Figure 7.2 shows the main layout of the WatchMan architecture (only the main components and blocks are shown here).

The user interacts with the framework through a "`Steering File`", a text-like file where the analyses can be defined in a simple way using `Python` dictionaries. The "`CutsLib`" module is a collection of common formulas and algorithms used in HEP analyses; the usage of those instead of implementing the same formula for

every analysis class, makes the analysis code virtually error-free and easier to validate: once the built-in formulas are validated, all the analysis code generated by WatchMan is automatically validated as well. In the "`Modular Interface`" the details of the specific framework or data format are provided. As written above in Section 7.1, three interfaces are provided with the package so far, and others can be added by users, as it will discussed in Section 7.4. The "`Parser`" module is the core of WatchMan. It is the code generator: it parses the user settings, combining them with the specific experiment interface and with the common algorithms, and it generates the final `Python` analysis code: the "`GeneratedAnalysisLib`" in the figure. Beside that, it also generates scripts to run the analysis on data — locally or on the GRID network — which contain instructions related to the specific data format in use.

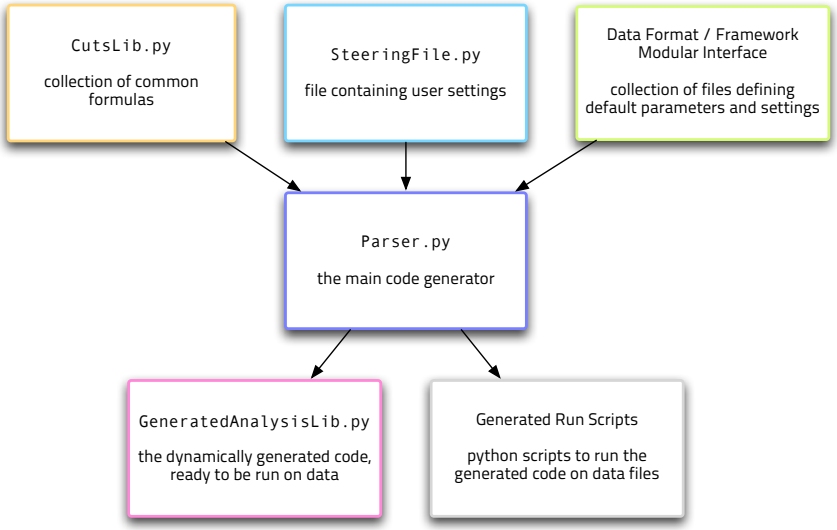


Figure 7.2: The main layout of the components of WatchMan. The "`Parser`" is the core engine of the framework, whose duty is to combine user settings with common code, to generate a complete analysis code ready to be run on data.

The whole framework has been written with the `Python` programming language, and it makes use of the `ROOT` framework [194] and its `Python` bindings `PyROOT` [188] to read the input data files and to store the output of the analysis.

WatchMan, at first, was born as a tool to handle code generation within the ATLAS experiment, where it was presented as "ATLASWatchMan". Then the software was reshaped, expanded and partially rewritten, adding the modular interface mechanism described in the next section. After the restructure of the code to make it publicly available as an open-source project, the ATLAS-related code has been moved in a separate interface plug-in module.

7.4 The modular interfaces

The core packages of WatchMan do not contain any code related to a specific experiment or data-format. All the code for a specific framework, is provided to the parser via a modular interface mechanism, whose layout is shown in Figure 7.3. For each interface a specific set of files has to be provided. Then the parser blends those information with user settings, building the final analysis code.

Three interfaces have been provided with the framework until now: an interface to the publicly available Delphes [182] data files, and another for two different data formats used in the ATLAS experiment: ESD/AOD and D3PD. Other additional interfaces can be set up by the user, in a modular way.

Each interface provides the specific instructions for a given system or format: for example the methods to call the external packages to set up the environment, or the list of the names of the containers storing the physics objects within the data file; or also the implementation of the methods that return the physical properties of the objects, for the given data format.

As said before, the interface can be written in `Python` or `C++`, according to the requirements of the data format or the experiment framework one has to interact with. When written in `C++`, `Python` bindings are built in an automated way, through dictionaries [189], automatically built with the tools provided by ROOT [194].

More details about the components of the interface, and instruction about adding new interfaces to the framework, are described in the WatchMan wiki [163].

7.5 The user's interface: the Steering File

The so-called "steering file" is the only final user interface, so far. It is a Python file containing only dictionaries and variables, where the final user inserts configuration parameters and custom code in a text-like way. The look&feel of such a `Python` file is that one of a simple text file with labels and fields to be filled, and the user, even if not so used to `Python` code, is not bothered with more complex syntax, like with other languages.

An example of steering file usage is shown in Figure 7.4, where a simple SUSY-like example analysis is defined as explained in Section 7.6[1].

Each steering file contains four main sections: a first part where global options are defined (like folders, file names, etc. . .); a second part where the analyses and the containers of the output file are chosen; a third part where the user can set the data files which the generated code will be run upon; and a last fourth part where the user can insert custom formulas to be used in the analyses.

[1] A complete example of a *steering file* can be found at [148]

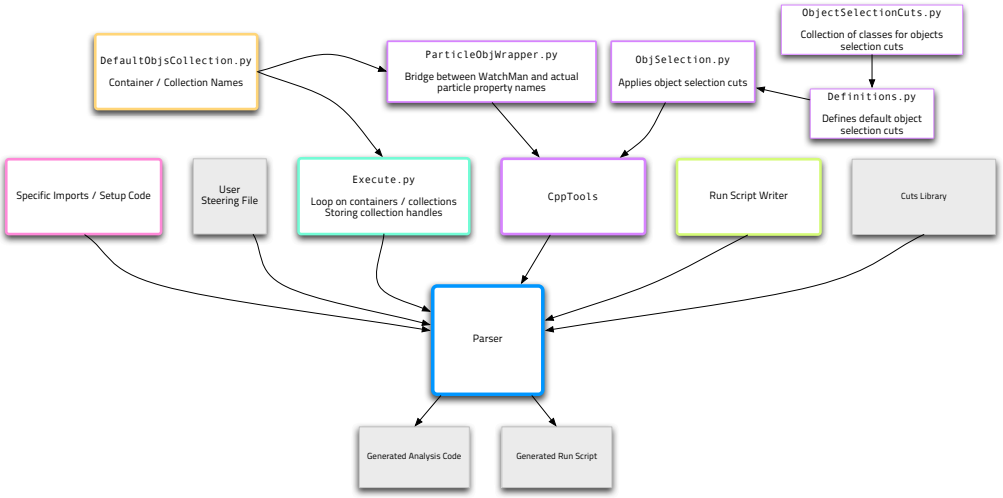


Figure 7.3: The main components of the WatchMan modular interface. Data-format or experiment-related settings are specified in the files of the modular interface. The parser blends that information with the user settings to build the analysis code.

In the next session an example of an implementation of a simple analysis with the WatchMan analysis code generator will be presented.

7.6 An example: how to implement a simple SUSY-like analysis with WatchMan

In the following an example of implementation of a simple cut-based SUSY-like analysis will be discussed.

Skipping the explanations about the first and third parts of the steering file, which are mainly lists of file-system paths and boolean values used for the setup — and whose instructions can be find in the WatchMan web site [163] —, the discussion here will focus on the description of the second and the fourth part of the steering file, concerning the implementation of the physics analysis.

The second part of the steering file is where the user defines the analyses, as shown in Figure 7.4.

The user starts by defining a channel, selecting a given number of particles of a certain type, with various properties; in this example the user defines a channels with one lepton, 3 hadronic jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the final state. And the user wants to apply a given number of cuts to the physical properties of such objects (like the p_{T}), or of the whole event (like the M_{eff}[2] or the *Sphericity*[3]).

[2]The M_{eff} is defined in Equation 6.2.

[3]The *Sphericity* is the measure of how spherical a collision event is, i.e. the measure of how

```
1  channels = {
2
3      '3j1lep': {
4
5            'channel': 'ljjjv',
6
7            'objSelection': {'electron':{'ptMin': 20.*Units.GeV,
8                                         'deltaR_ej': 0.25},
9                            },
10
11           'cuts': {  1:  { 'label': 'leptonPtCutsExclusive',
12                            'value': [20*Units.GeV]},
13
14                      2:  { 'label': 'jetPtCuts',
15                            'value': [100*Units.GeV, 40*Units.GeV,
16                                      40*Units.GeV],},
17
18                      3:  { 'label': 'jetPtVeto',
19                            'value': [40*Units.GeV] },
20
21                      4:  { 'label': 'missingEtCut',
22                            'value': 80*Units.GeV, },
23
24                      5:  { 'label':   'meff',
25                            'value':   100*Units.GeV,
26                            'formula': 'meff3JetsMetLeps',
27                            'custom':  True },
28                   },
29     },
30 }
31
32 #---
33 collections = {'electron':{'select':True}}
34 #---
35 userBranchesToFill = { 'meff4j0lep' : {
36           'label':'meff4j','type': 'float','formula': 'meff'} }
37 #---
38 dumpContainers = {'jet':{}, 'electron':{}, }
```

Figure 7.4: Example of physics analysis implementation using the *steering file*.

In the steering file, the container used for those definitions is the "`channels`" `Python` dictionary (at line 1 in the listing), whose main entries are the different physics analyses defined by the user, identified by a string label as key (the "`3j1lep`" in the example, at line 3). Each channel is defined by a label, by a set of cuts used for the object selection, and by a set of sorted cuts that are used for the event selection, and which are related to the particular channel one wants to analyze. Three main sections can be configured for that: respectively, "`channel`", "`objSelection`" and "`cuts`". The first one, "`channel`" (at line 5), is a string saying which objects (particles) are going to be analyzed in that specific channel, and it is then used for labels, plot names and some internal checks. The second key of the dictionary is the "`objSelection`" key (at line 7 in the listing). Here the user

spherical is the spatial distribution of all objects in the event.

can set the criteria used to select the objects inside data; in this simple example the user sets threshold values on physics quantities (here a certain energy for `electrons` is requested, and a certain angle for their trajectories[4]); many other options and requirements can be added. Objects passing such first-stage selection will be flagged as described in Section 7.7 and shown in Figure 7.7 on page 142.

The definition of the analysis continues then in the "`cuts`" sub-dictionary (starting at line 11 in Figure 7.4), which sets the second stage: the event selection. Here the user can enter as many selection criteria as wanted: each entry has an integer number as key, defining the order of application of each criterion. Here integers are used as hash-able keys, since the filters on the physical quantities have to be applied in a specific order. Practically the user sets a sorted sequence of actions, whose integer key sets the position in the sequence. The same order will then be used when generating the code, so that the actions will be taken on data in the right sequence. As you can see in the listing (for example at lines 14–16), each criterion has a key "`label`", whose value is a string (e.g. "`jetPtCuts`"); that string is the name of the formula, or the algorithm, which will be applied. That name is associated to a formula stored in the built-in library (the "`CutsLib`" module shown in Figure 7.2), and it is used simply by calling its name in the steering file. Arguments to the formula are passed by the user in the "`value`" entry (for example at lines 15–16 the energy threshold values are passed as arguments to the "`jetPtCuts`" formula).

Running WatchMan on this steering file, the complete Python analysis code is automatically generated, and ready to be run on data. In the Figure 7.5 a very simple example of the generated code is shown; more complex selection criteria trigger the generation of more complex and longer pieces of code. In Figure 7.5, on the right side, only the snippet directly generated from the settings shown on the left side is presented; but more code is actually generated, as already stated before: for example the code needed to connect to interact with the experiment framework and to access data.

The example steering file here, in Figure 7.4, is the direct implementation of the "napkin scribble" shown on Figure 7.1 on page 134; that is the goal of WatchMan: from an idea about a great physics analysis, to the complete analysis code, in few minutes!

It is worth to notice that a whole complete, yet very simple, analysis, like that one taken as example, would have required a certain — large — amount of lines of code to be written, without using WatchMan; and it would have also required also a certain — usually long — developing and debugging time. The user — who is a physicist — ideally should not need to know how to program the analysis; he should just have to plug in the strategy in a simple way and

[4]More in details, a value of at least 20 GeV is requested by the user for the transverse momentum of the electrons, and a angular distance $\Delta R > 0.25$ is requested between the trajectory of the electron and a jet.

```
'cuts': {
  1:{'label':'electronPtCut',
     'value': [10*Units.GeV]
    }
}
```

```
## ---------------------------------
## lepton number cut - Inclusive
## ---------------------------------
if len(leptons) < len( value ):
  if self.printLevel < 3:
     print "not passing leptons cut"
  cutPassed = False
if cutPassed:
  ## -----------------
  ## Cut on lepton Pts
  ## -----------------
  for i,lepPtCut in enumerate(value):
     if leptons[i]['Pt'] <= lepPtCut:
       if self.printLevel < 3:
          print "not pass lepPtCut"
       cutPassed = False
```

This line in the steering file... ...generates this snippet of analysis code.

Figure 7.5: How the generation of code works: on the left side a line of the analysis definition, set by the user in the steering file; on the right side the actual snippet of code automatically generated from that line. Actually the line on the left side triggers the generation of much more code, if including in the count all the code needed to connect to the experiment framework, to access data in data sets, to build histograms, to handle the output, and so forth.

obtain the results, without taking care of the software details underneath. And WatchMan is a first attempt to get that.

In the steering file the user can also decide what to read from the data files and what to store in the output file: which object collections or variables. In the same Figure 7.4 at line 33 the user decides to read only the "`electron`" particle collection from the input data sets. At lines 35–36 the user decides also to store the outcome of the custom formula "`meff`" (discussed here below) as a variable called "`meff4j`", and at line 38 it sets that also the collections of electrons and jet-particles have to be stored in the final output file where the results of the analyses defined in the steering file are saved.

As said before, the main and most common formulas and algorithms are built-in in WatchMan, and they simply can be used in the analysis definition. But the user can also expand the built-in library adding custom formulas. Those can be specified in the last section of the steering file, in the dictionary "`userFormula`": each formula fits in a dictionary entry, whose key is a string with the formula name, as shown in Figure 7.6. Let's go back for a moment. In Figure 7.4, the last field in the "`cuts`" section is an entry called "`meff`" (at lines 24–27 of the listing), which is flagged as "`custom`". The user flags it as "`custom`" if it is not part of the built-in library but it is a custom formula she or he wants to add. To do so the user must add the formula definition in the last part of the steering file. Custom formulas have to be written in `Python`, as shown in the listing 7.6. Once the formula is defined there, and provided that the "`custom`" flag is set to "`True`",

```
##--- User-Defined Formula
userFormula = {
# Meff formula: Highest Pt 4 Jets + MET + All Leptons
'meff': {
'position':3,
'formula':
"""
meff = 0.
if len(candidates['jet']) < 3: return meff
for i,jet in enumerate(candidates['jet']):
    if i >= 3: break
    meff += getVal(candidates['jet'][i], 'Pt')
    pass
for i,el in enumerate(candidates['electron']):
    meff += getVal(candidates['electron'][i], 'Pt')
    pass
for i,mu in enumerate(candidates['muon']):
    meff += getVal(candidates['muon'][i], 'Pt')
    pass
meff += MET_corrected(candidates)
return meff
"""},
}
```

Figure 7.6: Example of custom user-defined formula to be used for cuts or to fill containers in the output file.

the custom formula can be used like those contained in the built-in library: the "`meff`" formula in the example is used in the analysis definition, and values are passed to it in the same way as to built-in formulas, through the key "`value`". A same custom formula can be used in the definition of one or more analyses, or to fill a container in the output file.

Here the choice of the `Python` programming language helped a lot. Its dynamic and interpreted nature, in fact, let the user insert a custom formula in the steering file without having to strongly typed variables and without having to compile and linking it, like the user has to do when using other languages, like `C++`. So the user can enter a custom formula in the steering file and run the generation of the code directly, without passing through a compilation phase which would make the whole process more complex and slower.

Concerning security, so far there are no routines to check and sanitize the user's inputs to formulas. In fact, even if the tool is not intended to be exposed to other people different than the user, and even if it is run only on private workstations and not exposed on the web, a check on the formula inserted by the user should be performed anyway, in order to verify that the code is not accidentally harmful for data or for the underlain system. Checks on input formulas and values may be implemented in future releases of the package.

7.7 The WatchMan output file and the objects flagging mechanism

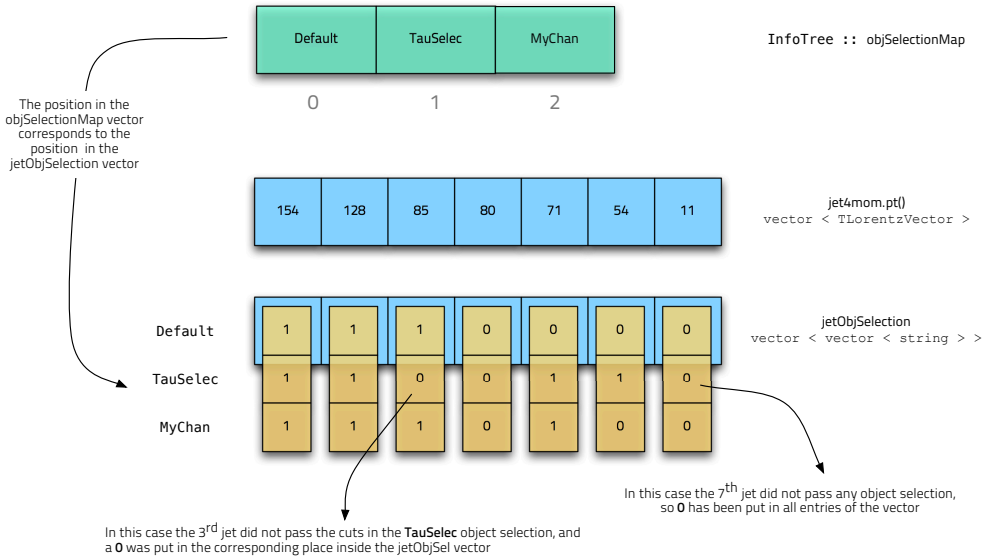


Figure 7.7: How the `ObjSelection` branch is filled for particle-like objects.

Once the data analysis code generated by WatchMan is run over the data, the output of all the analyses is collected in a unique file, to save storage space and to let the user easily combine together or compare the different analyses. But, as already said in the introduction, WatchMan also has to ensure that each analysis can be kept well separated from the others, particularly for validation purpose.

And here the choice of `Python` helped again: its dynamic-typed nature allowed the build of a flagging mechanism which flags all the objects in data, at run-time, without having to know the type of the object *a priori*.

Each object in data (e.g. particles, events, physics quantities, ...) can "belong" to one or more analyses, according to the requirements and selection criteria which have been applied in the algorithms. All objects are stored only once in the output file, but they are flagged according to the analysis or the analyses they belong. In Figure 7.7 an example of this flagging mechanism for the object selection is shown. In this example a collection of "jet-particles" is taken, and the energy associated with each of them is considered. Each particle is stored only once in the output file, but when it satisfies the different energy selection requirements of the different analyses, it is flagged accordingly, and the flags are saved with the object itself. More in details, in Figure 7.7 the middle row flagged as "`jet4mom.pt()`" is

the container storing the value of the energy for each jet-particle in one event: "`154`" is the value of the energy of the first jet-particle, "`128`" is the value of the second one, and so on. The upper row flagged as "`InfoTree::objSelectionMap`", instead, is the map storing all the selection criteria the user set in the steering file, grouped under the name given to a specific set of requirements. In Figure 7.7 the map that stores the object selections defined as example in the listing of Figure 7.8 is shown. There (as a part of a steering file) the user globally defined two sets of object-selection criteria, under two names: "`TauSelec`" and "`MyChan`" (at lines 3 and 8). Those two sets are used in the definition of the analyses; in the example the selection "`TauSelec`" is used by the analysis "`tau_channel`" (line 17–20 in the same Figure 7.8). The names of the two sets are then stored in the output file, inside the map shown on Figure 7.7 (top row), and they are used to flag the objects which have passed the requirements. In the same example figure the bottom row, flagged as "`jetObjSelection`", contains the flags for the corresponding particles; for example the first particle-jet (whose energy value is "`54`") passed both requirements the user defined for the selection "`TauSelec`" and the particle gets a "`1`" in the vector entry corresponding to the "`TauSelec`" position in the map. The same for the requirements set for the "`MyChan`" selection. If the particle does not pass the requirements — for example the last one, whose energy value, "`11`", is too low to pass the energy requirement "`ptMin`" — the corresponding flag is set to "`0`". "`Default`" is the default object selection requirements applied if the user does not specify custom ones.

This flagging mechanism let WatchMan handle and save all the analyses together at the same time, still keeping each analysis well separated from the others; with a light-weight and fast to be read output file.

The final output, which is produced running the generated code over the data, and which stores the results of the analyses, is stored in a "`.root`" file. This is a file format of the open-source `ROOT` analysis framework [194], which contains trees and branches of saved data. We chose this format mainly because it is the *de facto* standard within the HEP community, but also because it is a well suited format to study, compare and combine event-based analyses, like those commonly conceived and run in particle physics. The output "`.root`" files can be then read with the ROOT framework itself — via its interpreter or via `C++` code — or, more comfortably, through the Python binding `PyROOT` [188].

7.8 Conclusions

WatchMan implements a new idea in the HEP field, the usage of a *Computer-Aided Software Engineering* principles to build reliable, easy to maintain and easy to validate data analysis code, mainly aimed at analyzing new data from the experiments running at the LHC collider at CERN.

```
1  objectSelectionAndOverlap = {
2
3    'TauSelec': {'jet':{'ptMin': 40.*Units.GeV
4                        'etaMax': 2.5},
5                 'tau':{'applyOverlapRemoval':True},
6                },
7
8    'MyChan': {'jet':{'ptMin': 60.*Units.GeV,
9                      'etaMax': 2.5},
10             },
11   }
12
13
14 #---
15
16 channels = {
17     'tau_channel': {
18           'channel': 'ltX',
19           'objSelection': 'TauSelec'
20           ...
```

Figure 7.8: Definition of custom user-defined object-selection requirements in the *steering file*. In this way it is possible to globally define the requirements, and then using them in different analyses, as in the ‘‘`tau_channel`’’ in the example here above (lines 17–19). The object-selection sets get stored then in a map in the output file, to be used by the flagging mechanism, as explained in the text of Section 7.7.

WatchMan is an analysis code generation kit, written in `Python`, which handles many analyses at the same time, generating the final complete analysis code, ready to be run on data. The framework also takes care of the specific data format setup, relieving the final user of the need of learning the details of it. And it can be expanded with modular interfaces to work with new formats.

WatchMan is now an open-source project, with a first stable release, and it has been used with success to analyze the very first data of the LHC collider and to produce some scientific notes and contributions at the ATLAS experiment (among the public ones, see for example [101] and [154]).

Summary

In this Thesis, a new model-independent General Search for the ATLAS experiment at the LHC has been conceived, realized and presented. The motivations that leaded to the decision to implement such an analysis have been discussed, and the implementation of its strategy has been described and discussed.

The analysis is aimed at analyzing events with high-p_{T} objects, with a very loose selection, to be able to look for signs of new physics in as many physics channels as possible. The analysis has been compared to a standard ATLAS Supersymmetry analysis, in order to validate the selection cuts and the design choices. A new algorithm for an automated scan of data has been implemented as well, in order to search interesting discrepancies between data and the Standard Model background, and to quantify the significance through a statistical interpretation taking into account the "look-Elsewhere" effect introduced by the large number of final states that have been analyzed. In the end preliminary results from the first 2 fb^{-1} of data collected by the ATLAS experiment have been presented.

No evidence of new physics has been observed. Instead, the data showed a very good agreement with the Standard Model expectation, within statistical uncertainties, in the whole spectrum of the very many different final states analyzed, which are the signatures of physics processes very different one respect to the others. And this agreement is a valuable result: it reflects, in fact, the very good operation of the ATLAS experiment: from simulation to reconstruction, from data taking to calibration, everything looks like working in a very well coordinated way, giving very well results when comparing so many different final states together.

As already said, no evidence of new physics has been observed, but the strategy and the methodology of the new model-independent analysis for ATLAS have been set, defined and implemented, ready to be used in a future version of this General Search, aimed at analyzing a larger set of experimental data from the LHC.

Moreover, another problem has also been analyzed during this Thesis, and

a solution has been found and realized. From the need of a suitable software framework to handle very many analyses on different channels, the WatchMan package implements a new idea in the HEP field: using *Computer-Aided Software Engineering* (CASE) principles and techniques to automatically generate physics analysis code, from users' settings, in a reliable and easy way. WatchMan has been conceived and realized from scratch within this work, using different programming languages and software development techniques. At first integrated within the ATLAS Athena framework, WatchMan has been used to efficiently analyze the first data from LHC, proving its successful design. It has then be moved to a modular architecture, and presented at international Physics and Computer Science conferences as a new public open-source CASE framework for data analysis.

Bibliography

The numbers at the end of the headings reference the pages where the publications are cited

Books

[1] George E.P. Box and Norman R. Draper. *Empirical Model-Building and Responsive Surfaces*. Ed. by Wiley. 1987 (cit. on p. 79).

[2] Richard Feynman. *QED: The strange Theory of Light and Matter*. Ed. by Princeton University Press. 1985.

[3] Gian Francesco Giudice. *A Zeptospace Odyssey*. Oxford, 2009.

[4] D. Griffiths. *Introduction to Elementary Particles*. 2^{nd}. John Wiley & Sons, 2008 (cit. on pp. 5, 20).

[5] F. Mandl and G. G. Shaw. *Quantum Field Theory*. John Wiley & Sons, 1993 (cit. on pp. 5, 12).

[6] M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory*. 1995 (cit. on p. 20).

Journal articles

[7] Kaustubh Agashe et al. "Warped Gravitons at the LHC and Beyond". In: *Phys. Rev. D* 76 (2007). doi: arXiv:hep-ph/0701186 (cit. on p. 32).

[8] S. Agostinelli et al. "GEANT4: a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303 (cit. on p. 85).

[9] Benjamin C. Allanach, Sebastian Grab, and Howard E. Haber. "Supersymmetric monojets at the Large Hadron Collider". In: *Journal of High Energy Physics* (2011). doi: 10.1007/JHEP01(2011)138. url: http://arxiv.org/abs/1010.4261.

[10] G Altarelli. "Experimental Tests of Perturbative QCD". In: *Annual Review of Nuclear and Particle Science* 39.1 (1989), pp. 357–406. DOI: 10.1146/annurev.ns.39.120189.002041. eprint: http://www.annualreviews.org/doi/pdf/10.1146/annurev.ns.39.120189.002041. URL: http://www.annualreviews.org/doi/abs/10.1146/annurev.ns.39.120189.002041.

[11] G. Altarelli, R. Barbieri, and F. Caravaglios. "Electroweak precision tests: A concise review". In: *Int. J. Mod. Phys. A* 13 (1998) (cit. on p. 17).

[12] Luis Álvarez-Gaumé and John Ellis. "Eyes on a prize particle". In: *Nature Physics* (2011). DOI: doi:10.1038/nphys1874 (cit. on p. 13).

[13] Nima Arkani-Hamed, Savas Dimopoulos, and Gia Dvali. "The hierarchy problem and new dimensions at a millimeter". In: *Physics Letters B* 429 (1998), pp. 263–272. DOI: http://dx.doi.org/10.1016/S0370-2693(98)00466-3 (cit. on p. 30).

[14] ATLAS-Collaboration. "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV". In: *The European Physical Journal C - Particles and Fields* 73 (2013), p. 2304 (cit. on p. 58).

[15] ATLAS-Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716.CERN-PH-EP-2012-218 (2012). DOI: 10.1016/j.physletb.2012.08.020. URL: http://arxiv.org/abs/1207.7214 (cit. on pp. 13, 21).

[16] ATLAS-Collaboration. "Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS". In: *Eur.Phys.J. C* 72 (2012). DOI: 10.1140/epjc/s10052-011-1844-6. URL: http://arxiv.org/abs/1108.5602 (cit. on p. 60).

[17] ATLAS-Collaboration. "Search for Dark Matter Candidates and Large Extra Dimensions in Events with a Photon and Missing Transverse Momentum in pp Collision Data at s=7TeV with the ATLAS Detector". In: *Phys. Rev. Lett.* 110 (2012). DOI: 10.1103/PhysRevLett.110.011802. URL: http://link.aps.org/doi/10.1103/PhysRevLett.110.011802.

[18] ATLAS-Collaboration. "Search for Extra Dimensions in diphoton events using proton-proton collisions recorded at sqrt(s) = 7 TeV with the ATLAS detector at the LHC". In: *New J. Phys.* 15 (2012). DOI: arXiv:1210.8389 (cit. on p. 32).

[19] ATLAS-Collaboration. "Search for high-mass resonances decaying to dilepton final states in pp collisions at a center-of-mass energy of 7 TeV with the ATLAS detector". In: *JHEP* 1211 (2012). DOI: arXiv:1209.2535 (cit. on p. 32).

[20] ATLAS-Collaboration. "Search for new phenomena in the WW to l nu l' nu' final state in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector". In: *Phys. Lett. B* (2012). URL: http://arxiv.org/abs/1208.2880.

[21] ATLAS-Collaboration. "Search for second generation scalar leptoquarks in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector". In: *Eur.Phys.J. C* 72 (2012). DOI: arXiv:1203.3172 (cit. on p. 34).

[22] ATLAS-Collaboration. "Search for TeV-scale gravity signatures in final states with leptons and jets with the ATLAS detector at sqrt(s) = 7 TeV". In: *Phys. Lett. B* 716 (2012). DOI: 10.1016/j.physletb.2012.08.009. URL: http://arxiv.org/abs/1204.4646.

[23] V. E. Barnes et al. "Observation of a Hyperon with Strangeness Minus Three". In: *Phys. Rev. Lett.* 12 (8 1964), pp. 204–206. DOI: 10.1103/PhysRevLett.12.204. URL: http://link.aps.org/doi/10.1103/PhysRevLett.12.204.

[24] K. G. Begeman, A. H. Broeils, and R. H. Sanders. "Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics". In: *Mon. Not. Roy. Astron. Soc.* 249 (1991) (cit. on p. 21).

[25] G. Bertone, D. Hooper, and J. Silk. "Particle dark matter: Evidence, candidates and constraints". In: *Phys. Rept.* (2005). arXiv:hep-ph/0404175, pp. 279–390. URL: http://fr.arxiv.org/abs/hep-ph/0404175 (cit. on p. 22).

[26] Gianfranco Bertone. "The moment of truth for WIMP Dark Matter". In: *Nature* 468 (2010). arXiv:1011.3532 [astro-ph.CO]. DOI: 10.1038/nature09509. URL: http://arxiv.org/abs/1011.3532 (cit. on p. 27).

[27] Riccardo Maria Bianchi and Renaud Brunelière. "WatchMan project — A Python CASE framework for High Energy Physics data analysis in the LHC era". In: *Journal of Computational Science* 4 (2013), pp. 325–333. ISSN: 1877-7503. DOI: 10.1016/j.jocs.2012.04.005. URL: http://www.sciencedirect.com/science/article/pii/S1877750312000336 (cit. on p. 130).

[28] Nicola Cabibbo. "Unitary Symmetry and leptonic decays". In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533.

[29] M. Cacciari and G. P. Salam. "The anti-kt jet clustering algorithm". In: *Journal of High Energy Physics* 4 (2008), p. 63 (cit. on p. 58).

[30] S. Caron and B. Knuteson. "QUAERO@H1: an interface to high-pT HERA event data". In: *The European Physical Journal C - Particles and Fields* 53 (1 2008). 10.1140/epjc/s10052-007-0468-3, pp. 167–175. ISSN: 1434-6044. URL: http://dx.doi.org/10.1140/epjc/s10052-007-0468-3.

[31] CDF-Collaboration. "Model-independent and quasi-model- independent search for new physics at CDF-model- independent search for new physics at CDF". In: *Phys. Rev. D* 78 (1 2008), p. 012002. DOI: 10.1103/PhysRevD.78.012002. URL: http://link.aps.org/doi/10.1103/PhysRevD.78.012002 (cit. on pp. 64, 71–73).

[32] CDF-Collaboration. "Search for WW and WZ resonances decaying to electron, missing ET, and two jets in ppbar collisions at sqrt(s) = 1.96 TeV". In: *Phys. Rev. Lett.* 104 (2010). URL: http://arxiv.org/abs/1004.4946.

[33] M. Ciuchini et al. "2000 CKM — Triangle analysis: A Critical review with updated experimental inputs and theoretical parameters". In: *JHEP* (2001).

[34] Marco Ciuchini and Luca Silvestrini. "Introduction to flavour and CP violation in the Standard Model and beyond". In: *Comptes Rendus Physique* 13.2 (2012), pp. 115 –120. ISSN: 1631-0705. DOI: 10.1016/j.crhy.2011.11.001. URL: http://www.sciencedirect.com/science/article/pii/S1631070511002210 (cit. on p. 21).

[35] CMS-Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett. B* (2012). arXiv:1207.7235 [hep-ex]. URL: http://arxiv.org/abs/1207.7235 (cit. on pp. 13, 21).

[36] CMS-Collaboration. "Search for signatures of extra dimensions in the diphoton mass spectrum at the Large Hadron Collider". In: *Phys. Rev. Lett.* 108 (2011). DOI: arXiv:1112.0688 (cit. on p. 32).

[37] D0-Collaboration. "Quasi-Model-Independent Search for New High p_T Physics at D0". In: *Phys. Rev. Lett.* 86 (17 2001), pp. 3712–3717. DOI: 10.1103/PhysRevLett.86.3712. URL: http://link.aps.org/doi/10.1103/PhysRevLett.86.3712 (cit. on pp. 64–67).

[38] Abdelhak Djouadi. "The anatomy of electro-weak Symmetry breaking.I: The Higgs boson in the Standard Model". In: *Phys. Rept.* 457 (2008) (cit. on p. 14).

[39] Abdelhak Djouadi. "The anatomy of electro-weak Symmetry breaking.II: The Higgs boson in the Standard Model". In: *Phys. Rept.* 459 (2008), pp. 1–241.

[40] Alexander Friedland et al. "Probing nonstandard standard model backgrounds with LHC monojets". In: *Phys. Lett. B* 714 (2012), pp. 267–275.

[41] M. Gell-Mann. "A Schematic model of baryons and mesons". In: *Phys. Lett.* 8 (1964), pp. 214–215 (cit. on p. 9).

[42] Howard Georgi and S. L. Glashow. "Unity of All Elementary-Particle Forces". In: *Phys. Rev. Lett.* 32 (8 1974), pp. 438–441. DOI: 10.1103/PhysRevLett.32.438. URL: http://link.aps.org/doi/10.1103/PhysRevLett.32.438 (cit. on p. 21).

[43] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579 –588. ISSN: 0029-5582. DOI: 10.1016/0029-5582(61)90469-2. URL: http://www.sciencedirect.com/science/article/pii/0029558261904692.

[44] H1-Collaboration. "A general search for new phenomena in ep scattering at HERA". In: *Phys. Lett. B* hep-ex/0408044. DESY-04-140. DESY-2004-140 (2004).

[45] H1-Collaboration. "A general search for new phenomena in ep scattering at HERA". In: *Phys. Lett. B* 602.1-2 (2004), pp. 14 –30. ISSN: 0370-2693. DOI: DOI:10.1016/j.physletb.2004.09.057 (cit. on pp. 66, 68–70, 109).

[46] Howard E. Haber and G.L. Kane. "Gluino decays and experimental signatures". In: *Nuclear Physics B* 232 (1984), pp. 333–348. DOI: http://dx.doi.org/10.1016/0550-3213(84)90570-4.

[47] Howard E. Haber and G.L. Kane. "Signatures and possible evidence for Supersymmetry at the CERN collider". In: *Physics Letters* (1985). DOI: http://dx.doi.org/10.1016/0370-2693(84)91264-4.

[48] Howard E. Haber and G.L. Kane. "The search for Supersymmetry: probing physics beyond the Standard Model". In: *Physics Reports* 117 (1985). DOI: http://dx.doi.org/10.1016/0370-1573(85)90051-1 (cit. on p. 26).

[49] Thomas Hambye. "CP violation and the matter-antimatter asymmetry of the Universe". In: *Comptes Rendus Physique* 13.2 (2012), pp. 193 –203. ISSN: 1631-0705. DOI: 10.1016/j.crhy.2011.09.007. URL: http://www.sciencedirect.com/science/article/pii/S1631070511001873 (cit. on p. 21).

[50] F.J. Hasert et al. "Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment". In: *Physics Letters B* 46.1 (1973), pp. 138 –140. ISSN: 0370-2693. DOI: 10.1016/0370-2693(73)90499-1. URL: http://www.sciencedirect.com/science/article/pii/0370269373904991.

[51] F.J. Hasert et al. "Search for elastic muon-neutrino electron scattering". In: *Physics Letters B* 46.1 (1973), pp. 121 –124. ISSN: 0370-2693. DOI: 10.1016/0370-2693(73)90494-2. URL: http://www.sciencedirect.com/science/article/pii/0370269373904942.

[52] Peter W. Higgs. “Broken Symmetries and the masses of gauge bosons”. In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509 (cit. on p. 11).

[53] Peter W. Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Phys. Rev. Lett.* 12 (1964), pp. 132–133 (cit. on p. 11).

[54] I. Hinchliffe et al. “Precision SUSY measurements at CERN LHC”. In: *Physical Review D* 55.9 (1997), pp. 5520–5540. doi: 10.1103/PhysRevD.55.5520 (cit. on pp. 26, 83).

[55] Claus Horn. “A bottom-up approach to SUSY analyses”. In: *Journal of Physics G: Nuclear and Particle Physics* 36.10 (2009), p. 105005. url: http://stacks.iop.org/0954-3899/36/i=10/a=105005.

[56] N. Jarosik et al. “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, systematic errors, and basic results”. In: *Astrophysical Journal Supplement Series* (2010). doi: arXiv:1001.4744 (cit. on p. 21).

[57] M. Kobayashi and T. Maskawa. “CP violation in the renormalizable theory of weak interaction”. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657.

[58] Paul Langacker. “The Physics of Heavy Z' Gauge Bosons”. In: *Rev. Mod. Phys.* 81 (2008). doi: arXiv:0801.1345 (cit. on p. 33).

[59] C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell. “Observations on the Tracks of Slow Mesons in Photographic Emulsions”. In: *Nature* 160 (1947), pp. 486–492. doi: doi:10.1038/160486a0.

[60] C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell. “Observations on the Tracks of Slow Mesons in Photographic Emulsions”. In: *Nature* 160 (1947), pp. 453–456. doi: doi:10.1038/160453a0. url: http://www.nature.com/nature/journal/v160/n4066/pdf/160453a0.pdf.

[61] Christian Lippmann. “Particle identification”. In: *Nucl. Instrum. Meth. A* 66 (2012), pp. 148–172. doi: 10.1016/j.nima.2011.03.009. url: http://arxiv.org/abs/1101.3276 (cit. on p. 50).

[62] Hironari Miyazawa. “Baryon Number Changing Currents”. In: *Progress of Theoretical Physics* 36.6 (1966), pp. 1266–1276. doi: 10.1143/PTP.36.1266. url: http://ptp.ipap.jp/link?PTP/36/1266/ (cit. on p. 25).

[63] Emmy Noether. “Invariante Variationsprobleme”. In: *Nachr. d. Königl. Gesellsch. d. Wiss. zu Göttingen Math.-Phys. Klasse* (1918), pp. 235–257. url: http://www.physics.ucla.edu/~cwp/articles/noether.trans/german/emmy235.html (cit. on p. 8).

[64] S. Nussinov and R. Shrock. “Upper Limits on a Possible Gluon Mass”. In: *Phys. Rev. D* 82 (2010). url: http://arxiv.org/abs/1005.0850.

[65] J. Beringer et al. (Particle Data Group). "The Review of Particle Physics". In: *Phys. Rev. D* 86 (2012). URL: http://pdg.web.cern.ch/pdg/index.html (cit. on pp. 7, 58).

[66] J. Beringer et al. (Particle Data Group). "The Review of Particle Physics Booklet". In: *Phys. Rev. D* (2012). URL: http://pdg.lbl.gov/current/booklet.pdf (cit. on p. 7).

[67] K. Nakamura et al. (Particle Data Group). "Review of particle physics". In: *J. Phys. G* 37.075021 (2010). URL: http://pdg.lbl.gov/ (cit. on p. 7).

[68] L. Randall and R. Sundrum. "Large Mass Hierarchy from a Small Extra Dimension". In: *Phys. Rev. Lett.* 83 (1999) (cit. on p. 32).

[69] A. Salam and J.C. Ward. "Electromagnetic and weak interactions". In: *Physics Letters* 13.2 (1964), pp. 168 –171. ISSN: 0031-9163. DOI: 10.1016/0031-9163(64)90711-5. URL: http://www.sciencedirect.com/science/article/pii/0031916364907115.

[70] G. P. Salam. "Towards jetography". In: *The European Physical Journal C - Particles and Fields* 67.3 (2010), pp. 637–686 (cit. on p. 58).

[71] ERIC SATHER. "The Mystery of the Matter Asymmetry". In: *Beam Line* (1996). Spring/Summer (cit. on p. 21).

[72] Michael J. Wichura. "Algorithm AS 241: The Percentage Points of the Normal Distribution". In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 37 (1988), pp. 477–484. URL: http://www.jstor.org/stable/2347330.

Conference proceedings

[73] ATLAS-Collaboration. "A search for high mass resonances decaying to tau+tau- in the ATLAS detector". In: 36^{th} *International Conference on High Energy Physics*. Melbourne, Australia, 2012. URL: https://cds.cern.ch/record/1460263.

[74] Roger Barlow. "Systematic errors: facts and fictions". In: *Advanced Statistical Techniques in Particle Physics*. Ed. by Institute for Particle Physics Phenomenology. Institute for Particle Physics Phenomenology, Durham University. 2002. URL: http://www.ippp.dur.ac.uk/Workshops/02/statistics/proceedings/barlow.ps (visited on 2011).

[75] R M Bianchi. "Discovery potential of Supersymmetry and Universal Extra Dimensions in the ATLAS experiment at the Large Hadron Collider at CERN". In: *20th Hadron Collider Physics Symposium 2009, Evian-les-Bains, France, 16 - 20 Nov 2009*. Ed. by PoS Proceedings of Science. ATL-PHYS-PROC-2010-005. Evian-les-Bains, France, 2010, p. 66. URL: https://cdsweb.cern.ch/record/1234334.

[76] Riccardo Maria BIANCHI. "Recent updates of the Control and Configuration of the ATLAS Trigger and Data Acquisition System". In: *The 2011 Europhysics Conference on High Energy Physics, EPS-HEP 2011*. ATL-DAQ-PROC-2011-031. Grenoble, 2011. URL: http://cdsweb.cern.ch/record/1394278 (cit. on p. 57).

[77] Riccardo Maria Bianchi, Renaud Brunelière, and Sascha Caron. "WatchMan Project – Computer Aided Software Engineering applied to HEP Analysis Code Building for LHC". In: *ACAT 2010 - 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research*. Ed. by PoS Proceedings of Science. Jaipur, India, 2010 (cit. on p. 130).

[78] C.M. Buttar et al. "The Underlying Event". In: *HERA and the LHC – A workshop on the implications of HERA for LHC physics*. CERN. Geneva, 2008. URL: http://www.desy.de/~heralhc/proceedings/WG2-Buttar.pdf (cit. on p. 86).

[79] G. Dissertori. "The Strong Coupling Constant: An experimentalist's review". In: *9eme Rencontre Itzykson*. Saclay, 2004.

[80] John Ellis. "Searching for particle physics beyond the standard model at the LHC and elsewhere". In: *11th conference on "Frontiers of Fundamental Physics"*. arXiv:1102.5009. Paris, 2011. URL: http://arxiv.org/abs/1102.5009 (cit. on pp. 26, 27).

[81] Thomas Kittelmann et al. "The Virtual Point 1 Event Display for the ATLAS Experiment". In: *Proceedings of the 17th International Conference on Computing in High Energy Physics (CHEP09)*. Vol. Journal of Physics: Conference Series. 219. 2010 (cit. on pp. 18, 19).

[82] Greg Landsberg. "Black Holes at Future Colliders and Beyond - a Review". In: *SUSY 2002 Conference, DESY, Hamburg*. 2002. DOI: arXiv:hep-ph/0211043 (cit. on p. 31).

[83] Juan Maldacena. "Gravity, particle physics and their unification". In: *19th International Symposium on Lepton and Photon Interactions at High Energies*. Stanford, CA, USA, 1999, pp. 840–852. URL: http://arxiv.org/abs/hep-ph/0002092 (cit. on p. 20).

[84] Francesco Santanastasio. "Exotic Phenomena Searches at Hadron Colliders". In: *Proceedings of the PIC 2012*. 2013. DOI: arXiv:1301.2521 (cit. on pp. 30–32, 34).

Papers

[85] R.S. Thorne A. Sherstnev. *Parton Distributions for LO Generators*. Tech. rep. 0711.2473. 2008. DOI: 10.1140/epjc/s10052-008-0610-x. URL: http://arxiv.org/abs/0711.2473.

[86] I. J. R. Aitchison. *Supersymmetry and the MSSM: An Elementary introduction*. Tech. rep. hep-ph/0505105. 2005. URL: http://arxiv.org/abs/hep-ph/0505105 (cit. on p. 26).

[87] Johan Alwall, Philip Schuster, and Natalia Toro. *Simplified Models for a First Characterization of New Physics at the LHC*. Oct. 2008. eprint: 0810.3921v2. URL: http://arxiv.org/abs/0810.3921v2.

[88] Simone Amoroso, Riccardo Maria Bianchi, and Sascha Caron. *A general search for new phenomena in proton-proton collissions at a centre-of-mass energy of 7 TeV - supporting documentation*. ATLAS Internal Physics Communication. CERN, 2012. URL: https://cdsweb.cern.ch/record/1458546.

[89] S Asai et al. *Further search for squarks and gluinos using final states with jets and missing transverse momentum with the ATLAS experiment in s = 7 TeV proton-proton collisions: supporting documentation*. ATL-PHYS-INT-2011-085. Geneva: CERN, 2011 (cit. on p. 93).

[90] S Asai et al. *Search for Supersymmetry with jets and missing transverse momentum and one lepton at sqrt(s) = 7 TeV*. ATL-PHYS-INT-2011-082. Geneva: CERN, 2011 (cit. on pp. 90, 92, 93).

[91] ATLAS-Collaboration. *A general search for new phenomena with the ATLAS detector in pp collisions at $\sqrt{s}$ = 7 TeV*. ATLAS Public Conference Note. CERN, 2012. URL: http://cdsweb.cern.ch/record/1472686.

[92] ATLAS-Collaboration. *A general search for new phenomena with the ATLAS detector in pp collisions at sqrt(s)=7 TeV*. Tech. rep. CERN, 2012. URL: https://cds.cern.ch/record/1472686.

[93] ATLAS-Collaboration. *ATLAS computing : Technical Design Report*. Tech. rep. CERN-LHCC-2005-022, ATLAS-TDR-017. CERN, 2005 (cit. on p. 133).

[94] ATLAS-Collaboration. *ATLAS High-Level Trigger, Data Acquisition and Controls Technical Design Report*. Tech. rep. CERN, 2003. URL: http://cdsweb.cern.ch/record/616089/ (cit. on pp. 55, 57).

[95] ATLAS-Collaboration. *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*. Tech. rep. CERN-PH-EP-2011-117. CERN, 2011 (cit. on p. 59).

[96] ATLAS-Collaboration. *Expected electron performance in the ATLAS experiment*. Tech. rep. ATL-PHYS-PUB-2011-006. CERN, 2011 (cit. on p. 59).

[97] ATLAS-Collaboration. *Expected performance of the ATLAS experiment: detector, trigger and physics*. Tech. rep. CERN- OPEN-2008-020. CERN, 2009 (cit. on p. 59).

[98] ATLAS-Collaboration. *Luminosity Determination in pp Collisions at sqrt(s) = 7 TeV using the ATLAS Detector in 2011*. ATLAS-CONF-2011-116. Geneva: CERN, 2011. URL: https://cdsweb.cern.ch/record/1376384 (cit. on pp. 84, 86).

[99] ATLAS-Collaboration. *Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data*. Tech. rep. ATLAS-CONF-2013-013. CERN, 2013. URL: http://cds.cern.ch/record/1523699 (cit. on pp. 13–15).

[100] ATLAS-Collaboration. *Pile-up corrections for jets from proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS in 2011*. Tech. rep. ATLAS-CONF-2012-064. CERN, 2012 (cit. on p. 58).

[101] ATLAS-Collaboration. *Prospects for Supersymmetry and Universal Extra Dimensions discovery based on inclusive searches at a 10 TeV centre-of-mass energy with the ATLAS detector*. Tech. rep. ATL-PHYS-PUB-2009-084. http://cdsweb.cern.ch/record/1191916. Geneva: CERN, 2009 (cit. on p. 144).

[102] ATLAS-Collaboration. *Search for dark matter candidates and large extra dimensions in events with a jet and missing transverse momentum with the ATLAS detector*. Tech. rep. CERN-PH-EP-2012-210. CERN, 2012. URL: http://arxiv.org/abs/1210.4491.

[103] ATLAS-Collaboration. *Search for high mass dilepton resonances in pp collisions at sqrts = 7 TeV with the ATLAS experiment*. Tech. rep. ATLAS-CONF-2011-083. CERN, 2011 (cit. on p. 33).

[104] ATLAS-Collaboration. *Search for high-mass states with one muon plus missing transverse momentum in proton-proton collisions at $\sqrt{s} = 7TeV$ with the ATLAS detector*. Tech. rep. ATLAS-CONF-2011-082. CERN, 2011 (cit. on p. 33).

[105] ATLAS-Collaboration. *Search for New Phenomena in Monojet plus Missing Transverse Momentum Final States using 10fb-1 of pp Collisions at sqrts=8 TeV with the ATLAS detector at the LHC*. Tech. rep. CERN, 2012. URL: https://cds.cern.ch/record/1493486.

[106] ATLAS-Collaboration. *Search for Supersymmetry with jets and missing transverse momentum and one lepton at sqrt(s) = 7 TeV*. Tech. rep. ATL-COM-PHYS-2011-848. CERN, 2011 (cit. on p. 92).

[107] ATLAS-Collaboration. *Search for $W' \to t\bar{b}$ in proton-proton collisions at a centre-of-mass energy of $\sqrt{s}$ = 8 TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2013-050. CERN, 2013 (cit. on p. 33).

[108] ATLAS-Collaboration. *The ATLAS Inner Detector commissioning and calibration*. Tech. rep. Geneva: CERN, 2010. DOI: 10.1140/epjc/s10052-010-1366-7 (cit. on p. 49).

[109] P Bagnaia et al. *Calibration model for the MDT chambers of the ATLAS Muon Spectrometer*. Tech. rep. ATL-MUON-PUB-2008-004. CERN, 2008 (cit. on p. 52).

[110] T Baroncelli et al. *Study of MDT calibration constants using H8 testbeam data of year 2004*. Tech. rep. CERN, 2006 (cit. on p. 52).

[111] A J Barr et al. *Search for squarks and gluinos in multi-jet final state with the ATLAS experiment in sqrts=7TeV proton-proton collisions*. ATL-PHYS-INT-2011-087. Geneva: CERN, 2011.

[112] M Begel et al. *Jet energy scale and its systematic uncertainty in ATLAS for jets produced in proton-proton collisions at sqrts = 7 TeV*. ATL-COM-PHYS-2010-404. Geneva: CERN, 2010 (cit. on pp. 58, 86).

[113] Michael Begel. *2012 Jet Trigger*. Tech. rep. ATLAS, 2012. URL: https://indico.cern.ch/getFile.py/access?contribId=50&sessionId=2&resId=0&materialId=slides&confId=181541 (cit. on p. 89).

[114] M Bellomo et al. *Observation of W->munu and Z->munu in proton-proton collisions at sqrt(s) = 7TeV with the ATLAS Detector*. ATL-COM-PHYS-2010-265. Geneva: CERN, 2010.

[115] R M Bianchi. *Control And Configuration Of The ATLAS Trigger And Data Acquisition System During Data Taking Activities*. Tech. rep. ATL-DAQ-SLIDE-2011-344. CERN, 2011. URL: http://cdsweb.cern.ch/record/1368183 (cit. on p. 57).

[116] Volker Blobel. *Unfolding – Linear Inverse Problems*. 2010.

[117] J Boomsma and S Chekanov. *General searches for new particles*. ATL-COM-PHYS-2009-619. Geneva: CERN, 2009.

[118] J.M. Butterworth, J.R. Forshaw, and M.H. Seymour. *Multiparton Interactions in Photoproduction at HERA*. Tech. rep. hep-ph/9601371. 1996. URL: http://arxiv.org/abs/hep-ph/9601371 (cit. on p. 86).

[119] CDF-Collaboration, D0-Collaboration, and Tevatron Electroweak Working Group. *2012 Update of the Combination of CDF and D0 Results for the Mass of the W Boson*. Tech. rep. Fermilab, 2012. URL: http://arxiv.org/abs/1204.0042 (cit. on p. 7).

[120] CDF-Collaboration, D0-Collaboration, and Tevatron Electroweak Working Group. *Combination of CDF and DO results on the mass of the top quark using up to 5.8 fb-1 of data*. Tech. rep. Fermilab, 2011. URL: http://arxiv.org/abs/1107.5255 (cit. on p. 7).

[121] CMS-Collaboration. *CMS Physics Technical Design Report Volume I: Detector, Performance and Software*. Tech. rep. Geneva: CERN, 2006 (cit. on p. 40).

[122] CMS-Collaboration. *Model Unspecific Search for New Physics in pp Collisions at sqrt(s) = 7 TeV*. CMS-PAS-EXO-10-021. CERN, 2011 (cit. on pp. 64, 73–76).

[123] CMS-Collaboration. *Properties of the Higgs-like boson in the decay H to ZZ to 4l in pp collisions at sqrt s =7 and 8 TeV*. Tech. rep. CMS-PAS-HIG-13-002. CERN, 2013. URL: http://cds.cern.ch/record/1523767 (cit. on pp. 13, 14).

[124] ATLAS Collaboration. *Performance of the ATLAS Muon Trigger in p-p collisions at* $\sqrt{s} = 7TeV$. CERN, 2010.

[125] The ATLAS Collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider*. Tech. rep. CERN, 2008. DOI: dx.doi.org/10.1088/1748-0221/3/08/S08003 (cit. on pp. 39, 49).

[126] G. Corcella et al. *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*. Event Generator arXiv:hep-ph/0011363. arXiv, 2001. DOI: 10.1088/1126-6708/2001/01/010. URL: http://arxiv.org/abs/hep-ph/0011363 (cit. on pp. 81, 85, 86).

[127] D0-Collaboration. *Model independent search for new phenomena in pp(bar) collisions at sqrt(s) =1.96 TeV*. arXiv:1108.5362. FERMILAB-PUB-11-400-E. Comments: Submitted to PRD. 20 pages, 21 figures. Fermilab, 2011 (cit. on p. 65).

[128] E. Dobson. *Performance of the ATLAS trigger and data acquisition system*. Tech. rep. 2011 (cit. on p. 56).

[129] J. R. Ellis. *Supersymmetry for Alp Hikers*. Tech. rep. hep-ph/0203114. CERN, 2002. URL: http://arxiv.org/abs/hep-ph/0203114 (cit. on p. 26).

[130] Lyndon Evans and Philip Bryant. *LHC Machine*. Tech. rep. CERN, 2008. DOI: doi:10.1088/1748-0221/3/08/S08001 (cit. on p. 39).

[131] Dieter Haidt. *The Discovery of Weak Neutral Currents*. Tech. rep. DESY, 2003. URL: https://www.desy.de/~haidt/nc30_text.pdf.

[132] M. J. Herrero. *The Standard Model*. Tech. rep. arXiv: hep-ph/9812242v1. 1998 (cit. on pp. 12–16).

[133] Tom Junk. *Sensitivity, Exclusion and Discovery with Small Signals, Large Backgrounds, and Large Systematic Uncertainties*. Fermilab, 2007. URL: http://www-cdf.fnal.gov/~trj/mclimit/mclimit_csm.pdf (visited on 2007).

[134] S Klous. *Event streaming in the online system: Real-Time Organization of ATLAS Data*. ATL-COM-DAQ-2010-080. ATL-DAQ-SLIDE-2010-085. ATL-COM-DAQ-2010-047. Geneva: CERN, 2010.

[135] S Klous et al. *Event streaming in the ATLAS experiment*. ATL-DAQ-INT-2010-002. Geneva: CERN, 2010.

[136] H. Kroha et al. *New methods for the alignment of the ATLAS muon spectrometer*. Tech. rep. CERN, 2008, pp. 666–668. DOI: 10.1109/NSSMIC.2007.4436420 (cit. on p. 52).

[137] W. Lampl and al. *Calorimeter clustering algorithms: Description and performance*. Tech. rep. ATL-COM-LARG-2008-003. CERN, 2008 (cit. on p. 58).

[138] Victor Lendermann et al. *Combining Triggers in HEP Data Analysis*. Jan. 2009. eprint: 0901.4118v1. URL: http://arxiv.org/abs/0901.4118v1.

[139] M.L. Mangano et al. *ALPGEN, a generator for hard multiparton processes in hadronic collisions*. Tech. rep. hep-ph/0206293. 2003. DOI: 10.1088/1126-6708/2003/07/001. URL: http://arxiv.org/abs/hep-ph/0206293 (cit. on pp. 81, 86).

[140] Z. Marshall. *Re-defining the Standard QCD Di-Jet Samples*. Tech. rep. ATL-COM-PHYS-2011-992. Geneva: CERN, 2011. URL: https://cdsweb.cern.ch/record/1370089.

[141] Stephen P. Martin. *A Supersymmetry Primer*. 2008. eprint: hep-ph/9709356. URL: http://arxiv.org/abs/hep-ph/9709356 (cit. on pp. 22, 26–28, 30).

[142] Arthur M. Moraes. *Describing the underlying event at the LHC with JIMMY4.1*. Tech. rep. 2004. URL: http://indico.cern.ch/getFile.py/access?resId=1&materialId=0&contribId=s6t21&sessionId=6&subContId=1&confId=a044163 (cit. on p. 86).

[143] Michael E. Peskin. *Beyond the Standard Model*. Tech. rep. hep-ph/9705479. SLAC, 1997. URL: http://cds.cern.ch/record/326995 (cit. on p. 13).

[144] M. Schott, G. Duckeck, and M. Obermaier. *Study of the sagitta resolution of MDT chambers*. Tech. rep. ATL-MUON-PUB-2006-010. CERN, 2006 (cit. on p. 51).

[145] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. *PYTHIA 6.4 Physics and Manual*. Tech. rep. hep-ph/0603175. 2006. DOI: 10.1088/1126-6708/2006/05/026. URL: http://arxiv.org/abs/hep-ph/0603175 (cit. on pp. 81, 85, 86).

[146] Dan R. Tovey. *Measuring the SUSY Mass Scale at the LHC*. arXiv. arXiv:hep-ph/0006276: University of Sheffield, 2000 (cit. on p. 83).

Other resources

[147] http://www.scipy.org/Weave.

[148] *A complete exemple of WatchMan steering file implementing an analysis*. URL: https://svnweb.cern.ch/trac/WatchMan/browser/trunk/run/WatchMan_AnalysisSteeringFile_benchmarkChannels_Delphes.py (cit. on p. 136).

[149] University of Athens. *HY.P.A.T.I.A. – HYbrid Pupil's Analysis Tool for Interactions in ATLAS*. 2011. URL: http://hypatia.phys.uoa.gr/applet/en/help.html (cit. on p. 47).

[150] ATLAS-Collaboration. *ATLAS Experiment - Luminosity Public Results*. CERN. 2011. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults (cit. on p. 84).

[151] ATLAS-Collaboration. *ATLAS Experiment - Supersymmetry public results*. CERN. 2012. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults (cit. on p. 29).

[152] ATLAS-Collaboration. *ATLAS Jet Energy Resolution Provider class*. 2011. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetEnergyResolutionProvider (cit. on p. 86).

[153] ATLAS-Collaboration. *ATLAS Jet Trigger Menu*. URL: https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TrigJetMenu (cit. on p. 88).

[154] ATLAS-Collaboration. *ATLAS Plots on E_{CM} Dependence of Physics Reach*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/AtlasResultsEcmDependence. WatchMan was used to analyze data in order to produce the plots of the Supersymmetry discovery reach. Geneva, 2010 (cit. on p. 144).

[155] ATLAS-Collaboration. *Et Miss "RefFinal"*. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EtMissRefFinal (cit. on p. 60).

[156] ATLAS-Collaboration. *Event display of a Higgs boson decaying into four muons*. URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-092/fig_33.png (cit. on p. 19).

[157] ATLAS-Collaboration. *Event display of a Higgs boson decaying into two photons*. 2013. URL: http://atlas-vp1.web.cern.ch/atlas-vp1/screenshots/Hgg/evt_86694500/vp1_Hgg_run191426_evt86694500_hi-res.png (cit. on p. 18).

[158] ATLAS-Collaboration. *Extended Pileup Reweighting tool*. 2011. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ExtendedPileupReweighting (cit. on p. 85).

[159] ATLAS-Collaboration. *JetEtmiss approved 2011 Pileup Offset and JVF*. Ed. by ATLAS-Collaboration. 2011. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2011PileupOffsetAndJVF (cit. on p. 91).

[160] ATLAS-Collaboration. *Latest Results from ATLAS Higgs Search*. 2012. URL: http://www.atlas.ch/news/2012/latest-results-from-higgs-search.html (cit. on p. 13).

[161] ATLAS-Collaboration. *Missing Et in rel17*. URL: https://twiki.cern.ch/twiki/bin/view/Main/Rel17MissingEt (cit. on p. 60).

[162] ATLAS-Collaboration. *Virtual Point One (VP1) - Main website*. URL: https://atlas-vp1.web.cern.ch/atlas-vp1/ (cit. on pp. 18, 19).

[163] Riccardo Maria Bianchi. WatchMan – *An highly automated Analysis Code Generator*. https://twiki.cern.ch/twiki/bin/view/Main/WatchMan (cit. on pp. 129, 136, 137).

[164] Ian C. Brock. *Physics of and with Leptons WS 05/06*. Ed. by Universit at Bonn. 2005. URL: http://www-zeus.physik.uni-bonn.de/~brock/feynman/vtp_ws0506/chapter09/higgs_pot.jpg (cit. on p. 12).

[165] Jon Butterworth et al. *JIMMY Generator — Multiparton Interactions in HERWIG*. 2012. URL: http://jimmy.hepforge.org (cit. on pp. 81, 86).

[166] M. Cacciari and G. P. Salam. *FastJet*. URL: http://fastjet.fr/ (cit. on p. 58).

[167] *CASE – Computer-aided software engineering*. http://en.wikipedia.org/wiki/Computer-aided_software_engineering (cit. on p. 131).

[168] CDF-Collaboration. *CDF Top Quark Physics Public Results*. 2011. URL: http://www-cdf.fnal.gov/physics/new/top/public_mass.html (cit. on p. 7).

[169] CERN. *Computer-generated diagram of an LHC dipole*. 1998. URL: https://cds.cern.ch/record/39731 (cit. on p. 43).

[170] CERN. *Computer Generated image of the ATLAS calorimeter*. 2008. URL: https://cds.cern.ch/record/1095927 (cit. on p. 51).

[171] CERN. *Computer generated image of the ATLAS inner detector*. 2008. URL: https://cds.cern.ch/record/1095926 (cit. on p. 50).

[172] CERN. *Computer generated image of the ATLAS Muons subsystem*. 2008. URL: https://cds.cern.ch/record/1095929 (cit. on p. 53).

[173] CERN. *Computer generated image of the whole ATLAS detector*. 2008. URL: https://cds.cern.ch/record/1095924 (cit. on p. 48).

[174] CERN. *Cross section of an LHC dipole in the tunnel*. 2011. URL: https://cds.cern.ch/record/1365795 (cit. on p. 44).

[175] CERN. *Cross section of LHC dipole*. 1999. URL: https://cds.cern.ch/record/842530 (cit. on p. 43).

[176] CERN. *Event Cross Section in a computer generated image of the ATLAS detector*. 2008. URL: https://cds.cern.ch/record/1096081 (cit. on p. 48).

[177] CERN. *LHC Layout*. 1997. URL: https://cds.cern.ch/record/841573 (cit. on p. 40).

[178] CERN. *LHC Machine Outreach*. 2012. URL: http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/collisions.htm (cit. on p. 41).

[179] CERN. *The CERN accelerator complex*. 2008. URL: https://cds.cern.ch/record/1260465 (cit. on p. 45).

[180] CMS-Collaboration. *Observation of a New Particle with a Mass of 125 GeV*. 2012. URL: http://cms.web.cern.ch/news/observation-new-particle-mass-125-gev (cit. on p. 13).

[181] *Cython – C extensions for Python*. http://cython.org.

[182] *Delphes – A framework for fast simulation of a generic collider experiment*. http://projects.hepforge.org/delphes/ (cit. on pp. 131, 136).

[183] Thorsten Dietzsch. "Search for New Phenomena in Dijet Angular Distributions Measured with the ATLAS Detector in Proton-Proton Collisions at $\sqrt{s} = 7TeV$". PhD thesis. Ruperto-Carola-University of Heidelberg, 2013 (cit. on p. 89).

[184] Niels Van Eldik. "The ATLAS muon spectrometer: calibration and pattern recognition". PhD thesis. University of Amsterdam, 2007 (cit. on p. 54).

[185] Stefano Frixione et al. *The MC@NLO Package*. 2012. URL: http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO/ (cit. on pp. 81, 86).

[186] Jeremiah Jet Goodson and Robert McCarthy. "Search for Supersymmetry in States with Large Missing Transverse Momentum and Three Leptons including a Z-Boson". PhD thesis. Stony Brook U., 2012 (cit. on pp. 52, 56).

[187] *Instant inlining of C/C++ code in Python*. http://heim.ifi.uio.no/~kent-and/software/Instant/doc/Instant.html.

[188] W. Lavrijsen. *PyROOT – A Python-ROOT Bridge*. http://root.cern.ch/drupal/content/how-use-use-python-pyroot-interpreter (cit. on pp. 135, 143).

[189] Wim Lavrijsen. *PyROOT Glossary - Dictionaries*. URL: http://wlav.web.cern.ch/wlav/pyroot/glossary.html#gloss-dictionary (cit. on p. 136).

[190] LEP. *LEP Working groups*. 2000. URL: http://delphiwww.cern.ch/offline/lepwgs.html (cit. on p. 39).

[191] LHC. *LHC: Facts and Figures*. 2008. URL: http://public.web.cern.ch/public/en/lhc/Facts-en.html (cit. on p. 42).

[192] Sven Menke. *ATLAS Data & Analysis Model*. contribution at Seminar Uni-Wuppertal. 2009. URL: http://www.mppmu.mpg.de/~menke/Atlas/wuppertal29012009.pdf.

[193] CERN Press Office. *CERN experiments observe particle consistent with long-sought Higgs boson*. 2012. URL: http://press.web.cern.ch/press-releases/2012/07/cern-experiments-observe-particle-consistent-long-sought-higgs-boson (cit. on p. 13).

[194] *ROOT – An Object-Oriented Data Analysis Framework*. http://root.cern.ch/drupal/ (cit. on pp. 135, 136, 143).

[195] *SFrame – A ROOT data analysis framework*. http://sourceforge.net/projects/sframe/.

[196] Martin Wessels. "General Search for New Phenomena in ep Scattering at HERA". PhD thesis. Rheinisch-Westfalischen Technischen Hochschule Aachen, 2004 (cit. on pp. 67, 68, 109).

[197] Wikipedia. *Eightfold Way*. URL: http://en.wikipedia.org/wiki/Eightfold_Way_%28physics%29.

[198] Wikipedia. *Multiset*. Ed. by Wikipedia. 2012. URL: http://en.wikipedia.org/wiki/Multiset (cit. on p. 81).

[199] Wikipedia. *Pseudorapidity definition*. 2012. URL: http://en.wikipedia.org/wiki/Pseudorapidity (cit. on p. 47).

Index

This Thesis has been defended on the 18^{th} of March 2014
at the Physikalisches Institut of the Albert-Ludwigs-Universität
in Freiburg, Germany

This Thesis has been typeset with LaTeX,
using the memoir *class and* BibLaTeX