QCD background studies for inclusive SUSY searches in the 1-lepton channel

Diplomarbeit der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Jonas Zacharias Will
geboren in München

München, den 17.03.2009

Erstgutachterin: Prof. Dr. Dorothee Schaile
Zweitgutachter: Prof. Dr. Wolfgang Dünnweber

Zusammenfassung

Der Large Hadron Collider (LHC) am CERN wird dieses Jahr in Betrieb gehen und im Jahr 2010 voraussichtlich 200 pb^{-1} von hochenergetischen Proton-Proton-Kollisionen geliefert haben. Nach dieser ersten Phase, die bei einer Schwerpunktsenergie von 10 TeV stattfinden wird, soll die Schwerpunktsenergie auf 14 TeV erhöht werden. ATLAS, einer der beiden Universaldetektoren, wird versuchen das Higgsboson, das letzte unentdeckte Teilchen des Standardmodells, nachzuweisen und nach neuer Physik auf der TeV-Skala suchen. Bis zum Zeitpunkt der ersten realen Kollisionen basieren die Studien auf Daten, die durch einen Monte-Carlo-Generator erzeugt werden und anschließend den ATLAS-Detektor in simulierter Form durchlaufen.

Supersymmetrie (SUSY) ist ein vielversprechender Kandidat für die Beschreibung von Physik jenseits des Standardmodells. Eine Möglichkeit nach Hinweisen auf SUSY zu suchen ist die Untersuchung von Ereignissen, die genau ein Elektron oder Myon im Endzustand besitzen. Diese Form der Suche nach SUSY bezeichnet man als den 1-Lepton-Kanal. Obwohl in den Wechselwirkungen der starken Kraft keine Leptonen auftreten können, enthalten die in dieser Arbeit analysierten QCD-Ereignisse dennoch eine kleine Anzahl an rekonstruierten Elektronen und Myonen, so genannte "Fakes". Daher zählen auch diese QCD-Ereignisse, die sich durch viele, zum Teil hochenergetische, Jets auszeichnen, zum Untergrund bei der 1-Lepton-SUSY-Suche.
Die Schnitte der 1-Lepton-SUSY-Analyse, einer Standard-Cut-Analyse, wirken so effektiv auf die QCD-Untergrundereignisse, dass nach ihrem Anwenden kein QCD-Ereignis übrig bleibt. Wegen der vergleichsweise geringen Anzahl an simulierten QCD-Ereignissen führt dies zu einer großen Unsicherheit in der Effizienz der Schnitte. Deshalb wird im Folgenden eine Strategie zur Abschätzung der Schnitteffizienz für den QCD-Untergrund entwickelt, die eine geringere statistische Unsicherheit aufweist, dafür aber Korrelationen zwischen mehr als zwei Schnittvariablen vernachlässigt.
Des Weiteren werden die rekonstruierten Elektronen und Myonen des QCD-Untergrundes genauer studiert. Dazu wird eine Zuordnungsmethode entwickelt, die den rekonstruierten Leptonen ein Lepton oder einen Jet auf Generatorebene zuweist. Rekonstruierten Leptonen, denen ein gleichartiges Lepton auf Generatorebene zugeordnet wird, werden als "nonprompt" bezeichnet, da sie aus Zerfällen schwerer Quarks stammen, die nach der anfänglichen starken Wechselwirkung von Quarks oder Gluonen stattfinden. Im Falle einer Zuordnung zwischen rekonstruierten Leptonen und generierten Jets handelt es sich um sog. "jet-faked" Leptonen. In der vorliegenden Arbeit werden sowohl die integrierten Fake-Raten als auch die Fake-Raten in Abhängigkeit von kinematischen Größen bestimmt. Dies geschieht für unterschiedliche Lepton-Isolationskriterien. Für das üblicherweise verwendete Lepton-Isolationskriterium dominieren die jet-faked Elektronen die Gesamt-Fake-Rate. Dagegen ist bei diesem Lepton-Isolationskriterium der Anteil der jet-faked Myonen an der Gesamt-Fake-Rate sehr gering. Abschließend wird die Wahrscheinlichkeit eines generierten Jets oder eines generierten Leptons bestimmt als Lepton rekonstruiert zu werden.

abstract

The Large Hadron Collider (LHC) at CERN will start running this year, providing 200 pb^{-1} of high-energetic proton-proton collisions in 2010. After this first run at 10 TeV, the center-of-mass energy should be increased to 14 TeV. ATLAS, a multi-purpose detector, will analyze the signatures of these collisions to look for evidence of the Higgs boson, the last undetected particle of the standard model, and for indications of new physics at the TeV scale. Until then, studies rely on collision events that are Monte-Carlo generated and are passed through a simulated version of the ATLAS detector.

Supersymmetry (SUSY) is a promising candidate to describe physics beyond the standard model. The 1-lepton SUSY analysis is one possible search mode being developed to look for hints of SUSY in events containing exactly one electron or muon in the final state. Although in strong interactions no leptons can occur, a small fraction of the simulated QCD multijet events analyzed in this thesis contain reconstructed electrons and muons, so-called fake leptons, and therefore represent one background component for this SUSY search channel.
The cuts of the 1-lepton SUSY analysis, a standard cut analysis, act very efficiently on the QCD background samples, reducing the number of surviving events to zero. However, the rather low statistics of these samples lead to a big uncertainty on the cut efficiency. Therefore, a strategy is developed to estimate the QCD background cut efficiency which is based on neglecting correlations between more than two cut variables.
Furthermore, the reconstructed electrons and muons contained in the QCD background samples are studied. A matching strategy is designed, assigning reconstructed leptons to a lepton or to a jet at generator level. If a reconstructed lepton is paired with a generated lepton of the same flavour, it is indicated as nonprompt since it originates from a heavy flavour decay following the initial strong interaction of quarks or gluons. A reconstructed lepton matched by a jet at generator level is referred to as jet-faked lepton. In this thesis, the integrated fake rates as well as the fake rates as functions of kinematic variables are calculated for different lepton isolation criteria. For the standard lepton isolation criterion, jet-faked electrons dominate the overall fake rate, whereas the contribution of jet-faked muons to the overall fake rate is marginal. Finally, the probability of a generated jet or lepton to be reconstructed as a lepton is determined.

Contents

Chapter 1

Introduction

This autumn, the Large Hadron Collider (LHC) at CERN, the European Organisation for Nuclear Research, will start operations and prepare the ground for the largest experiment worldwide. After a short run at a center-of-mass energy of 10 TeV at the beginning, the LHC will bring two beams of protons to collision with a unique center-of-mass energy of 14 TeV, one order of magnitude above the current maximal energy probed, and open a new era in experimental particle physics. Two general-purpose experiments at the LHC will search in particular for the Higgs boson, the last undetected particle of the standard model of particle physics, and for experimental indications of theories describing physics beyond the standard model, like supersymmetry.
The present section starts with an overview of the LHC and the ATLAS detector followed by an introduction to the standard model and to supersymmetry.

1.1 The LHC and ATLAS

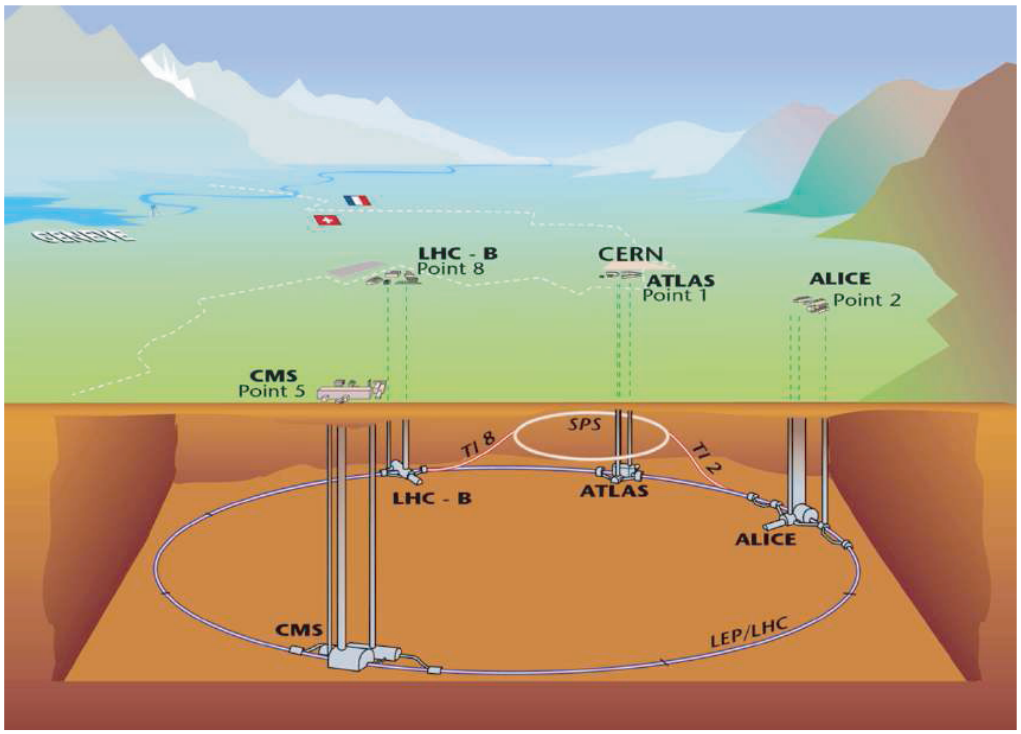
The LHC has four collision points where the different experiments reside. The largest experiments are ATLAS and CMS, two huge multi-purpose detectors, that are built to analyze the products of the high-energetic proton collisions for a precise determination of standard model parameters, for signatures of the Higgs boson and for new phenomena potentially detectable at the TeV scale. To allow for confirmations, it is important to have more than one experiment in case of discoveries. LHCb, a medium-size experiment, is specialized in the study of the differences between matter and antimatter by comparing the decays of two fundamental particles, the b quark and the anti-b quark. Besides the potential to collide protons, the LHC is also equipped to collide heavy ions. ALICE, a further medium-size experiment, will analyze the collisions of lead ions which will hopefully generate a particular state of matter, the quark-gluon plasma. Finally, two small experiments, TOTEM and LHCf, are located near CMS and ATLAS, respectively, and focus on particles that are only slightly deflected in the collisions of protons or ions [1].

1.1.1 The LHC at CERN

The LHC is a hadron accelerator and collider installed in the 26.7 km long tunnel that was originally built for the CERN LEP machine, an electron-positron collider which operated from 1989 to 2000 [2]. The tunnel lies between 45 m and 170 m below the surface on a plane inclined at 0.8°. Two transfer tunnels exist with a length of around 2.5 km each that link the LHC to the CERN accelerator complex acting as pre-accelerator and injector [3].
In the CERN accelerator complex, protons, obtained by removing electrons from hydrogen atoms, are at first accelerated by a linear accelerator (LINAC2). Afterwards, they pass three further pre-accelerators, the PS Booster, the Proton Synchrotron (PS) and the Super Proton

Synchrotron (SPS), before being finally injected into the LHC. Here, counter-rotating bunches of up to 10^{11} protons will collide 40 million times per second providing high-energetic proton-proton collisions of 14 TeV at a design luminosity of 10^{34} cm^{-2}s^{-1} [1].
The pre-acceleration of lead ions, gained by a source of vaporized lead, varies only in the first two acceleration steps with respect to protons. Instead of LINAC2, ions are at first accelerated by LINAC3, a further linear accelerator, and afterwards injected into the Low Energy Ion Ring (LEIR). In the LHC, pairs of ions collide with a center-of-mass energy of 5.5 TeV at a design luminosity of 10^{27} cm^{-2}s^{-1} [1].
Figure 1.1 illustrates CERN with the LHC at the border between France and Switzerland at the foot of the Jura mountains. It shows the LHC ring with its final pre-accelerator, the SPS, and the caverns of the four experiments: ATLAS, CMS, ALICE and LHCb.

Figure 1.1: The LHC at CERN [1] and its geographical environment.



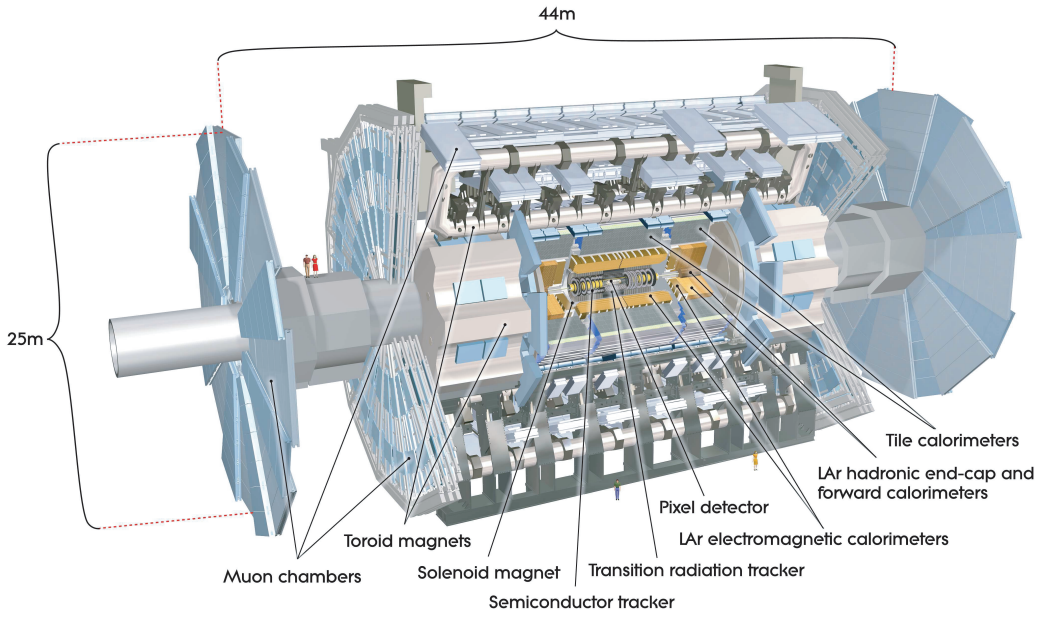
1.1.2 ATLAS at the LHC

With a height of 25 m and a length of 44 m, ATLAS (A Toroidal LHC ApparatuS), illustrated in Figure 1.2, is the largest detector of the LHC, although its weight of 7000 t falls below the one of the CMS detector. As a multi-purpose detector, ATLAS is characterized by different detector components surrounding the interaction point: the inner detector, the calorimeter system and the muon spectrometer. The magnet configuration comprises a superconducting solenoid located around the inner detector and three large superconducting toroid systems (one barrel and two end caps) surrounding the calorimeters. Information about the general functionality of detectors can be found in [4]. The information about ATLAS is extracted from [5].

The ATLAS detector is forward-backwards symmetric with respect to the nominal interaction point representing the origin of the right-handed coordinate system. The beam direction defines the z-axis and the $x-y$ plane is transverse to it. The positive x-axis points from the interaction point to the center of the LHC ring and the y-axis is defined as pointing

upwards. The azimuthal angle ϕ (the polar angle θ) is measured around the z-axis (the x-axis) in the range $[-\pi, +\pi]$ ($[0, 2\pi]$). $\phi = 0$ ($\theta = 0$) corresponds to the positive x-axis (positive z-axis) and increases clockwise looking into the positive z-direction (negative x-direction). The pseudorapidity η is defined as $\eta = -\ln\tan(\theta/2)$. $\eta = +\infty$ ($\eta = -\infty$) corresponds to the positive (negative) z-direction. $\eta = 0$ points in the positive y-direction. Transverse variables, such as the transverse momentum, p_T, of an object or its transverse energy, E_T, are related to the $x - y$ plane of the detector. In the pseudorapidity-azimuthal space, the distance is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Figure 1.2: Cut-away view of the ATLAS detector [5]. To demonstrate the dimensions of the detector, some people are also pictured.



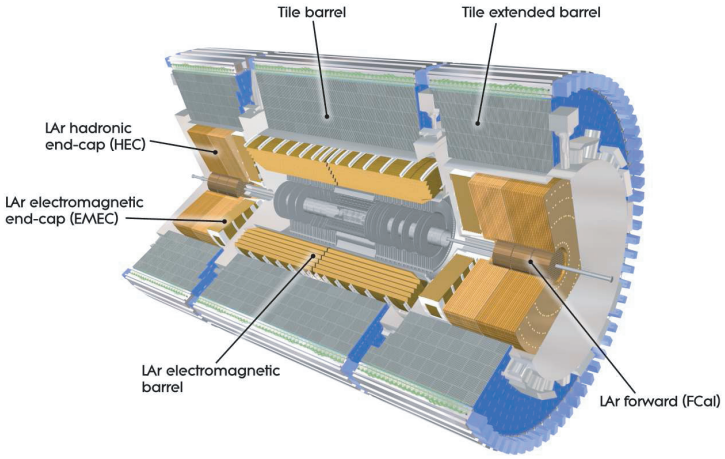
The ATLAS detector was designed to optimize its response to new physics processes that could possibly occur at the TeV scale. In the following, the main features of the detector components are described:

- inner detector: the inner detector is immersed in the 2T magnetic field of the central solenoid. It is designed for high-precision momentum and vertex measurements while handling the very large track density due to the around 1000 particles which will emerge every 25 ns. To satisfy these requirements, the inner detector comprises in its innermost part pixel detectors and semiconductor trackers for precision track measurement, the former having 80.4 million and the latter, 6.3 million read-out channels. Both components cover the region $|\eta| < 2.5$. The outer part consists of straw-tube tracking detectors in which transition radiation is generated and detected. It is restricted to $|\eta| < 2.0$.

- calorimeter system (for a detailed illustration, see Figure 1.3): by absorbing energy from traversing electrons, photons, taus or hadrons, the calorimeter system allows for precise measurements of their energy and position. Its overall coverage is $|\eta| < 4.9$. It is divided into an electromagnetic and a hadronic part.
  The electromagnetic calorimeter system consists of liquid-argon (LAr) sampling calorimeters that use lead plates as absorber and are characterized by accordion-shaped electrodes. It is divided into a barrel part surrounding the central solenoid and a sepa-

rated end-cap wheel on each side, providing precise measurements of electromagnetic showers of electrons and photons. The barrel component covers the region $|\eta| < 1.52$ and the end cap, the region $1.375 < |\eta| < 3.2$.
Two different constructions are used for the hadronic calorimeter measuring hadronic showers. The hadronic tile calorimeter is a sampling calorimeter featured by scintillating tiles and absorbing steel. It comprises a barrel and an extended barrel covering the regions $|\eta| < 1.0$ and $0.8 < |\eta| < 1.7$, respectively. The LAr hadronic end-cap calorimeter with copper absorber and a coverage of $1.5 < |\eta| < 3.2$ is located directly behind the electromagnetic end cap.
Finally, the calorimeter system includes a LAr forward calorimeter for measuring both the electromagnetic and hadronic showers with copper absorber and tungsten absorber, respectively. It covers $3.1 < |\eta| < 4.9$.
Due to gaps and insensitive ("dead") material between different calorimeter components, the performance of the calorimeter depends on η. Especially worth mentioning is the region $1.37 < |\eta| < 1.52$, characterized by the transition from the electromagnetic barrel to the electromagnetic end cap and referred to as crack region in the following, and the region $|\eta| \approx 1.1$ where the transition from the barrel to the extended barrel of the hadronic tile calorimeter is located. In both regions, the reconstruction performance for electrons is degraded.

Figure 1.3: The calorimeter system of ATLAS [5].



- muon spectrometer: the muon system surrounds the calorimeter and determines the overall dimensions of ATLAS. By means of the large bending power of the toroid magnets, muon tracking chambers perform high-precision measurements of the muon momentum for $|\eta| < 2.7$.
  In the barrel region, the magnetic field is mainly provided by the eight radially aligned coils building the barrel toroid. For larger $|\eta|$ values, the importance of the two end-cap toroids, which are inserted at the ends of the barrel toroid, grows. In the barrel region, the muon chambers are aligned in three layers around the beam axis, whereas in the end-cap region three layers of chambers are located perpendicular to the beam. The muon system comprises four types of chambers: Monitored Drift Tubes ($|\eta| < 2.7$) and Cathode Strip Chambers ($2.0 < |\eta| < 2.7$) are developed for a precise measurement of the muon coordinate along the principal bending direction of the magnetic field. Resistive Plate Chambers ($|\eta| < 1.05$) and Thin Gap Chambers ($1.05 < |\eta| < 2.7$) are mainly devoted to triggering and to the measurement of the muon coordinate which is orthogonal to the one measured by the precision tracking chambers.

At the design luminosity of 10^{34} cm^{-2}s^{-1}, the expected interaction rate of proton-proton collisions is approximately 1 GHz. By contrast, the recording of the event data is limited to around 200 Hz. Therefore, a rejection factor of $5 \cdot 10^6$ is necessary to single out the events of interest. This is realized by different trigger levels: the hardware-based Level-1 (L1) trigger system uses a part of the total detector information to decide if the recorded event should be further processed or discarded. L1 reduces the data rate to about 75 kHz and defines Regions of Interest (RoI's), i.e. regions of the detector where interesting features are recognized in the event. The subsequent two levels, denoted as high-level trigger, are software based and comprise the Level-2 (L2) trigger and the event filter. L2 uses the detailed event information in the RoI's and reduces the trigger rate to approximately 3.5 kHz. The event filter finally reduces the rate to around 200 Hz by using offline analysis procedures.

1.2 From the standard model to supersymmetry

Since the 1970s, the standard model of particle physics, shortly described in Section 1.2.1, is the state of the art for the description of fundamental particles and their interactions. Although its predictive power was successfully tested by precision measurements on the order of at least $\mathcal{O}(10^{-3})$, this theory bears lots of deficiencies [6]. Some of them are described in Section 1.2.2. One interesting candidate for a theory beyond the standard model is supersymmetry. Its main features are brought up in Section 1.2.3, while Section 1.2.4 briefly describes how it can solve some of the standard model problems.

1.2.1 The standard model of particle physics

The standard model of particle physics (SM) is a consistent, renormalizable quantum field theory, describing the interactions of the fundamental fermions which are the constituents of all ordinary matter [6]. It comprises six quarks and six leptons, each falling into three generations with increasing mass, and their antiparticles [7]. Table 1.1 depicts the three generations of fermions and their possible interactions.

Table 1.1: The three generations of fundamental fermions of the SM and their interactions ("e.m." stands for electromagnetic).

	quarks	e.m.	weak	strong	leptons	e.m.	weak	strong
1^{st}	u (up)	+	+	+	e^- (electron)	+	+	-
	d (down)	+	+	+	ν_e (electron neutrino)	-	+	-
2^{nd}	c (charm)	+	+	+	μ^- (muon)	+	+	-
	s (strange)	+	+	+	ν_μ (muon neutrino)	-	+	-
3^{rd}	t (top)	+	+	+	τ^- (tau)	+	+	-
	b (bottom)	+	+	+	ν_τ (tau neutrino)	-	+	-

Four fundamental interactions exist: the gravitation, the electromagnetism, the weak force and the strong force. The SM neglects the gravitation due to its weakness compared to the other forces. The latter are described by two quantum field theories: the electroweak theory for electromagnetic and weak interactions and the quantum chromodynamics (QCD) for strong interactions. These quantum field theories are based on a common principle, the gauge principle: to describe interactions between particles, the basic equations are required to be invariant under local gauge transformations, thereby introducing local symmetries in the theory. The SM has the following gauge symmetry: $U(1) \otimes SU(2) \otimes SU(3)$, where $U(1) \otimes SU(2)$ represents the symmetry group of the electroweak theory and $SU(3)$, the one

of QCD. Consequences are the emergence of a spin-1 gauge boson, the quantized mediator of the considered interaction, and the appearance of charge. To solve concrete problems in quantum field theory, one relies on pertubation theory, known as the Feynman calculus [7]. However, the gauge symmetry holds only for massless particles. Since all quarks and leptons have mass, this symmetry is broken. The favoured mechanism for the introduction of mass is the Higgs mechanism, based on the principle of spontaneous symmetry breaking. This mechanism predicts one Higgs boson, the last undetected particle of the SM [7].
In the realistic case of massive particles, the electromagnetic and weak interactions are described separately and the gauge bosons and charges are represented by: the massless photon, γ, and the electromagnetic charge for the electromagnetic interaction, three massive bosons, $W^{\pm}$ and Z, and the weak isospin for the weak interaction and eight massless gluons, g, and the colour charge for the strong force. Only particles bearing a given charge can participate in the related interaction.
In contrast to leptons, no isolated quarks (or gluons) have been observed up to now. They are confined in colourless hadrons, which are groups of two (mesons) or three (baryons) quarks. As the interaction strength increases with increasing distance between the interacting partners, no perturbation theory can be applied to this problem [7]. Therefore, one relies on models describing the hadronization, i.e. the transformation of quarks or gluons into hadrons which are detectable in high-energetic particle collisions. The different existing models are based on an asymptotic approach: the high-energetic initial quark q_0 emits a pair of quarks $q_1\overline{q}_1$, such that $q_0\overline{q}_1$ builds a meson and q_1 is left. q_1 is later on included in the meson $q_1\overline{q}_2$ of the produced pair $q_2\overline{q}_2$, and so on. The group of hadrons originating from the initial quark or gluon is called jet. Event generators, such as PYTHIA, rely on hadronizing models to describe the step between the interaction of fundamental quarks or gluons and the interaction of the produced hadrons with the detector material [8].

1.2.2 Shortcomings of the standard model

Despite having successfully passed many experimental tests, the SM has insufficiencies. Among them are the following [6]:

- Although the observed neutrino oscillations strongly suggest that neutrinos are massive particles, they are considered massless in the SM.
- Elaborate studies of the cosmic microwave background conclude that only a small fraction of the universe consists of matter describable by the SM. The bulk of the energy density seems to consist of a mysterious dark matter and dark energy.
- The SM neglects the existence of gravitation.
- The SM has a large number of free parameters (18) [9]. Many of these parameters are a posteriori introduced to *describe* experimental observations without deeper *understanding* of the circumstances.
- The SM suffers from a fine-tuning problem. When considering the one-loop corrections for the estimation of the Higgs mass, one ends up with the relation

$$m^2_{H_{SM}}(phys) \simeq m^2_{H_{SM}} + \frac{c}{16\pi^2}\Lambda^2, \tag{1.1}$$

  where $m_{H_{SM}}(phys)$ denotes the physical Higgs mass and $m_{H_{SM}}$, the SM Higgs mass parameter; c is a coefficient depending on the SM coupling constants and Λ is the cut-off parameter representing the energy scale where one assumes the end of the SM validity. This cut-off parameter can be set at the TeV scale if one considers exotic theories such

as large extra dimensions, but the Planck scale $\mathcal{O}(10^{19}$ GeV) is usually assumed as, at this scale, gravitation can not be neglected anymore. Theoretical arguments postulate that the physical Higgs mass has to be smaller than $\mathcal{O}(1$ TeV). This requirement leads to extreme fine tuning between the two terms on the right hand side of Equation (1.1).

1.2.3 Introduction to supersymmetry

Supersymmetry (SUSY) is one of the favoured theories describing physics beyond the SM. It introduces a new symmetry between fermions and bosons: each spin-1/2 fermion of the SM is paired with a new spin-0 boson and each spin-1 gauge boson of the SM is assigned to a new spin-1/2 fermion. If the symmetry is perfectly realized in nature, these so-called sparticles should have the same mass and the same gauge quantum numbers as their SM partners. However, since no supersymmetric particle has been observed up to now, supersymmetry must be broken, the sparticle masses being beyond the reach of former collider experiments [6].

Restricting the theory to a minimal amount of new particles and interactions leads to the Minimal Supersymmetric Standard Model (MSSM). Besides the contents of the SM (except for the postulated single Higgs boson), the MSSM comprises the particles as depicted in Table 1.2. In contrast to the SM, the MSSM has two Higgs doublets leading to five Higgs particles. Squarks and sleptons represent the supersymmetric partners of quarks and leptons. By mixing higgsinos and gauginos, the supersymmetric gauge eigenstates of the Higgs doublets and the weak gauge bosons, into mass eigenstates, one ends up with four neutral neutralinos and four charged charginos. Furthermore, the gluinos are the supersymmetric partners of the gluons. The requirement of a local invariance under supersymmetric transformations results in a new gauge boson, the spin-2 graviton, and its supersymmetric partner, the spin-3/2 gravitino.
To prevent the violation of the lepton and baryon numbers in the MSSM, which could result in a fast decay of the proton, one introduces a further symmetry, the R-parity. It is defined as $P_R = (-1)^{3(B-L)+2S}$ with the lepton number, L, the baryon number, B, and the spin quantum number, S. Hence, SM particles and the five Higgs bosons have $P_R = (+1)$, while all sparticles have $P_R = (-1)$. As consequence of the R-parity conservation, sparticles can only be produced in pairs and the lightest sparticle (LSP) is stable. Since the breaking mechanism of supersymmetry is unknown, the MSSM has a large number of free parameters (124) [10].

To reduce the number of SUSY parameters and to get better insight into the SUSY breaking mechanism, different supersymmetric models were developed based on the following picture: the origin of the SUSY breaking is unknown and takes place in a "hidden sector"; this symmetry breaking is mediated by messenger fields to the "visible sector", the MSSM. The different models vary in the choice of messenger fields.
The most popular scenario is mSUGRA (minimal SUper GRAvity). In this model, the mediator of the SUSY breaking is represented by the graviton. Due to the unification of parameters at an energy scale of $\approx \mathcal{O}(10^{17}$ GeV), mSUGRA has only five parameters at this scale: the mass of squarks and sleptons, m_0, the gaugino mass, $m_{1/2}$, the tri-linear coupling, A_0, the ratio of the two Higgs vacuum expectation values, $\tan(\beta)$, and the sign of the Higgs mass parameter, $\text{sign}(\mu)$. The complete mass spectrum and all mixing angles of the MSSM can be predicted by these five parameters. On the other hand, since no sparticle has been detected up to now, the 5-dimensional parameter space of mSUGRA has been constrained by the experimental exclusion of sparticle mass ranges, by Higgs or rare decay searches and by cosmological constraints.
Due to the conservation of R-parity in mSUGRA, all sparticles decay into the LSP, usually

Table 1.2: The undetected particles of the MSSM [10], [11]. The squarks and sleptons of the 3^{rd} generation are indicated by "1" or "2" instead of "L" or "R" since, for the 3^{rd} generation, the mixture of gauge eigenstates in mass eigenstates is not negligible anymore.

name	spin	P_R	mass eigenstates
Higgs boson	0	+1	h^0 H^0 A^0 $H^\pm$
squarks	0	-1	$\tilde{u}_L$ $\tilde{u}_R$ $\tilde{d}_L$ $\tilde{d}_R$ $\tilde{c}_L$ $\tilde{c}_R$ $\tilde{s}_L$ $\tilde{s}_R$ $\tilde{t}_1$ $\tilde{t}_2$ $\tilde{b}_1$ $\tilde{b}_2$
sleptons	0	-1	$\tilde{e}_L$ $\tilde{e}_R$ $\tilde{\nu}_e$ $\tilde{\mu}_L$ $\tilde{\mu}_R$ $\tilde{\nu}_\mu$ $\tilde{\tau}_1$ $\tilde{\tau}_2$ $\tilde{\nu}_\tau$
neutralinos	1/2	-1	$\tilde{\chi}_1^0$ $\tilde{\chi}_2^0$ $\tilde{\chi}_3^0$ $\tilde{\chi}_4^0$
charginos	1/2	-1	$\tilde{\chi}_1^\pm$ $\tilde{\chi}_2^\pm$
gluinos	1/2	-1	$\tilde{g}$
gravitinos	3/2	-1	$\tilde{G}$

the lightest neutralino. Since neutralinos interact only weakly, they can not be directly detected by a multi-purpose detector like ATLAS [6], [10].

1.2.4 Supersymmetry as a solution of standard model problems

Besides the aesthetic appeal, the theory of supersymmetry is able to solve some of the SM problems [6]:

- Due to the postulated existence of sparticles, the quadratic loop corrections in Equation (1.1) are cancelled as the contributions of SM fermions and the contributions of SUSY bosons bear opposite signs. Hence, the fine tuning problem can be resolved by SUSY if the mass difference between the partners is at most $\mathcal{O}(1\ \text{TeV})$.
- In the case of mSUGRA, the LSP is usually a neutralino which interacts only weakly. Therefore, it is an excellent candidate for cold dark matter.
- A local supersymmetry introduces the graviton and the gravitino in the particle list. Thus, gravitation is also considered in SUSY. However, as a 4-dimensional theory of gravitation, it is not renormalizable and is therefore no candidate for a fundamental theory. One possible solution is to consider particles not as point-like objects, but as extended ones; this is the basic concept of string theories. At the moment, SUSY is a necessary ingredient to guarantee the consistency of string theories.

Chapter 2

The search for supersymmetry in the 1-lepton channel

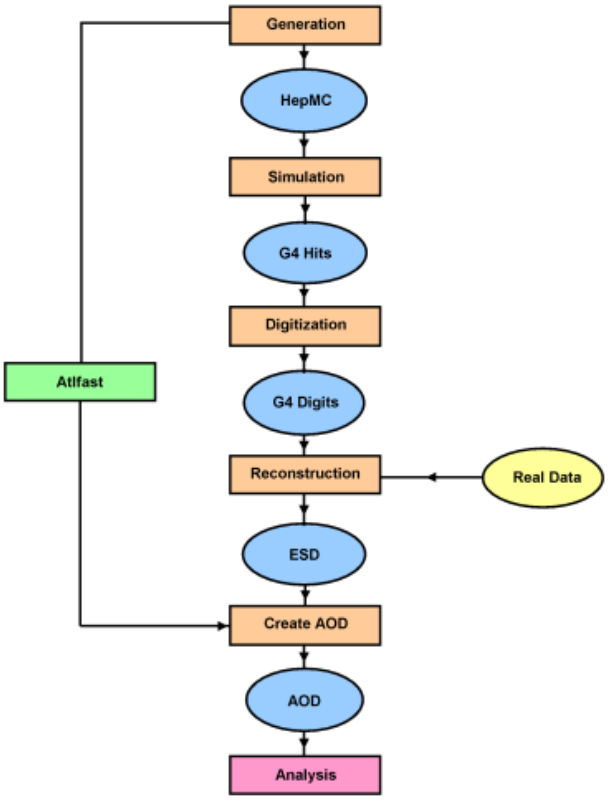
Since LHC is not yet running, no data is available up to now from real p-p collisions. Therefore, all analysis work, such as developing search strategies for theories beyond the standard model, relies on simulated data. A summary of the efforts done in this direction for the theory of supersymmetry is published in [12].
The present section gives a short review of how the simulated data is produced followed by a description of the reconstruction and definition of the considered objects. Finally, an introduction to the analysis of the 1-lepton SUSY channel and its QCD background is given.

2.1 Event generation and ATLAS software

The production and simulation of events is performed within the ATLAS software framework Athena, a skeleton providing and connecting a large amount of software tools necessary to manage and treat the real data of proton-proton collisions or simulated data from event generators [13].
In the latter case, the events produced by an event generator run through different stages of a computing chain, depicted in Figure 2.1, until they have achieved the format used for the physics analysis. After generation, the events are passed through a GEANT4 simulation of the ATLAS detector. GEANT4 is a software toolkit describing the interaction of particles with matter [14]: the interactions of the previously generated particles with the different ATLAS detector components are simulated and stored in this step in the form of Hits. These GEANT4 Hits are subsequently digitized. Thereafter, the GEANT4 Digits are run through the reconstruction where the information of the GEANT4 Digits is assembled in particle tracks and energy deposits, for instance, and the identification of the objects traversing the detector takes place. The reconstruction and the subsequent steps are done analogously for simulated data and real data. The data format containing very detailed information about the reconstructed event is called ESD (Event Summary Data). In a final step, these ESD are summarized to AOD (Analysis Object Data), a more user-friendly data format.
The AOD can now be directly analyzed or, before the investigation, further slimmed and adjusted to the purposes of the planned analysis by the Athena EventView package [15], for example. By means of this tool, the range of variables, like the number of reconstructed jets in an event, can already be restricted depending on the topic of the planned physics analysis. EventView exists in different ”flavours”, like HighPtView or TopView, providing specific tools for the different research areas. The desired information extracted from the AOD can now be stored in a data format called Ntuples produced by EventView. These Ntuples can be easily accessed by ROOT, an object-oriented data analysis framework devoted to the

needs of data analysis in high energy physics [16]. ROOT is developed and maintained by physicists for physicists: along with a C++ interpreter, it provides a large amount of classes written in C++ to support the data analysis for users.

Figure 2.1: Computing chain of the ATHENA software framework [13].



The outlined full chain can be circumvented by Atlfast, the Atlas Fast Simulation package [17], producing AOD directly from generator level. Atlfast is four to five orders of magnitude faster than running the full chain. This reduction of production time is achieved by a parametrization of the detector's response which is based on studies done with the full simulation.
One main computing difference between real data and simulated data is that for simulated events, the information about the reconstructed events (Reco) is complemented by the information about the generator level (Truth). The comparison of Reco and Truth can be used for studies of the reconstruction algorithm performance as well as for studies of fakes (see Section 4).

2.2 Reconstruction and definition of jets and leptons

The reconstructed objects studied in this thesis and in the analysis of the 1-lepton SUSY channel are jets and leptons (electrons, muons), where the terms "electron" and "muon" design both particle and antiparticle. These three objects are reconstructed using the following Athena algorithms [12]:

- the standard eGamma algorithm [18] is used for electron reconstruction. It is seeded by the transverse energy deposited in the electromagnetic calorimeter and attempts to match one of the reconstructed tracks of the inner detector to this deposit. For the

electron identification, various discrimination variables are defined which are mainly based on the shower shape produced in the calorimeter. Depending on how many criteria the electron candidate has to fulfill to be denoted as electron, the passing electron is labeled loose, medium or tight. For the 1-lepton channel SUSY analysis as well as for this thesis medium electrons are used.

- muons are reconstructed by the STACO (STAtistical COmbination) algorithm [19]. It statistically merges the independent track measurements in the inner detector and in the muon spectrometer. As matching criterion, the χ^2 is required to be less than 100 between an inner detector and a muon spectrometer track. If more than one inner detector track satisfies this requirement for one spectrometer track, only the inner detector track with the best matching is considered.

- the Cone4Tower algorithm [20] is used for jet reconstruction. It defines a jet as a set of constituents that lie within a cone of radius $R = 0.4$. Thereby, the axis of the cone has to be aligned with the jets 4-momentum. The algorithm is seeded by high-p_T constituents found in the calorimeter system around which it iteratively searches for stable configurations.

When running the outlined reconstruction algorithms, every object identified as an electron is in general also identified as a jet. To remove this ambiguity, a so-called overlap removal is performed: if jets and electrons are located within a cone of $R < 0.2$ (approximately the core of a jet), the jet is rejected.
Further criteria are introduced to improve the isolation of the reconstructed leptons:

- the *etcone*20 criterion is imposed to get an energetically isolated lepton: the variable *etcone*20 is calculated by subtracting the estimated lepton energy deposited in the calorimeter system from the overall deposited energy within a cone of radius $R = 0.2$. Thus, *etcone*20 is a quantity describing the activity inside the calorimeter around the lepton. By asking *etcone*20 < 10 GeV, for instance, the energy around the lepton is limited to 10 GeV.

- the $\Delta\mathrm{R}(l, j)$ requirement is applied to provide a minimal distance between reconstructed leptons and reconstructed jets. For muons, $\Delta\mathrm{R}(l, j) < 0.4$ means that all reconstructed muons whose distance to a reconstructed jet falls below $\Delta\mathrm{R} < 0.4$ are removed. For electrons, due to the overlap removal, $\Delta\mathrm{R}(l, j) < 0.4$ means that all reconstructed electrons which are within a distance $0.2 \leqslant \Delta\mathrm{R} < 0.4$ to a reconstructed jet are rejected. In Section 4, the $\Delta\mathrm{R}(l, j)$ requirement is only applied in combination with an energy isolation of *etcone*20 < 10 GeV.

The reconstructed objects used in the 1-lepton channel SUSY analysis have to fulfill the following cuts [12]:

- electrons and muons:
  - $p_T > 20$ GeV
  - $|\eta| < 2.5$
  - *etcone*20 < 10 GeV
  - $\Delta\mathrm{R}(l, j) < 0.4$

- jets:
  - $p_T > 20$ GeV
  - $|\eta| < 2.5$

In the rest of this thesis, when no *etcone*20 cut or $\Delta R(l, j)$ requirement is applied on the reconstructed leptons, they are denoted as non-isolated leptons. The isolation consisting of *etcone*20 < 10 GeV and $\Delta R(l, j) > 0.4$ is the one used in the SUSY Computing System Challenge (CSC) Notes, published as a section of [12], and is therefore called CSC isolation below.
All events containing non-isolated electrons with $p_T > 10$ GeV aiming at the crack region of the calorimeter ($1.37 < |\eta| < 1.52$) are rejected (vetocrack) because this region suffers from a high level of electron misidentification (see Section 1.1.2) [12].

2.3 The 1-lepton SUSY analysis

Besides the tau mode and the b-jets mode, the bulk of R-parity conserved SUSY search strategies being developed for the ATLAS detector consists of the different lepton modes [12], which are defined by the number of reconstructed leptons in the final state, from 0 to 3. Here, we will concentrate on the mode asking for exactly 1 lepton in the final state of the event. It is an inclusive SUSY search meaning that the decay channel from which the lepton originates is not specified. The analysis is designed as a classical cut study: at first, a couple of convenient variables have to be found which take different values for the SUSY signal and the standard model background. Afterwards, proper values for these quantities have to be chosen so that cutting on these variables leads to an excess of SUSY signal over the background.
Several hard cuts are applied to the simulated SUSY signal samples as well as to the background samples to discriminate signal from background:

1. exactly 1 isolated lepton (electron or muon)
2. at least 4 jets: 1 with $p_T > 100$ GeV and 3 with $p_T > 50$ GeV
3. $\not{E}_T > \max(100\,\text{GeV}, 0.2 \cdot M_{eff})$
4. $S_T > 0.2$
5. $M_T > 100$ GeV
6. $M_{eff} > 800$ GeV

The first cut asks for events having 1 CSC-isolated lepton with $p_T > 20$ GeV and no further CSC-isolated lepton with $p_T > 10$ GeV.
The transverse sphericity, S_T, is an event-shape variable ranging from 0 to 1 describing the degree of isotropy of an event within the x-y plane of the detector. It is the projection of the event shape variable sphericity, S, from three to two dimensions [21]. For instance, an event with $S_T = 0$ contains only one axis realized by the jets or leptons, whereas $S_T = 1$ belongs to an ideally isotropic distribution of objects.
The missing transverse energy, $\not{E}_T$, indicates in general the occurrence of a neutrino or, in the case of SUSY, of a neutralino: since these particles interact only weakly, they are not directly detectable in ATLAS. As the sum of the transverse energy E_T in an event has to amount to ≈ 0, $\not{E}_T$ should belong to an undetected particle in case of a proper energy measurement.
The transverse mass, M_T, and the effective mass, M_{eff}, are further observables defined as follows [22]:

$$M_T = \sqrt{2 p_T^{lep} \not{E}_T (1 - \cos\phi(\vec{\not{p}}_T, \vec{p}_T^{\,lep}))}, \tag{2.1}$$

$$M_{eff} = \sum_{i=1}^{4} p_T^{jet,i} + p_T^{lep} + \not{E}_T. \tag{2.2}$$

For the gauge boson W decaying leptonically into a lepton and a neutrino, M_T corresponds to the mass of the decaying particle. M_{eff} gives an idea of the mass of the initially produced sparticles as it is defined as a sum over the main transverse energy components which originate from the decay of the initial sparticles.
The choice of variables used for the cuts reflects the characteristics of SUSY events (mSUGRA provided):

- many rather hard jets due to the cascade decays of squarks and gluinos, which are the most abundantly produced sparticles in p-p collisions if they are light enough [12]
- large $\not{E}_T$ due to the two escaping LSP s
- large S_T because of the isotropic shape of SUSY events

The QCD background from strongly interacting quarks or gluons is mainly reduced by the first, the third and the fourth cut: the efficiency of the first cut can be explained by the fact that there can be no lepton produced in the initial strong interaction of QCD events since leptons have no colour charge. The source of the small fraction of reconstructed leptons occurring anyhow in QCD events is the topic of Section 4. Furthermore, the cut on $\not{E}_T$ should effectively reduce the QCD background since, if the jet energies are measured well enough, only small amounts of $\not{E}_T$ occur in QCD events which originate from heavy-flavour decays. As the considered QCD events are mainly oriented back-to-back and therefore have low S_T values, the cut $S_T > 0.2$ further reduces the background from QCD.
By means of the cut $M_T > 100$ GeV, the background from $t\bar{t}$ can be effectively reduced: the top quarks decay into a b quark and a W boson, which can in turn decay into a lepton and a neutrino, leading to a M_T value peaking at the W mass.
Further background comes from events containing at least one gauge boson decaying leptonically: W + jets, Z + jets and Dibosons (WW, ZZ, WZ). Just as $t\bar{t}$ background, the background containing a W can be particularly suppressed by the M_T cut.

Table 2.1: Cut flow for the SUSY signal sample at benchmark point SU3 ($m_0 = 100$ GeV, $m_{1/2} = 300$ GeV, $A_0 = -300$ GeV, $\tan\beta = 6$, $sign(\mu) = +$) and for the main backgrounds, normalized to $L = 1$ fb^{-1} [22].

	SU3	$t\bar{t}$	W+j	Z+j	Diboson	QCD
Total	27680	450000	19190	15110	55940	1644146
1.cut	4000	183994	5909	3448	29332	3937
2.cut	1491	15897	1206	473	31	1024
3.cut	995	2030	422	22	7	< 113
4.cut	768	1549	316	16	5	< 113
5.cut	451	132	14	1	1	< 113
6.cut	363	36	5	0.2	0	< 113

Table 2.1 shows the cut flow for the signal sample at benchmark point SU3 as well as for the main standard model backgrounds. The samples are normalized to an integrated luminosity $L = 1$ fb^{-1}. Already after the second cut no QCD event is left. Therefore, an upper limit at 90% confidence level corresponding to approximately 113 events is stated for the remaining cuts (for a definition of classical confidence levels, see Section 3.3). Applying all cuts leads to a signal significance of $\frac{S}{\sqrt{B}} = \frac{363}{\sqrt{36+5+0.2+0+113}} \approx 29$, including the large uncertainty on QCD background.

2.4 QCD background in the 1-lepton SUSY channel

Section 2.3 demonstrated the strong impact of the cuts of the 1-lepton SUSY analysis on QCD background: no single event is left after the second cut. Hence, the expected QCD background for real data, consisting of a much larger amount of events than the available Monte Carlo samples, can be predicted only in a very coarse way. One possible solution to this problem would be the simulation of a larger sample. However, the number of produced QCD events is relatively large. The problem is that it needs the simulation of an immense number of QCD events to get a reasonable integrated luminosity as the considered QCD events have huge cross sections, except for the extremely high-energetic range (see Table 3.1). Therefore, in Section 3 a strategy for a better estimation of the QCD background will be developed.
The second intention of this thesis is to get a better understanding of the QCD background for the 1-lepton SUSY channel. Therefore Section 4 will deal with the origin and the rate of reconstructed leptons occurring in QCD background.

Chapter 3

Cut efficiency estimation on the QCD background

The aim of the present study is to get a better prediction of the QCD background for the 1-lepton SUSY analysis by an approach based on loosening the cuts and neglecting the correlations between more than two variables.
The figures that are shown in this section contain only statistical uncertainties.

3.1 Properties of the Monte Carlo samples

This cut study on the QCD background for the 1-lepton SUSY searches was performed within the analysis framework ROOT (version 5.17). It is based on five QCD multijet samples named J4, J5, J6, J7 and J8 with the CSC ID numbers 008090, 008091, 008092, 008093 and 008094, respectively [23]. The same samples were used in Section 2.3 to evaluate QCD background. They were produced, simulated and reconstructed within the Athena framework release 12.0.6.1. PYTHIA (release 6.403) [8] linked to Athena was used for the event generation. PYTHIA is a general-purpose event generator describing proton-proton, electron-electron or proton-electron collisions. As SUSY events within the mSUGRA framework are expected to have lots of high momentum jets and high missing transverse energy $\not{E}_T$ (see 2.3), the QCD background samples J4, J5 and J6 were produced with the following filters at generator level [23]:

- Events are required to have at least 2 high-p_T jets (the highest jet transverse momentum has to lie above 80 GeV, the second, above 40 GeV) with $|\eta| < 5$
- $\not{E}_T \gtrsim 100$ GeV [24]

The high-p_T samples J7 and J8 were produced without any event filter as their cross section is sufficiently low to obtain a sizable integrated luminosity.
Moreover, the five samples differ in ckin(3) and ckin(4) (See Table 3.1) [23] [25]. These are PYTHIA variables describing the minimum and maximum p_T values in the $2 \rightarrow 2$ subprocess of strongly interacting quarks or gluons [8]. Furthermore, one can gather from Table 3.1 that all considered samples, except J7 and J8, possess high cross sections σ compared to the number of generated events N_{sample}. This leads to very small generated integrated luminosities $L = \frac{N_{sample}}{\sigma}$. For J4, J5 and J6, the listed cross sections are effective cross sections, i.e. the cross sections referring to the part of the phase space passing the event filter. The event filter efficiencies EF_{eff} are also listed in Table 3.1 [12]. For the simulation step, the full Geant4 simulation linked to Athena [14] was applied. The transformation of the AOD into Ntuples, that can be analyzed by ROOT and that are especially suited for analysis in the field

Table 3.1: The four columns display the effective cross sections σ, the event filter efficiency EF_{eff}, the generated integrated luminosity L and the PYTHIA variables ckin(3,4), respectively.

sample	σ [fb]	EF_{eff} [%]	L [fb^{-1}]	ckin(3)/ckin(4) [GeV]
J4	916400	0.29	0.03	140/280
J5	655000	2.85	0.14	280/560
J6	67424	19.6	0.40	560/1120
J7	5300	100	0.66	1120/2240
J8	22.1	100	192.31	2240/-
QCD	1644146.1	-	1.00	140/-

Table 3.2: Cut flow of the samples J4 - J8 when applying the 1-lepton SUSY analysis cuts (see Section 2.3), normalized to $L = 1$ fb^{-1} [25].

sample	N_{total}	lepton cuts	jet cuts	$\not{E}_T$ cut	S_T cut	M_T cut
J4	916400	2543.5	448.8	0	0	0
J5	655000	1294.8	520.8	0	0	0
J6	67424	89.9	49.9	0	0	0
J7	5300	9.1	4.5	0	0	0
J8	22.1	0	0	0	0	0
QCD	1644146.1	3937.3	1024.2	0	0	0

of supersymmetry, was realized by the Athena package SUSYView (release 12.0.7.2) [26]. The Ntuple files produced by SUSYView contain information about the generated event (Truth) and the reconstructed event with the full detector simulation (Reco). For this study, the reconstructed data is analyzed.
The five given samples are merged together after normalizing each of them to $L = 1$ fb^{-1}: the outcome is one sample comprising 1644146.1 events. However, due to the normalizing procedure, this sample contains lots of identical events, a fact that has to be taken into account when evaluating statistical errors later on.

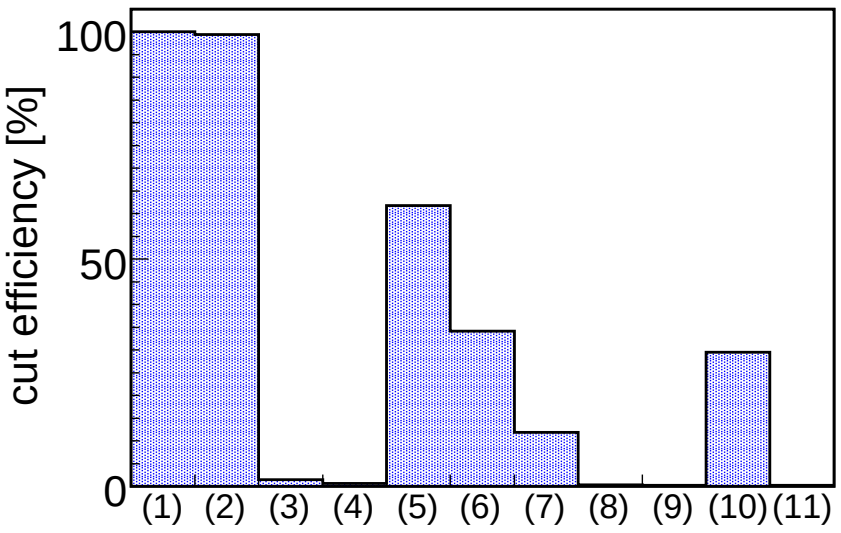
3.2 Evaluating cut efficiencies under different constraints

As already mentioned in Section 2.3, if one applies all the cuts of the 1-lepton SUSY analysis on the QCD samples, one encounters a statistical problem: no event is left after cutting on $\not{E}_T$ (see Table 3.2). Therefore, a strategy has to be developed to increase statistics after cutting on leptons and jets, in order to obtain an upper limit on events passing the mentioned cuts. As a first step, the different cuts, outlined in Section 2.3, are separated. However, it is not possible to separate them completely as the following quantities, which cuts are applied on, need preceding cuts (pre-cuts) before their evaluation:

- at least 4 jets and exactly 1 lepton are necessary to evaluate M_{eff}. For events with more than 4 reconstructed jets, the sum over the jet momenta comprises the 4 jets with the highest p_T (see Equation (2.2)).
- at least 1 jet or at least 1 lepton is necessary to evaluate S_T
- exactly 1 lepton is necessary to evaluate M_T (see Equation (2.1))

The cut efficiencies after a maximal cut separation are shown in Figure 3.1. These cut efficiencies are calculated by dividing the number of events passing a cut by the number of events before *any* cut.

Figure 3.1: Cut efficiencies of different cuts applied separately on the merged QCD samples. The labels on the x-coordinate stand for the following cuts: (1) no cut, (2) vetocrack (v), (3) v + 1 lepton, (4) v + 1 HighPtLepton, (5) v + $\geq$ 4 Jets, (6) v + $\geq$ 4 HighPtJets, (7) v + $\not{E}_T \geq$ 100 GeV, (8) v + M_{eff} $\geq$ 800 GeV [$\geq$4J][1l], (9) v + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l], (10) v + $S_T \geq 0.2$ [$\geq$1J || $\geq$1l], (11) v + $M_T \geq$ 100 GeV [1l]. "[$\geq$4J][1l]", "[1l]" and "[$\geq$1J || $\geq$1l]" denote the additional necessary pre-cuts for the evaluation of M_{eff}, M_T and S_T, respectively, see Section 3.2. "1 lepton" asks for exactly one lepton with $p_T > 10$ GeV. "1 HighPtLepton" requires exactly one lepton with $p_T > 20$ GeV and no further with $p_T > 10$ GeV. Compared to the lepton definition given in Section 2.2, the modified definition is used as outlined in Section 3.2. "$\geq$ 4 Jets" requires at least four jets which are defined as in Section 2.2. "$\geq$ 4 HighPtJets" asks for at least four jets, where four jets have to fulfill $p_T > 50$ GeV and the one with the maximal p_T has to satisfy $p_T > 100$ GeV.



Since the observables cuts are applied on are not independent of each other, one is not allowed to simply multiply chosen cut efficiencies of Figure 3.1 to get the percentage of events passing the selected cuts. Thus, one is forced to study the influence of preceding cuts on a given cut efficiency. For this analysis, the following approximations and simplifications are made:

- Since statistics is still high enough after the lepton and jet cuts (see Table 3.2), these cuts are directly applied on the sample. Thus, the focus is put on evaluating the $\not{E}_T$, S_T, M_T and M_{eff} cut efficiencies.
- Only correlations between *two* quantities are taken into consideration.
- For studying the $\not{E}_T$ cut efficiency, the cut on $\not{E}_T$ is simplified to $\not{E}_T \geq 100$ GeV instead of using $\not{E}_T > \max(100\,\text{GeV}, 0.2 \cdot M_{eff})$ because of the necessary pre-cuts for evaluating M_{eff} that severely decrease statistics. By contrast, when considering the $\not{E}_T$ cut as a pre-cut for S_T, M_T and M_{eff}, two pre-cut definitions involving $\not{E}_T$ are studied: one time defined as $\not{E}_T \geq 100$ GeV and the second time defined as $\not{E}_T \geq (0.2 \cdot M_{eff})$. These two simplifications only make the study more pessimistic.
- Changes are done in the definition of reconstructed leptons to increase statistics: in contrast to Section 2.2, leptons have to fulfill $p_T > 10$ GeV instead of $p_T > 20$ GeV. Furthermore, the isolation cut $etcone20 < 10$ GeV is removed. The reconstructed jets are defined as in Section 2.2.

The behaviour of the $\not{E}_T$, S_T, M_T and M_{eff} cut efficiencies under different pre-cuts is shown in the Tables 3.3, 3.4, 3.5 and 3.6. This time, the cut efficiency represents the ratio of the remaining number of events left after a considered cut and the number of events before the given cut (after the pre-cuts). The statistical uncertainties of the cut efficiencies are calculated as outlined in Section 3.3. The number of events passing the cuts for each individual QCD sample can be found in the Tables A.1, A.2, A.3 and A.4.

Table 3.3: $\not{E}_T$ cut efficiency influenced by several pre-cuts. The uncertainty added to the calculated value represents the upper limit at 68.27%C.L. (value in brackets). For the definition of the pre-cuts, see the caption of Figure 3.1.

pre-cut	$\not{E}_T$ cut efficiency [%]
v	11.91 + 0.13 (12.04)
v + 1 lepton	11.74 + 1.06 (12.80)
v + 1 HighPtLepton	12.42 + 1.62 (14.04)
v + $\geq$ 4 Jets	7.15 + 0.13 (7.28)
v + $\geq$ 4 HighPtJets	5.28 + 0.14 (5.42)
v + $M_{eff} \geq 800$ GeV [$\geq$4J][1l]	8.62 + 1.23 (9.85)
v + $S_T \geq 0.2$ [$\geq$1J $\|\|$ $\geq$1l]	7.17 + 0.20 (7.37)
v + $M_T \geq 100$ GeV [1l]	27.89 + 5.25 (33.14)

Table 3.4: S_T cut efficiency influenced by several pre-cuts. The uncertainty added to the calculated value represents the upper limit at 68.27%C.L. (value in brackets).

pre-cut	S_T cut efficiency [%]
v + [≥1J || ≥1l]	29.72 + 0.18 (29.90)
v + [≥1J || ≥1l] + 1 lepton	34.40 + 1.54 (35.94)
v + [≥1J || ≥1l] + 1 HighPtLepton	37.77 + 2.41 (40.18)
v + [≥1J || ≥1l] + ≥ 4 Jets	41.03 + 0.23 (41.26)
v + [≥1J || ≥1l] + ≥ 4 HighPtJets	52.54 + 0.28 (52.82)
v + [≥1J || ≥1l] + $\not{E}_T \geq 100$ GeV	17.90 + 0.47 (18.37)
v + [≥1J || ≥1l] + $M_{eff} \geq 800$ GeV [≥4J][1l]	37.71 + 1.96 (39.67)
v + [≥1J || ≥1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J][1l]	51.51 + 9.59 (61.10)
v + [≥1J || ≥1l] + $M_T \geq 100$ GeV [1l]	29.64 + 5.36 (35.00)

Table 3.5: M_T cut efficiency influenced by several pre-cuts. The uncertainty added to the calculated value represents the upper limit at 68.27%C.L. (value in brackets).

pre-cut	M_T cut efficiency [%]
v + [1l]	5.70 + 0.76 (6.46)
v + [1l] + HighPtLepton	12.59 + 1.64 (14.23)
v + [1l] + ≥ 4 Jets	5.28 + 0.87 (6.15)
v + [1l] + ≥ 4 HighPtJets	3.55 + 0.80 (4.35)
v + [1l] + $\not{E}_T \geq 100$ GeV	13.55 + 2.68 (16.23)
v + [1l] + $M_{eff} \geq 800$ GeV [≥4J]	5.35 + 1.00 (6.35)
v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	16.37 + 6.51 (22.88)
v + [1l] + $S_T \geq 0.2$	4.91 + 0.98 (5.89)

Table 3.6: M_{eff} cut efficiency influenced by several pre-cuts. The uncertainty added to the calculated value represents the upper limit at 68.27%C.L. (value in brackets).

pre-cut	M_{eff} cut efficiency [%]
v + [≥4J][1l]	38.92 + 1.64 (40.56)
v + [≥4J][1l] + HighPtLepton	40.14 + 2.59 (42.73)
v + [≥4J][1l] + HighPtJets	53.99 + 2.45 (56.44)
v + [≥4J][1l] + $\not{E}_T \geq 100$ GeV	45.32 + 6.56 (51.88)
v + [≥4J][1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$	1.02 + 4.57 (5.59)
v + [≥4J][1l] + $S_T \geq 0.2$	32.60 + 2.38 (34.98)
v + [≥4J][1l] + $M_T \geq 100$ GeV	39.43 + 7.73 (47.16)

3.3 Statistical uncertainties on the cut efficiencies

As one can gather from the Tables A.1, A.2, A.3 and A.4, often less than 10 events pass a cut, and in some cases only the pre-cut brings the number of events down to this level. For calculating the statistical uncertainties, Poisson statistics is used when up to 10 events are left; Gaussian statistics is used when more than 10 events are left.
As the aim is an estimation of the upper limit of *background* for a given signal, an upper confidence level of 68.27% is used, which corresponds to one standard deviation when using Gaussian statistics.
For Poisson statistics, the upper confidence limit is calculated by the Neyman's construction [27]. To determine the upper limit by this method, one sums up all Poisson probabilities in the range from 0 up to the number of events retrieved from the sample for a fixed value of the Poisson mean μ. For a given number of events, this cumulative probability P decreases when increasing μ from 0 to 20 in small steps. When $P = (100 - 68.27)$ %, the corresponding μ represents the upper limit at 68.27% confidence level. For the range from 0 to 10 events at 68.27% confidence level, the upper limits are shown in Table B.1.
As an upper *classical* (frequentist) confidence level, these limits denote that the probability that the confidence region calculated with our merged sample covers the true value of the Poisson mean μ, is 68.27% [27]. By contrast, a Bayesian confidence level would be interpreted as follows [27]: the probability that the true value lies within the calculated interval is 68.27%. Unlike for the classical confidence levels, a prior probability distribution of the true mean μ_T would be necessary to determine a Bayesian confidence level.
As already mentioned in Section 3.1, many events of the merged sample are identical due to normalization. Thus, the statistical uncertainties on the cut efficiencies $\Delta\varepsilon$ shown in the Tables 3.3, 3.4, 3.5 and 3.6 are calculated using Gaussian error propagation as approximation in which the errors on the five samples J4 - J8 are combined:

$$(\Delta\varepsilon)^2 = \left(\frac{\varepsilon}{N}\right)^2 \sum_{i=1}^{5} A_i^2 \left[\left(\frac{N_{nc}}{N_c}\right)^2 (\Delta N_{c\,i})^2 + (\Delta N_{nc\,i})^2 \right] \tag{3.1}$$

$\Delta\varepsilon$: statistical error on the cut efficiency
ε: cut efficiency $\varepsilon = \frac{N_c}{N}$
N: merged number of events before the cut (after previous cuts)
$N = \sum_{i=1}^{5} A_i N_i = N_{nc} + N_c$
N_i: number of events before the cut (after previous cuts) in sample i
N_{nc}: merged number of events not passing the cut
N_c: merged number of events passing the cut
A_i: scaling factor of sample i to normalize it to $L = 1\ \text{fb}^{-1}$
$A_i = \frac{\sigma_i}{N_{sample\ i}}$
$\Delta N_{c\,i}$: uncertainty on the number of events passing the cut for sample i
$\Delta N_{nc\,i}$: uncertainty on the number of events not passing the cut for sample i

3.4 Determining the combined cut efficiency

As mentioned in Section 3.2, cut efficiencies can only be influenced by cuts that are applied *before*. Therefore, one has to decide for a special cut order of $\not{E}_T$, S_T, M_T and M_{eff} as a first step (the cuts on leptons and jets are applied before without modifications). For instance, one takes the cut sequence outlined in Section 2.3. For this default cut sequence, one chooses the worst cut efficiency upper limits (calculated values plus uncertainties) on $\not{E}_T$, S_T, M_T and M_{eff} by analyzing the Tables 3.3, 3.4, 3.5 and 3.6:

- a $\not{E}_T$ cut efficiency upper limit of 14.04% after requiring 1 HighPtLepton
- a S_T cut efficiency upper limit of 61.10% after requiring $\not{E}_T \geq (0.2 \cdot M_{eff})$, besides the cuts necessary to calculate S_T and M_{eff}
- a M_T cut efficiency upper limit of 22.88% after requiring $\not{E}_T \geq (0.2 \cdot M_{eff})$, besides the cuts necessary to calculate M_T and M_{eff}
- a M_{eff} cut efficiency upper limit of 56.44% after requiring $\geq$ 4 HighPtJets, besides the cuts necessary to calculate M_{eff}

For the choice of worst cut efficiencies as well as for our further calculations, one uses as selection rule the upper limits of the efficiencies instead of the efficiencies themselves as it is hardly predictable how the errors propagate under the correlations of the variables.
The influence of the pre-cuts applied on the distribution of the corresponding observables is depicted in Figures 3.2, 3.3, 3.4 and 3.5. To compare their shapes, the distributions are normalized to 1.
The $\not{E}_T$ histograms, shown in Figure 3.2, hardly changes under the 1 HighPtLepton pre-cut. As one can see from this figure, the PYTHIA event filter on $\not{E}_T$ decreases the number of events with low $\not{E}_T$ values but is not able to exclude them: the event filter approximates the $\not{E}_T$ value at generator level, while the measured $\not{E}_T$ comes from complex interactions and detector effects. Indeed, the undetectable neutrinos of weak particle decays and energy mismeasurement are sources for $\not{E}_T$.
As one is confronted with very low statistics after applying the $\not{E}_T$ pre-cut on the S_T distribution, only a few bins can be used to draw its distribution. Nevertheless, the difference between the two S_T distributions, shown in Figure 3.3, is clearly visible. The request for $\geq$ 4 Jets necessary for the M_{eff} calculation is probably responsible for the different shapes as events containing many jets have on average higher S_T values (see Figure 3.8).
Low statistics also affects the two M_T histograms, shown in Figure 3.4. The correlation between $\not{E}_T$ and M_T follows directly from the definition of M_T (see Equation (2.1)).
The impact of requiring HighPtJets on the M_{eff} distribution, depicted in Figure 3.5, follows from the M_{eff} definition (see Equation (2.2)).
The chosen upper limits for the cut efficiencies of $\not{E}_T$, S_T, M_T and M_{eff} are now considered to be uncorrelated in this worst case scenario, so that one simply multiplies these upper limits of cut efficiencies to finally get a combined upper limit $\varepsilon_{combined}$ for the group of cuts on $\not{E}_T$, S_T, M_T and M_{eff}. For the default cut order, one obtains $\varepsilon_{combined} = 14.04\% \cdot 61.10\% \cdot 22.88\% \cdot 56.44\% = 1.11\%$. By multiplying this combined upper limit with the number of events left after the lepton and jet cuts (1024.2 events), one retrieves an upper limit of events passing all cuts. Thus, if using the default cut order, $1024.2 \cdot 1.11\% = 11.4$ events survive the cuts.

Since the choice of the worst cut efficiencies of $\not{E}_T$, S_T, M_T and M_{eff} depends on the cut order that was fixed before their selection (*relative* worst cut efficiencies), the combined cut efficiency, $\varepsilon_{combined}$, also depends on this order. By contrast, the permutation of the default cut sequence has no influence on the number of events passing all cuts if the cuts are applied directly one after another. Thus, from this point of view, it is necessary to calculate $\varepsilon_{combined}$ for all $4 \cdot 3 \cdot 2 \cdot 1 = 24$ permutations of the cuts on $\not{E}_T$, S_T, M_T and M_{eff} in order to choose the worst $\varepsilon_{combined}$.
For the selection of the relative worst cut efficiencies for different cut permutations, a number of observations can be made with the help of the Tables 3.3, 3.4, 3.5 and 3.6. Independently of the permutation of the $\not{E}_T$, S_T, M_T and M_{eff} cuts, the M_{eff} cut efficiency of 56.44% ($\geq$4 high-p_T jets pre-cut) is chosen as relative worst cut efficiency, as it is the absolute worst M_{eff} cut efficiency and as the corresponding pre-cut is independent of $\not{E}_T$, S_T and M_T. Unlike

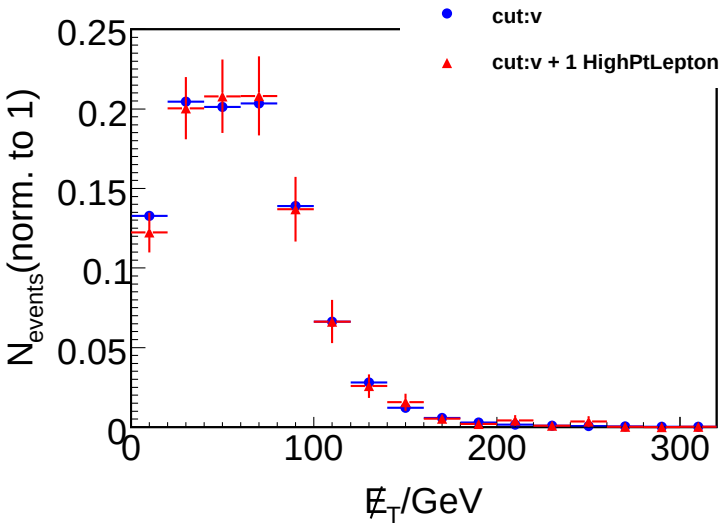


Figure 3.2: Change in the $\not{E}_T$ distribution caused by requiring 1 HighPtLepton.

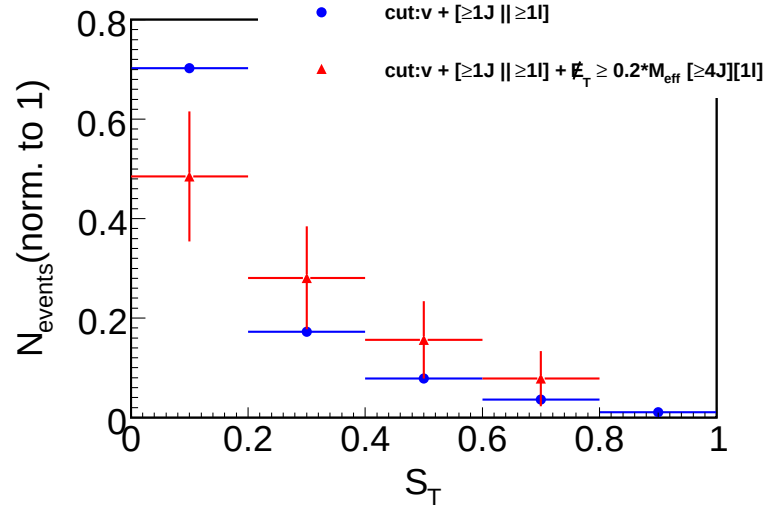


Figure 3.3: Change in the S_T distribution caused by requiring $\not{E}_T \geq (0.2 \cdot M_{eff})$.

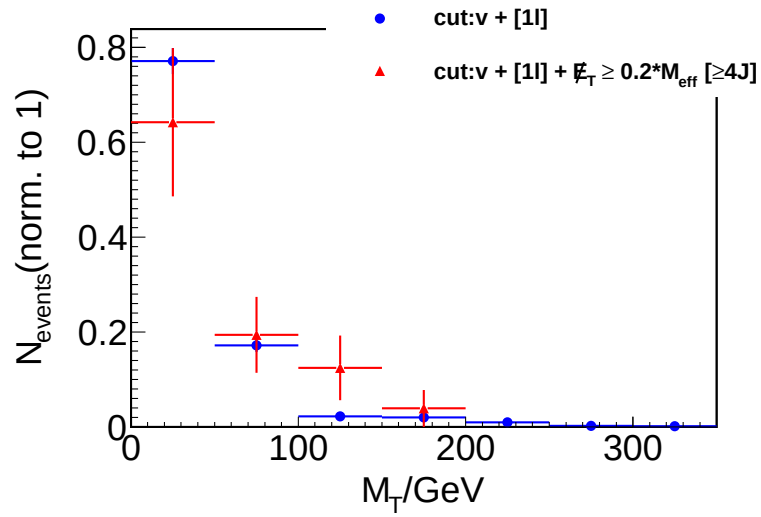


Figure 3.4: Change in the M_T distribution caused by requiring $\not{E}_T \geq (0.2 \cdot M_{eff})$.

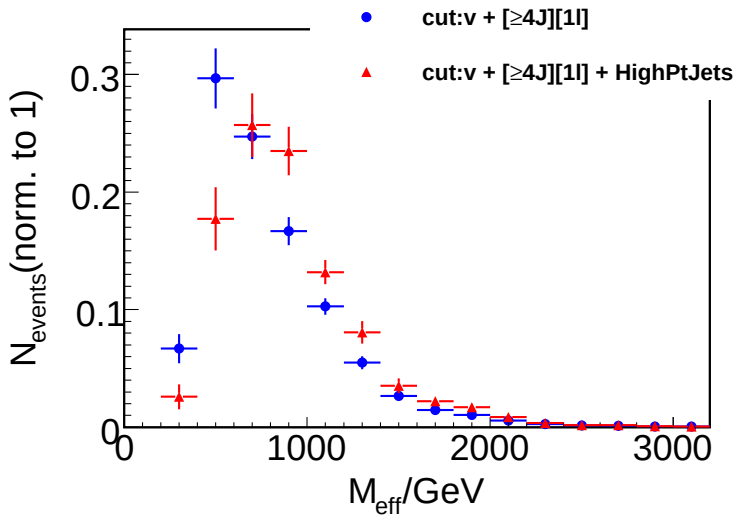


Figure 3.5: Change in the M_{eff} distribution caused by requiring $\geq$ 4 HighPtJets.

the variable M_{eff}, the choice of the relative worst cut efficiencies for $\not{E}_T$, S_T and M_T depends on the given permutation. For these quantities, the pre-cut causing the absolute worst cut efficiency depends on M_T or $\not{E}_T$. However, the next-to-worst cut efficiency for these quantities is independent of $\not{E}_T$, S_T, M_T and M_{eff}.
With this knowledge, it is quite easy to handle all 24 permutations: for a given permutation and a given variable ($\not{E}_T$, S_T, M_T), one has to find out if the cut causing the absolute worst cut efficiency is applied before or not. In the first case, one chooses the absolute worst cut efficiency for the given variable, in the second case, one simply selects the next-to-worst one:

- $\not{E}_T$: If the cut on M_T is applied before the cut on $\not{E}_T$, one has a relative worst cut efficiency upper limit of 33.14% for $\not{E}_T$. Otherwise, the upper limit is 14.04% with a pre-cut on 1 HighPtLepton.
- S_T: If the cut on $\not{E}_T$ is applied before the cut on S_T, one has a relative worst cut efficiency upper limit of 61.10% for S_T. Otherwise, the upper limit is 52.82% with a pre-cut on $\geq$ 4 HighPtJets.
- M_T: If the cut on $\not{E}_T$ is applied before the cut on M_T, one has a relative worst cut efficiency upper limit of 22.88% for M_T. Otherwise, the upper limit is 14.23% with a pre-cut on 1 HighPtLepton.

This has to be performed for all permutations and the variables $\not{E}_T$, S_T and M_T. As already mentioned, the relative worst cut efficiency for M_{eff} is 56.44% ($\geq$ 4 HighPtJets pre-cut) for all permutations.
The Figures 3.2 to 3.7 and 3.9 show the distributions of $\not{E}_T$, S_T, M_T and M_{eff} and their change under the pre-cuts just mentioned.
Figure 3.6 shows a clear shift of $\not{E}_T$ to higher values after the pre-cut on M_T, a behaviour explained by the definition of M_T (see Equation (2.1)).
Concerning the S_T distribution shown in Figure 3.7 on a logarithmic scale, the cut on $\geq$ 4 Jets already removes very efficiently events at low S_T since the isotropy of the event tends to increase with an increasing number of jets. This explains the important decrease in the S_T cut efficiency after the $\geq$ 4 Jets pre-cut. The correlation between S_T and the number of reconstructed jets can be better seen in the two dimensional histogram of Figure 3.8.
The 1 HighPtLepton pre-cut shifts the M_T distribution to higher values, as Figure 3.9 illustrates on a logarithmic scale. This behaviour is due to the definition of M_T, where the square root of the lepton p_T plays the role of a multiplicative factor (see Equation (2.1)).

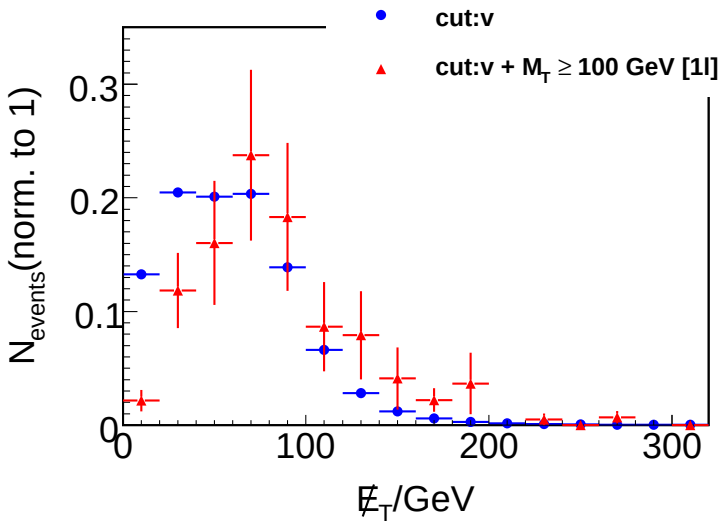


Figure 3.6: Change of the $\not{E}_T$ distribution caused by requiring $M_T \geq 100$ GeV.

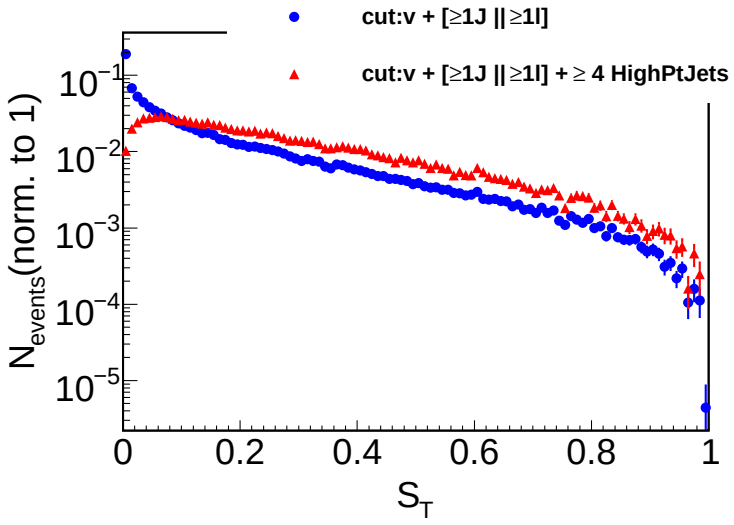


Figure 3.7: Change of the S_T distribution caused by requiring $\geq$ 4 HighPtJets. Since statistics is not seriously reduced by the considered cuts, a fine histogram binning is possible.

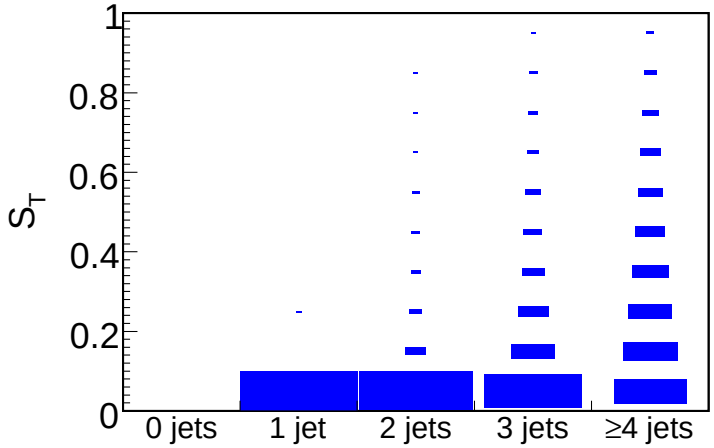


Figure 3.8: Correlation between S_T and the number of reconstructed jets. The surface of the rectangles is proportional to the histogram content.

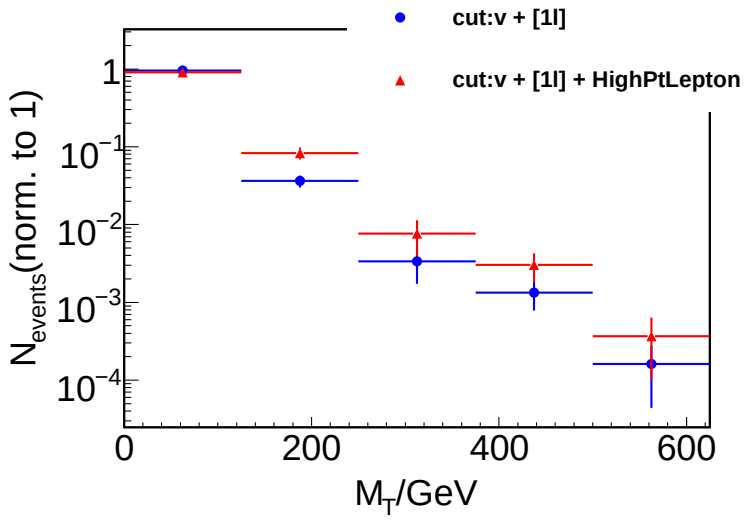


Figure 3.9: Change of the M_T distribution caused by requiring 1 HighPtLepton.

The detailed results of the study of all 24 permutations are listed in Appendix C. One obtains only 4 different results for $\varepsilon_{combined}$: 4 times $\varepsilon_{combined} = 0.96\%$, 8 times $\varepsilon_{combined} = 1.11\%$, 8 times $\varepsilon_{combined} = 1.41\%$ and 4 times $\varepsilon_{combined} = 1.63\%$. Thus, one retrieves a mean value $\overline{\varepsilon_{combined}} = 1.27\%$ (corresponding to 13.0 passing events) and an upper limit $\varepsilon_{combined\ max} = 1.63\%$ (corresponding to 16.7 passing events).

3.5 Alteration of the SUSY signal significance

With these results for the QCD background, one can again estimate the statistical discovery significance $\frac{S}{\sqrt{B}}$ of the 1-lepton SUSY signal over the main backgrounds. Using the numbers listed in Table 2.1 and the upper limit of 16.7 passing events for the QCD background, one gets a statistical significance of $\frac{S}{\sqrt{B}} = \frac{363}{\sqrt{36+5+0.2+0+16.7}} \approx 48$ for $L = 1$ fb^{-1} compared to $\frac{S}{\sqrt{B}} \approx 29$ in Section 2.3. However, one has to notice that the result for $\frac{S}{\sqrt{B}}$ obtained here is calculated using a 68.27% confidence level. This contrasts with the 90% confidence level employed in Section 2.3. When adapting the significance of Section 2.3 to include a 68.27% confidence level on the QCD background, one retrieves $\frac{S}{\sqrt{B}} = \frac{363}{\sqrt{36+5+0.2+0+56}} \approx 37$. Hence, by the estimation of the QCD background, one can improve the significance by a factor ≈ 1.3.
Furthermore, one has to consider that, up to now, only statistical uncertainties were taken into account. The main sources of systematic uncertainties at the generator level are the parton density functions (PDF), which are not very well known for high-p_T QCD events, and the description of the underlying event [12]. At detector level, the systematic uncertainties are dominated by the determination of the jet energy scale [12]. For a coarse estimation, the following values for the systematic uncertainties are used: 50% for the QCD sample and 20% for the other background samples [12]. Considering the systematic uncertainties σ_{syst}, the significance drops from $\frac{S}{\sqrt{B}} \approx 48$ to $\frac{S}{\sqrt{B+\sigma_{syst}^2}} = \frac{363}{\sqrt{(36+5+0.2+0+16.7)+(51.8+1+0+0+69.7)}} \approx 27$. Without the QCD estimation developed in this section, one obtains $\frac{S}{\sqrt{B+\sigma_{syst}^2}} = \frac{363}{\sqrt{(36+5+0.2+0+56)+(51.8+1+0+0+784)}} \approx 12$.

At first glance, it seems unnecessary to improve the significance of the SUSY signal by an estimation of the QCD background since the significance is already very high despite the large uncertainty on the QCD background. However, the given significance is the one for the benchmark point SU3. For different points in the mSUGRA parameter space,

the signal significance is worse and an estimation of the QCD background becomes more important than for SU3.

Chapter 4

Studies on leptons originating from the QCD background

The considered QCD background for SUSY is made of multijet events having a SUSY-like signature but coming from strong interactions of quarks and gluinos. Events containing top quarks are not included here but treated separately. No leptons are produced in the initial interaction of quarks or gluons. Nevertheless, the simulated detection of generated QCD events shows that, in some of these events, reconstructed leptons occur, leading to the second important background for the 1-lepton SUSY search according to the study outlined in Section 3. The present chapter will discuss the origin of the leptons, their rate in reconstructed events and the probability of a generated particle to fake a reconstructed lepton.
As events containing more than one reconstructed lepton are also considered in this study, the latter is not exclusively devoted to the 1-lepton channel. However, as will be shown later, QCD events having more than one reconstructed lepton are very rare.
This study was performed within ROOT (version 5.18) and was based on QCD Ntuples defined as in Section 3.1. However, the samples used in this section were produced in release 13.0.40.5 of the Athena framework and the generation of Ntuples out of AOD's was done by the Athena program package HighPtView release 00-01-15 [28] (see Section 2.1). The produced Ntuples J4, J5, J6, J7 and J8 have an integrated luminosity of 0.05 fb^{-1}, 0.14 fb^{-1}, 0.71 fb^{-1}, 0.83 fb^{-1} and 92.71 fb^{-1}, respectively.
As in Section 3, only statistical uncertainties are considered in the shown figures. Empty histogram bins are not labelled by the corresponding histogram marker.

4.1 Particle matching

To get to know the source of the reconstructed leptons and to calculate the fake rates afterwards, a matching between the objects at the generator level (Truth) and the reconstructed level (Reco) is necessary. The geometrical matching procedure used in this study is described in this section.

4.1.1 The matching strategy

The aim of the matching procedure is to find a partner object at the generator level for the reconstructed electrons and muons. For this task, an asymptotic approach is chosen consisting of two main steps. In the first step, the program tries to find an appropriate Truth electron (Truth muon) for a Reco electron (Reco muon), assuming that a Reco lepton originates most frequently from a Truth lepton of the same flavour. In the second step, it attempts to assign

a Truth jet to the remaining Reco leptons. The electromagnetic part of the jet can be wrongly associated with an inner track and thus be misidentified. The misidentification of a true jet as a muon can, for example, be the result of a punch through, denoting the scenario of a high-energetic jet reaching the muon system and being reconstructed as a muon. After these two steps, only a tiny fraction of $\approx 2\%$ ($\lesssim 1\%$) of the reconstructed electrons (muons) cannot be paired either to a Truth lepton or a Truth jet.
As an example, we will detail the matching between Reco and Truth electrons. Note that the procedure for muon matching would be the same: for each event, the analysis program loops over the Reco and Truth electrons in parallel, calculating for each combination of Reco and Truth electrons the distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. These values are then stored in a matrix of dimension $N_{Reco\ electrons} \times N_{Truth\ electrons}$, where $N_{Reco\ electrons}$ ($N_{Truth\ electrons}$) denotes the number of Reco (Truth) electrons in the considered event. By searching for the matrix entry with the minimal value, the first Reco-Truth pair is assembled. Afterwards, the column and the row of the matrix belonging to the selected Reco-Truth pair is removed and the next Reco-Truth pair is attributed to the entry with the minimal value of ΔR in the reduced matrix. The cancellation of the row and the column leads to a perfectly symmetric matching between Reco and Truth electrons preventing double counting. Each Reco as well as each Truth electron is only used for one Reco-Truth pair at maximum. The procedure is repeated until the number of Reco or Truth electrons is exhausted. One ends up with a vector $V_{\Delta R}$ of length $\min(N_{Reco\ electrons}, N_{Truth\ electrons})$ containing the values of ΔR for all Reco-Truth pairs. In order to determine which Reco-Truth pair is to be considered as matched, one needs a cut-off parameter, called here matching parameter $d_{Reco\ object,Truth\ object}$ (e.g. $d_{elec,elec}$ for electron-electron matching), which corresponds to the highest ΔR value acceptable for a match. In this way, one retrieves the number of Reco electrons matched to Truth electrons in one event of the sample looped over.
For the following step, i.e. the matching between Reco leptons and Truth jets, the procedure is altered in two aspects. First, one considers only the Reco leptons that were not matched to Truth leptons in the previous step. Second, the column of the matrix containing the ΔR values of the different combinations of Reco leptons and Truth jets is not cancelled during the Reco-Truth pairing. This allows the matching of more than one Reco lepton to a single Truth jet. In fact, the latter modification of the matching procedure leads to a significant decrease of $\lesssim 50\%$ ($\lesssim 70\%$) in the number of Reco electrons (muons) that are not matched to Truth leptons nor to Truth jets, depending on the isolation imposed to the Reco lepton.
Before each matching step, one has to determine well-motivated values for the matching parameter d. The determination of this parameter will be done in Section 4.1.2.
In summary, to get the number of Reco leptons matched to Truth leptons/Truth jets, the following steps have to be carried out:

1. define the matching parameter $d_{Reco\ lepton,Truth\ lepton}$
2. perform the Reco lepton to Truth lepton matching
3. define the matching parameter $d_{Reco\ lepton,Truth\ jet}$ (using only the remaining Reco leptons)
4. perform the Reco lepton to Truth jet matching (using only the remaining Reco leptons)

The following denominations will be used in the rest of this thesis: the generic term for any reconstructed lepton in QCD samples is fake lepton, as these leptons do not occur in the initial strong interactions. Reco leptons matched by Truth leptons of the same flavour are called nonprompt leptons. Reco leptons matched by Truth jets are denoted as jet-faked leptons.

4.1.2 Specification of the matching parameters

To determine the value of the matching parameter $d_{lepton,lepton}$, the distribution of ΔR values of the corresponding $V_{\Delta R}$ vectors (see 4.1.1) and the relative p_T difference $\Delta p_T(relative) = \frac{p_T^{Truth} - p_T^{Reco}}{p_T^{Truth}}$ were studied.
The following object definitions were used to have adequate statistics and comparable values for Reco and Truth objects (the isolation criteria are defined in Section 2.2):

- Reco leptons:
  - $p_T > 10$ GeV
  - $|\eta| < 2.5$
  - no $etcone20$ isolation
  - no $\Delta R(l, j)$ requirement
- Truth leptons:
  - $p_T > 10$ GeV
  - $|\eta| < 2.5$

The left part of Figure 4.1 shows the distribution of $\log_{10}(\Delta R)$ for the electron-electron matching, normalized to $L = 1$ fb^{-1}. In the right part one can see the two-dimensional scatter plot for the variables $\log_{10}(\Delta R)$ and $\Delta p_T(relative)$. Both plots can be separated into three regions. The first region is the broad maximum in the $\log_{10}(\Delta R)$ distribution at small $\log_{10}(\Delta R)$ values, where the Reco-Truth pairs have a good geometrical matching and a proper p_T matching. Only for a few pairs does the p_T of the Reco electron considerably exceed the one of the assigned Truth electron. The second region corresponds to the maximum at high $\log_{10}(\Delta R)$ values. In this region, Reco-Truth pairs are neither correlated geometrically nor energetically. In between those two peaks is a low populated transition region characterized by a $\Delta p_T(relative)$ spread similar to the first region. The general shape of the $\log_{10}(\Delta R)$ distribution is probably caused by the event shape of the QCD events: as most of the events are back-to-back, the Reco electrons which do not have a proper Truth partner in the region of "their" jet are assigned to a Truth electron in the opposite jet region, leading to high $\log_{10}(\Delta R)$ values (for $\Delta\eta = 0$: $\log_{10}(\Delta R) \approx 0.5$ corresponds to $\Delta\phi \approx \pi$). For the present study, a value of $\log_{10}(d_{elec,elec}) = -2.4$ is chosen ($d_{elec,elec} \approx 0.004$). Hence only the population of Reco-Truth pairs possessing a very good agreement in their coordinates is selected.

Figure 4.1: The $\log_{10}(\Delta R)$ distribution (left plot) and $\log_{10}(\Delta R)$ against $\Delta p_T(relative)$ (right plot) for the electron-electron matching, normalized to $L = 1$ fb^{-1}.

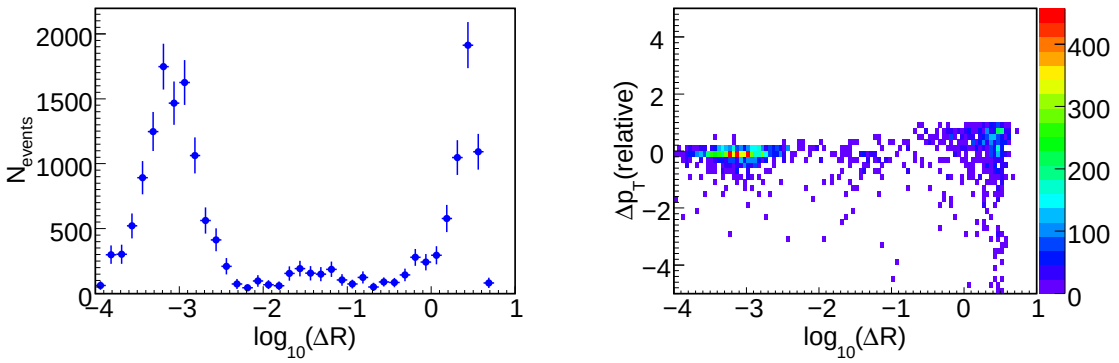


Figure 4.2: The $\log_{10}(\Delta R)$ distribution (left plot) and $\log_{10}(\Delta R)$ against $\Delta p_T(relative)$ (right plot) for the muon-muon matching, normalized to $L = 1$ fb^{-1}.

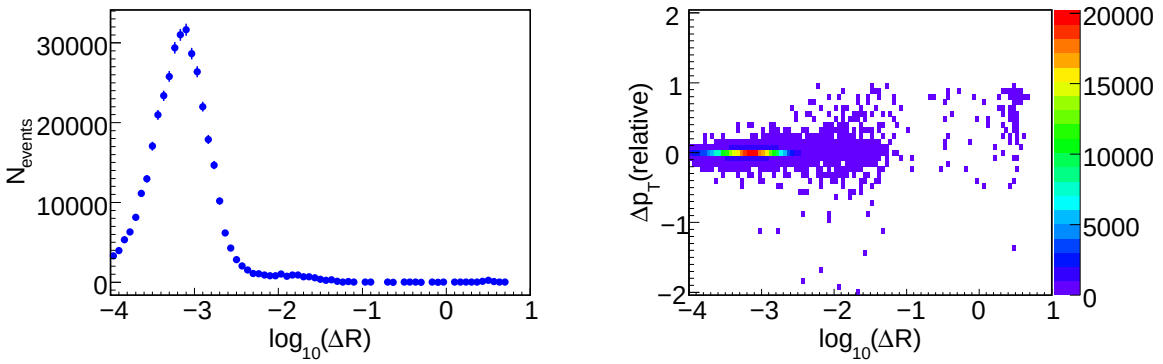
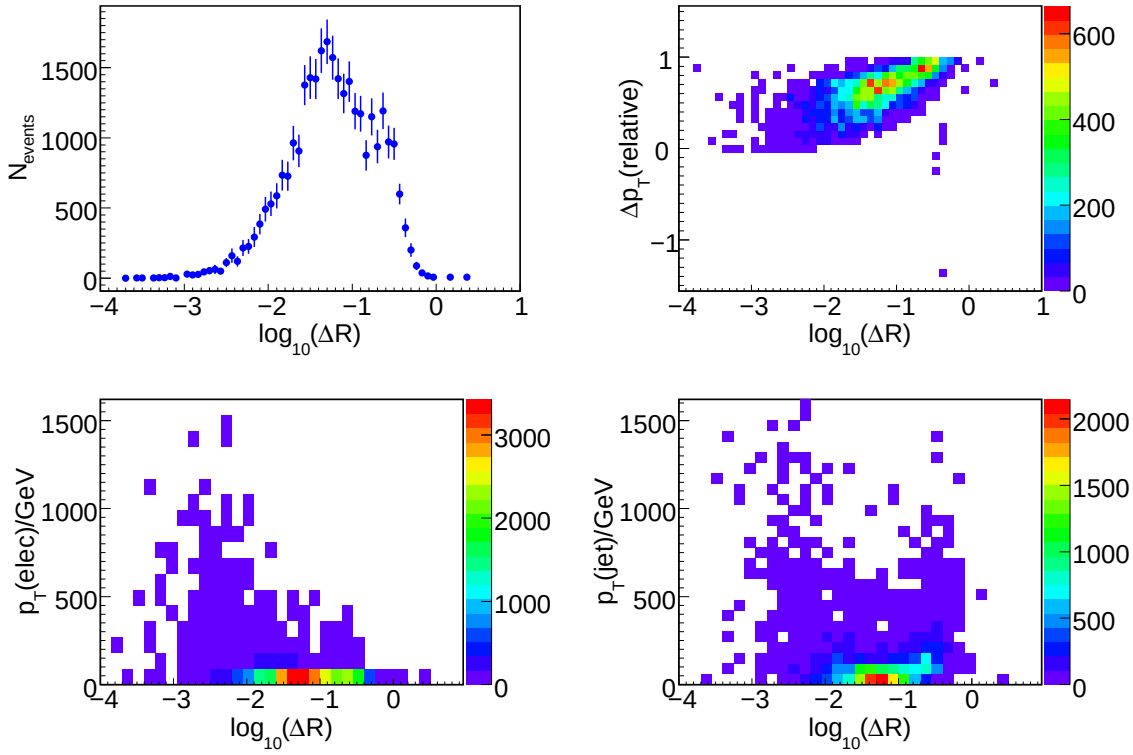


Figure 4.2 shows the analog plots for the muon-muon matching. Although the separation is not as clear as it is for the electron-electron matching, also here three different populations can be detected: the maximum for very small $\log_{10}(\Delta R)$ values with an excellent p_T matching, a tail for $-2.5 \lesssim \log_{10}(\Delta R) \lesssim -1.5$ with good p_T matching and finally a small region at high $\log_{10}(\Delta R)$ values, where no correlation between the paired Truth and Reco muon is expected. Again, a tight value is chosen for the muon-muon matching $\log_{10}(d_{muon,muon}) = -2.4$, for the same reason as for electrons and incidentally coinciding with the one for the electron-electron matching.
One can note that the distributions for the electron-electron matching have much lower statistics than the ones for muons. This is linked to the different isolations of Reco muons compared to Reco electrons and will be outlined in Section 4.2.1. Secondly, Figure 4.2 does not show a distinct maximum at high $\log_{10}(\Delta R)$ values as is seen in Figure 4.1. This difference illustrates that the ratio of nonprompt muons to jet-faked muons is much larger than the corresponding ratio for electrons, as explained below. Finally, the p_T matching for the muon-muon matching is much better than for the electron-electron matching. For electrons, large negative values of $\Delta p_T(relative)$ occur. This could be explained by Reco electrons incidentally matched to low-p_T Truth electrons but not correlated to them. These Reco electrons would be faked by jets and the matched low-p_T Truth electrons would come from decays inside these jets. As the ratio of nonprompt muons to jet-faked muons is much higher than the one for electrons, as will be shown in Section 4.3.1, this phenomenon predominantly occurs for electrons.
The tight $d_{lepton,lepton}$ values chosen in this study are much smaller than the ones stated in [12], where $d_{elec,elec} = 0.1$ ($d_{muon,muon} = 0.05$) is used.

Having performed the lepton-lepton matching, whose results are the topic of Section 4.2, the remaining leptons are used to fix $d_{lepton,jet}$ before the subsequent lepton-jet matching.
The Truth jets are constructed by the same algorithm as the Reco jets, namely a Cone4 algorithm. In contrast to the Reco objects, no overlap removal is employed for the Truth objects (see Section 2.2). For the determination of $d_{lepton,jet}$, $|\eta| < 2.5$ and $p_T > 10$ GeV are requested for Truth jets.
Figure 4.3 shows four plots: the upper plots are analog to the Figures 4.1 and 4.2, while the bottom two depict the p_T of the Reco electron and of the Truth jet against $\log_{10}(\Delta R)$, respectively.
The $\log_{10}(\Delta R)$ distribution has a maximum at $\log_{10}(\Delta R) \approx -1.3$ followed by a shoulder that drops abruptly at $\log_{10}(\Delta R) \approx -0.4$, corresponding to $\Delta R \approx 0.4$. This sudden decrease can be explained by the Cone4 jet algorithm: as the radius of the jets is $\Delta R = 0.4$, most electrons

Figure 4.3: The $\log_{10}(\Delta R)$ distribution (upper left), $\log_{10}(\Delta R)$ against $\Delta p_T(relative)$ (upper right), $\log_{10}(\Delta R)$ against the p_T of the Truth jet (lower right) and $\log_{10}(\Delta R)$ against the p_T of the Reco electron (lower left) for the electron-jet matching, normalized to $L = 1$ fb^{-1}.

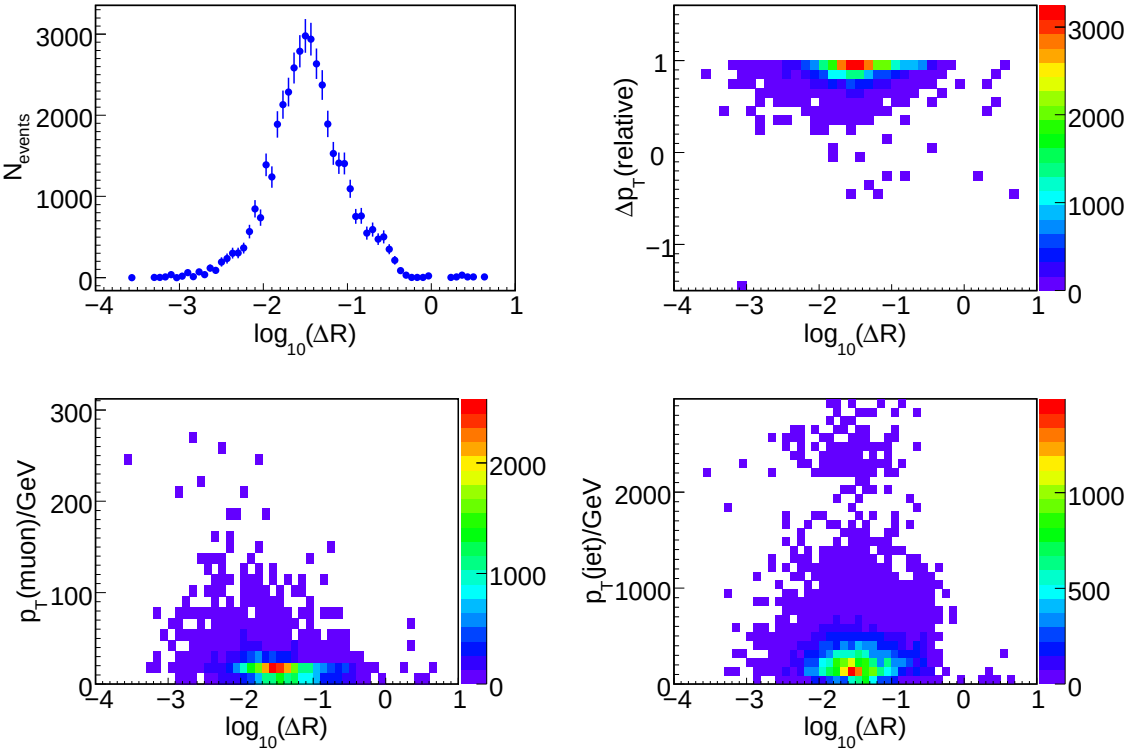


matched by jets have a distance between 0 and 0.4 to the jet axis. The second plot shows an increase of $\Delta p_T(relative)$ from 0 for very small distances to 1 for large distances between Reco electrons and Truth jets. The two further plots can help with the interpretation of this behaviour: for very small $\log_{10}(\Delta R)$ values, the complete jet is misidentified as an electron, leading to a good p_T matching; for increasing distances, only a part of the jet is misidentified as an electron, causing a larger and larger discrepancy between the jet p_T and the electron p_T. In other words, the larger the $\log_{10}(\Delta R)$ value, the smaller is the p_T of the electron matched by a jet on average. To include low-p_T Reco electrons assigned to a Truth jet, a value of $\log_{10}(d_{elec,jet}) = -0.4$ is chosen ($d_{elec,jet} \approx 0.4$). This value is twice as large as the one used for the electron performance studies in [12].

Figure 4.4 shows the same plots as in Figure 4.3, but for muon-jet matching. The maximum of the $\log_{10}(\Delta R)$ distribution is at $\log_{10}(\Delta R) \approx -1.5$, slightly shifted to lower values compared to the corresponding distribution for electron-jet matching, and the shoulder is much less distinct. The plot for $\Delta p_T(relative)$ looks completely different than in the electron-jet case: even for very small distances between Reco muons and Truth jets, $\Delta p_T(relative)$ is dominated by the p_T of the jet. In fact, the second row of plots in Figure 4.4 illustrates that the p_T of the matching jets exceeds the p_T of muons by one order of magnitude while for electron-jet matching, the p_T of the electron and the p_T of the assigned jet are of the same order, see Figure 4.3. This difference between jet-faked muons and jet-faked electrons can be explained in the following way: to fake a muon, the jet has to reach the muon system after being almost completely absorbed in the calorimeter (the idea of punch through provided). Therefore, these jets have very high p_T values but for the fake muons only a tiny p_T fraction is left which is measured in the muon spectrometer using the deflection in the toroid magnets. By contrast, both electrons and jets are detected in the calorimeter system, although electrons

and jets are reconstructed by different algorithms. For muon-jet matching, the same value as for electron-jet matching is used: $\log_{10}(d_{muon,jet}) = -0.4$. As in [12] no matching between Reco muons and Truth jets is performed, no reference value can be quoted here.

Figure 4.4: The $\log_{10}(\Delta R)$ distribution (upper left), $\log_{10}(\Delta R)$ against $\Delta p_T(relative)$ (upper right), $\log_{10}(\Delta R)$ against the p_T of the Truth jet (lower right) and $\log_{10}(\Delta R)$ against the p_T of the Reco muon (lower left) for the muon-jet matching, normalized to $L = 1\ \text{fb}^{-1}$.



4.1.3 Behaviour of the matching parameters under different lepton definitions

In the last section, the determination of the matching parameters, $d_{Reco\ object,Truth\ object}$, was performed using a lepton definition that maximized the statistics, i.e. non-isolated leptons fulfilling $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$. As in the Sections 4.2 and 4.3 different isolation requirements and p_T cuts will be studied, it is important to know if the same $d_{Reco\ object,Truth\ object}$ can be used for different lepton definitions.
Therefore, the $\log_{10}(\Delta R)$ distributions were recalculated for lepton-lepton and lepton-jet matching adding one of the following requirements to the initial lepton definition:

- $p_T > 20\,\text{GeV}$
- $etcone20 < 10\,\text{GeV}$
- $etcone20 < 10\,\text{GeV}$ and $\Delta R(l, j) > 0.4$

Figure 4.5 (4.6) shows the $\log_{10}(\Delta R)$ distributions, normalized to 1, for electron-electron (muon-muon) and electron-jet (muon-jet) matching for the different lepton definitions.
For lepton-lepton matching, the distributions for different lepton definitions coincide quite well within their statistical uncertainties. Hence, the values for $d_{Reco\ object,Truth\ object}$ determined in Section 4.1.2 can also be used for different lepton criteria. However, slight systematic differences can be seen between the distributions for the lepton-jet matching: for leptons

with $p_T > 20$ GeV, the $\log_{10}(\Delta R)$ distribution is slightly shifted to lower values, whereas for isolated leptons, the plots show the opposite tendency. The first aspect can be explained by Figures 4.3 and 4.4: the leptons reconstructed from a Truth jet have a higher p_T when they are reconstructed closer to the jet axis because a bigger fraction of the jet's energy is then taken into account by the electron reconstruction algorithm. By contrast, the isolation requirement $etcone20 < 10$ GeV forces the leptons to be located at the periphery of the assigned jet and so shifts the $\log_{10}(\Delta R)$ distribution to higher values. Applying additionally $\Delta R(l, j) > 0.4$ to the $etcone20 < 10$ GeV cut does not cause any significant change in the distribution. It could seem surprising that the requirement $\Delta R > 0.4$ does not affect the distribution at small $\log_{10}(\Delta R)$ values. However, one has to take into account that the Reco jet might also be some distance away from the corresponding Truth jet (a tight matching value for $d_{jet,jet}$ is $\log_{10}(\Delta R) \approx -1.2$), which can lead to a $\Delta R(l, j) > 0.4$. Furthermore, only Reco jets with $p_T > 20$ GeV are considered, leaving many Truth jets unreconstructed. Finally, the overlap removal in the range $0 < \Delta R < 0.2$ ($\infty < \log_{10}(\Delta R) < -0.7$) between Reco electrons and Reco jets further explain the little effect of the $\Delta R(l, j)$ requirement since it could only apply for $0.2 \leqslant \Delta R < 0.4$.

Figure 4.5: The $\log_{10}(\Delta R)$ distribution for the electron-electron matching (left) and the $\log_{10}(\Delta R)$ distribution for the electron-jet matching (right), both normalized to 1, are shown for four different definitions of Reco electrons.

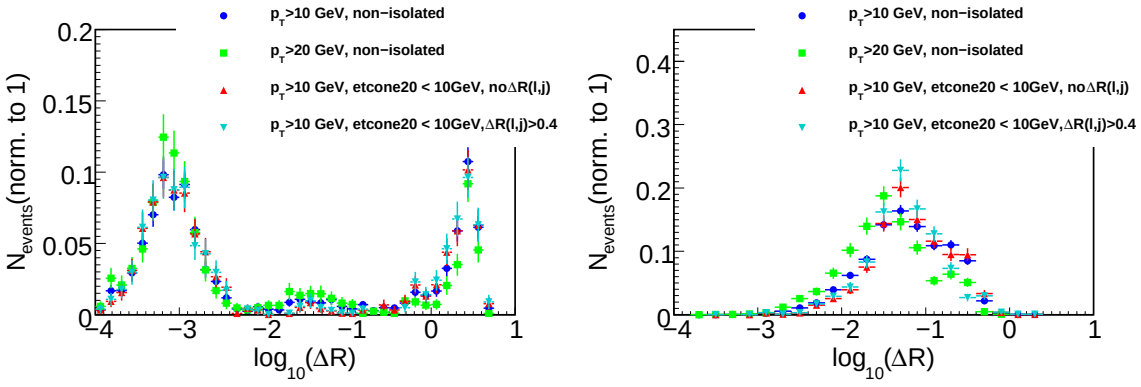
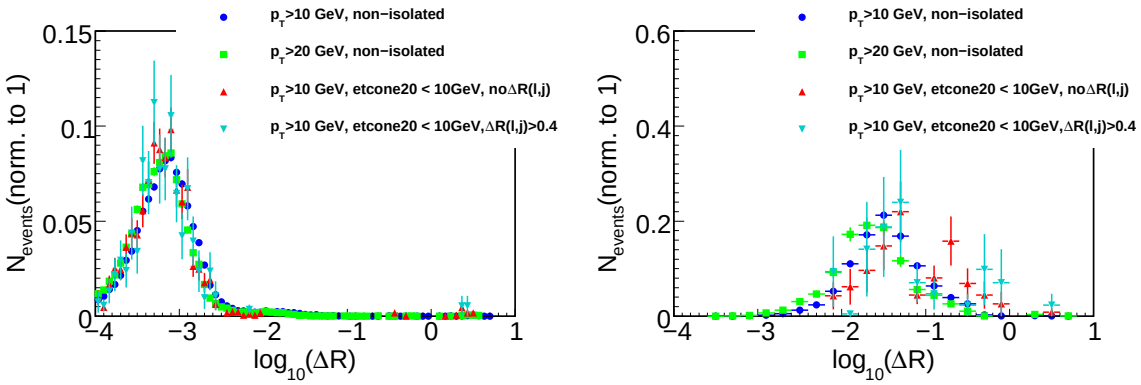


Figure 4.6: The $\log_{10}(\Delta R)$ distribution for the muon-muon matching (left) and the $\log_{10}(\Delta R)$ distribution for the muon-jet matching (right), both normalized to 1, are shown for four different definitions of Reco muons.



4.2 Detailed investigation of nonprompt leptons

The two main sorts of reconstructed leptons in QCD samples are nonprompt leptons and leptons faked by jets. The following section focusses on characterizing the nonprompt leptons found by the matching procedure outlined in the previous section.

4.2.1 Parents of the nonprompt leptons

As we have seen in Section 4.1, Truth leptons can be found in the QCD samples. Since those leptons were not created during the initial strong interaction, they must occur afterwards, during the decay of unstable particles. Therefore, these leptons are called nonprompt leptons. Figure 4.7 and Figure 4.8 depict, for different lepton definitions, the nature of the parent particles decaying into an electron or a muon, respectively. Three columns are shown in these plots, which are normalized to 1. The first (second) one, labeled "charm" ("bottom"), represents all mesons and baryons containing one charm (bottom) quark. The last column labeled "other" stands for all sources of leptons, as τ leptons and light mesons. The τ leptons, leptons themselves, derive mainly from bottom meson or baryon decays except for a small fraction coming from the charmed meson $D_s^{\pm}$.
Figure 4.7 shows that for all considered electron definitions, around 75% of all Truth electrons matching Reco electrons stem from bottom quark decays. The rest comes mainly from charm decays. For muons, the situation is different. According to Figure 4.8, the fraction between Truth muons coming from charm decays and bottom decays is well balanced for non-isolated muons. Applying $etcone20 < 10$ GeV leads to a bottom to charm ratio of ≈ 3. This ratio can be further increased to more than 5 by requiring additionally $\Delta R(l, j) > 0.25$. However, tightening $\Delta R(l, j) > 0.25$ to $\Delta R(l, j) > 0.4$ does not cause any further significant change.
How can this alteration in the ratio bottom/charm be explained? By imposing $etcone20 < 10$ GeV, one selects the leptons which are at the border of a jet, where the energy density is not as high as in the jet center. These leptons must thus have a large angle with respect to the jet axis, i.e. a high $p_{T\ rel}$ value, where $p_{T\ rel}$ is the momentum component which is orthogonal to the jet axis. In the majority of cases, a b quark decays into a c quark, whereas a c quark mostly decays into a d quark. As the mass difference between b and c quarks is larger than the one between c and d quarks, the mass that can be transformed into momentum, for instance $p_{T\ rel}$, during the decay is larger for bottom hadrons than for charmed hadrons [25], [29]. Hence, the bottom/charm ratio goes up with the $etcone20 < 10$ GeV cut. An explicit minimal distance criterion between Reco muons and Reco jets like $\Delta R(l, j) > 0.25$ or $\Delta R(l, j) > 0.4$ further increases the bottom/charm ratio for the same reason.
Comparing the electron and muon parents' distributions of Figures 4.7 and 4.8 for non-isolated leptons and CSC-isolated leptons (see Section 2.2), respectively, one can see that before any isolation cut, the distributions differ quite a lot, whereas for isolated leptons the distributions coincide within their statistical uncertainty. Non-isolated electrons are already much better isolated than non-isolated muons, as shown in Figure 4.9. This figure depicts the $etcone20$ distributions, normalized to $L = 1\ \text{fb}^{-1}$, for reconstructed non-isolated nonprompt electrons and muons. Two differences are clearly visible. First, the overall number of nonprompt muons is much higher than the one of electrons. Second, the muon distribution is shifted to higher $etcone20$ energies explaining why the isolation cuts do not shift the electrons' parent distribution as they do for muons. Electrons, as well as jets, are reconstructed using the calorimeter system. Hence, the electron reconstruction algorithm searches for reasonable isolated electron candidates to limit the number of jets being misidentified as electrons. Due to the muon spectrometer, such a tight initial isolation is not necessary for the muon identification.

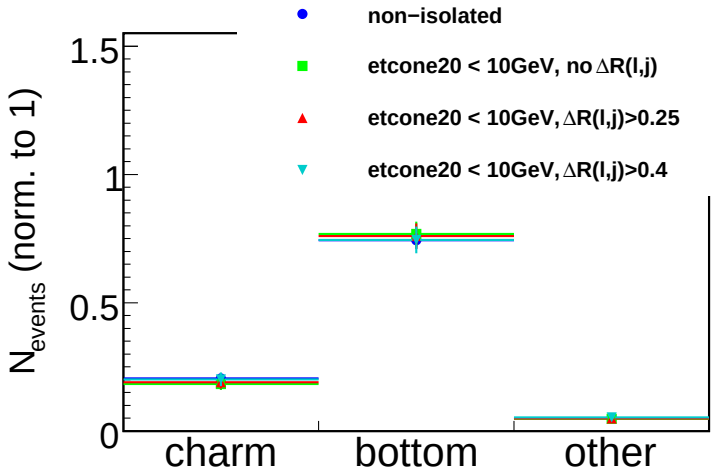


Figure 4.7: The parents of nonprompt electrons, normalized to 1, shown for four definitions of Reco electrons.

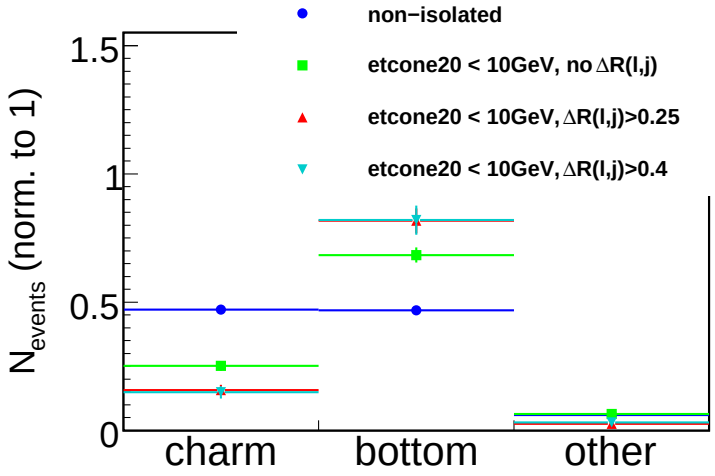


Figure 4.8: The parents of nonprompt muons, normalized to 1, shown for four definitions of Reco muons.

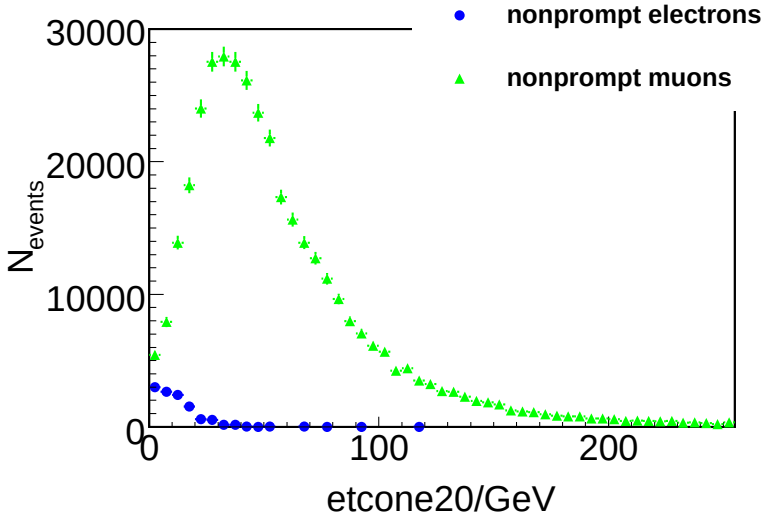


Figure 4.9: *etcone*20 distribution, normalized to $L = 1\,\mathrm{fb}^{-1}$, shown for non-isolated leptons.

4.2.2 Behaviour of the integrated fake rates under different lepton isolations

This section will present the integrated fake rate of nonprompt leptons, where the integrated fake rate is defined as:

$$F_{Reco\ object,Truth\ object} = \frac{N(\text{Reco object matched by Truth object})}{N(\text{Reco jet})} \tag{4.1}$$

$F_{Reco\ object,Truth\ object}$ is calculated by applying at first the matching procedure (see Section 4.1.1) to all events of the five samples J4 - J8. For each sample, the number of matched Reco objects is summed up and merged to get an integrated luminosity of $L = 1$ fb^{-1}. Adding these values for all samples gives N(Reco object matched by Truth object). In parallel, the number of all reconstructed jets, as defined in Section 2.2, is calculated to get N(Reco jet). The Truth leptons have to fulfill $p_T > 5$ GeV and $|\eta| < 3.0$, while the Reco leptons have to fulfill $p_T > 10$ GeV and $|\eta| < 2.5$ and are studied for different isolation criteria.

Table 4.1: $F_{elec,elec}$ for selected definitions of Reco electrons.

$F_{elec,elec}/10^{-4}$	no $\Delta R(l,j)$ requirement	$\Delta R(l,j) > 0.4$
no *etcone*20 cut	15.5 ± 0.6	-
*etcone*20 $<$ 10 GeV	7.9 ± 0.4	6.9 ± 0.4
*etcone*20 $<$ 5 GeV	4.2 ± 0.3	3.7 ± 0.3

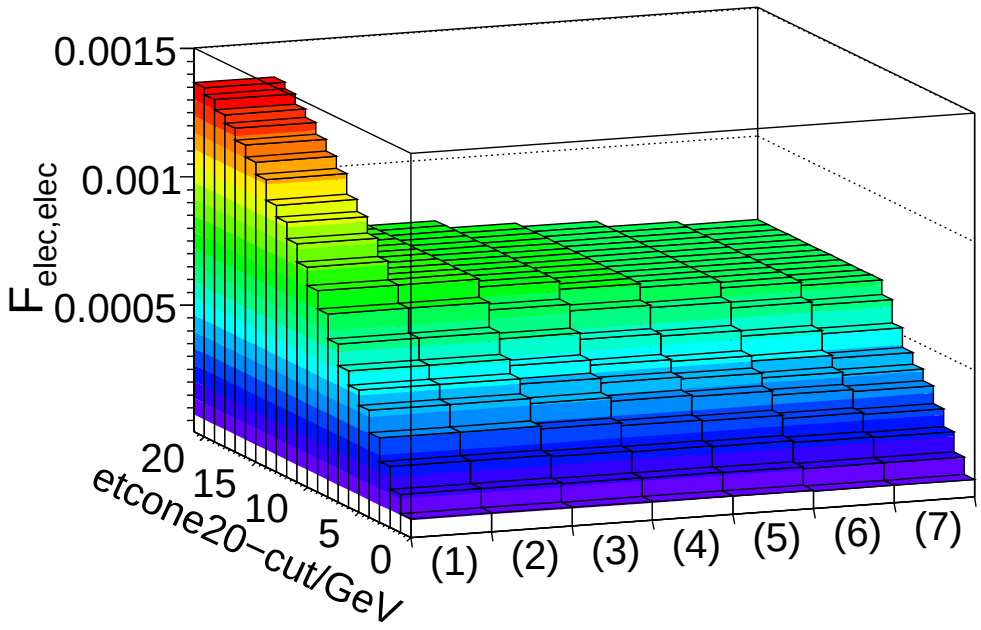
Table 4.2: $F_{muon,muon}$ for selected definitions of Reco muons.

$F_{muon,muon}/10^{-4}$	no $\Delta R(l,j)$ requirement	$\Delta R(l,j) > 0.4$
no *etcone*20 cut	522.3 ± 3.6	-
*etcone*20 $<$ 10 GeV	18.5 ± 0.7	5.8 ± 0.4
*etcone*20 $<$ 5 GeV	7.5 ± 0.4	4.4 ± 0.3

Table 4.1 (4.2) lists values of $F_{elec,elec}$ ($F_{muon,muon}$) for some specific isolation cuts: the three rows are characterized by different *etcone*20 cuts, while the two columns represent different $\Delta R(l,j)$ requirements. Among the different isolations resides also the CSC isolation, the constraint imposed to the element located in the second row, second column. The fake rate in the first row, second column of Table 4.1 (4.2) is missing, as in Section 4 the $\Delta R(l,j)$ requirement is combined with *etcone*20 $<$ 10 GeV. The tables show that the fake rates are on the order of $\mathcal{O}(10^{-4}) - \mathcal{O}(10^{-3})$ for nonprompt electrons as well as for nonprompt muons except for the case of non-isolated muons.
The influence of the *etcone*20 cut and the $\Delta R(l,j)$ requirement on $F_{elec,elec}$ ($F_{muon,muon}$) is shown in detail in Figure 4.10 (4.11). In this figure, $F_{elec,elec}$ ($F_{muon,muon}$) is depicted for different isolations of the Reco leptons: each column represents $F_{elec,elec}$ ($F_{muon,muon}$) for a special combination of the *etcone*20 cut (left axis) and $\Delta R(l,j)$ requirement (right axis). In the first two bins of the $\Delta R(l,j)$ axis, no $\Delta R(l,j)$ requirement is applied. Hence, both bins represent the fake rate as a function of the cut on *etcone*20 only. Asking for *etcone*20 $<$ 10 GeV is the only difference between the first and the second bin: the distribution of the second bin is a constant function for *etcone*20 $\geq$ 10 GeV, as a looser *etcone*20 cut is redundant for this range. For the remaining five bins, the $\Delta R(l,j)$ requirement is tightened from $\Delta R(l,j) > 0.25$ to $\Delta R(l,j) > 0.45$. The plots illustrate that $F_{elec,elec}$ and $F_{muon,muon}$ can be reduced very efficiently by a proper cut on *etcone*20. By contrast, adding a $\Delta R(l,j)$ requirement to *etcone*20 $<$ 10 GeV does not cause a further remarkable reduction, except for $F_{muon,muon}$, for which applying the $\Delta R(l,j)$ requirement or not changes the fake rate drastically. It is surprising that the step between the second and third bin on the $\Delta R(l,j)$ axis in Figure 4.11 is significantly larger than

for the following bins. The cut $etcone20 < 10$ GeV means that the deposited energy around the muon in a cone of radius $R = 0.2$ has to be less than 10 GeV, while the Reco jets have $p_T > 20$ GeV in a cone of radius $R = 0.4$. Hence, it seems to be very unlikely that a Reco jet is located closer than $\Delta R = 0.2$ to an isolated Reco muon, but this is what the step in Figure 4.11 suggests. In fact, further investigations show that the fraction of Reco jets lying in a radius of $\Delta R = 0.2$ around a Reco muon with $etcone20 < 10$ GeV is not negligible. Furthermore, in some cases these jets even have a p_T of more than 50 GeV. This phenomenon is induced by the imprecise energy measurement of either the jets or the $etcone20$ variable in certain regions of the detector. It will be addressed in more details in Section 4.2.3.

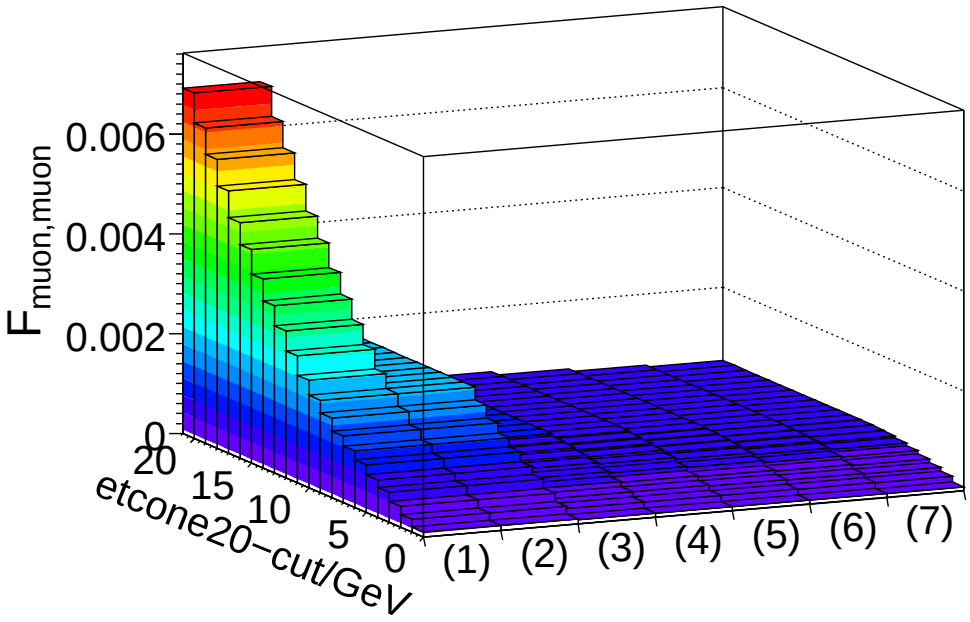
Figure 4.10: $F_{elec,elec}$ as a function of the $etcone20$ cut and the $\Delta R(l, j)$ requirement. The labels on the right axis stand for different isolation stages of Reco electrons (for the $\Delta R(l, j)$ requirement the cut $etcone20 < 10$ GeV is provided): (1) non-isolated, (2) $etcone20 < 10$ GeV, (3) $\Delta R(l, j) > 0.25$, (4) $\Delta R(l, j) > 0.30$, (5) $\Delta R(l, j) > 0.35$, (6) $\Delta R(l, j) > 0.40$, (7) $\Delta R(l, j) > 0.45$.



A projection of the Figures 4.10 and 4.11 on the plane formed by $F_{elec,elec}$ ($F_{muon,muon}$) and the $\Delta R(l, j)$ axis is shown in Figure 4.12. This figure compares the development of $F_{elec,elec}$ and $F_{muon,muon}$ for a tightening $\Delta R(l, j)$ requirement after the cut $etcone20 < 10$ GeV. For nonprompt muons it illustrates again the unexpected drop of $F_{muon,muon}$ between the first and second bin when asking $\Delta R(l, j) > 0.25$. In the second bin, the values of $F_{elec,elec}$ and $F_{muon,muon}$ are comparable. The following bins show that, when strengthening the $\Delta R(l, j)$ requirement, $F_{muon,muon}$ decreases faster than $F_{elec,elec}$. Hence, for the CSC isolation, $F_{muon,muon}$ lies significantly below $F_{elec,elec}$.

As the less energetic regions of a jet are located far from the jet axis, it is interesting to know to what extent the cut $etcone20 < \delta$ ($\delta \leq 10$ GeV) and the $\Delta R(l, j)$ requirement (provided $etcone20 < 10$ GeV) are redundant. Figure 4.13 (4.14) depicts the distributions of $F_{elec,elec}$ ($F_{muon,muon}$) as a function of the $etcone20$ cut for several $\Delta R(l, j)$ requirements. All distributions are normalized to 1. For electrons the distributions coincide very well, meaning that applying a $\Delta R(l, j)$ requirement in addition to $etcone20 < \delta$ ($\delta \leq 10$ GeV) has the same small effect on $F_{elec,elec}$ regardless of the chosen δ value. In other words, no

Figure 4.11: $F_{muon,muon}$ as a function of the $etcone20$ cut and the $\Delta R(l,j)$ requirement. The labels on the right axis stand for different isolation stages of Reco muons (for the $\Delta R(l,j)$ requirement the cut $etcone20 < 10$ GeV is provided): (1) non-isolated, (2) $etcone20 < 10$ GeV, (3) $\Delta R(l,j) > 0.25$, (4) $\Delta R(l,j) > 0.30$, (5) $\Delta R(l,j) > 0.35$, (6) $\Delta R(l,j) > 0.40$, (7) $\Delta R(l,j) > 0.45$.



correlation between these cuts is noticeable for electrons. For muons, for which the $\Delta R(l,j)$ requirement is more effective than for electrons, the situation is different: the stronger the $\Delta R(l,j)$ requirement, the flatter in $etcone20$ the $F_{muon,muon}$ distribution. For harder and harder cuts on $etcone20$, an additional $\Delta R(l,j)$ requirement has less and less impact. However, this tendency is only statistically significant when comparing the distribution with no $\Delta R(l,j)$ requirement and the one using $\Delta R(l,j) > 0.25$.

For the SUSY search in the 1-lepton channel, only events containing exactly 1 Reco lepton are analyzed. As QCD events containing more than one reconstructed lepton are rare, the fake rates decrease only slightly when considering only 1-lepton events: the corrected value of $F_{elec,elec}$ ($F_{muon,muon}$) amounts to $\approx 98\%$ ($\approx 95\%$) of the value given in Table 4.1 (4.2) for CSC-isolated leptons.

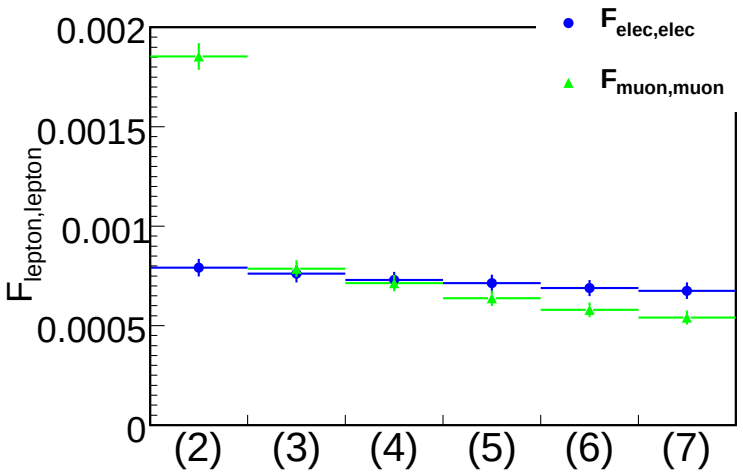


Figure 4.12: $F_{elec,elec}$ and $F_{muon,muon}$ as a function of the $\Delta R(l, j)$ requirement. The label definition of the x-coordinate is the same as in the Figures 4.10 and 4.11.

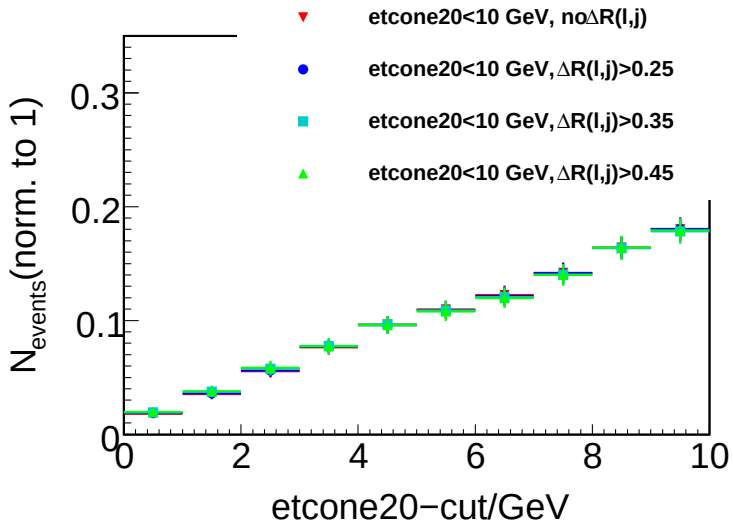


Figure 4.13: $F_{elec,elec}$, normalized to 1, as function of the *etcone*20 cut for different $\Delta R(l, j)$ requirements.

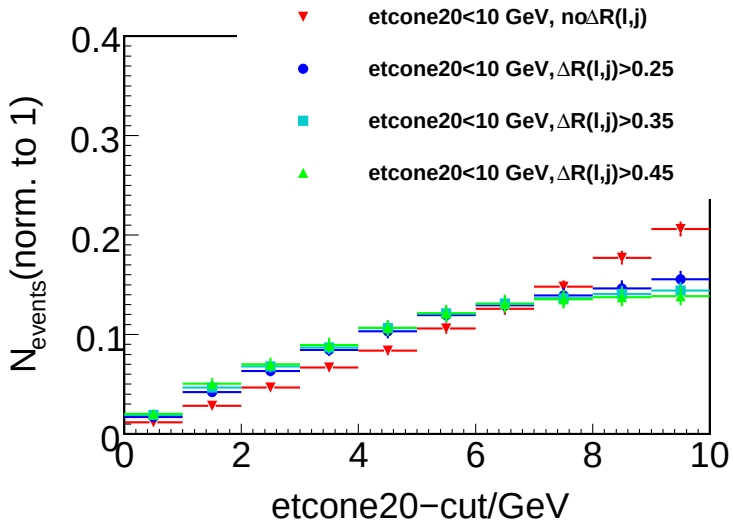


Figure 4.14: $F_{muon,muon}$, normalized to 1, as function of the *etcone*20 cut for different $\Delta R(l, j)$ requirements.

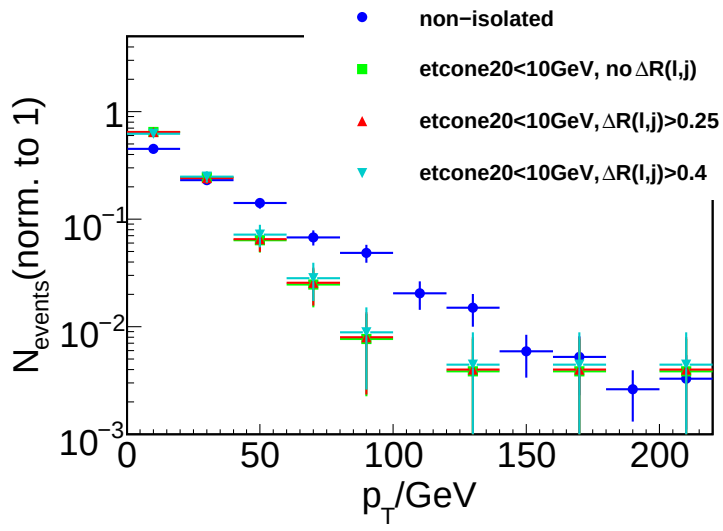


Figure 4.15: $F_{elec,elec}$, normalized to 1 and on a logarithmic scale, as a function of p_T for different definitions of isolated Reco electrons.

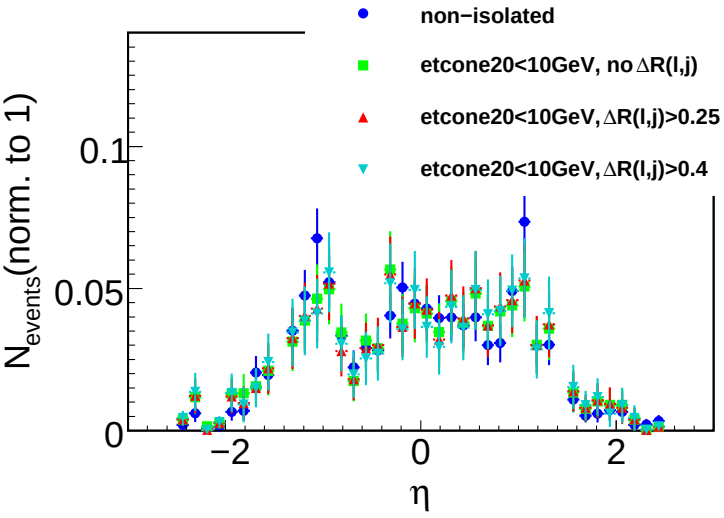


Figure 4.16: $F_{elec,elec}$, normalized to 1, as a function of η for different definitions of isolated Reco electrons. The asymmetry of the distribution at $|\eta| \approx 0.7$ does not seem to be statistically significant.

4.2.3 Fake rates as a function of the kinematic variables p_T, η and ϕ for different lepton isolations

After the study of integrated fake rates for nonprompt leptons in Section 4.2.2, the present section deals with $F_{elec,elec}$ and $F_{muon,muon}$ as a function of p_T, η and ϕ. Therefore, Equation (4.1) has to be modified so that the numerator of $F_{Reco\ object,Truth\ object}$, N(Reco object matched by Truth object), becomes a function of the corresponding variables. Thereby, the variable value of the Reco object is taken. However, the denominator, N(Reco jet), is further on considered as constant. As in Section 4.2.2, $p_T > 10$ GeV is required for the Reco leptons to enhance statistics. Several lepton isolations are used in parallel:

- no $etcone20$ cut, no $\Delta R(l, j)$ requirement
- $etcone20 < 10$ GeV, no $\Delta R(l, j)$ requirement
- $etcone20 < 10$ GeV, $\Delta R(l, j) > 0.25$
- $etcone20 < 10$ GeV, $\Delta R(l, j) > 0.4$ (CSC isolation)

The Figures 4.15, 4.16 and 4.17 (4.18, 4.19 and 4.20) depict the p_T, η and ϕ distributions of $F_{elec,elec}$ ($F_{muon,muon}$). Since the integrated fake rates are known from Section 4.2.2 and

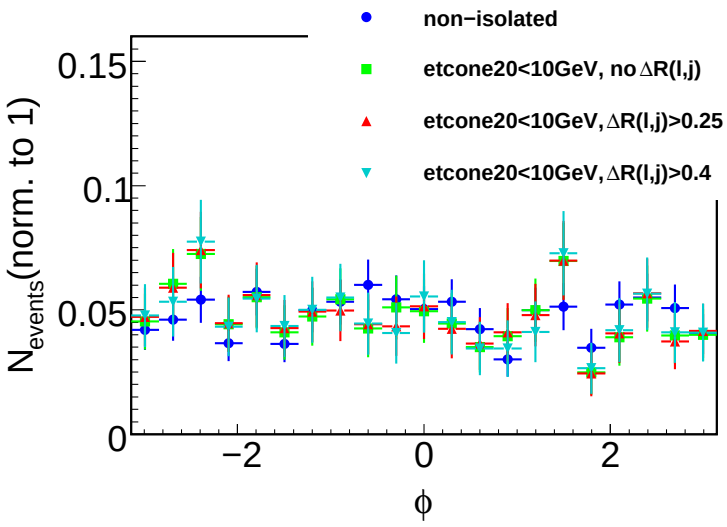


Figure 4.17: $F_{elec,elec}$, normalized to 1, as a function of ϕ for different definitions of isolated Reco electrons.

the distributions for different lepton isolations are to be compared, the distributions are normalized to 1 and overlaid for the different lepton definitions. All these distributions are strongly influenced by the initial p_T, η and ϕ distributions of Reco electrons (Reco muons), i.e. the distributions of all Reco electrons (Reco muons).

Table 4.3: Remaining fraction of $F_{elec,elec}$ after tightening the p_T cut from $p_T > 10$ GeV to $p_T > 20$ GeV or $p_T > 30$ GeV for two different definitions of Reco electrons.

$F_{elec,elec}/(F_{elec,elec}$ for $p_T > 10$ GeV)	$p_T > 20$ GeV	$p_T > 30$ GeV
no $etcone20$ cut, no $\Delta R(l,j)$ requirement	55%	40%
CSC isolation	38%	20%

Table 4.4: Remaining fraction of $F_{muon,muon}$ after tightening the p_T cut from $p_T > 10$ GeV to $p_T > 20$ GeV or $p_T > 30$ GeV for two different definitions of Reco muons.

$F_{muon,muon}/(F_{muon,muon}$ for $p_T > 10$ GeV)	$p_T > 20$ GeV	$p_T > 30$ GeV
no $etcone20$ cut, no $\Delta R(l,j)$ requirement	60%	38%
CSC isolation	15%	1%

The p_T distributions of nonprompt electrons (Figure 4.15) as well as of nonprompt muons (Figure 4.18) feature an exponential decrease when going to higher values, a behaviour coming from the initial p_T distributions of Reco electrons and Reco muons. Since isolation selects nonprompt leptons which have a large angle with respect to the axis of the corresponding jet, the isolated nonprompt leptons usually have a lower contribution to the jet p_T. Therefore, imposing an isolation criterion to the reconstructed leptons leads to a shift of the p_T distributions to lower values. Hence, combining an isolation cut with a hard cut on the lepton p_T reduces the fake rates $F_{elec,elec}$ and $F_{muon,muon}$ very efficiently. The percentage of the fake rates remaining after asking for Reco leptons fulfilling $p_T > 20$ GeV and $p_T > 30$ GeV, respectively, are shown in Table 4.3 (4.4) for non-isolated electrons (muons) as well as for CSC-isolated ones. $F_{muon,muon}$ can especially be suppressed by applying a CSC isolation and tightening the p_T cut.
The η distributions of nonprompt electrons (Figure 4.16) reflect the different transition regions of the calorimeter system as described in Section 1.1.2. In the crack region $1.37 < |\eta| < 1.52$, the transition between the EM barrel and the EM endcap calorimeter,

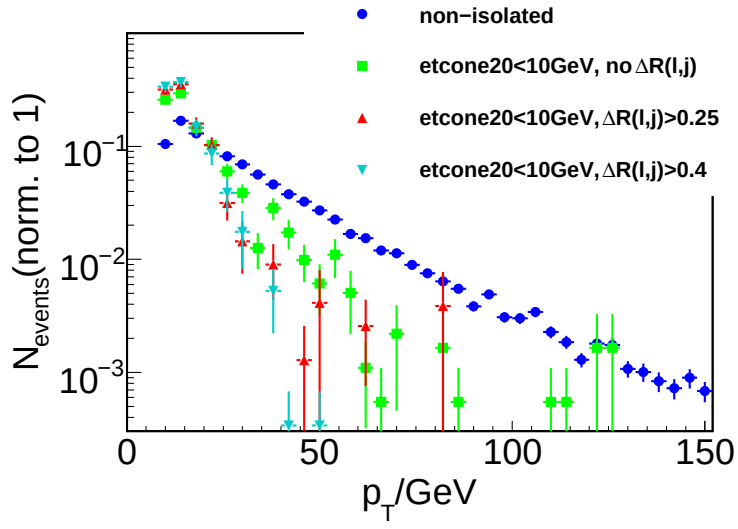


Figure 4.18: $F_{muon,muon}$, normalized to 1 and on a logarithmic scale, as a function of p_T for different definitions of isolated Reco muons.

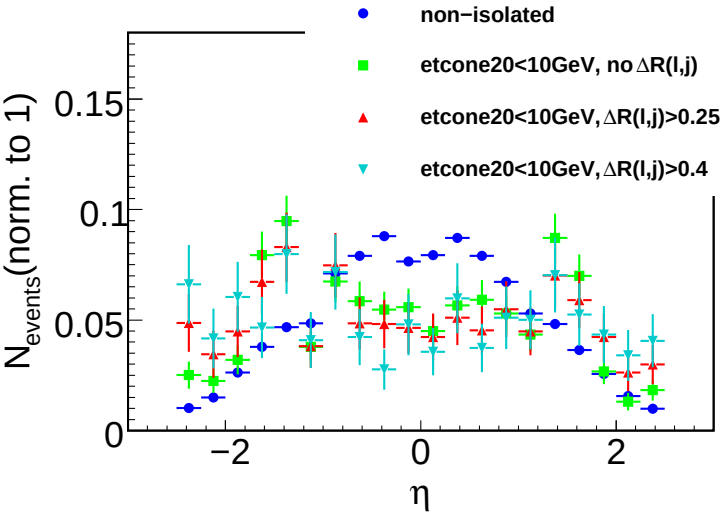


Figure 4.19: $F_{muon,muon}$, normalized to 1, as a function of η for different definitions of isolated Reco muons.

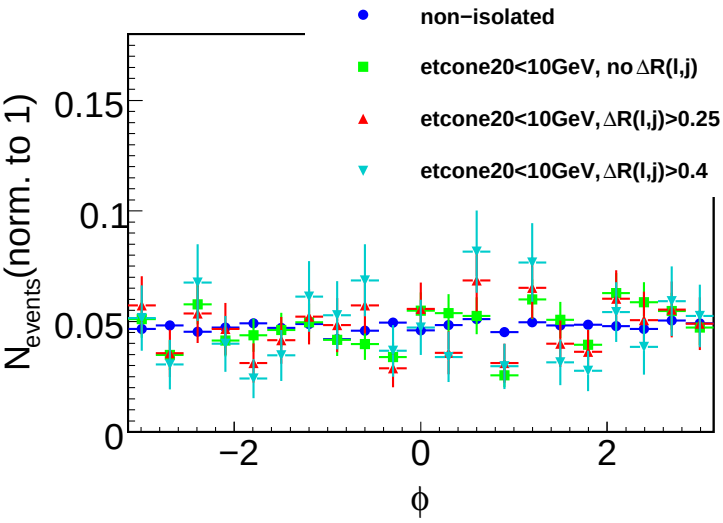


Figure 4.20: $F_{muon,muon}$, normalized to 1, as a function of ϕ for different definitions of isolated Reco muons.

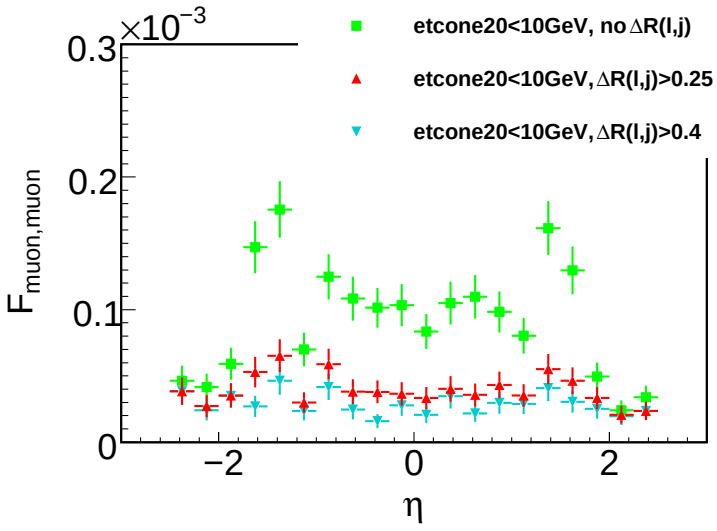


Figure 4.21: $F_{muon,muon}$ as a function of η for different definitions of isolated Reco muons.

the bin is empty because of the veto applied on this region (see Section 2.2). Furthermore, $F_{elec,elec}$ is increased at $|\eta| \approx 1.1$: in this region, the transition between the hadronic barrel and the hadronic extended barrel takes place. As the energy deposited in the first layer of the hadronic calorimeter plays a role in the identification of reconstructed electrons, this transition region affects the reconstruction efficiency and consequently, the η distribution of $F_{elec,elec}$. For the different electron definitions, the η distributions coincide quite well. However, non-isolated electrons are slightly more influenced by the transition region at $|\eta| \approx 1.1$. Imposing a veto on the latter region ($1 \lesssim |\eta| \lesssim 1.2$) reduces the fake rate of nonprompt electrons to $\approx 80\%$ ($\approx 85\%$) for non-isolated (CSC-isolated) electrons.

When comparing the η distributions of nonprompt electrons (Figure 4.16) with the ones of nonprompt muons (4.19) the differences are clearly visible. Non-isolated nonprompt muons are focused at small $|\eta|$ values, although there is a dip around $|\eta| \approx 0$ because of a detector gap for service work in this region. For isolated nonprompt muons, the η distributions change quite a lot: $F_{muon,muon}$ is distributed more uniformly in η. In addition, the nonprompt muons isolated with $etcone20 < 10$ GeV and no $\Delta R(l, j)$ requirement peak in the crack region $1.37 < |\eta| < 1.52$. This effect can be better seen in Figure 4.21. In contrast to Figure 4.19, this figure shows the absolute η distributions of $F_{muon,muon}$ for isolated muons. It illustrates that tightening the isolation by an additional $\Delta R(l, j)$ requirement decreases $F_{muon,muon}$ especially in the crack region $1.37 < |\eta| < 1.52$. Further studies confirm that the crack region is the region for which the probability of finding a jet closer than $\Delta R = 0.2$ to an isolated ($etcone20 < 10$ GeV) Reco muon is the highest. Almost 50% of all Reco muons, isolated by $etcone20 < 10$ GeV, have a Reco jet which is closer than $\Delta R = 0.2$ to the muon. This is understood to be caused by the muon $etcone20$ calculation algorithm which returns a flawed value in this region. This also causes the unexpected drop in $F_{muon,muon}$ which occurs when asking $\Delta R(l, j) > 0.25$ and which has been discussed in Section 4.2.2 (see Figure 4.11). Figure 4.17 (4.20) illustrates $F_{elec,elec}$ ($F_{muon,muon}$) as function of ϕ. For both nonprompt electrons and muons, the ϕ distributions coincide for the different lepton isolations and have, as expected, a nearly uniform shape.
Further investigations demonstrate that for CSC-isolated leptons, cutting on $p_T > 10$ GeV or $p_T > 20$ GeV does not cause statistically significant differences in the η and ϕ distributions of $F_{elec,elec}$ and $F_{muon,muon}$.

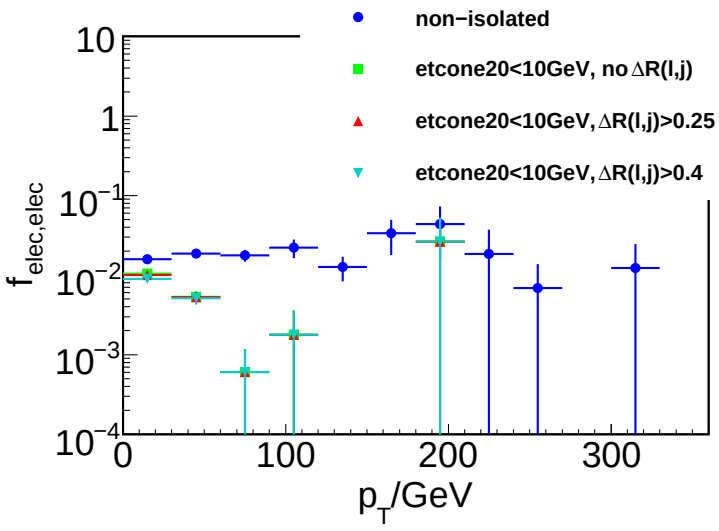


Figure 4.22: $f_{elec,elec}$, on a logarithmic scale, as a function of p_T for different definitions of Reco electrons.

4.2.4 Characteristics of the fake rates as a function of generator particle kinematics

The aim of the previous section was to study the p_T, η and ϕ distributions of $F_{elec,elec}$ and $F_{muon,muon}$ for high-energetic QCD events in the ATLAS detector. The Truth level was only used for the matching procedure to label the reconstructed leptons as nonprompt or jet-faked leptons. By contrast, the present section focusses on the Truth level to find the probability, given a Truth lepton, to match it to a Reco one. For this purpose, one defines the fraction between the number of matched Reco-Truth pairs and the number of Truth objects, both normalized to $L = 1\ \text{fb}^{-1}$:

$$f_{Reco\ object,Truth\ object} = \frac{N(\text{Reco object matched by Truth object})}{N(\text{Truth object})} \tag{4.2}$$

For the Reco leptons, different isolations are imposed with $p_T > 10\,\text{GeV}$. Only Truth electrons (Truth muons) that have $p_T > 6$ GeV ($p_T > 9$ GeV) and $|\eta| < 2.5$ ($|\eta| < 2.5$) are considered in this section as Truth electrons (Truth muons) outside these cuts do not lead to a successful matching. These kinematic cuts at generator level lead to a reasonable calculation of the denominator N(Truth object). $f_{Reco\ object,Truth\ object}$ as a function of p_T, η or ϕ is calculated in the following way: both the numerator and the denominator in Equation (4.2) are considered as functions of the chosen kinematic variable and divided bin by bin. The variable value of the Truth object is used for the numerator. The denominator N(Truth object) represents the kinematic distribution of all Truth objects. This initial distribution can be considered as a probability distribution of the merged QCD samples: it describes the probability that a Truth object has a fixed value of p_T, η or ϕ. Hence, by this bin-by-bin division, $f_{Reco\ object,Truth\ object}$ becomes independent of the initial distribution and can be interpreted as the probability that a *single* Truth lepton, having a fixed value of p_T, η or ϕ, matches a Reco lepton. This contrasts with Section 4.2.3, in which only the numerator was considered as a function of the kinematic variable in the calculation of $F_{elec,elec}$ or $F_{muon,muon}$.

Figure 4.22 (4.23) illustrates $f_{elec,elec}$ ($f_{muon,muon}$), on a logarithmic scale, as a function of p_T for different lepton definitions. The bin content represents the probability of matching a Reco lepton for a given Truth lepton whose p_T falls within this bin. Due to the bin-by-bin division, the bins are independent of each other and the sum over all bins does *not* represent the integrated probability for a Truth lepton to match a reconstructed one. For nonprompt electrons, the bin content increases slightly in the range 5 GeV $< p_T <$ 200 GeV and

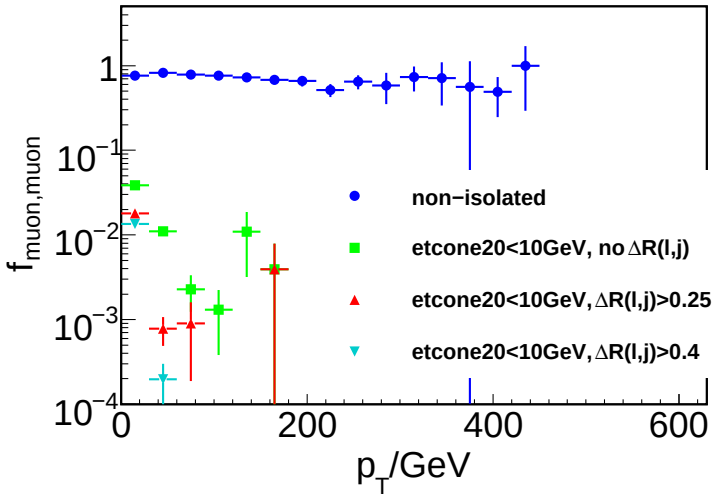


Figure 4.23: $f_{muon,muon}$, on a logarithmic scale, as a function of p_T for different definitions of Reco muons.

decreases afterwards. For isolated electrons, though, $f_{elec,elec}$ peaks at low p_T values and drops very quickly to reach $\mathcal{O}(10^{-3})$ at $p_T \approx 80$ GeV. Non-isolated muons have an almost uniform bin content in the range $5 < p_T < 400$ GeV. Imposing isolation criteria leads to a clear maximum at low p_T values followed by a fast decrease. Compared to nonprompt electrons this behaviour is even more distinctive as significant changes are also noticeable between different isolation steps. For CSC-isolated nonprompt muons, $f_{muon,muon}$ disappears already at $p_T \approx 50$ GeV.

As non-isolated leptons have much higher statistics than isolated ones, it would be convenient to concentrate on this lepton definition and extend the conclusions to isolated leptons. However, the investigation of the variable p_T shows that the behaviour of nonprompt leptons depends a lot on the lepton definition considered. Therefore, for the rest of this section, CSC-isolated leptons are used, as this definition is the recommended one for SUSY analysis. However, in contrast to the 1-lepton analysis, $p_T > 10$ GeV is used for Reco leptons instead of $p_T > 20$ GeV in order to maintain reasonable statistics.
To notice possible correlations between the p_T and η variables of a Truth lepton matching a Reco lepton, $f_{lepton,lepton}$ is now investigated as a function of η for given p_T bins of Truth leptons. Therefore, the procedure is carried out for each p_T bin separately: the η distribution of Truth leptons having a p_T within a fixed bin and matching a Reco lepton is divided by the initial η distribution of this p_T bin. For statistical reasons, only two p_T bins are defined: for nonprompt electrons (nonprompt muons) the first bin covers 6 GeV $< p_T <$ 20 GeV (9 GeV $< p_T <$ 15 GeV) and the second, the range $p_T >$ 20 GeV ($p_T >$ 15 GeV). Figure 4.24 presents the results for nonprompt electrons: in this figure, a given point describes the probability that a Truth electron whose p_T and η values fall within the corresponding bins matches a Reco electron. The analog approach for ϕ leads to Figure 4.25 for nonprompt electrons. For both Figures 4.24 and 4.25 $f_{elec,elec}$ for the low-p_T bin exceeds slightly the values for the high-p_T bin. This means that, independently of η or ϕ, low-p_T electrons have a higher probability of matching a Reco electron than high-p_T ones, CSC isolation provided. Since between the two p_T bins no obvious difference in the shapes of $f_{elec,elec}$ as a function of η or ϕ is noticeable, the correlation between p_T and η or ϕ does not seem to be very strong. For both p_T bins, $f_{elec,elec}$ is quite constant over the ranges $-2.5 < \eta < 2.5$ and $-\pi < \phi < \pi$ respectively. The Figures 4.26 and 4.27 show the analog plots for nonprompt muons. The probability that a Truth muon matches a Reco muon is significantly higher in the low-p_T bin than in the high-p_T bin, a tendency which can also be seen in Figure 4.23. Furthermore,

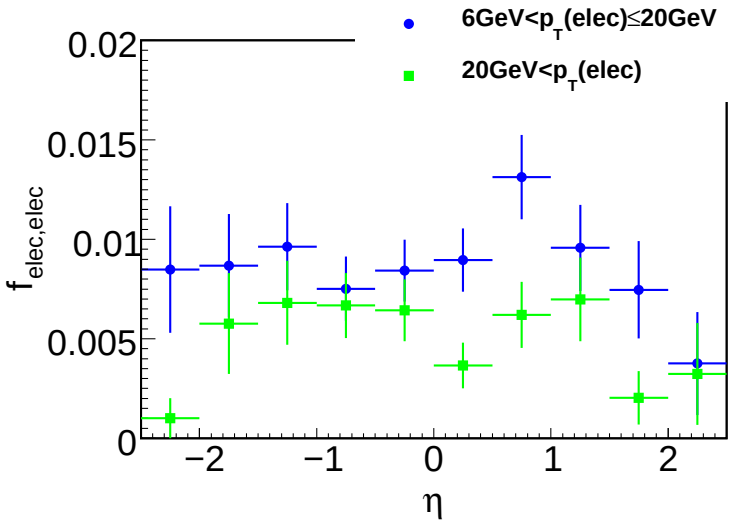


Figure 4.24: $f_{elec,elec}$ as a function of η for two different p_T bins, shown for CSC-isolated Reco electrons.

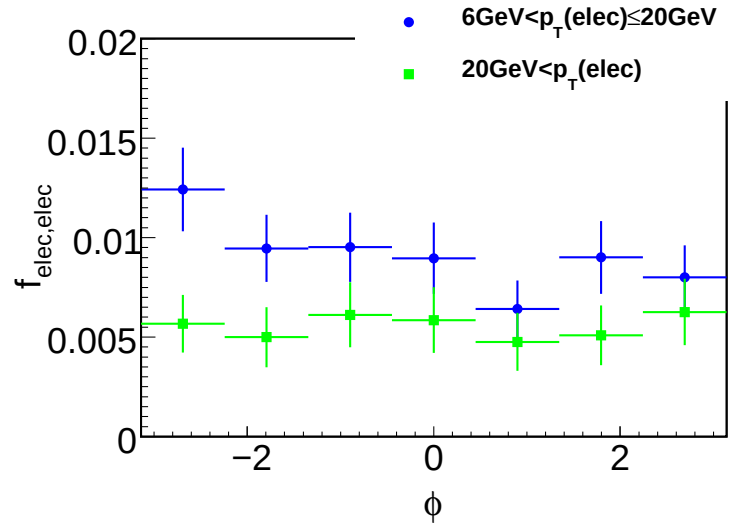


Figure 4.25: $f_{elec,elec}$ as a function of ϕ for two different p_T bins, shown for CSC-isolated Reco electrons.

Figure 4.26 illustrates that $f_{muon,muon}$ has a minimum at low $|\eta|$ values and increases at higher $|\eta|$ values. This behaviour seems to be even more pronounced for Truth muons in the low-p_T bin than in the high-p_T bin. So the variables $|\eta|$ and p_T seem to be significantly correlated: the lower the p_T of the Truth muon, the stronger the probability of matching a Reco muon at high $|\eta|$ values. The minimum at low $|\eta|$ values could be explained by the concentration of jets in this region. The higher the jet density as a function of $|\eta|$, the higher the probability of a Reco muon to be rejected by the $\Delta R(l, j)$ requirement, although the Reco muon and the corresponding Reco jet are not correlated. The η distribution of reconstructed jets is depicted in Figure D.1.
The ϕ distribution of $f_{muon,muon}$ depicted in Figure 4.27 is quite uniform for $-\pi < \phi < \pi$ for both p_T bin: for nonprompt muons, no correlation between p_T and ϕ is observable.

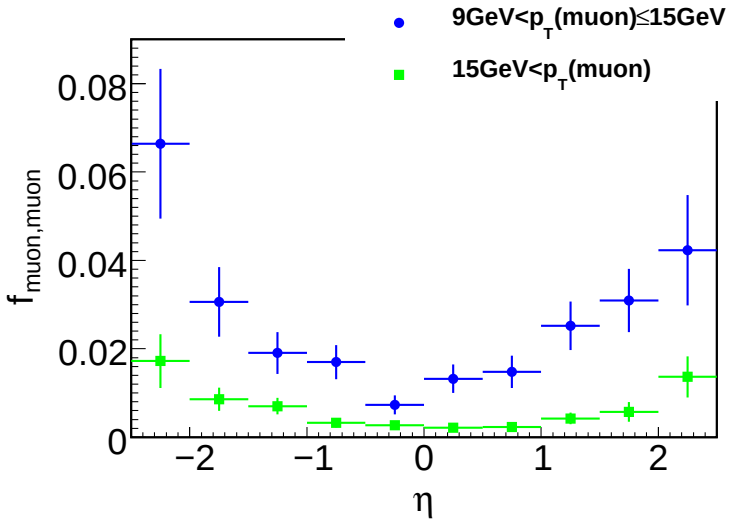


Figure 4.26: $f_{muon,muon}$ as a function of η for two different p_T bins, shown for CSC-isolated Reco muons.

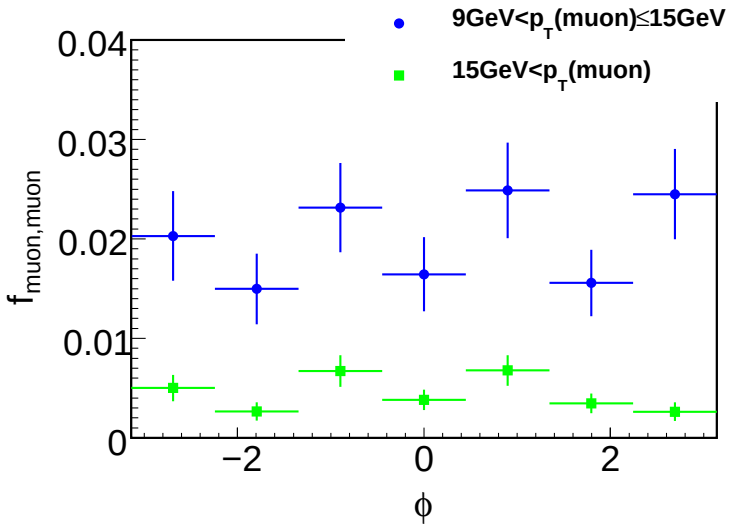


Figure 4.27: $f_{muon,muon}$ as a function of ϕ for two different p_T bins, shown for CSC-isolated Reco muons.

4.3 Detailed investigation of leptons faked by jets

The following section will focus on jet-faked leptons; the Sections 4.3.1 to 4.3.3 will be structured as the nonprompt lepton Sections 4.2.2 to 4.2.4. As the variables describing fakes were already introduced in the last section, their definitions will not be repeated here; the focus will be on the results and the comparison between nonprompt and jet-faked leptons. Section 4.3.4 will deal briefly with the connection between jet-faked leptons and missing transverse energy $\not{E}_T$.

4.3.1 Behaviour of the integrated fake rates under different lepton isolations

When the attempt of matching a Truth lepton to a Reco lepton fails, the matching procedure tries to find an adequate Truth jet for the Reco lepton (see Section 4.1.1). Table 4.5 (4.6) presents some values for the fake rates due to jet-faked leptons, $F_{elec,jet}$ ($F_{muon,jet}$), for Reco electrons (Reco muons) having $p_T > 10$ GeV, $|\eta| < 2.5$ and different isolation criteria. Just as the rates for nonprompt electrons, $F_{elec,jet}$ is on the order of $\mathcal{O}(10^{-4}) - \mathcal{O}(10^{-3})$. For muons, however, $F_{muon,jet}$ is around 10 times less than $F_{muon,muon}$. Applying the CSC isolation is especially effective in decreasing $F_{muon,jet}$.

Table 4.5: $F_{elec,jet}$ for different definitions of Reco electrons.

$F_{elec,jet}/10^{-4}$	no $\Delta R(l,j)$ requirement	$\Delta R(l,j) > 0.4$
no *etcone*20 cut	40.4 ± 0.9	-
*etcone*20 < 10 GeV	17.4 ± 0.6	15.3 ± 0.6
*etcone*20 < 5 GeV	6.7 ± 0.4	5.9 ± 0.3

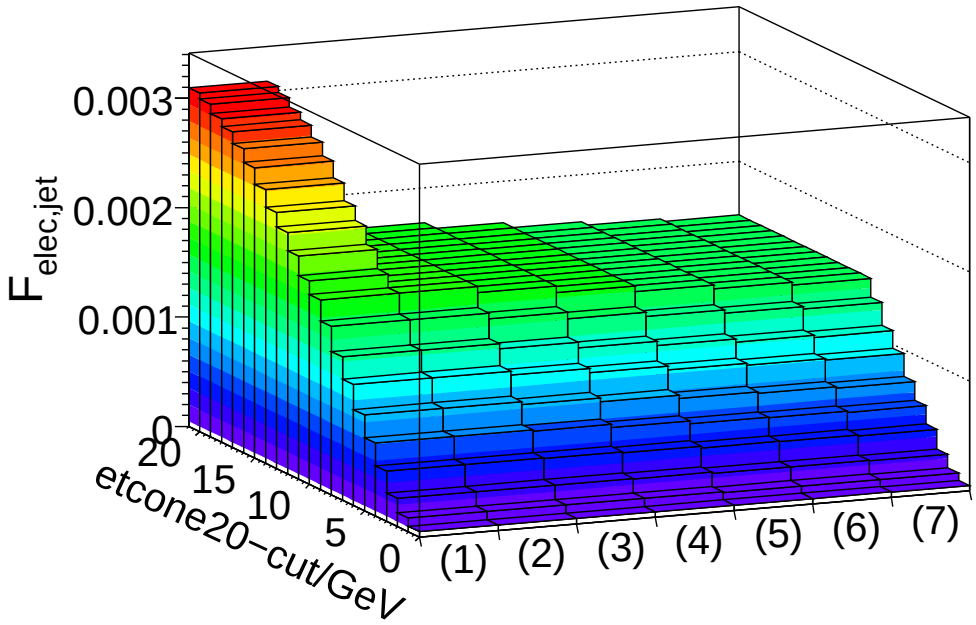
Table 4.6: $F_{muon,jet}$ for different definitions of Reco muons.

$F_{muon,jet}/10^{-4}$	no $\Delta R(l,j)$ requirement	$\Delta R(l,j) > 0.4$
no *etcone*20 cut	56.8 ± 1.0	-
*etcone*20 < 10 GeV	1.1 ± 0.2	0.4 ± 0.1
*etcone*20 < 5 GeV	0.5 ± 0.1	0.3 ± 0.09

Figure 4.28 (4.29) details the influence of the *etcone*20 cut and the $\Delta R(l,j)$ requirement on $F_{elec,jet}$ ($F_{muon,jet}$). Both rates $F_{elec,jet}$ and $F_{muon,jet}$ show a behaviour under different isolations which is similar to the corresponding rates of nonprompt leptons $F_{elec,elec}$ and $F_{muon,muon}$, respectively (see Section 4.2.2). For fake electrons, the $\Delta R(l,j)$ requirement (provided *etcone*20 $<$ 10 GeV) acts very inefficiently compared to a tightening of the *etcone*20 cut. However, for fake muons isolated by *etcone*20 < 10 GeV, the impact of a tighter *etcone*20 cut or applying a $\Delta R(l,j)$ requirement is comparable due to the flawed measurement of *etcone*20 particularly in the crack region. For fake muons, the cut *etcone*20 < 10 GeV has more influence on the fake rate as for fake electrons: for both nonprompt and jet-faked electrons, the maxima of the *etcone*20 distributions, shown in Figure 4.30, are located at much lower *etcone*20 values than the ones of fake muons. This different initial isolation for reconstructed electrons and muons originates from the different reconstruction of these objects (see Section 4.2.1).
A direct comparison of the four types of fakes can be seen in Figure 4.31, showing the nonprompt as well as the jet-faked lepton rates for a tightening $\Delta R(l,j)$ requirement. When *etcone*20 < 10 GeV is the only isolation requirement, $F_{muon,muon}$ exceeds all other fake rates. However, imposing in addition $\Delta R(l,j) > 0.25$ or tighter efficiently reduces $F_{muon,muon}$, leav-

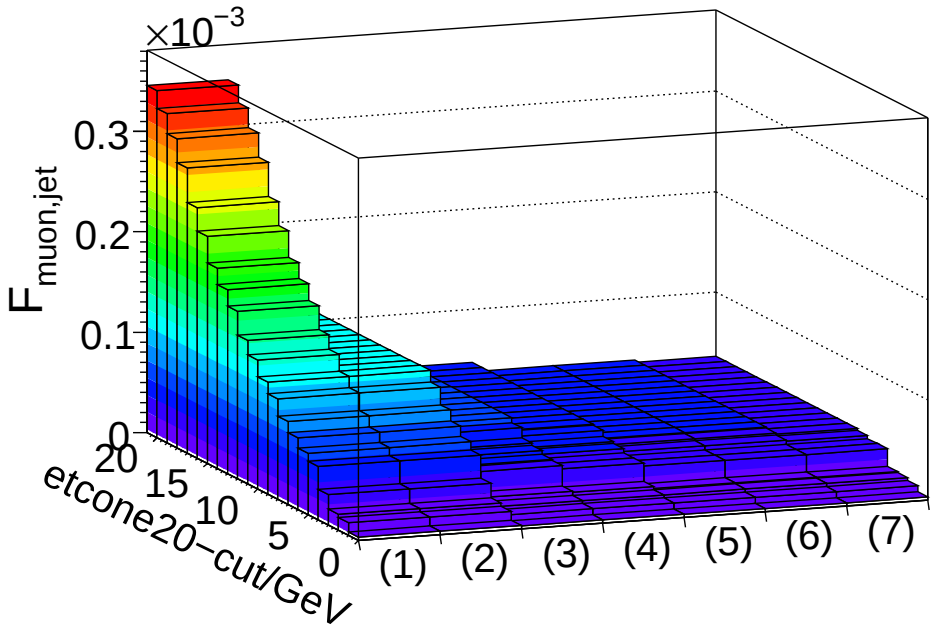
ing $F_{elec,jet}$ as the bulk of the overall fake rate. For CSC-isolated leptons, $F_{elec,jet}$ is more than twice as large as $F_{elec,elec}$ or $F_{muon,muon}$ and almost 40 times as large as $F_{muon,jet}$.

Figure 4.28: $F_{elec,jet}$ as a function of the *etcone*20 cut and the $\Delta R(l,j)$ requirement. The labels on the right axis stand for different isolation stages of Reco electrons (for the $\Delta R(l,j)$ requirement the cut *etcone*20 $<$ 10 GeV is provided): (1) non-isolated, (2) *etcone*20 $<$ 10 GeV, (3) $\Delta R(l,j) > 0.25$, (4) $\Delta R(l,j) > 0.30$, (5) $\Delta R(l,j) > 0.35$, (6) $\Delta R(l,j) > 0.40$, (7) $\Delta R(l,j) > 0.45$.



For the investigation of the redundancy between the *etcone*20 cut and the $\Delta R(l,j)$ requirement, Figure 4.32 (4.33) depicts $F_{elec,jet}$ ($F_{muon,jet}$) as a function of *etcone*20 $< \delta$ ($\delta \leq 10$ GeV), normalized to 1, for different $\Delta R(l,j)$ requirements. Just like for nonprompt electrons, no correlation between the cut on *etcone*20 and the $\Delta R(l,j)$ requirements is observable for jet-faked electrons. Nonprompt and jet-faked muons also behave similarly in that a harder $\Delta R(l,j)$ requirement means a lower efficiency gained by tightening the cut on *etcone*20. However, because of insufficient statistics, this phenomenon is not statistically significant for $F_{muon,jet}$. For $F_{muon,muon}$, a significant difference is only perceptible when comparing $\Delta R(l,j) > 0.25$ with no $\Delta R(l,j)$ requirement (see Figure 4.14 in Section 4.2.2).
Events containing exactly 1 Reco lepton are analyzed for the 1-lepton channel SUSY search. When restricting the fake rates to such events, they decrease marginally: for CSC-isolated leptons, the corrected value of $F_{elec,jet}$ ($F_{muon,jet}$) amounts to $\approx 92\%$ ($\approx 99\%$) of the value given in Table 4.5 (4.6).

Figure 4.29: $F_{muon,jet}$ as a function of the *etcone*20 cut and the $\Delta R(l, j)$ requirement. The labels on the right axis stand for different isolation stages of Reco muons (for the $\Delta R(l, j)$ requirement the cut *etcone*20 $<$ 10 GeV is provided): (1) non-isolated, (2) *etcone*20 $<$ 10 GeV, (3) $\Delta R(l, j) > 0.25$, (4) $\Delta R(l, j) > 0.30$, (5) $\Delta R(l, j) > 0.35$, (6) $\Delta R(l, j) > 0.40$, (7) $\Delta R(l, j) > 0.45$.



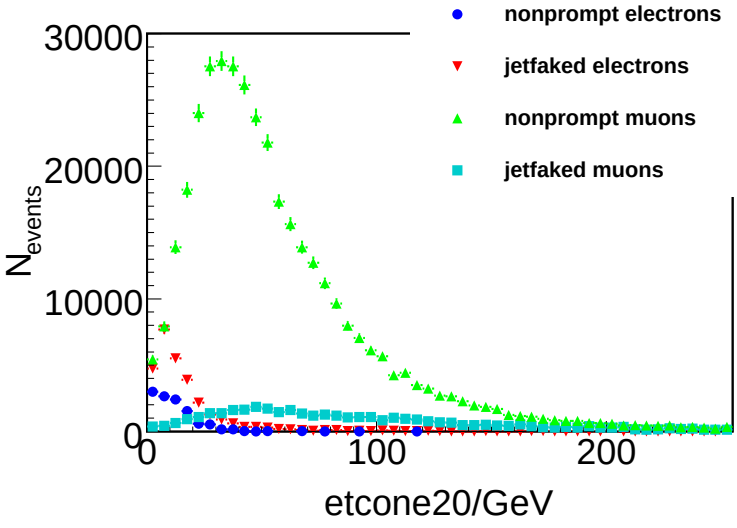


Figure 4.30: *etcone*20 distribution of nonprompt and jetfaked leptons, normalized to $L = 1$ fb^{-1}, for non-isolated Reco leptons.

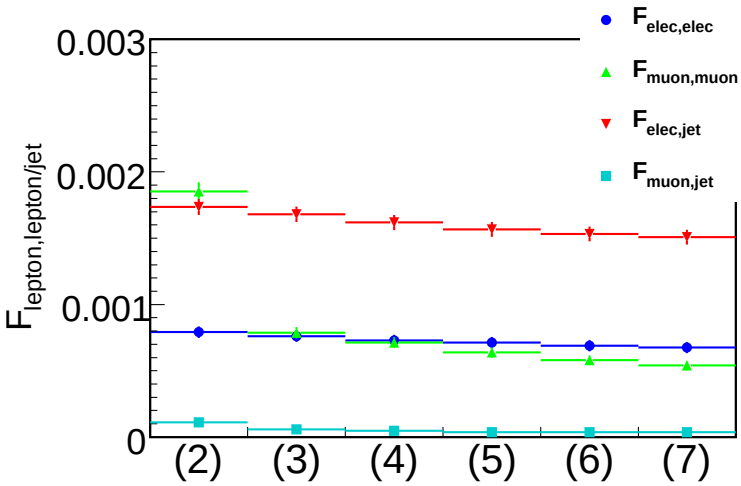


Figure 4.31: $F_{elec,jet}$, $F_{muon,jet}$, $F_{elec,elec}$ and $F_{muon,muon}$ as a function of the $\Delta R(l,j)$ requirement. The label definition of the x-coordinate is the same as in the Figures 4.28 and 4.29.

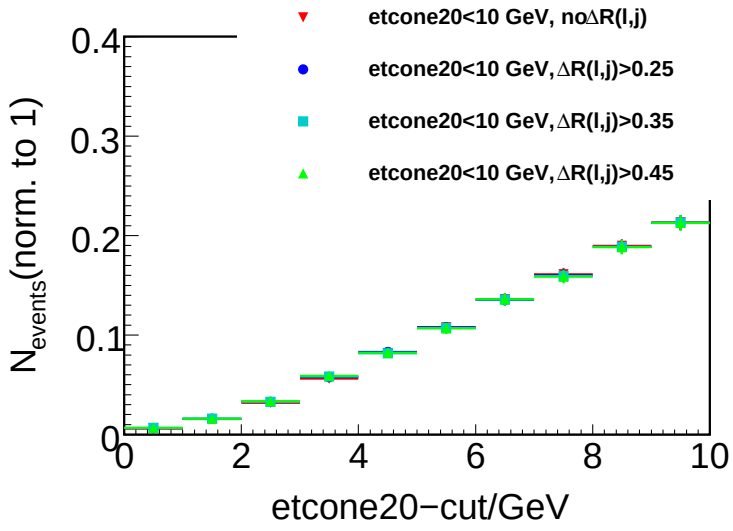


Figure 4.32: $F_{elec,jet}$, normalized to 1, as function of the *etcone*20 cut for different $\Delta R(l,j)$ requirements.

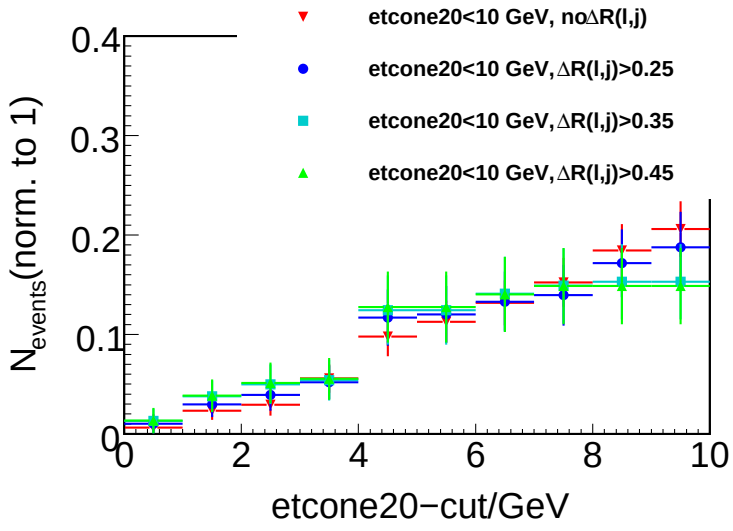


Figure 4.33: $F_{muon,jet}$, normalized to 1, as function of the *etcone*20 cut for different $\Delta R(l,j)$ requirements.

4.3.2 Fake rates as a function of the kinematic variables p_T, η and ϕ for different lepton isolations

Analogously to $F_{elec,elec}$ and $F_{muon,muon}$ in Section 4.2.3, $F_{elec,jet}$ and $F_{muon,jet}$ are investigated in this section as a function of p_T, η and ϕ, stressing the comparison between different lepton isolations. Figure 4.34 (4.37) illustrates $F_{elec,jet}$ ($F_{muon,jet}$) as a function of p_T, normalized to 1, for different lepton definitions all fulfilling $p_T > 10$ GeV. As for nonprompt leptons, the p_T distribution of jet-faked leptons is also shifted to lower values when asking for $etcone20 < 10$ GeV. Isolated Reco leptons have the tendency to be faked by only a part of the Truth jet (see Section 4.1.3) and therefore usually possess less energy than non-isolated ones. Contrarily to the nonprompt fake rate distribution, the jet-faked rate distribution is concentrated at low p_T values for all lepton definitions. Hence, a cut on $p_T > 20$ GeV or $p_T > 30$ GeV has a greater impact on jet-faked leptons than on nonprompt ones. The remaining fractions of $F_{elec,jet}$ ($F_{muon,jet}$) left after tightening the p_T cut to $p_T > 20$ GeV or $p_T > 30$ GeV are listed in Table 4.7 (4.8) for non-isolated and CSC-isolated Reco electrons (muons).

Table 4.7: Remaining fraction of $F_{elec,jet}$ left after tightening the p_T cut to $p_T > 20$ GeV or $p_T > 30$ GeV for two different definitions of Reco electrons.

$F_{elec,jet}/(F_{elec,jet}$ for $p_T > 10$ GeV)	$p_T > 20$ GeV	$p_T > 30$ GeV
no $etcone20$ cut, no $\Delta R(l, j)$ requirement	42%	30%
CSC isolation	22%	10%

Table 4.8: Remaining fraction of $F_{muon,jet}$ left after tightening the p_T cut to $p_T > 20$ GeV or $p_T > 30$ GeV for two different definitions of Reco muons. Due to the limited statistics, the values for $F_{muon,jet}$ using isolated muons can not be taken too seriously.

$F_{muon,jet}/(F_{muon,jet}$ for $p_T > 10$ GeV)	$p_T > 20$ GeV	$p_T > 30$ GeV
no $etcone20$ cut, no $\Delta R(l, j)$ requirement	25%	12%
CSC isolation	0%	0%

$F_{muon,jet}$ can be efficiently suppressed by raising the p_T threshold: as all CSC-isolated jet-faked muons contained in the studied samples have a p_T between 10 GeV and 20 GeV, $F_{muon,jet}$ is reduced to 0 after asking for $p_T > 20$ GeV within the statistical uncertainty.
Figure 4.35 (4.38) shows $F_{elec,jet}$ ($F_{muon,jet}$) as a function of η. For jet-faked electrons, the distributions coincide quite well for the different definitions of Reco electrons. Only in the transition region at $|\eta| \approx 1.1$ is $F_{elec,jet}$ significantly larger for non-isolated electrons than for isolated ones. Furthermore, this η region is in general more populated for jet-faked electrons than for nonprompt electrons (see Figure 4.16). Hence, a cut on this η region ($1 \lesssim |\eta| \lesssim 1.2$) can decrease $F_{elec,jet}$ to $\approx 70\%$ of its nominal value for non-isolated electrons and to $\approx 80\%$ for CSC-isolated electrons. For jet-faked non-isolated muons, the η distribution looks similar to the one for nonprompt non-isolated muons. Unfortunately, statistics is too low for isolated muons to draw conclusions about the distribution.
Finally, Figure 4.36 (4.39) depicts the ϕ distribution for $F_{elec,jet}$ ($F_{muon,jet}$). For both jet-faked electrons and muons, the distributions do not reveal any significant discrepancies between the different lepton definitions and are rather uniform. For CSC-isolated electrons, cutting on $p_T > 20$ GeV instead of $p_T > 10$ GeV does not cause statistically significant differences in the η and ϕ distributions of $F_{elec,jet}$, as further studies show. No conclusions can be made for jet-faked muons on this aspect as there is no jet-faked CSC-isolated muon with $p_T > 20$ GeV in the sample.

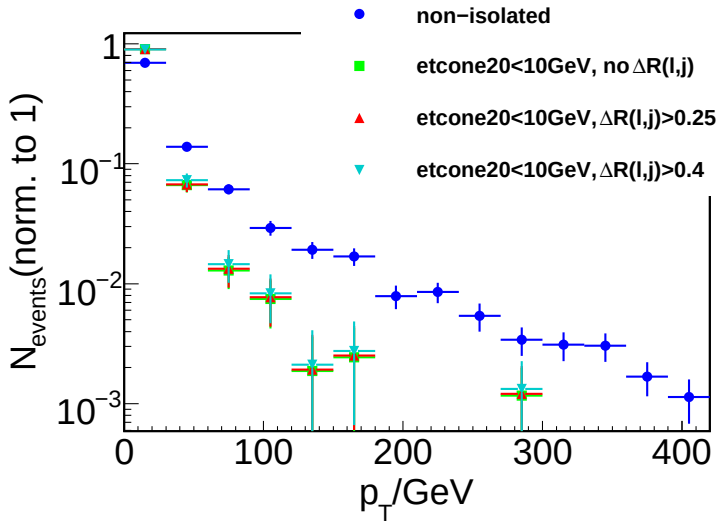


Figure 4.34: $F_{elec,jet}$ as a function of p_T, normalized to 1 and on a logarithmic scale, for different definitions of Reco electrons.

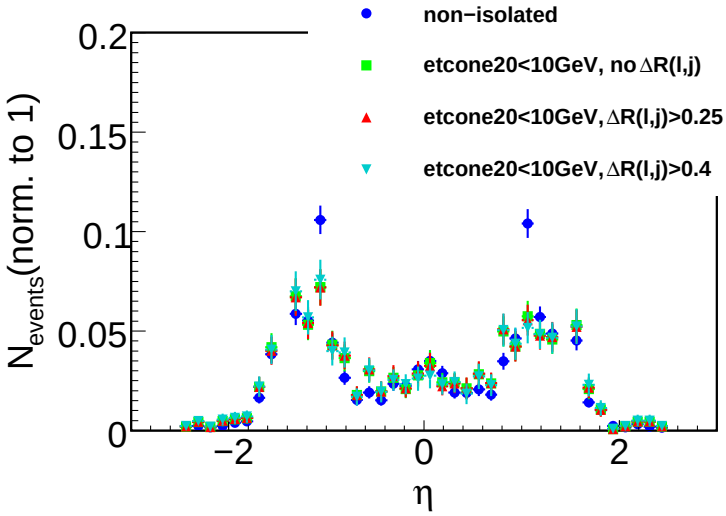


Figure 4.35: $F_{elec,jet}$ as a function of η, normalized to 1, for different definitions of Reco electrons.

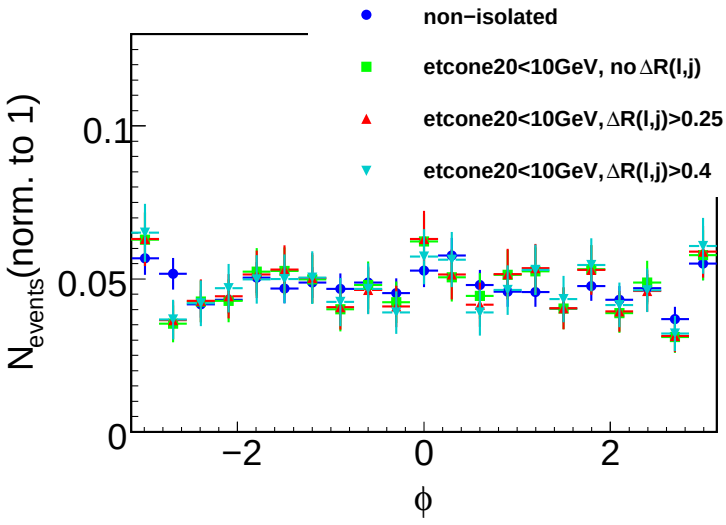


Figure 4.36: $F_{elec,jet}$ as a function of ϕ, normalized to 1, for different definitions of Reco electrons.

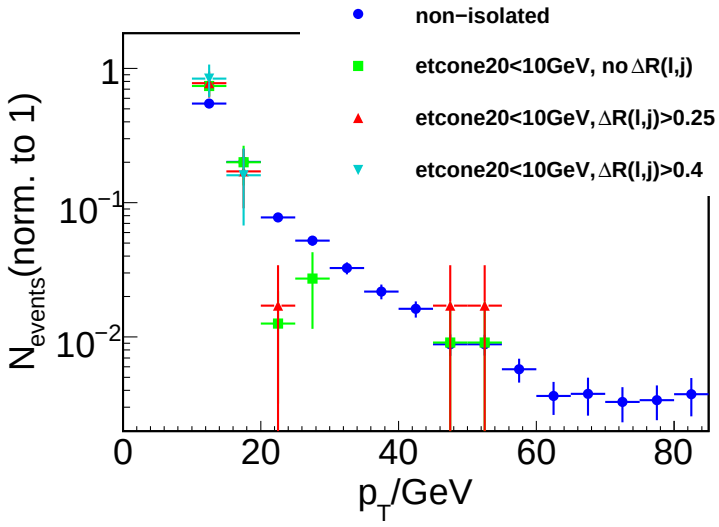


Figure 4.37: $F_{muon,jet}$ as a function of p_T, normalized to 1 and on a logarithmic scale, for different definitions of Reco muons.

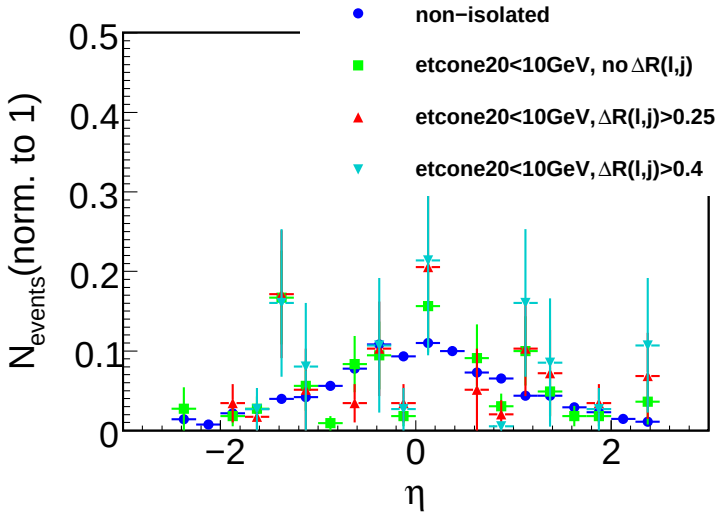


Figure 4.38: $F_{muon,jet}$ as a function of η, normalized to 1, for different definitions of Reco muons.

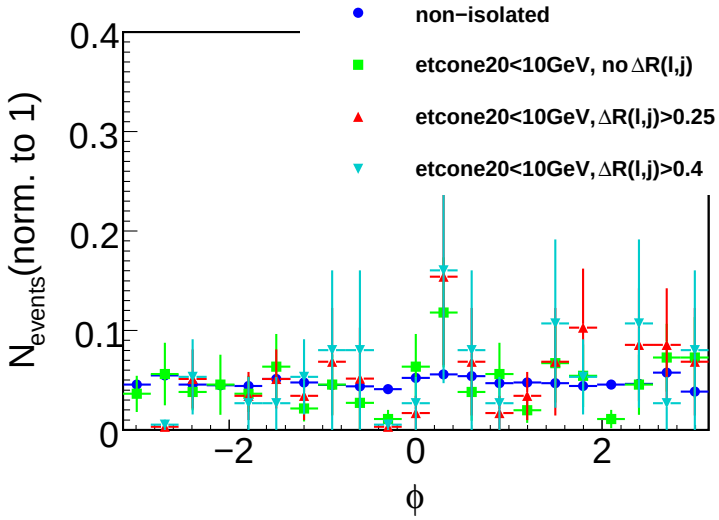


Figure 4.39: $F_{muon,jet}$ as a function of ϕ, normalized to 1, for different definitions of Reco muons.

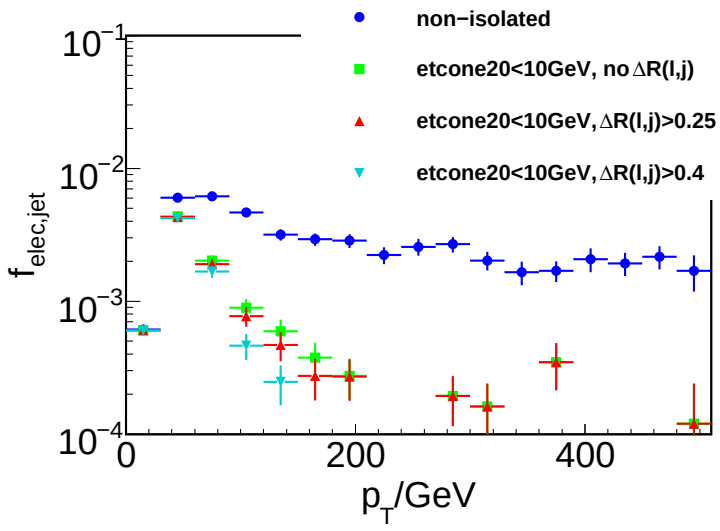


Figure 4.40: $f_{elec,jet}$, on a logarithmic scale, as a function of p_T for different definitions of Reco electrons.

4.3.3 Characteristics of the fake rates as a function of generator particle kinematics

As done in Section 4.2.4 for nonprompt leptons, the focus of this section is directed at the Truth level to estimate the probability of a Truth jet to fake a Reco electron or Reco muon. When studying the minimal $|\eta|$ and p_T values of Truth jets faking a Reco electron or Reco muon, one retrieves the following results used as requirements for the Truth jets below: $|\eta| < 2.7$ and $p_T > 10$ GeV.
Figure 4.40 (4.41) depicts $f_{elec,jet}$ ($f_{muon,jet}$) as a function of the Truth jet p_T for different lepton definitions. For low-p_T Truth jets (10 GeV $\lesssim p_T \lesssim$ 30 GeV), $f_{elec,jet}$ is rather small ($\approx 6 \cdot 10^{-4}$) and coincides for all kinds of Reco electrons. By contrast, Truth jets having a p_T value at $\approx$ 50 GeV have the highest probability of faking a Reco electron for all kinds of electron isolations, i.e. $f_{elec,jet} \approx 6 \cdot 10^{-3}$ for non-isolated electrons and $f_{elec,jet} \approx 4 \cdot 10^{-3}$ for isolated electrons. When going to higher p_T values, $f_{elec,jet}$ decreases exponentially: the stronger the isolation of the considered Reco electrons, the more important the drop. When switching to Truth jets faking Reco muons, the p_T scale is increased by around one order of magnitude. For non-isolated Reco muons, $f_{muon,jet}$ rises from $\approx 3 \cdot 10^{-3}$ to $\approx 5 \cdot 10^{-2}$ when the Truth jet p_T increases from $\approx$ 10 GeV to $\approx$ 3000 GeV. This behaviour can be considered as a further hint for the interpretation of punch through (see 4.1.1): the higher the jet p_T, the higher the probability of not being completely absorbed in the calorimeter and thus reaching the muon spectrometer. For isolated Reco muons, $f_{muon,jet}$ decreases from $\mathcal{O}(10^{-4}) - \mathcal{O}(10^{-5})$ for Truth jets having $p_T \lesssim$ 300 GeV to $\mathcal{O}(10^{-5}) - \mathcal{O}(10^{-6})$ for the p_T range 300 GeV $\lesssim p_T \lesssim$ 600 GeV.

As for nonprompt leptons, the Figures 4.40 and 4.41 demonstrate that the p_T behaviour of jet-faked leptons depends a lot on the isolation of the Reco leptons. Hence, it is not possible to apply conclusions obtained for jet-faked non-isolated leptons to jet-faked isolated leptons. Therefore, it is necessary to investigate the jet-faked CSC-isolated leptons, the recommended isolation, directly, although the available statistics is much worse than for non-isolated leptons. Since for CSC-isolated muons statistics is too low to analyze any correlation between p_T and η or ϕ, only jet-faked CSC-isolated electrons are considered to study this aspect.
Three different p_T bins are defined: 10 GeV $< p_T <$ 30 GeV, 30 GeV $< p_T <$ 50 GeV and $p_T >$ 50 GeV. Figure 4.42 shows $f_{elec,jet}$ as a function of η within the defined p_T bins. As expected, $f_{elec,jet}$ in the $30 < p_T < 50$ GeV bin dominates $f_{elec,jet}$ found in the lower and higher p_T bin (see Figure 4.40). The same behaviour can be seen in Figure 4.43 showing

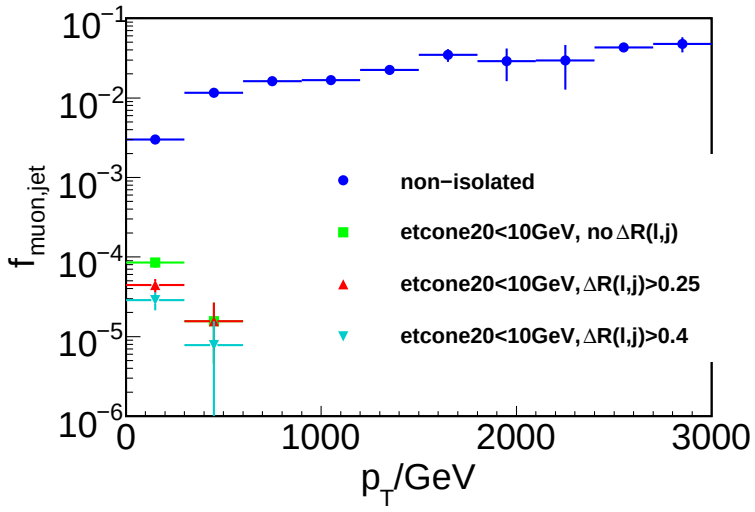


Figure 4.41: $f_{muon,jet}$, on a logarithmic scale, as a function of p_T for different definitions of Reco muons.

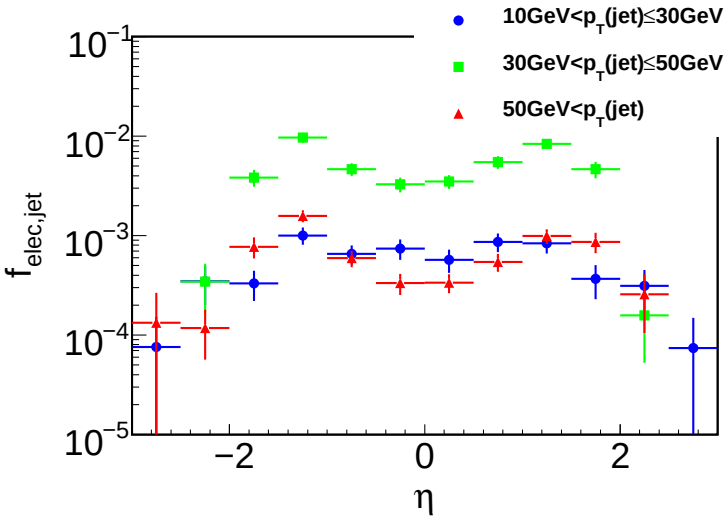


Figure 4.42: $f_{elec,jet}$ as a function of η for three different p_T bins using CSC-isolated Reco electrons.

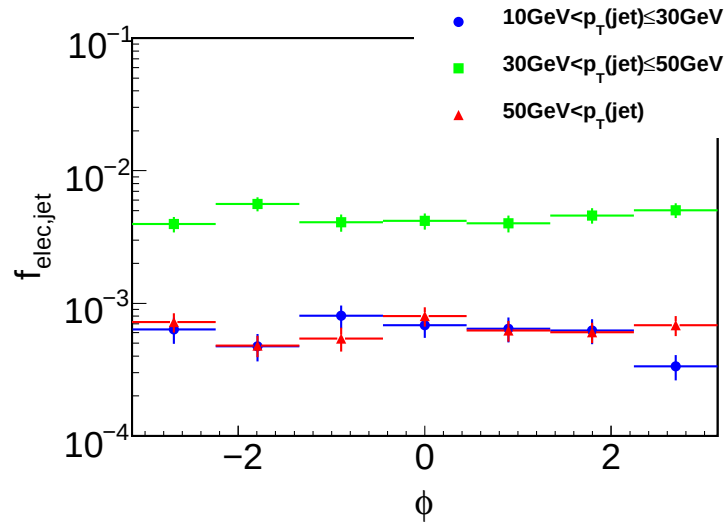


Figure 4.43: $f_{elec,jet}$ as a function of ϕ for three different p_T bins using CSC-isolated Reco electrons.

$f_{elec,jet}$ as a function of ϕ for the above defined p_T bins. As no significant difference in the shapes of the η or ϕ distributions corresponding to the three p_T bins is noticeable, the variables p_T and η or ϕ of the Truth jets faking Reco electrons do not seem to be correlated.

4.3.4 Fake leptons and missing transverse energy

As could be seen in the Figures 4.3 and 4.4, a large difference can occur between the p_T of the faking Truth jet and the one of the fake Reco lepton. This holds especially for jet-faked muons. For jet-faked electrons, the p_T differences are not so important and appear mainly for Reco electrons that are faked by only a part of the matched Truth jet.

Figure 4.44: The distribution of $|\phi_{lepton} - \phi_{\not E_T}|$, normalized to $L = 1$ fb^{-1}, for jet-faked electrons (left plot) and jet-faked muons (right plot).

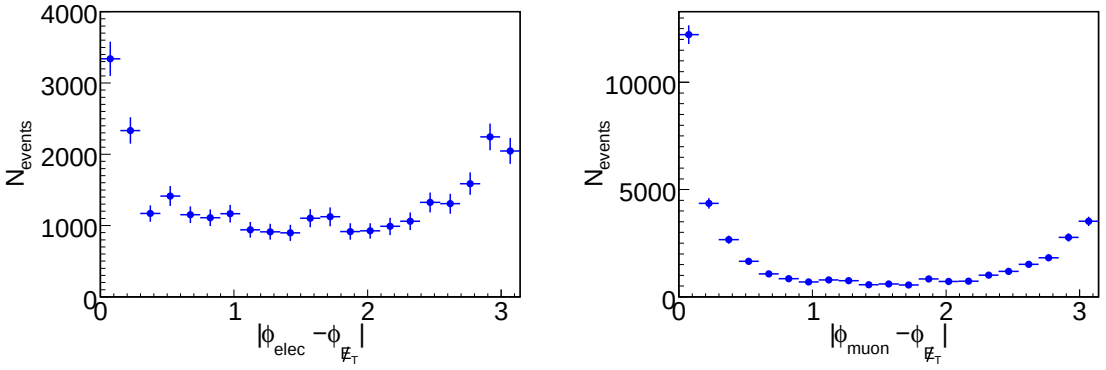
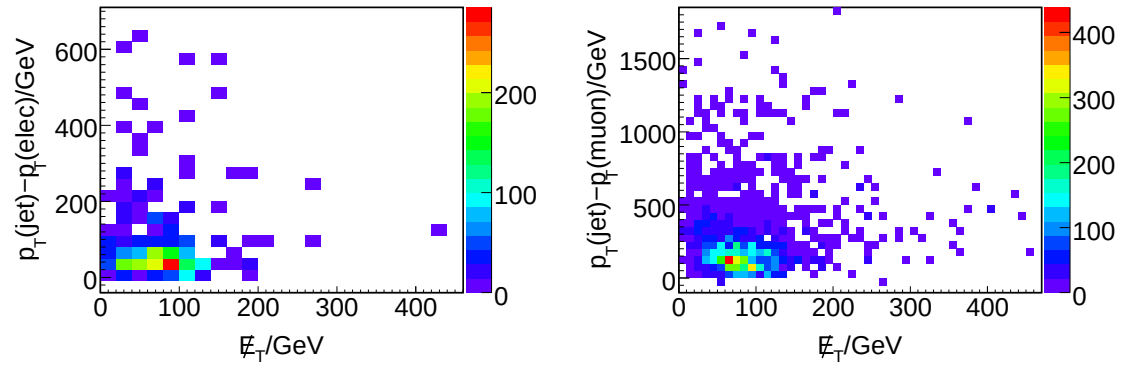


Figure 4.45: $\not E_T$ plotted against the lepton-jet pair p_T difference for jet-faked electrons (left plot) and jet-faked muons (right plot). Both figures are normalized to $L = 1$ fb^{-1}.



The question is to find out if there is any obvious connection between the phenomenon of a jet-faked lepton and the occurrence of missing transverse energy $\not E_T$. To perform this investigation, non-isolated leptons with $p_T > 10$ GeV are used to have enough statistics available.
Figure 4.44 shows the ϕ difference between jet-faked leptons and $\not E_T$, $|\phi_{lepton} - \phi_{\not E_T}|$, normalized to $L = 1$ fb^{-1}, for jet-faked electrons and jet-faked muons, respectively. As the QCD multijet events have a back-to-back structure, one would expect a symmetric distribution with equal maxima at low and high values of $|\phi_{lepton} - \phi_{\not E_T}|$ if there was no correlation between jet-faked leptons and $\not E_T$. However, the maximum at low values of $|\phi_{lepton} - \phi_{\not E_T}|$ is more distinctive than the one at high values, especially for jet-faked muons. This could be a hint for a connection between the direction of jet-faked leptons and $\not E_T$.

Figure 4.45 shows the amount of $\not\!\!E_T$ against the lepton-jet pair p_T difference, $p_T(\text{jet}) - p_T(\text{lepton})$, normalized to $L = 1$ fb^{-1}, for jet-faked electrons and jet-faked muons, respectively. To make this plot, only the events in which the jet-faked leptons and $\not\!\!E_T$ are oriented collinearly ($|\phi_{lepton} - \phi_{\not\!\!E_T}| < 0.1$) were selected. Even though, for jet-faked electrons as well as for jet-faked muons, $\not\!\!E_T$ and the p_T difference of the lepton-jet pair are of the same order of magnitude, Figure 4.45 does not reveal an approximately linear dependency between these two quantities. It is possible that the jet-faked lepton causes just a part of the $\not\!\!E_T$ in the event. In this case, no obvious connection between the lepton-jet pair p_T difference and $\not\!\!E_T$ is expected since $\not\!\!E_T$ is a vectorial sum.

Chapter 5

Summary and outlook

In the first study outlined in Section 3, a worst-case strategy was developed to estimate the QCD background for the 1-lepton SUSY channel. A combined cut efficiency for the cuts on $\not{E}_T$, S_T, M_T and M_{eff} was calculated, while sufficient statistics allowed a direct application of the cuts on lepton and jets.
At first, the cuts were separated as much as possible. However, a strict separation was not possible since most of the cut variables can not be evaluated without first applying some pre-selection cuts. To increase statistics, the cut on $\not{E}_T$, $\not{E}_T > \max(100\text{ GeV}, 0.2 \cdot M_{eff})$, was simplified to $\not{E}_T \geq 100\text{ GeV}$ and the lepton isolation was reduced. Neglecting the correlation between more than two quantities, the cut efficiencies of $\not{E}_T$, S_T, M_T and M_{eff} were evaluated at 68.27%C.L. after having applied different pre-cuts. Multiplying the worst upper limit on efficiency of each cut variable led to $\varepsilon_{combined}$. However, $\varepsilon_{combined}$ depends on the given cut order of $\not{E}_T$, S_T, M_T and M_{eff} as cut efficiencies can only be influenced by cuts that are imposed *before*. Therefore, $\varepsilon_{combined}$ was evaluated for all permutations of the cuts on $\not{E}_T$, S_T, M_T and M_{eff}. A mean value of $\overline{\varepsilon_{combined}} = 1.27\%$ (13.0 passing events) and a maximal value of $\varepsilon_{combined\ max} = 1.63\%$ (16.7 passing events) were obtained.
This method to evaluate the QCD background predicts that the SUSY signal significance should be improved by a factor 1.3 with respect to [22].

Section 4 focussed on the lepton fakes in the QCD background samples. A geometrical matching procedure between Reco leptons and Truth leptons or jets which classifies the Reco leptons as nonprompt or jet-faked was developed.
Nonprompt leptons originate from heavy-quark decays inside jets. It was shown that the ratio of b quarks to c quarks responsible for a nonprompt lepton depends on the isolation of the Reco lepton. More isolated nonprompt Reco leptons tend to come more frequently from a b-quark decay.
When the calorimeter energy deposition of a jet is incidentally related to an inner detector track, the jet can be misidentified as an electron and a jet-faked electron emerges.
High-energetic jets reaching the muon spectrometer are possibly responsible for jet-faked muons.
The integrated fake rates of non-isolated leptons can be efficiently reduced by applying further isolation criteria. For fake electrons isolated by $etcone20 < 10$ GeV, a tightening of the $etcone20$ cut is much more effective to further reduce the fake rate than applying a $\Delta R(l, j)$ requirement. For fake muons, however, the impact of tightening the $etcone20$ cut and imposing a $\Delta R(l, j)$ requirement are comparable, this being mainly due to an inaccurate $etcone20$ calculation for muons in the crack region. For CSC-isolated leptons, $F_{elec,jet}$ represents the bulk of the overall fake rate with a value of $F_{elec,jet} = 1.5 \cdot 10^{-3}$. It is twice as large as the rates for nonprompt leptons and forty times as large as $F_{muon,jet}$. When considering only events containing exactly 1 lepton, the fake rates decreases only in the

percent range.
The fake rates can be effectively suppressed by a tightening of the p_T cut since the p_T distributions of the fake rates drop exponentially. This holds in particular for isolated leptons. For jet-faked electrons, a cut on the transition region between the hadronic barrel and the hadronic extended-barrel calorimeter also acts efficiently. For CSC-isolated electrons, a cut on this region can reduce this dominating fake rate to $\approx 80\%$.
The study at the generator level showed that the probability of a Truth lepton or jet to fake a reconstructed lepton decreases exponentially as a function of p_T, CSC-isolation provided. However, for non-isolated nonprompt electrons and non-isolated fake muons, this behaviour is quite different: the p_T distribution of $f_{elec,elec}$ shows a slight maximum at 200 GeV, $f_{muon,muon}$ is rather constant as function of p_T and $f_{muon,jet}$ even increases. For nonprompt muons, the variables p_T and η show a significant correlation, whereas p_T and ϕ do not seem to be correlated. For fake electrons, no significant correlation between p_T and η or ϕ could be observed.

With regard to the first LHC run at a center-of-mass energy of 10 TeV instead of 14 TeV, these studies should be repeated for the modified conditions.
Furthermore, one could also examine the influence of the event shape and the number of jets on the fake rate. If the fake rate was found to be independent of these variables, one could assign to each QCD jet of given p_T, η and ϕ a fixed fake probability. This would be useful to parametrize the fakes in large QCD multijet samples produced with Atlfast instead of running the full simulation chain (see Section 2.1).
Finally, the lepton definition needs to be optimized with respect to the signal sample and further important background samples. A convenient isolation criterion needs to be found that efficiently reduces QCD fakes and leptons from other background sources but does not cut away too much signal.

Appendix A

Cut behaviour of the individual QCD samples

Table A.1: Listed for each sample are the number of events passing the pre-cut followed by the number of events passing the additionally applied cut on $\not{E}_T$, $\not{E}_T \geq 100$ GeV.

pre-cut	J4	J5	J6	J7	J8
v	24348/3461	90126/7750	26898/3149	3486/936	4241/2434
v + 1 lepton	401/53	1248/114	355/42	35/13	19/8
v + 1 HighPtLepton	185/23	497/60	142/19	19/8	10/5
v + $\geq$ 4 Jets	12445/1174	67835/3191	22564/1982	1639/410	1665/950
v + $\geq$ 4 HighPtJets	5413/419	43920/1456	16019/1239	919/225	936/517
v + $M_{eff} \geq 800$ GeV [$\geq$4J][1l]	20/4	665/44	307/29	21/5	11/5
v + $S_T \geq 0.2$ [$\geq$1J $\|\|$ $\geq$1l]	7599/690	26061/1115	5104/358	119/32	62/29
v + $M_T \geq 100$ GeV [1l]	25/6	52/20	42/10	5/2	5/3

Table A.2: Listed for each sample are the number of events passing the pre-cut followed by the number of events passing the additionally applied cut on S_T, $S_T > 0.2$ (the necessary requirement for the evaluation of S_T, [$\geq$1J $||$ $\geq$1l], is provided and not listed in this table).

pre-cut	J4	J5	J6	J7	J8
v	24330/7599	90123/26061	26898/5104	3486/119	4241/62
v + 1 lepton	401/153	1248/372	355/69	35/2	19/0
v + 1 HighPtLepton	185/77	497/161	142/28	19/1	10/0
v + $\geq$ 4 Jets	12445/6071	67835/24459	22564/4918	1639/100	1665/52
v + $\geq$ 4 HighPtJets	5413/3584	43920/20712	16019/4429	919/84	936/46
v + $\not{E}_T \geq 100$ GeV	3453/690	7750/1115	3149/358	936/32	2434/29
v + $M_{eff} \geq 800$ GeV [$\geq$4J][1l]	20/11	665/251	307/67	21/2	11/0
v + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l]	24/13	8/1	1/0	0/0	0/0
v + $M_T \geq 100$ GeV [1l]	25/8	52/14	42/8	5/1	5/0

Table A.3: Listed for each sample are the number of events passing the pre-cut followed by the number of events passing the additionally applied cut on M_T, $M_T > 100$ GeV.

pre-cut	J4	J5	J6	J7	J8
v + [1l]	401/25	1248/52	355/42	35/5	19/5
v + [1l] + HighPtLepton	185/24	497/50	142/42	19/5	10/4
v + [1l] + $\geq$ 4 Jets	233/14	943/34	307/36	21/2	11/3
v + [1l] + $\geq$ 4 HighPtJets	96/3	620/17	218/28	11/1	8/1
v + [1l] + $\not{E}_T \geq 100$ GeV	53/6	114/20	42/10	13/2	8/3
v + [1l] + $M_{eff} \geq 800$ GeV [$\geq$4J]	20/1	665/29	307/36	21/2	11/3
v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J]	24/4	8/1	1/0	0/0	0/0
v + [1l] + $S_T \geq 0.2$	153/8	372/14	69/8	2/1	0/0

Table A.4: Listed for each sample are the number of events passing the pre-cut followed by the number of events passing the additionally applied cut on M_{eff}, $M_{eff} > 800$ GeV.

pre-cut	J4	J5	J6	J7	J8
v + [$\geq$4J][1l]	233/20	943/665	307/307	21/21	11/11
v + [$\geq$4J][1l] + HighPtLepton	110/12	382/293	122/122	11/11	5/5
v + [$\geq$4J][1l] + HighPtJets	96/18	620/474	218/218	11/11	8/8
v + [$\geq$4J][1l] + $\not{E}_T \geq 100$ GeV	18/4	63/44	29/29	5/5	5/5
v + [$\geq$4J][1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$	24/0	8/1	1/1	0/0	0/0
v + [$\geq$4J][1l] + $S_T \geq 0.2$	124/11	352/251	67/67	2/2	0/0
v + [$\geq$4J][1l] + $M_T \geq 100$ GeV	14/1	34/29	36/36	2/2	3/3

Appendix B

Upper limits of the Poisson distribution at 68.27% C.L.

Table B.1: Upper limits of the Poisson distribution at 68.27% C.L., calculated by the Neyman's construction.

number of events	upper limit [68.27%C.L.]
0	1.15
1	2.36
2	3.52
3	4.65
4	5.77
5	6.87
6	7.97
7	9.06
8	10.14
9	11.22
10	12.29

Appendix C

Combined cut efficiencies for 24 permutations

Table C.1: In the first column, the different permutations of the cuts on $\not{E}_T$, S_T, M_T and M_{eff} are depicted. The second column lists the pre-cuts leading to the combined cut efficiency $\varepsilon_{combined}$ for a given permutation. The values of the corresponding relative worst cut efficiencies are shown in the third column. The fourth column eventually describes the percentage of events passing the different stages of the cut sequence. Its values are retrieved by multiplying successively the relative worst cut efficiencies of the third column. The final value of the fourth column corresponds to $\varepsilon_{combined}$ for a given permutation. When multiplying $\varepsilon_{combined}$ with 1024.2 (the number of events left after cutting on leptons and jets), one obtains the number of events passing all cuts for an integrated luminosity of $L = 1$ fb^{-1}.

	pre-cut	relative worst cut efficiency	cut flow [%]
$\not{E}_T$	v + 1 HighPtLepton	14.04	14.04
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l]	61.10	8.58
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J]	22.88	1.96
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	1.11 ($\varepsilon_{combined}$)
$\not{E}_T$	v + 1 HighPtLepton	14.04	14.04
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l]	61.10	8.58
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	4.84
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J]	22.88	1.11 ($\varepsilon_{combined}$)
$\not{E}_T$	v + 1 HighPtLepton	14.04	14.04
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J]	22.88	3.21
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l]	61.10	1.96
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	1.11 ($\varepsilon_{combined}$)
$\not{E}_T$	v + 1 HighPtLepton	14.04	14.04
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J]	22.88	3.21
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	1.81
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l]	61.10	1.11 ($\varepsilon_{combined}$)

Table C.2: continuation of Table C.1

	pre-cut	relative worst cut efficiency	cut flow [%]
$\not{E}_T$	v + 1 HighPtLepton	14.04	14.04
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	7.92
S_T	v + [≥1J || ≥1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J][1l]	61.10	4.84
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	22.88	1.11 ($\varepsilon_{combined}$)
$\not{E}_T$	v + 1 HighPtLepton	14.04	14.04
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	7.92
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	22.88	1.81
S_T	v + [≥1J || ≥1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J][1l]	61.10	1.11 ($\varepsilon_{combined}$)
S_T	v + [≥1J || ≥1l] + ≥ 4 HighPtJets	52.82	52.82
$\not{E}_T$	v + 1 HighPtLepton	14.04	7.42
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	22.88	1.70
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	0.96 ($\varepsilon_{combined}$)
S_T	v + [≥1J || ≥1l] + ≥ 4 HighPtJets	52.82	52.82
$\not{E}_T$	v + 1 HighPtLepton	14.04	7.42
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	4.19
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	22.88	0.96 ($\varepsilon_{combined}$)
M_T	v + [1l] + HighPtLepton	14.23	14.23
$\not{E}_T$	v + $M_T \geq 100$ GeV [1l]	33.14	4.72
S_T	v + [≥1J || ≥1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J][1l]	61.10	2.88
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	1.63 ($\varepsilon_{combined}$)
M_T	v + [1l] + HighPtLepton	14.23	14.23
$\not{E}_T$	v + $M_T \geq 100$ GeV [1l]	33.14	4.72
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	2.66
S_T	v + [≥1J || ≥1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J][1l]	61.10	1.63 ($\varepsilon_{combined}$)
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	56.44
$\not{E}_T$	v + 1 HighPtLepton	14.04	7.92
S_T	v + [≥1J || ≥1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J][1l]	61.10	4.84
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	22.88	1.11 ($\varepsilon_{combined}$)
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	56.44
$\not{E}_T$	v + 1 HighPtLepton	14.04	7.92
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	22.88	1.81
S_T	v + [≥1J || ≥1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J][1l]	61.10	1.11 ($\varepsilon_{combined}$)
S_T	v + [≥1J || ≥1l] + ≥ 4 HighPtJets	52.82	52.82
M_T	v + [1l] + HighPtLepton	14.23	7.52
$\not{E}_T$	v + $M_T \geq 100$ GeV [1l]	33.14	2.49
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	1.41 ($\varepsilon_{combined}$)
S_T	v + [≥1J || ≥1l] + ≥ 4 HighPtJets	52.82	52.82
M_{eff}	v + [≥4J][1l] + HighPtJets	56.44	29.81
$\not{E}_T$	v + 1 HighPtLepton	14.04	4.19
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [≥4J]	22.88	0.96 ($\varepsilon_{combined}$)

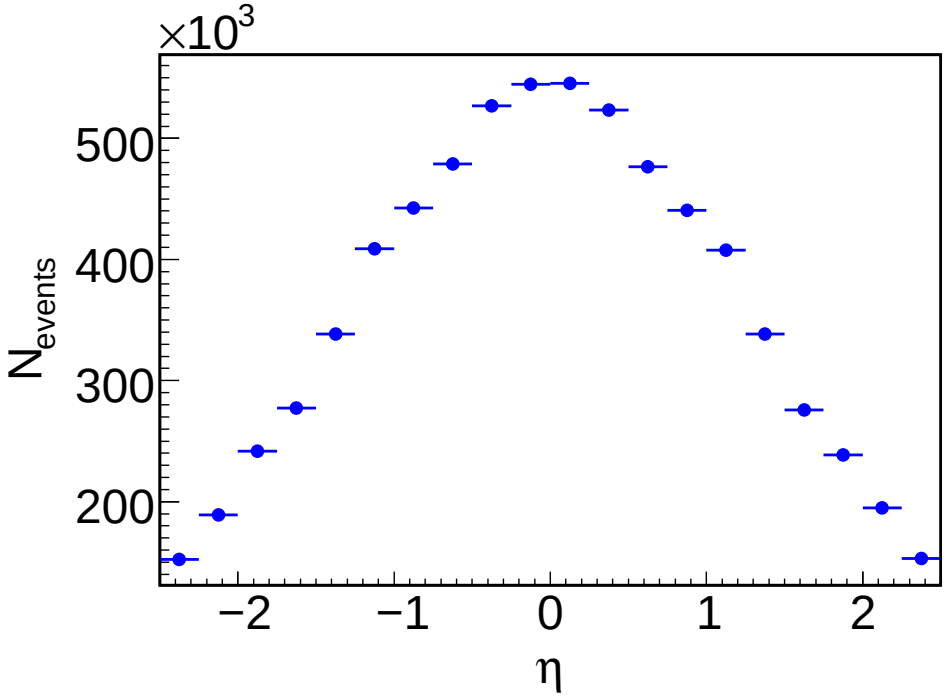
Table C.3: continuation of Table C.1

	pre-cut	relative worst cut efficiency	cut flow [%]
M_T	v + [1l] + HighPtLepton	14.23	14.23
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	7.52
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	2.49
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	1.41 ($\varepsilon_{combined}$)
M_T	v + [1l] + HighPtLepton	14.23	14.23
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	8.03
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	2.66
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l]	61.10	1.63 ($\varepsilon_{combined}$)
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	56.44
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	29.81
$\not{E}_T$	v + 1 HighPtLepton	14.04	4.19
M_T	v + [1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J]	22.88	0.96 ($\varepsilon_{combined}$)
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	56.44
M_T	v + [1l] + HighPtLepton	14.23	8.03
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	2.66
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\not{E}_T \geq (0.2 \cdot M_{eff})$ [$\geq$4J][1l]	61.10	1.63 ($\varepsilon_{combined}$)
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	52.82
M_T	v + [1l] + HighPtLepton	14.23	7.52
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	4.24
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	1.41 ($\varepsilon_{combined}$)
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	52.82
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	29.81
M_T	v + [1l] + HighPtLepton	14.23	4.24
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	1.41 ($\varepsilon_{combined}$)
M_T	v + [1l] + HighPtLepton	14.23	14.23
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	7.52
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	4.24
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	1.41 ($\varepsilon_{combined}$)
M_T	v + [1l] + HighPtLepton	14.23	14.23
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	8.03
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	4.24
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	1.41 ($\varepsilon_{combined}$)
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	56.44
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	29.81
M_T	v + [1l] + HighPtLepton	14.23	4.24
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	1.41 ($\varepsilon_{combined}$)
M_{eff}	v + [$\geq$4J][1l] + HighPtJets	56.44	56.44
M_T	v + [1l] + HighPtLepton	14.23	8.03
S_T	v + [$\geq$1J $\|\|$ $\geq$1l] + $\geq$ 4 HighPtJets	52.82	4.24
$\not{E}_T$	v + $M_T \geq$ 100 GeV [1l]	33.14	1.41 ($\varepsilon_{combined}$)

Appendix D

The η distribution of reconstructed jets

Figure D.1: The η distribution of reconstructed jets, normalized to $L = 1$ fb^{-1}. The reconstructed jets fulfill $p_T > 20$ GeV and $|\eta| < 2.5$.



Bibliography

[1] *The CERN homepage*. http://public.web.cern.ch/public/.

[2] *The DELPHI Experiment homepage*. http://delphiwww.cern.ch/Welcome.html.

[3] L. Evans and P. Bryant (editors). *LHC Machine*. JINST 3. 2008.

[4] D. Perkins. *Hochenergiephysik*. Addison-Wesley. 1990.

[5] ATLAS collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider*. JINST 3. 2008.

[6] H. Baer and X. Tata. *Weak Scale Supersymmetry*. Cambridge University Press. 2006.

[7] D. Griffiths. *Introduction to elementary particles*. John Wiley & Sons. 1987.

[8] T. Sjstrand, S. Mrenna, and P. Skands. *PYTHIA 6.4 Physics and Manual*. hep-ph/0603175. 2006. See also http://home.thep.lu.se/∼torbjorn/Pythia.html.

[9] O. Biebel. *Standardmodell der Teilchenphysik und Erweiterungen*. Lecture notes. The lecture was held at the Ludwig-Maximilians-Universität München in the summer terms 2005 and 2006.

[10] K. Müller. *Einführung Supersymmetrie*. Universität Zürich. 2002.

[11] S. Martin. *A Supersymmetry Primer*. arXiv:hep-ph/9709356v4. 2006.

[12] The ATLAS Collaboration. *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*. December 2008. CERN-OPEN-2008-020.

[13] S. LLoyd. *The ATLAS Computing Workbook*. https://twiki.cern.ch/twiki/bin/view/Atlas/WorkBook. (Last accessed: March 12th, 2008).

[14] S. Agostinelli et al. *Geant4: A simulation toolkit*. Nucl. Instrum. Meth. A506:250-303. 2003. See also http://geant4.web.cern.ch/geant4/.

[15] *EventView*. https://twiki.cern.ch/twiki/bin/view/Sandbox/EventView. (Last accessed: March 11th, 2008).

[16] R. Brun and F. Rademakers. *ROOT - An Object-Oriented Data Analysis Framework*. http://root.cern.ch/.

[17] S. Dean and P. Sherwood. *ATLFAST - the ATLAS fast simulation package*. http://www.hep.ucl.ac.uk/atlas/atlfast/.

[18] *Standard electron reconstruction with the eGamma algorithm*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronReconstruction#Standard_electron_cluster_based. (Last accessed: March 11th, 2008).

[19] *Muon reconstruction with the STACO algorithm*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MuonRecoPedia#STACO. (Last accessed: March 11th, 2008).

[20] *Jet reconstruction with the Cone algorithm*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetAlgorithms. (Last accessed: March 11th, 2008).

[21] G. Dissertori, I. Knowles, and M. Schmelling. *Quantum Chromodynamics, High Energy Experiments and Theory*. Clarendon Press Oxford. 2003.

[22] M.-H. Genest, J. de Graat, G. Krobath, C. Kummer, D. Schaile, C. Cerfon, R. Ströhmer, and M. Wichmann. *Search for SUSY in the 1-Lepton Channel with ATLAS at the LHC*. Jahresbericht 2007, chair D. Schaile LMU München.

[23] *Definition of background samples for SUSY searches*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SusyCscMcProduction#background_samples. (Last accessed: March 11th, 2008).

[24] R. Prabhu, K. Desch, and P. Wienemann. *Estimating Missing Transverse Energy at the Generator Level*. http://doc.cern.ch//archive/electronic/cern/others/atlnot/Communication/phys/com-phys-2007-019.pdf. 2007.

[25] M.-H. Genest. Private communication.

[26] *SUSYView in 12.0.x*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SusyView. (Last accessed: March 11th, 2008).

[27] G. Feldmann and R. Cousins. *A Unified Approach to the Classical Statistical Analysis of Small Signals*. Phys.Rev. D57 (1998) 3873-3889. 1999.

[28] *HighPtView*. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HighPtView. (Last accessed: March 11th, 2008).

[29] T. Nunnemann. Private communication.

Danksagungen

An dieser Stelle möchte ich herzlichst denjenigen danken, die zum Gelingen der vorliegenden Arbeit beigetragen haben:

- Prof. Dr. Dorothee Schaile für die Aufnahme an ihren Lehrstuhl. Erst dadurch wurde mir diese Arbeit ermöglicht und konnte ich ein Teil der ATLAS-Kollaboration werden. Darüber hinaus danke ich ihr für das Erstellen des Erstgutachtens.
- Prof. Dr. Wolfgang Dünnweber für das Erstellen des Zweitgutachtens.
- Mein besonderer Dank gilt Dr. Marie-Hélène Genest für die ausgezeichnete Betreuung dieser Arbeit, die Hilfestellung in allen wichtigen Fragen und das Korrekturlesen. Merci beaucoup!
- Prof. Dr. Otmar Biebel, PD Dr. Thomas Nunnemann und im Besonderen PD Dr. Raimund Ströhmer für die wertvollen Diskussionen.
- Prof. Dr. Dorothee Schaile, PD Dr. Raimund Ströhmer, Dr. Marie-Hélène Genest, Dr. Cristina Galea, Dr. Cédric Serfon, Julien de Graat, Christian Kummer und Michael Wichmann für die vielen fachlichen Impulse im SUSY-Gruppen-Meeting.
- Johannes Ebke, Albert Engl, Klaus Herrmann, Markus Lichtnecker und Dr. Raphael Mameghani für die interessanten Gespräche auch über die Physik hinaus.
- Allen weiteren noch nicht genannten Mitgliedern des Lehrstuhls danke ich für die angenehme Arbeitsatmosphäre: Stefanie Adomeit, Sebastian Becker, Dr. Philippe Calfayan, Dr. Günter Duckeck, Dr. Johannes Elmsheuser, Herta Franz, Alexander Grohsjean, Dr. Petra Haefner, Dr. Ralf Hertenberger, Dr. John Kennedy, Dr. Gernot Krobath, Thomas Langer, Dr. Tariq Mahmoud, Dr. Doris Merkl, Christoph Anton Mitterer, Thomas Müller, Dr. Felix Rauscher, Benjamin Ruckert, Dr. Michiel Sanders, Attila Varga und Dr. Rod Walker.
- Ich danke meinen Freunden und Mitstreitern Johannes Erdmann, Florian Seilmeier und Johann Weber für die gegenseitige Unterstützung während des gesamten Studiums.
- Dankbarkeit empfinde ich gegenüber meinem Opa, Friedrich Müller, und in Erinnerung an meine Oma, Anna Müller, und meine Großeltern, Gertrud und Friedrich Will.
- Von Herzen danke ich meinen Eltern, Sigrid und Hermann Will, und meinen Geschwistern, Johanna und Rebekka Will, für den Rückhalt und die Unterstützung, die ich während meiner Studienzeit erfahren durfte.
- M.K. - In Liebe.

Selbständigkeitserklärung

Ich versichere hiermit, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Jonas Will

München, den 17. März 2009