Signal Region Optimisation Studies Based on BDT and Multi-Bin Approaches in the Context of Supersymmetry Searches in Hadronic Final States with the ATLAS Detector

Master Thesis

UNIVERSITY OF FREIBURG

Department of Mathematics and Physics

Manuel Guth

Supervisor: Prof. Dr. Gregor Herten

Submission date: March 16, 2018

Acknowledgement

Every scientific thesis is a challenge which requires the support of experienced people. I would like to give my sincere gratitude to Prof. Dr. Gregor Herten, who gave me the opportunity to write my Masterthesis as well as my Bachelorthesis in his group.
Foremost I would like to give special thanks to Nikola Makovec who was my supervisor during my stay at Laboratoire de l'Accélérateur Linéaire in Orsay. He guided me through the whole work for this thesis. Also Zuzana Rúriková was a great help during all the work for the thesis. Moreover, I would like to thank Andrea Knue for her support. In addition I would like to thank everyone in the analysis team from Freiburg: Fabio Cardillo, Simone Curcio, Veronika Magerl, Peter Tornambè, Andrea Di Simone and all the guys from LAL, especially Konie Al Khoury, Aishik Gosh and Antoine Laudrain who always helped me in case of any problems.
Thanks also to all who reviewed the thesis and gave me helpful suggestions.

Abstract

The searches for supersymmetric phenomena are mostly based on simple Cut & Count methods. One example is the search for squarks and gluinos in final states with multiple jets, missing transverse momentum and without leptons. This analysis, based on $36.1\,\mathrm{fb}^{-1}$ of pp collision data at $\sqrt{s} = 13\,\mathrm{TeV}$ recorded with the ATLAS detector, uses Cut & Count based methods in the signal regions. In order to improve the analysis sensitivity, the use of sophisticated techniques, such as boosted decision trees (BDT) and Multi-Bin, is being investigated in this thesis. The focus of the study lies on squarks and gluino searches. These techniques are evaluated using Monte Carlo simulation. The goal is to find a new approach which is on the one hand simple but allows for a significant improvement. A gain up to approximately 200 GeV in the neutralino mass and an enhancement of about 200 GeV in the squark and gluino mass is achieved with these new techniques.

Kurz-Zusammenfassung

In der Suche nach supersymmetrischen Phänomenen werden häufig einfache schnitt-basierte Signalregionen verwendet. Ein Beispiel dafür ist die Suche nach supersymmetrischen Phänomenen in den Endzuständen mit mehreren Jets, fehlendem transversem Impuls und ohne Leptonen. Diese Suche, beruhend auf Proton-Proton-Kollisionsdaten mit $36.1\,\mathrm{fb}^{-1}$ und einer Schwerpunktsenergie von $\sqrt{s} = 13\,\mathrm{TeV}$, aufgenommen mit dem ATLAS Detektor nutzte ausschließlich schnitt-basierte Signalregionen. Um mögliche Verbesserungen zu untersuchen, werden komplexere Analyseverfahren, wie der Multi-Bin-Ansatz und Boosted Decision Trees (BDT), in dieser Arbeit getestet. Dabei liegt der Fokus auf Squark- und Gluino-Suchen. Diese neuen Techniken werden mit Hilfe von simulierten Daten ausgewertet. Das Ziel ist, eine möglichst hohe Signifikanz zu erhalten, ohne die Komplexität der Analyse dabei zu stark zu erhöhen. Diese Methoden erlauben eine Verbesserung in der Größenordnung von 200 GeV in der Neutralinomasse als auch der Squark- und Gluinomasse.

Contents

1 Introduction

"We are trying to prove ourselves wrong as quickly as possible,
because only in that way can we find progress."
Richard Feynman

Physics strives to understand the fundamental principles of nature. Especially, the field of particle physics tries to understand the smallest building blocks of our universe and the interactions between them. To get a glimpse behind the scenes of elementary physics, complex experiments are necessary.

Several of these large experiments are located at the Large Hadron Collider (LHC) at CERN (Centre européenne pour la recherche nucléaire). Running these experiments and analysing the recorded data represents a major challenge for physicists.
There are four main experiments situated around the LHC (ATLAS, CMS, LHCb and Alice) among which ATLAS and CMS are general purpose detectors which carry a wide range of physics: from probing the Standard Model of particle physics to the search for Supersymmetry (SUSY) and extra dimensions. After the discovery of the Higgs boson in 2012 [1, 2] the main focus lies on precision measurements of Standard Model particles and on the search for new physics.

The Standard Model represents a very successful model to describe the elementary particles and their interactions up to a high precision. After the discovery of the the Higgs boson, there are still several phenomena which cannot be explained by the Standard Model such as dark matter, the hierarchy problem or inclusion of gravity. Consequently, a new theory or rather an extension of the theory is necessary.

There are a large number of models trying to describe these phenomena. One of them is a model which adds another symmetry to the Standard Model, called *supersymmetry*. This symmetry relates fermions to bosons and vice versa and consequently each SM particle is related to a superpartner which has identical quantum numbers except the spin.
The search for supersymmetry (SUSY) is an important area of research taking place at the ATLAS experiment. Many of the supersymmetric models were already investigated using the data from LHC RUN I (2009-2012). However, no evidence for physics beyond the Standard Model was found. The ongoing LHC RUN II which started in 2015 performs at a higher centre-of-mass energy, increased from $\sqrt{s} = 8\,\text{TeV}$ to $\sqrt{s} = 13\,\text{TeV}$. This allows to access new phase-space regions and therefore provides new opportunities to search for SUSY.

SUSY analyses are divided in different subanalyses depending on the signal processes and their final states. The analysis which will be further discussed and presented in this thesis is called *SUSY 0-lepton* analysis. It is focused on the search for squarks and gluinos (supersymmetric partners of quarks and gluons) which are pair produced at the LHC and decay either directly or via cascades into SM particles and the lightest SUSY particle (LSP). Due to a better background suppression, purely hadronic final states are considered, labelling it therefore *0-lepton*.

The latest published 0-lepton results from ATLAS were presented at Moriond 2017 [3]. This analysis used only cut-based event selections. For the next update of the analysis, more sophisticated analysis techniques are needed to achieve significant improvements.
One of these techniques are boosted decision trees (BDT) and the Multi-Bin approach. These two methods are studied in this thesis in order to improve the signal region definitions of squark and gluino pair production with subsequent direct decays.
The goal is to find an approach which is simple but also allows for significant improvements.

In the chapters 2 and 3 of this thesis an overview of the theory is given. The technical setup of the ATLAS experiment is explained in chapter 4 and a brief introduction to the 0-lepton analysis is presented in chapter 5. The idea of multivariate analysis techniques is introduced in chapter 6 while in chapter 7, the signal region improvement studies are performed for the squark pair production followed by the gluino pair production in chapter 8.

2 Theoretical Overview

After decades of research in high energy physics, the last puzzle piece of the Standard Model was found in 2012 at the Large Hadron Collider (LHC) at CERN, the Higgs boson [1, 2]. All phenomena seen at colliders are pretty well described by this theory. Nevertheless, there are indications that physics beyond the Standard Model is present, e.g. from astro physics , one of these models is called Supersymmetry (SUSY).
The following chapter will give a brief overview of the Standard Model and possible extensions.

2.1 The Standard Model

The most successful theory in particle physics is the Standard Model of particle physics (SM). It describes all known elementary particles and their interactions with a high precision. The construction of the SM is inspired by two main principles: simplicity and symmetries.
Generally, the SM involves all fundamental forces namely the electromagnetic, weak and strong force but does not include gravity. Furthermore, the SM is a relativistic quantum field theory (QFT) combining the classical field theory and quantum mechanics in a relativistic manner [4].

It is a gauge theory based on the non-Abelian symmetry group

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

where Y refers to the hypercharge of the particles, L to left handed fields (these two symmetries are the basis of the electroweak theory), and C the colour charge of the quarks being the symmetry of the strong theory also called quantum chromodynamics (QCD).

The SM lagrangian $\mathcal{L}_{SM}$ is a scalar function which fully determines the dynamics and interactions of the model and can be grouped into different parts

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{EW}} + \mathcal{L}_{\mathrm{QCD}} + \mathcal{L}_{\mathrm{H}} + \mathcal{L}_{\mathrm{YU}}, \tag{2.2}$$

which are in detail the electroweak (EW), the strong (QCD) and the Higgs (H) sectors with the Yukawa (YU) sectors of the Higgs-fermion interaction. They will be described in more detail below.

Currently, the known SM particle content consists of twelve fermions and their antiparticles, twelve gauge bosons which mediate the forces and one scalar boson.
The fermions are sorted into three generations where each higher generation is a heavier copy of the previous one. In addition, the fermions are further divided into leptons and quarks as shown in table 2.1. All visible matter is almost exclusively made of first generation fermions, while all higher generation particles are unstable and decay into other particles.
The twelve gauge bosons are mediator particles of spin one, listed in table 2.2. The strong

	Leptons			Quarks		
Generation	Flavour	Charge [e]	Mass [MeV]	Flavour	Charge [e]	Mass [MeV]
I	e	-1	0.511	u	$2/3$	$2.2^{+0.6}_{-0.4}$
	ν_e	0	$< 2 \cdot 10^{-6}$	d	$-1/3$	$4.7^{+0.5}_{-0.4}$
II	μ	-1	105.7	c	$2/3$	$1,280 \pm 30$
	ν_μ	0	< 0.19	s	$-1/3$	96^{+8}_{-4}
III	τ	-1	$1,776.8$	t	$2/3$	$173,210 \pm 510 \pm 710$
	ν_τ	0	< 18.2	b	$-1/3$	$4,660^{+40}_{-30}$

Table 2.1: The mass eigenstates of the SM fermions grouped in three generations and divided into leptons and quarks [5].

interaction is mediated by eight massless gluons (which differ in their colour charge), the weak interaction by the massive $W^\pm$ and Z bosons and the electromagnetism is transmitted by the massless photon. Additionally, the scalar spin zero particle, the Higgs boson, is also part of the SM.

Gauge bosons					
	Mass[GeV]	Spin	Charge [e]	Coupling	Range [fm]
gluon g	0	1	0	$\alpha_S(m_Z) \sim 1/10$	~ 1
photon γ	$< 10^{-24}$	1	$< 10^{-35}$	$\alpha_{EM}(m_Z) \sim 1/127$	∞
$W^\pm$	80.385 ± 0.015	1	± 1	$\alpha_W \sim 10^{-6}$	$< 2 \cdot 10^{-3}$
Z	91.1876 ± 0.0021	1	0		

Table 2.2: Gauge bosons of the Standard Model [5].

2.1.1 Quantum Electrodynamics

Quantum electrodynamics (QED) is the theory describing the interaction of electrically charged particles via the electromagnetic force. Thereby, the Dirac Lagrangian describes massive spin $1/2$ particles

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(\mathrm{i}\not{\partial} - m)\psi, \tag{2.3}$$

where $\not{\partial} = \gamma_\mu \partial^\mu$ denotes the contraction with the Dirac matrices γ_μ, m the mass and ψ a free spinor field.
Comparing the Dirac equation with the relativistic Maxwell equations

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} = 0, \tag{2.4}$$

$$\partial_\mu F^{\mu\nu} = j^\nu, \tag{2.5}$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the field strength tensor and A_μ a vector field, one finds that the Maxwell equations are invariant under the gauge transformation $A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \alpha$, but the

Dirac Lagrangian is not (e is the electrical charge). The corresponding symmetry group is the $U(1)$ which includes a local phase transformation

$$\psi(x) \to e^{-\mathrm{i}\alpha(x)}\psi(x), \qquad \bar{\psi}(x) \to \bar{\psi}(x)e^{\mathrm{i}\alpha(x)}, \tag{2.6}$$

with α the electromagnetic coupling constant, which leads to an additional term $\bar{\psi}\partial\!\!\!/\alpha\psi$ in the Lagrangian under a local $U(1)$ transformation. In order to achieve an invariant Lagrangian, one can introduce a coupling between the Dirac fermion and the vector field A_μ in the form of a covariant derivative

$$D_\mu = \partial_\mu + \mathrm{i}eA_\mu. \tag{2.7}$$

Combining the free field dynamics described by equation (2.5), the Dirac Lagrangian in equation (2.3) and the new coupling from (2.7) one ends up with the QED Lagrangian

$$\mathcal{L}_{\text{QED}} = \bar{\psi}\left(\mathrm{i}\gamma^\mu D_\mu - m\right)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{2.8}$$

2.1.2 Quantum Chromodynamics

The theory of quantum chromodynamics (QCD) describes the strong interaction of particles carrying colour charge. This theory is constructed analogous to QED. The underlying symmetry group is $SU(3)$ and the gauge bosons are the gluons. A quark colour field $q(x)$ transforms under a local $SU(3)$ like

$$q(x) \to e^{\mathrm{i}\alpha_a T_a}q(x), \tag{2.9}$$

where T_a are the generators of the $SU(3)$ group and can be written as

$$T_a = \frac{1}{2}\lambda_a, \tag{2.10}$$

$$[T_a, T_b] = \mathrm{i}f_{abc}T_c, \tag{2.11}$$

with λ_a the Gell-Mann matrices [6], f_{abc} the completely anti-symmetric structure constant and the colour index a running from $1-8$. The coupling between quarks and gluons is introduced as a covariant derivative similar to QED:

$$D_\mu = \partial_\mu + \mathrm{i}g_s\frac{\lambda_a}{2}G^a_\mu, \tag{2.12}$$

where G^a_μ are the eight gluon field strength tensors and g_s is the coupling strength. The final QCD Lagrangian then reads

$$\mathcal{L}_{\text{QCD}} = \bar{q}\left(i\gamma^\mu D_\mu - m\right)q - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a. \tag{2.13}$$

As a consequence of the non-abelian structure of the $SU(3)$ group, in the Lagrangian (2.13) terms of the order g_sG^3 and $g_s^2G^4$ appear which describe three and four gluon self-interactions. Thus, gluons carry a colour charge themselves, which influences the behaviour of colour charged particles. If one separates two coloured particles, the force between these particles increases since the potential is proportional to the distance. After a certain point it becomes energetically more favourable for a quark anti-quark pair to appear and the resulting state is colourless. This effect is also called colour confinement and results in the formations of colourless states, called jets. This process is called *hadronisation*. With increasing energy (small distances) the coupling strength between quarks decreases and they behave like free particles. This is called *asymptotic freedom*.

2.1.3 Electroweak Unification

The third force included in the SM is the weak force. In the 1960s Glashow [7], Salam [8] and Weinberg [9] already combined the weak and the electromagnetic theories to a unified electroweak theory. This theory is invariant under the gauge group $SU(2)_L \otimes U(1)_Y$ whereby the $SU(2)_L$ part induces three W^j_μ bosons and the $U(1)_Y$ part gives rise to another boson B_μ. The W_μ bosons couple to the weak isospin T_3 and the B_μ boson to the weak hypercharge Y obeying the following Gell-Mann–Nishijima formula [10]

$$Y = 2Q - 2T_3, \tag{2.14}$$

with Q the electric charge.
However, these fields are not corresponding to the observed bosons but the mass eigenstates are a linear combination of them. The resulting mixing for the charged bosons $W^\pm$ is given as

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp W^2_\mu\right), \tag{2.15}$$

and the neutral ones are formed as follows

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W, \tag{2.16}$$

$$Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W, \tag{2.17}$$

where A_μ describes the photon and Z_μ the Z boson, with θ_W the weak mixing angle

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}. \tag{2.18}$$

Hence g is the coupling corresponding to the $SU(2)$ group and g' analogously to the $U(1)$ group. The fermions interacting via the electroweak force are shown in table 2.3 including their quantum numbers.

2.1.4 Electroweak Symmetry Breaking - Higgs Sector

The mediators of the electroweak theory as described above are by construction massless since any mass term bilinear in the fields would violate the gauge invariance. However, the observed particles carry mass and a new mechanism has to be introduced called the Brout-Englert-Higgs mechanism [11, 12]. The idea is to add an additional complex scalar $SU(2)_L$ doublet $\Phi = (\phi^+, \phi^0)^T$ of hypercharge $Y = 1$ to the SM

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi) = (D_\mu\Phi)^\dagger(D^\mu\Phi) + \mu^2(\Phi^\dagger\Phi) - \frac{\lambda}{4}(\Phi^\dagger\Phi)^2, \tag{2.19}$$

Generations			Charges		
I	II	III	T_3	Y	Q
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$+1/2$ $-1/2$	-1 -1	0 -1
e_R	μ_R	τ_R	0	-2	-1
$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$+1/2$ $-1/2$	$1/3$ $1/3$	$+2/3$ $-1/3$
u_R	c_R	t_R	0	$4/3$	$+2/3$
d_R	s_R	b_R	0	$-2/3$	$-1/3$

Table 2.3: The fermions are grouped in singlets according to the $U(1)$ symmetry group and in doublets corresponding to the $SU(2)$ symmetry group. The shown quantum numbers are the third component of the weak isospin (T_3), the hypercharge (Y) and the electric charge (Q).

where μ and λ are the parameters of the potential obeying $\mu^2 < 0$ and $\lambda > 0$. As a consequence, the gauge symmetry is broken for perturbations around the vacuum whereas the entire Lagrangian retains that symmetry (spontaneous symmetry breaking). By minimising the potential V (shown in fig. 2.1) with respect to $\Phi^\dagger\Phi$ in equation (2.19) one ends up with the vacuum expectation value

$$v = 2\sqrt{\frac{\mu^2}{\lambda}}. \tag{2.20}$$

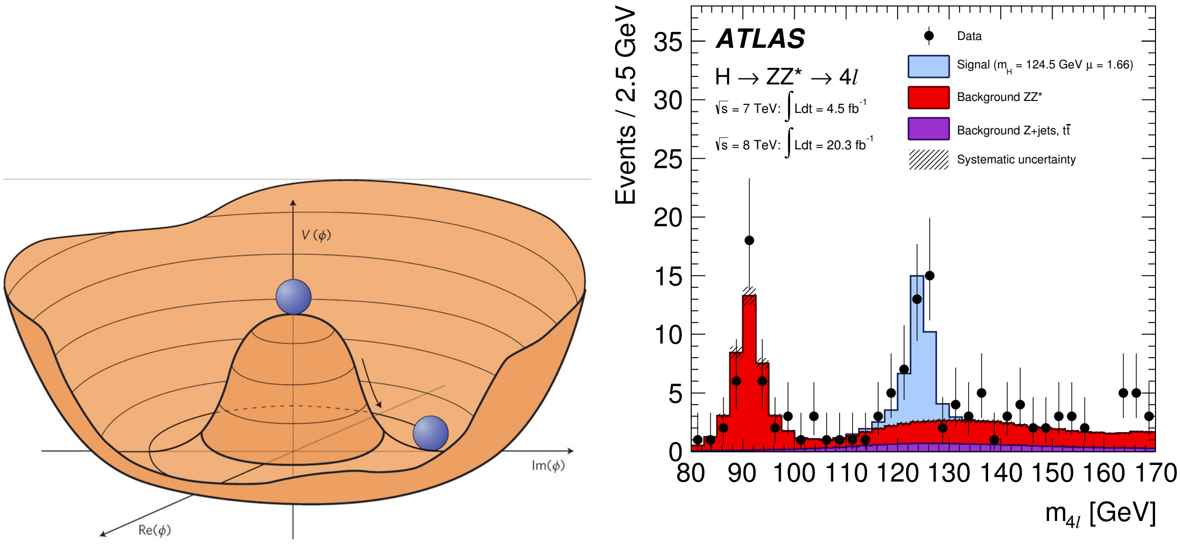
In addition, a new scalar particle occurs from the self-interaction of the Higgs field (the so-called Higgs boson) and the problem of observing massive weak gauge bosons is solved by this mechanism. The masses of these particles yield to

$$m_W = \frac{gv}{2} \qquad m_Z = \frac{m_W}{\cos\theta_W} \qquad m_H = \sqrt{2\mu^2}. \tag{2.21}$$

Besides the gauge bosons, the Higgs couples via Yukawa couplings to fermions which also explains their mass.

2.1.5 An Incomplete Theory

Despite its success, the SM fails to describe some phenomena observed in recent years. First of all, it does not include gravity. In addition, it fails to explain neutrino masses. From neutrino oscillations [15] however, it has been found that at least two out of three neutrinos need to have mass. From astrophysical observations the existence of so-called dark matter/dark energy is known [16, 17]. A particle contributing to dark matter needs to be massive, but only interacting weakly. In the SM, there is no particle which could explain this dark matter. Also, the matter-antimatter asymmetry, the number of quark and lepton generations and their large mass differences ($m_{\mathrm{top}}/m_{\mathrm{e}} \sim 3 \cdot 10^5$) represent open questions. Another aspect is the hierarchy problem occurring in the large separation of scales ($\Lambda_{\mathrm{QCD}} < \Lambda_{\mathrm{weak}} \lll \Lambda_{\mathrm{Planck}}$). Since scalars receive mass corrections proportional to Λ, this yields to a large instability of the Higgs mass and requires an extreme fine tuning.



a) b)

Figure 2.1: a) Higgs potential in two dimensions with $\lambda > 0$ and $\mu^2 < 0$ [13] and b) the distribution of the four-lepton invariant mass for the combined 7 TeV and 8 TeV data samples recorded with the ATLAS experiment showing the SM Higgs boson signal for $m_H = 124.5\,\text{GeV}$ normalised to the measured signal strength [14].

2.2 Supersymmetry (SUSY)

One possible solution to answer the open questions mentioned above, is to introduce a new symmetry which relates fermions to bosons and vice versa:

$$Q\,|fermion\rangle = |boson\rangle\,, \tag{2.22}$$
$$Q\,|boson\rangle = |fermion\rangle\,. \tag{2.23}$$

This symmetry is called supersymmetry, with Q the supersymmetric operator, and doubles the degrees of freedom of the SM. These new degrees of freedom combine into physical superpartners of the SM particles thus the divergences of the Higgs mass corrections get under control and thereby the hierarchy problem gets solved. Additionally, the lightest supersymmetric particle (LSP) provides a candidate for dark matter and some models also include gravity.

The naming convention of the SUSY particles can be shortly summarised

fermion → **s**fermion,

boson → bos**ino**,

so e.g. the superpartner of the tau is the stau and the gluino is the superpartner of the gluon.

2.2.1 R-Parity

In order to take into account phenomenological constraints e.g. the suppression of the proton decay a new quantum number denoted as R-parity is introduced and its conservation can be assumed

$$R = (-1)^{3(B-L)+2S}, \tag{2.24}$$

where B is the baryon number, L the lepton number and S the spin of the particle. Without R-parity conservation, the proton could decay since baryon and lepton numbers would not be conserved. One possible decay scheme of the proton is shown in figure 2.2 where the proton decays via an antistrange squark into an antielectron and an up-quark. All Standard Model particles have R-parity 1 whereas all SUSY particles have R-parity -1. Due to the introduction of the R-parity, the LSP is stable and therefore presents a dark-matter candidate.

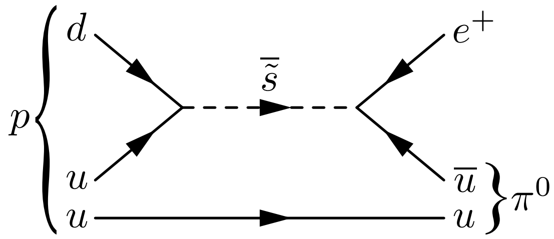


Figure 2.2: Feynman diagram of a possible proton decay mediated by the superpartner of the strange antiquark.

As another consequence of the conservation of the R-parity, the SUSY particles have to be produced in pairs at colliders with an initial state of SM particles.

2.2.2 Supersymmetry Breaking

If supersymmetry would be an exact symmetry, the masses of the superpartners would be equal to their Standard Model equivalent. Consequently, there should be a bosonic electron with a mass of 511 keV which has not been observed. Therefore the symmetry should be spontaneously broken and heavier SUSY particles are possible.

2.2.3 Particles in the MSSM

The most basic SUSY model is the Minimal Supersymmetric Model (MSSM) which is the renormalisable counterpart of the Standard Model containing the minimal number of fields and couplings (124 free parameters). The R-parity conservation is included as well as the soft SUSY breaking. In order to give mass to the up- and down-type quarks while conserving the supersymmetry, two Higgs doublets are necessary. The summary of all gauge/chiral eigenstates and mass eigenstates of the SUSY particles and the Higgs bosons are presented in table 2.4. Another feature of the MSSM is the unification of the three forces at a high scale ($\sim 10^{16}$ GeV).

2.2.4 The pMSSM

As a result of the mechanisms introduced in MSSM, 124 independent parameters are necessary (18 from SM, 1 Higgs parameter and 105 new parameters from the model).

Names	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	+1	$H_u^0\ H_d^0\ H_u^+\ H_d^-$	$h^0\ H^0\ A^0\ H^\pm$
squarks	0	−1	$\tilde{u}_L\ \tilde{u}_R\ \tilde{d}_L\ \tilde{d}_R$ $\tilde{s}_L\ \tilde{s}_R\ \tilde{c}_L\ \tilde{c}_R$ $\tilde{t}_L\ \tilde{t}_R\ \tilde{b}_L\ \tilde{b}_R$	(same) (same) $\tilde{t}_1\ \tilde{t}_2\ \tilde{b}_1\ \tilde{b}_2$
sleptons	0	−1	$\tilde{e}_L\ \tilde{e}_R\ \tilde{\nu}_e$ $\tilde{\mu}_L\ \tilde{\mu}_R\ \tilde{\nu}_\mu$ $\tilde{\tau}_L\ \tilde{\tau}_R\ \tilde{\nu}_\tau$	(same) (same) $\tilde{\tau}_1\ \tilde{\tau}_2\ \tilde{\nu}_\tau$
neutralinos	1/2	−1	$\tilde{B}^0\ \tilde{W}^0\ \tilde{H}_u^0\ \tilde{H}_d^0$	$\tilde{\chi}_1^0\ \tilde{\chi}_2^0\ \tilde{\chi}_3^0\ \tilde{\chi}_4^0$
charginos	1/2	−1	$\tilde{W}^\pm\ \tilde{H}_u^+\ \tilde{H}_d^-$	$\tilde{\chi}_1^\pm\ \tilde{\chi}_2^\pm$
gluino	1/2	−1	$\tilde{g}$	(same)
goldstino (gravitino)	1/2 (3/2)	−1	$\tilde{G}$	(same)

Table 2.4: Particle content of the MSSM (with sfermion mixing for the first two families assumed to be negligible) [18].

Since it is almost impossible to search for SUSY with such an amount of parameters, a phenomenological MSSM is introduced (pMSSM). The characteristics of this new model is that there is no new source of CP violation considered, no flavour changing neutral currents (FCNC) allowed and a first and second generation universality is assumed. Consequently, there are 19 input parameters left on top of the parameters of the Standard Model:

- $\tan\beta$: ratio of the vacuum expectation values (vev) of the two Higgs doublets,
- m_A: mass of pseudoscalar Higgs boson,
- μ: higgsino mixing mass,
- M_1, M_2, M_3: bino, wino & gluino masses,
- 10 sfermion masses (5 for 1$^{\text{st}}$ & 2$^{\text{nd}}$ generations and 5 for 3$^{\text{rd}}$)
- 3 third generation trilinear couplings ($\mathrm{A_b}$,$\mathrm{A_t}$,A_τ).

This model represents the basis for SUSY searches but there might be other assumptions added in specific searches. Since the 19 parameter space is still very complex, it is preferred to predominantly study so-called *simplified models* which are easier to interpret. These simplified models involve only a couple of new particles and interactions by considering one production process and the corresponding decay chains. The masses of all other particles are set such that the available phase-space is not sufficient to produce them. In the framework of simplified models the experimental analyses can easier develop search strategies which are also applicable in the context of pMSSM.

3 Physics at Hadron Colliders

In contrast to e^+-e^- colliders, hadron colliders have no well-defined initial state and the substructure of the colliding hadrons has to be taken into account.

3.1 Parton Distribution Function

Since protons are composite particles, each constituent (also called parton) carries a certain fraction of the proton momentum and hence any parton can interact in a proton-proton collision. At this point the so called Parton Distribution Function (PDF) comes into play. It describes the momentum distribution of a parton i inside a proton at a given energy scale Q^2 and will be denoted as $f_i(x_1, Q^2)$ with x_1 being the momentum fraction of the proton. The resulting distribution is shown in figure 3.1. Due to the factorisation theorem it is possible to measure the PDF at a certain energy scale and extrapolate to different energies. Importantly, the PDFs have to be measured and cannot be predicted by theory.
So, the cross section of a proton-proton collision can be written as:

$$\sigma = \sum_{i,j} \int_0^1 \mathrm{d}x_1 \mathrm{d}x_2 f_i(x_1, Q^2) f_j(x_2, Q^2) \hat{\sigma}_{i,j}, \tag{3.1}$$

with $\hat{\sigma}_{i,j}$ being the partonic cross section which can be perturbatively calculated.

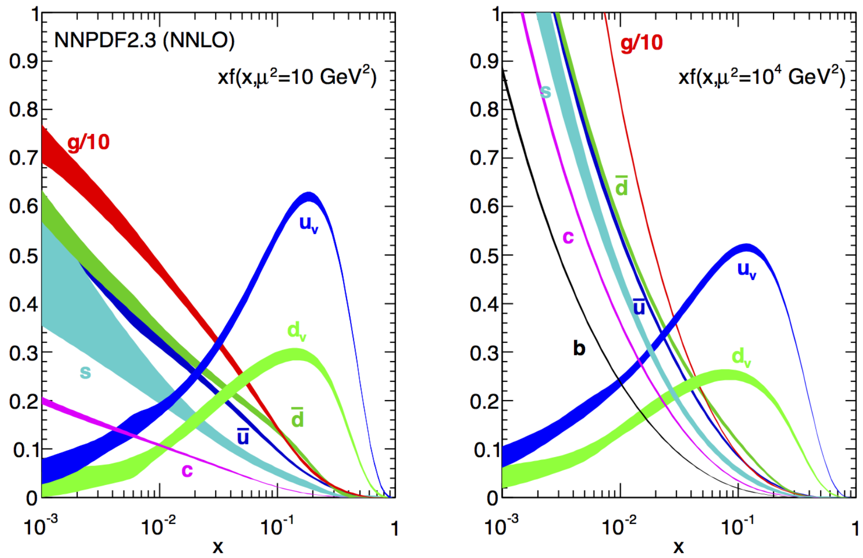
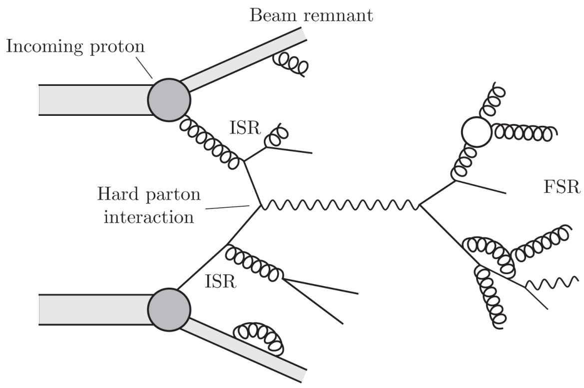


Figure 3.1: The NNPDF2.3 parton distribution functions of gluons, valence and sea quarks inside the proton at NNLO at two different scales [19].

3.2 Monte Carlo Simulation

Apart from the cross section calculations, several different phenomena have to be considered for proton-proton collisions e.g. parton showers, hadronisation and underlying events. However, it is not possible to treat them analytically (with some exceptions e.g. matrix elements). Therefore one uses the mathematical method Monte Carlo (MC) which is based on a random event generation.

The operating principle of MC generators is shown in figure 3.2. First the initial state composition and substructure is simulated as well as initial state radiations of gluons and photons. Afterwards the hard process and resonance decays are considered. Next, final state radiation and accompanying semi-hard processes are taken into account. And finally the hadronisation and further decays are simulated.



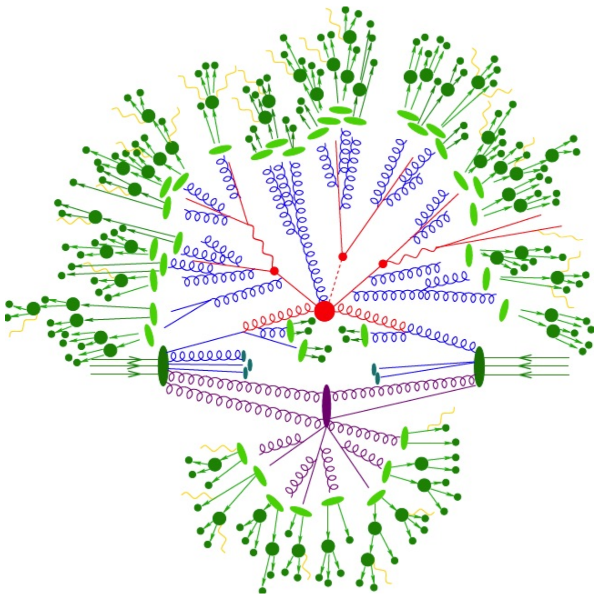


Figure 3.2: Left: Schematic overview of a pp-collision [20]. Right: MC simulation of a pp-interaction. The hard scattering process is simulated by matrix element generators shown in red. Parton showers are highlighted in the blue lines. Any secondary interactions - or rather underlying events - are marked in purple. The hadronisation is shown in green and hadron decays in dark green, the QED bremsstrahlung in yellow [21].

Mostly the generators are divided into two parts: matrix element generator and a parton shower and hadronisation modelling. There exist full MC generators which involve matrix element generation, parton shower, underlying event and fragmentation simulations like Pythia and Herwig. Those use mainly leading order matrix elements which is not sufficient. However, there exist many matrix element generators with higher order corrections which can be interfaced with the full MC generators.

The MC generator setups used for this thesis are summarised in table 3.1 with the corresponding order of calculation, PDF, parton showering (PS) tunes and generator versions. The used signals are direct gluino and squark production generated with the matrix element event generator MG5_aMC@NLO [22] and interfaced with Pythia8 [24]. Moreover the events for the W, Z and diboson backgrounds were simulated with the MC generator Sherpa [26].

Physics process	Generator	Cross section normalisation	PDF set	Parton shower	Tune
SUSY processes	MG5_aMC@NLO 2.2.2(3) [22]	NLO+NLL	NNPDF2.3LO [23]	Pythia 8.186 [24]	A14 [25]
$W(\to \ell\nu)$+jets	Sherpa 2.2.1 [26]	NNLO	NNPDF3.0NNLO	Sherpa	Sherpa default
$Z/\gamma^*(\to \ell\nu)$+jets	Sherpa 2.2.1	NNLO	NNPDF3.0NNLO	Sherpa	Sherpa default
$t\bar{t}$	Powheg-Box v2 [27]	NNLO+NNLL	CT10	Pythia 6.428 [28]	Perugia2012 [29]
Single top (Wt-channel)	Powheg-Box v1	NNLO+NNLL	CT10	Pythia 6.428	Perugia2012
Single top (s-channel)	Powheg-Box v1	NLO	CT10	Pythia 6.428	Perugia2012
Single top (t-channel)	Powheg-Box v1	NLO	CT10f4 [30]	Pythia 6.428	Perugia2012
Single top (Zt-channel)	MG5_aMC@LO 2.2.1	LO	CTEQ6L1 [31]	Pythia 6.428	Perugia2012
$t\bar{t} + W/Z/WW$	MadGraph5	NLO	NNPDF2.3LO	Pythia 8.186	A14
WW, WZ, ZZ	Sherpa 2.1.1	NLO	CT10	Sherpa	Sherpa default
Multi-jet	Pythia 8.186	LO	NNPDF2.3LO	Pythia 8.186	A14

Table 3.1: List of Monte Carlo generators, cross section normalisations, PDFs, Parton shower generators and tuning parameters used in the SUSY 0-lepton analysis for SM background and SUSY signals [32].

Additionally, mainly Powheg-Box [27] was used to simulate the SM background from top quarks. Last, the Multi-jet or QCD background samples were produced with Pythia.

4 The ATLAS Experiment at the Large Hadron Collider

The search for new physics in the field of particle physics requires higher and higher energy scales. In order to achieve such energy ranges, complex experiments are necessary like the Large Hadron Collider (LHC) at CERN. One of the big multi-purpose experiment located at the LHC is the ATLAS experiment [33]. The following chapter will give a brief overview of the LHC, ATLAS and the corresponding infrastructure.

4.1 CERN and the Large Hadron Collider

The *Centre Européene de la Recherche Nucléaire* (CERN) was one of the first joint research organisations in Europe founded in 1954. Today, it is one of the largest institutions for high energy physics. It provides the infrastructure for a wide variety of different experiments.

CERN has an outstanding accelerator complex as shown in figure 4.1. This large complex grew historically beginning with the Synchrocyclotron in 1957 and the Proton Synchrotron (PS) in 1959 which is still in use. The Large Hadron Collider is currently the largest accelerator. As is shown in figure 4.1, the LHC depends on a pre-accelerator chain, starting at the LINAC where the protons are being produced from hydrogen atoms. They are then further accelerated over the Booster, PS and Super Proton Synchrotron, before entering the LHC with an energy of 450 GeV.
The LHC is a circular particle accelerator with a circumference of 27 km. It is designed to collide protons with a centre-of-mass energy up to $\sqrt{s} = 14\,\text{TeV}$, though it is currently running with $\sqrt{s} = 13\,\text{TeV}$. The protons are grouped into bunches and the LHC design allows to collide them every 25 ns. In addition, heavy ion interactions are studied using Pb-Pb or p-Pb collisions.
There are four main experiments situated around the LHC ring as indicated in figure 4.1:

- ATLAS and CMS [35] the two large general-purpose detectors,
- ALICE [36], which is designed to study quark gluon plasmas from heavy ion collisions and
- LHCb [37], which is specialised on B-physics.

Besides the centre-of-mass energy, the delivered statistics is one of the main features characterising a collider. The number of scattering events N in the LHC for a specific physics process is given by the scattering cross section σ and the integrated luminosity L

$$N = \sigma \cdot \int \mathcal{L}\ \mathrm{d}t = \sigma \cdot L. \tag{4.1}$$

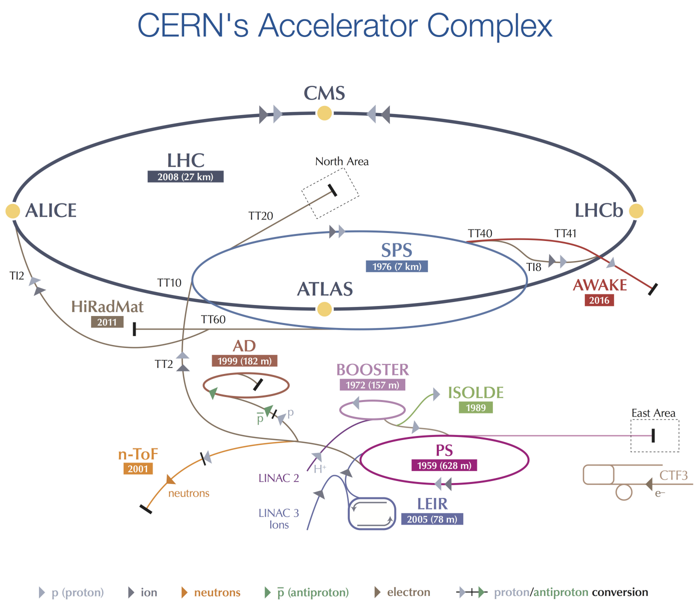


Figure 4.1: Accelerator complex and LHC detectors at CERN [34].

The luminosity reflects the experimental parameters of a collider and is given for two colliding beams as

$$\mathcal{L} = \frac{N_1 \cdot N_2 \cdot n_b \cdot f_{\text{rev}}}{4\pi\sigma_x\sigma_y}, \tag{4.2}$$

with N_i number of particles per bunch, n_b number of bunches, f_{rev} the revolution frequency and $\sigma_{x,y}$ the vertical and horizontal beam width assuming Gaussian beam shapes. While the cross section is measured in units of an area $1\,\text{b} = 10^{-24}\,\text{cm}^2$, the unit of the integrated luminosity is given as the inverse of an area.
The performance of the LHC is increasing with its runtime as illustrated in the right plot of figure 4.2 showing the ATLAS online luminosity of each year since 2011 with different centre-of-mass energies. Moreover the left plot in figure 4.2 shows the evolution of the integrated luminosity for $\sqrt{s} = 13\,\text{TeV}$ whereas the important number is the achieved luminosity *good for physics* which can be used for physics analyses yielding $80\,\text{fb}^{-1}$ by end of 2017.

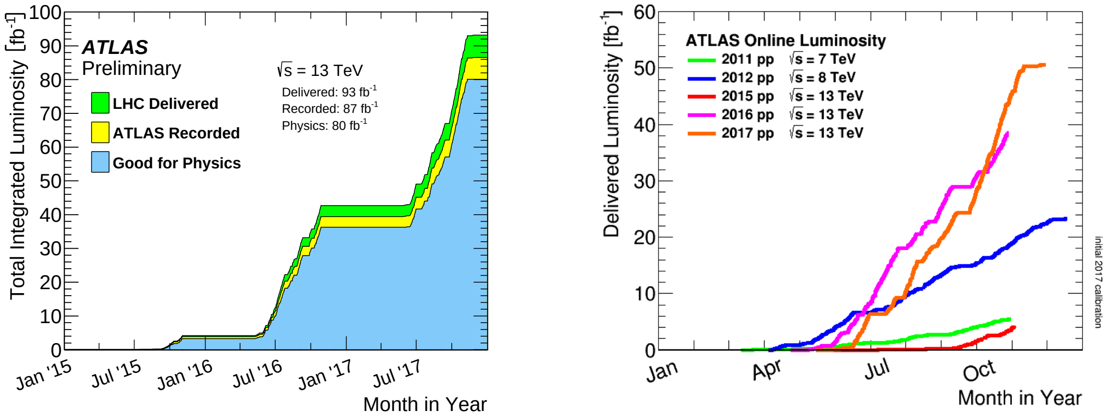


Figure 4.2: Integrated luminosity of the LHC collected by the ATLAS experiment. The left plot shows the evolution of the luminosity collected with $\sqrt{s} = 13\,\text{TeV}$ in the years 2015 to 2017. The online luminosity of ATLAS is shown on the right separated in the different physics runs of each year [38].

4.2 The ATLAS Experiment

ATLAS[1] is one of the two general-purpose particle detectors located at the LHC. The ATLAS collaboration represents with its 3000 scientific authors one of the largest collaborative efforts in science. After a long planning and construction period, starting in 1992, ATLAS took its first physics data in 2009. The detector is designed such that it can cover a wide range of physics, from probing the Standard Model to searches for new physics. Clearly, it is a big challenge to design such a detector and a complex hardware and software system is necessary. The detector is located 100 m below ground near Meyrin, Switzerland. The experiment is 46 m long, 25 m in diameter and has a weight of 7000 t.
A schematic overview of the ATLAS detector is shown in figure 4.3 with its four cylindrical structured main subsystems. The purpose of these systems is to record the path, momentum and energy of the particles. In order to measure the momentum, powerful magnets are necessary. Especially the toroid magnets are a prominent mark of the ATLAS detector. In order to deal with higher luminosities of the planned LHC upgrade in 2025 (Phase-II Upgrade) the detector is progressively upgraded as well. Below, the individual detector components are described in more detail.

The coordinate system used for the ATLAS detector is a right-handed one with its origin at the nominal interaction point. The z-axis points along the beam pipe, the y-axis points vertically upwards and the x-axis points towards the middle of the LHC ring. The azimuthal angle ϕ is measured from the x-axis in the plane perpendicular to the beam while the polar angle θ is measured from the beam axis. Though, instead of the polar angle θ one uses mainly the pseudorapidity defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{4.3}$$

it is a high energy approximation of the rapidity. The advantage of the rapidity is that its

[1] **A** **T**oroidal **L**HC **A**paratu**S**

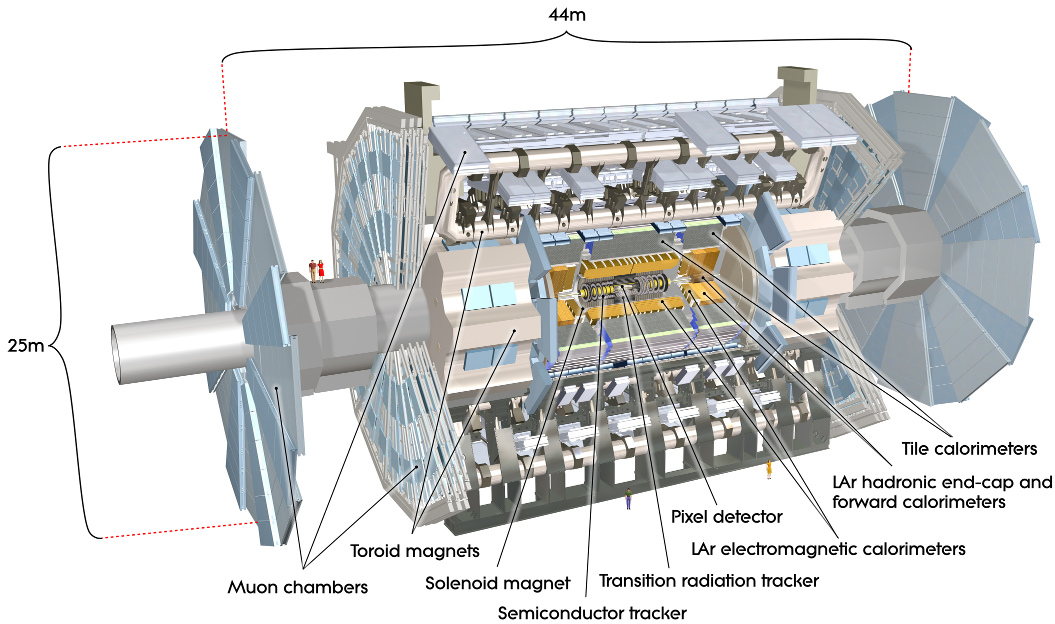


Figure 4.3: Schematic overview of the ATLAS detector [39].

difference is Lorentz invariant. Especially for high relativistic objects in high energy physics, calculations can be simplified this way.
The distances in the detector are typically calculated via

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{4.4}$$

Inner Detector

The inner detector is the innermost detector module and encloses directly the interaction point. The layout of this detector is illustrated in figure 4.4 and shows a separation into three subsystems namely the pixel detector, the semi-conductor tracker (SCT) and the transition radiation tracker (TRT).

The pixel detector has four layers: three cylindrical silicon sensor layers and the insertable B-layer (IBL) [41]. They have a very good spatial resolution of 10 μm to 115 μm, are 50.5 mm-122.5 mm away from the interaction point and cover a pseudorapidity range of $|\eta| < 2.5$.
The SCT is arranged in concentric cylinders around the pixel detector and has four double layers of long narrow silicon strips instead of pixels. This design provides a very good spatial resolution of around 70 μm in the (r, ϕ)-plane and about 580 μm in z-direction. Hence it contributes significantly to the momentum reconstruction by providing accurate track information.
The outer part of the inner detector is the TRT consisting of straw tubes with the purpose to measure photons from transition radiation yielding information for the electron identification. However, the TRT will be replaced in the Phase-II Upgrade by an all silicon tracker.

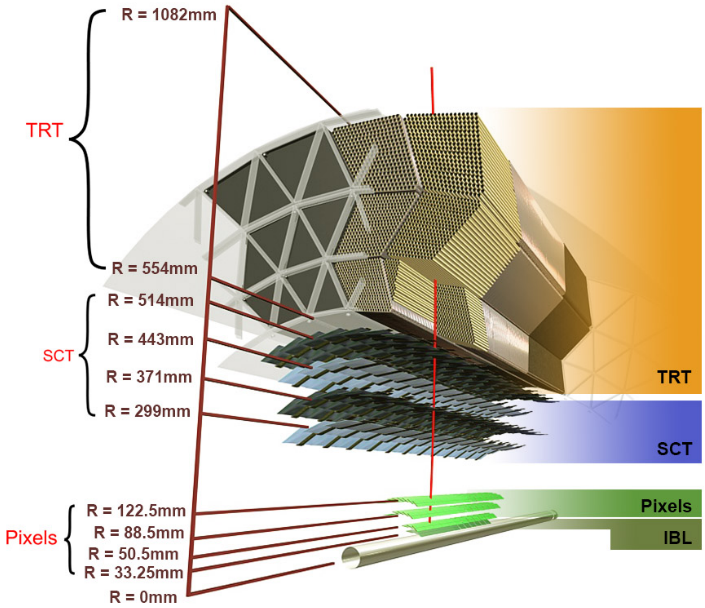


Figure 4.4: Overview of the inner detector of the ATLAS experiment divided into three sub-detectors (Pixels, SCT and TRT) [40].

Calorimeter System

The calorimeter system is responsible for the measurement of particle energies by absorbing them as well as measuring the shower properties to allow for particle identification. The calorimetry is mainly divided into two subsystems, the electromagnetic and the hadronic calorimeter. Both are sampling calorimeters which means that they have alternating layers of active materials and absorbers as indicated in the schematic cut-away view in figure 4.5. In the ATLAS case mainly liquid argon is used as active material except in the hadronic tile calorimeter where plastic scintillators are utilised. Furthermore, copper, iron, tungsten and lead are used as absorbers in the different parts of the calorimeter (see figure 4.5). The whole calorimeter system covers an η range up to a far forward region $|\eta| < 4.9$, the exact values for each part are shown in figure 4.5.

The electromagnetic calorimeter has a barrel and a end-cap module in which charged particles and photons are detected with the charged particles interacting via bremsstrahlung and undergo subsequently pair production. An important characteristic of a calorimeter is its energy resolution, in the EM calorimeter case $\sigma_E/E = 10\%/\sqrt{E/\text{GeV}} \oplus 0.7\%$.
As the name implies the hadronic calorimeter measures the energy deposition of hadrons and surrounds the EM calorimeter. The resolution of the hadronic tile and the hadronic end-cap is $\sigma_E/E = 50\%/\sqrt{E/\text{GeV}} \oplus 3\%$ while the resolution of the forward calorimeter is worse with $\sigma_E/E = 100\%/\sqrt{E/\text{GeV}} \oplus 10\%$.
Even though the calorimeters stop most of the energetic particles, only a fraction of their energy is deposited in the active material. To compensate for this and other effect a calorimeter calibration is performed to better estimate the full energy deposition.

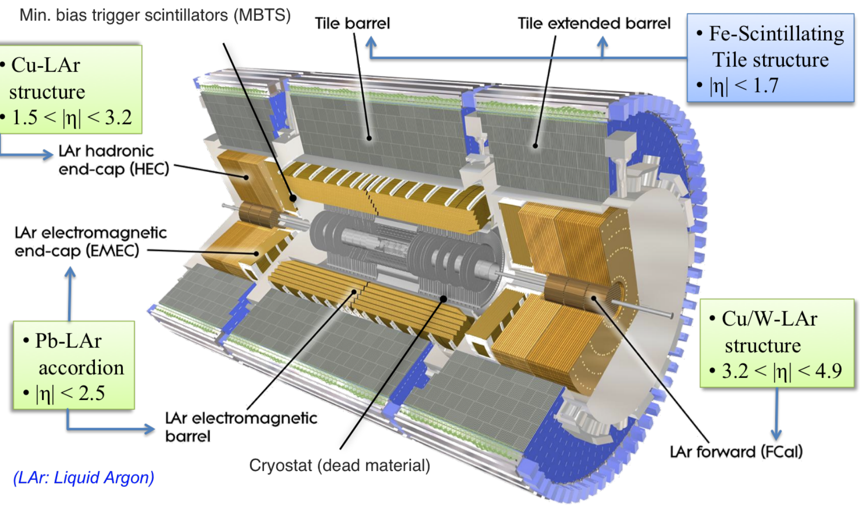


Figure 4.5: Cut-away view of the calorimeter system of the ATLAS experiment [42].

Muon Spectrometer

The outermost part of the ATLAS experiment is the muon system. It measures the coordinates, charge and momenta of muon tracks. To improve the momentum and charge measurement the muon system is embedded in a toroidal magnetic field. In figure 4.6 one can see the different components of the muon spectrometer.

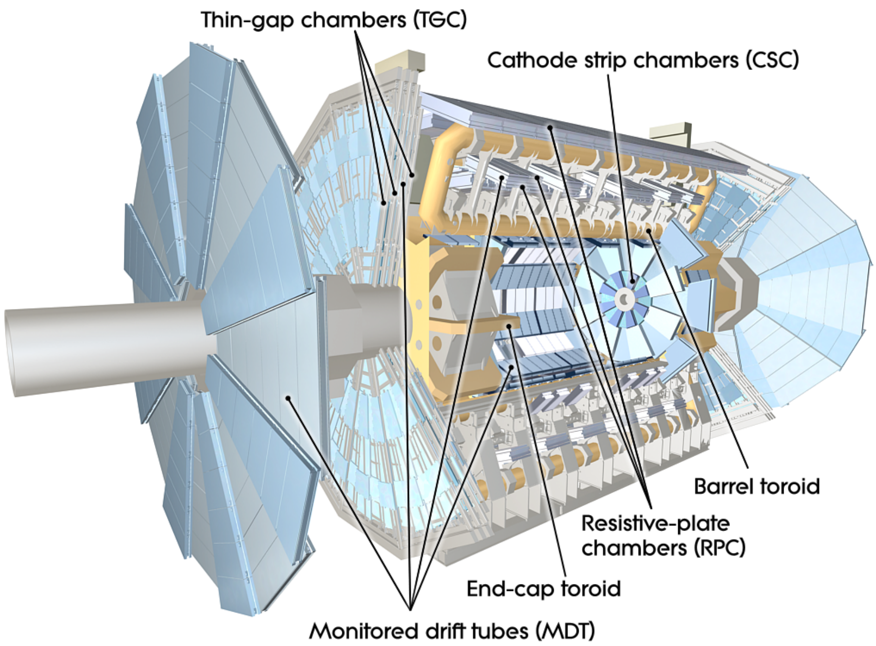


Figure 4.6: Layout of the ATLAS muon spectrometer [43].

Each part has a different technology. First, the thin gap chambers (TGC) are responsible

for triggering and the second coordinate measurement at the ends of the detector. Secondly, the resistive plate chambers (RPC) with the same tasks as the TGC but in the barrel region. Thirdly, the monitored drift tubes (MDT) which measure the curves of the tracks with a resolution of 80 μm. Last, the cathode strip chambers (CSC) which measure the precision coordinates in the forward region with a resolution of 60 μm.
The momentum resolution of the muon spectrometer is $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 2 - 4\%$ for the p_T-range 10-200 GeV and $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 10\%$ for 1 TeV.

Magnet system

As mentioned above, the magnets are needed to bend the trajectory of the charged particles to measure their momenta. There are two systems installed. First, the central solenoid magnet responsible for the reconstruction of low energy and fast decaying particles in the tracking section. It is located between the inner detector and the calorimetry. This magnet system is super-conductive and can achieve an axial magnetic field up to 2 T.
Secondly, the three super-conductive air-core toroid magnets consisting of eight coils each are located at the outermost part of the detector. They can generate a magnetic field up to 3.9 T in the barrel region and up to 4.1 T in the end-caps.

Trigger System and Data Acquisition

The enormous data flow provided by the LHC with a design frequency for proton-proton collisions of 40 MHz corresponds to roughly 52 TB/s of data which the ATLAS experiment would record. Undoubtedly, it is not possible to store all these data and thus an advanced trigger system is necessary to reduce the fraction written to tape.
The trigger system redesigned for Run II consists of a hardware-based first level trigger (Level-1) and a software-based high level trigger (HLT) shown in figure 4.7.
The Level-1 trigger consists of electronics responsible to figure out the region of interest based on the coarse granularity calorimeter and the muon detector information within a decision time of 2.5 μs. After this first trigger step the resulting rate is about 100 kHz considering the current bunch-crossing frequency of the LHC of around 30 MHz.
Next, the HLT uses sophisticated selection algorithms including the full granularity detector information in either the region of interest or the whole detector to further reduce the amount of data. This happens within 200 ms reducing the rate to 1 kHz.

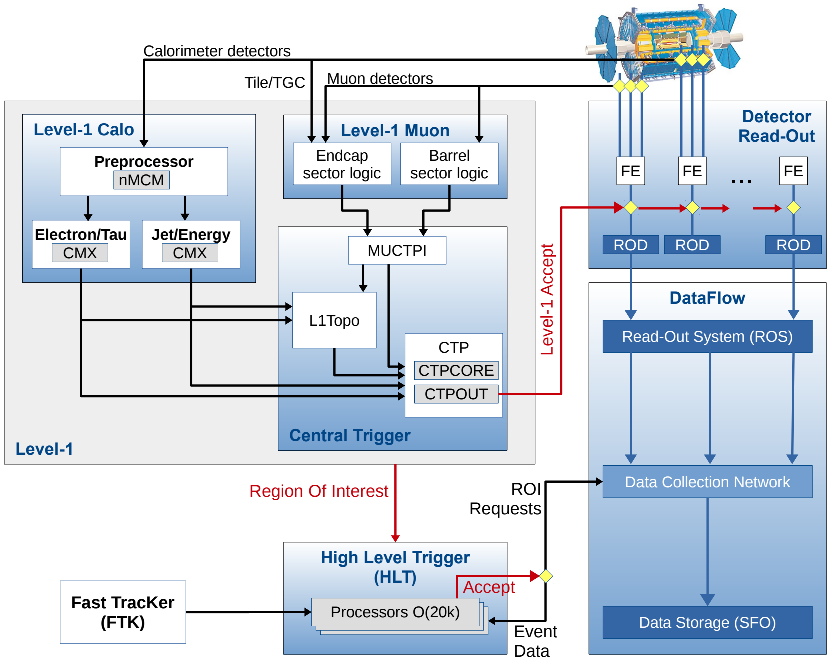


Figure 4.7: Schematic view of the ATLAS trigger and data acquisition system in Run II [44].

5 The SUSY 0-lepton Analysis

Since there are many models for supersymmetry with a large parameter space, the SUSY analyses are divided in several subanalyses defined by the decay topology of the particles under study. One of these analyses is the 0-lepton analysis which covers the SUSY processes as shown in figure 5.1.

This analysis focuses on the search for squarks and gluinos, in exclusively hadronic final states containing hadronic jets, missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ and no leptons. The analysed data was recorded in 2015 and 2016 by the ATLAS experiment at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ yielding an integrated luminosity of $36.1\,\mathrm{fb}^{-1}$. The analysis results were presented at Moriond 2017 [3] and were also published in a paper [32].

If SUSY is present at the TeV scale, squarks and gluinos may be produced at the LHC. The focus of this search is the strong production since its cross section is several orders of magnitude higher than in the electroweak case. If R-parity is conserved, the squarks and gluinos are always pair produced at the LHC and can either decay directly into SM particles and the LSP (see figure 5.1 (a,d)) which is assumed to be the lightest neutralino $\tilde{\chi}_1^0$, or via cascades (figure 5.1 (b,c,e,f,g)).

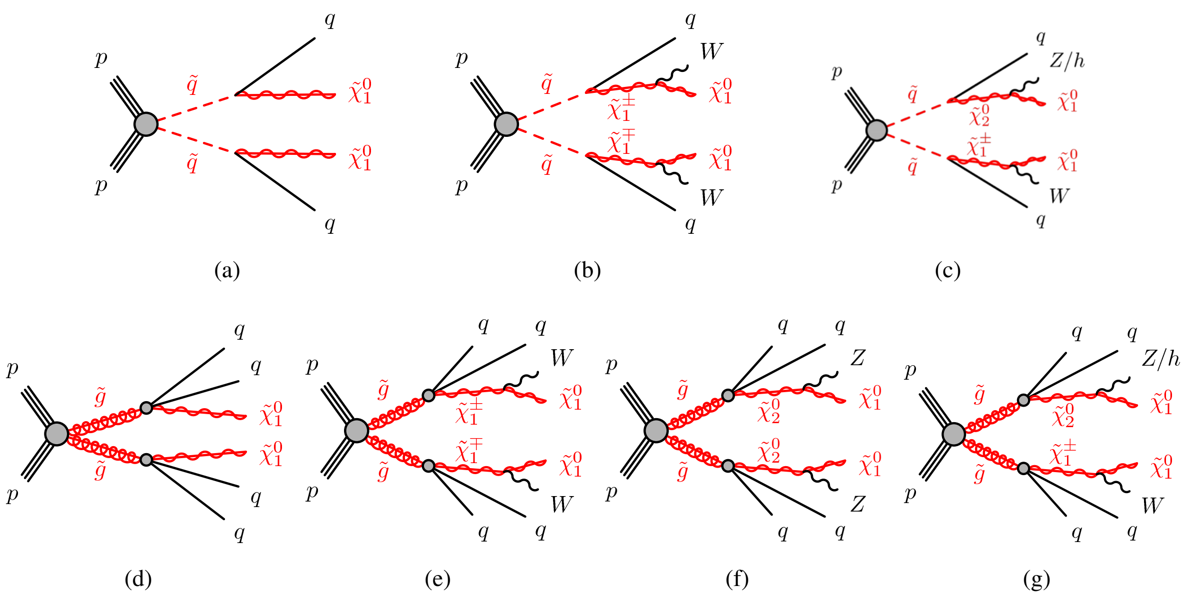


Figure 5.1: The decay topologies of (a,b,c) squark-pair production and (d, e, f, g) gluino-pair production, in the simplified models with (a) direct or (b,c) one-step decays of squarks and (d) direct or (e, f, g) one-step decays of gluinos [3].

The final states of these decays show the following characteristics in the detector:

- at least two jets, originating from quarks or W, Z, Higgs decays,
- missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$), coming from the LSP and W,Z decay,
- and leptons from leptonic W, Z decays.

As mentioned above, this analysis takes only fully hadronic final states and includes a veto on identified leptons to suppress SM background processes.

The basic analysis strategy is to first define signal regions (SRs) to maximise the expected signal yield relative to the SM background for the considered model. This is done using Monte-Carlo (MC) samples of SUSY signals and SM background processes.
The most important backgrounds are Z+jets, W+jets, Top ($\mathrm{t\bar{t}}$ and single top), diboson and multi-jet production. The $Z(\rightarrow \nu\nu)$+jets background is an irreducible background since it has the same signature in the detector as the signal. In contrast to this, the W+jets and Top background is mainly originating from unidentified leptons whereby the W+jets background is dominated by $\mathrm{W} \rightarrow \tau\bar{\nu}$ production. The multi-jet background originates principally from misreconstructions of jet energies.
In order to estimate the SM background for each signal region, four control regions (CRs) are defined for W, Top, Z and multi-jet. They are defined such that they are orthogonal to the signal regions to provide independent data samples enriched in particular background sources. The results from the CRs are converted via likelihood fits into background estimates for the SRs. In the limit where the background purity is 100%, the estimated background is given by:

$$N_{\mathrm{SR}}^{\mathrm{bkg,\ pred}} = N_{\mathrm{SR}}^{\mathrm{bkg,\ MC}} \frac{N_{\mathrm{CR}}^{\mathrm{data}}}{N_{\mathrm{CR}}^{\mathrm{bkg,\ MC}}}, \tag{5.1}$$

where the Ns are the number of events in the specified regions. Additionally, to cross check the background estimations, validation regions (VRs) are defined. They are chosen such that they have no overlap with the SRs and only a small overlap with the CRs.

Since there is a large variety of squark and gluino production signal processes, several SRs are considered to ensure a high sensitivity for each of them. There are two techniques to chose the selection requirements for each SR, first the m_{eff}-based and secondly the RJR-based method. They represent two complementary approaches. Both are described in more detail below.

The important variables for the current analysis are presented in the following. In general only jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $|\eta| < 2.8$ are considered. The jets are clustered using the anti-k_T algorithm [45] with the jet radius parameter 0.4.

$\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$ the missing transverse momentum is defined as

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{jet}} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{electron}} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{photon}} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{muon}} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{softterm}}. \tag{5.2}$$

H_T is defined as the scalar sum of the jet momenta

$$H_\mathrm{T} = \sum_i^n |\boldsymbol{p}_\mathrm{T}^{(i)}| \tag{5.3}$$

and is a measure for the jet activity in an event. The index n runs in general over all jets in an event (general cut of $p_\mathrm{T} > 50\,\mathrm{GeV}$) and in this case H_T is sometimes quoted as H_T(incl.). Also, if n runs only up to a specific number of jets it is indicated as $H_\mathrm{T}(N_j)$.

Effective mass is the scalar sum of the transverse jet momenta and the $E_\mathrm{T}^\mathrm{miss}$

$$m_\mathrm{eff} = H_\mathrm{T} + E_\mathrm{T}^\mathrm{miss}. \tag{5.4}$$

Again, the naming depends on the number of considered jets in the calculation of H_T. So for the sum over all jets in an event one uses m_eff(incl.) and for a defined number of jets $m_\mathrm{eff}(N_j)$.

$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$ also called dPhi measures the alignment of jets and the $E_\mathrm{T}^\mathrm{miss}$. It is a powerful variable to reject fake $E_\mathrm{T}^\mathrm{miss}$ due to jet mismeasurements or neutrinos in jets which can be the case e.g. in jets with semi-leptonic decays of heavy flavour hadrons.

$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$ or dPhiR is complementary to $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$ but describes the alignment for more than three jets.

$E_\mathrm{T}^\mathrm{miss}/\sqrt{H_\mathrm{T}}$ is mainly used in models with $\tilde{q}\tilde{q}$ production contributing to an enhanced sensitivity and sometimes also called metSig.

$E_\mathrm{T}^\mathrm{miss}/m_\mathrm{eff}(N_j)$ has an additional reduction power after the $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$ cut in higher jet multiplicities ($N_j \geq 4$).

Aplanarity is an event shape variable defined as $Ap = 3/2\lambda_3$, with λ_3 the smallest eigenvalue of the normalised momentum tensor of the jets. It is used to distinguish spherical events from planar or linear events.

In addition, some further variable definitions are shown which are not used in the current analysis but they will be used in the studies performed in the following chapters.

Eta is a simplified variable and defined as

$$\mathrm{Eta} = \max(|\eta|_{i=1,2}). \tag{5.5}$$

EtaR is similar to Eta but for a higher jet multiplicity

$$\mathrm{EtaR} = \max(|\eta|_{i=1-4}). \tag{5.6}$$

$N_\mathrm{trk}^{(i)}$ is the number of tracks in a jet. It is a useful variable to distinguish gluon and quark jets since gluon jets have usually a larger number of tracks than quark jets (background: gluon dominated, signal: quark dominated).

N_{trk} combines several jets

$$N_{\text{trk}} = \sum_{i=1}^{n} N_{\text{trk}}^{(i)}, \tag{5.7}$$

where n is typically equals to 4, if not it will be indicated.

The m_{eff}-based method uses the effective mass $m_{\text{eff}}(\text{incl.})$ as main discriminating variable motivated by the expected high mass scale of the considered SUSY model. Afterwards the signal regions are optimised using mainly $m_{\text{eff}}(\text{incl.})$ and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$. In addition, other criteria based on different variables such as the jet momenta, pseudorapidity, $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ are optimised and detector based and cleaning cuts are applied, all with the goal to get the best signal yield in each SR. Since the optimisation is only based on applying different cuts, this method is denoted as Cut & Count. A summary of the SR definitions are shown in table 5.1 for the direct decays of squarks and in table 5.2 for direct gluino decays. The aspiration of the work in this thesis is to find new approaches to improve the SR definitions.

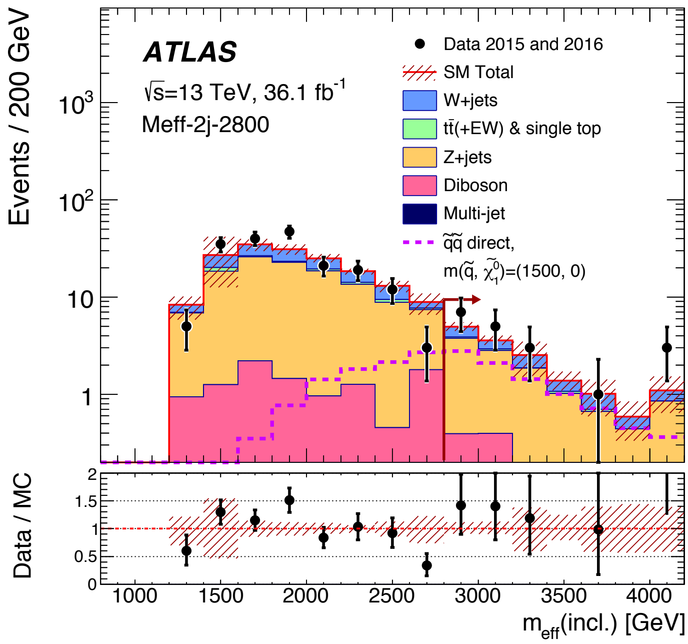


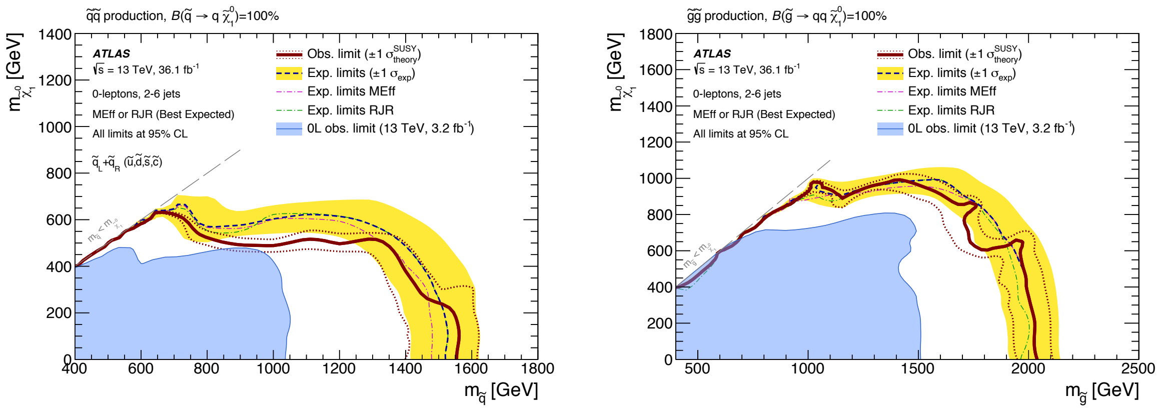
Figure 5.2: The $m_{\text{eff}}(\text{incl.})$ distribution in the Meff-2j-2800 signal region with the signal point $m(\tilde{q}, \tilde{\chi}_1^0) = (1500, 0)$ GeV [32].

The second method is the Recursive Jigsaw Reconstruction technique (RJR) which is based on defining kinematic variables on an event-by-event level. By assuming some kinematic conditions the RJR uses a different reference frame to define new variables to discriminate between signal and background.

Besides the statistical uncertainties there are several systematic uncertainties to be considered. The main source of of systematic uncertainties arises from the extrapolation factors relating observations in the CRs to background expectations in the SRs. A large fraction of these systematics originate from uncertainties on the theoretical modelling of the background processes, the jet energy scale (JES) and jet energy resolution (JER) uncertainties and statistical uncertainties of the CRs.

In the plot of figure 5.2 the Meff-2j-2800 signal region is shown. The name of this region reads as the following: Meff means that the m_{eff}-based method is used, 2j implies that at least 2 jets are required and 2800 represents the applied m_{eff}(incl.) cut. The signal shown in the plot assumes a squark mass of 1500 GeV and a neutralino mass of 0 GeV. It is nicely visible that the m_{eff}(incl.) is a good discriminant variable, it peaks around 2800 GeV in the signal while the backgrounds falls continuously.

In the presented results there is no excess occurring in any signal region. Therefore, limits are set to 95% confidence level (CL) by comparing the observed and expected event yields. The resulting limits are presented in the contour plots of figure 5.3 as a function of the squark mass (left) and the gluino mass (right).



(a) (b)

Figure 5.3: Exclusion limits for direct production of (a) first- and second-generation squark pairs with decoupled gluinos and (b) gluino pairs with decoupled squarks. Gluinos (first- and second-generation squarks) are required to decay to two quarks (one quark) and a neutralino LSP. Exclusion limits are obtained by using the signal region with the best expected sensitivity at each point. The blue dashed lines show the expected limits at 95% CL, with the light (yellow) bands indicating the 1σ excursions due to experimental and background-only theoretical uncertainties. Observed limits are indicated by brown curves where the solid contour represents the nominal limit, and the dotted lines are obtained by varying the signal cross-section by the renormalisation and factorisation scale and PDF uncertainties. Results are compared with the observed limits obtained by the previous ATLAS searches with jets, missing transverse momentum, and no leptons [32].

All the results are interpreted in terms of simplified models only considering light-flavour squarks or gluinos and the neutralino as LSP, all other sparticles are considered to be too heavy for LHC physics. Finally several mass regions for squarks, gluinos and neutralinos are excluded at the 95% confidence level which allows a further exclusion of the parameter region compared to previous results.

<table>
<tr><td>Targeted signal</td><td colspan="8">$\tilde{q}\tilde{q}, \tilde{q} \to q\tilde{\chi}_1^0$</td></tr>
<tr><td rowspan="2">Requirement</td><td colspan="8">Signal Region [Meff-]</td></tr>
<tr><td>2j-1200</td><td>2j-1600</td><td>2j-2000</td><td>2j-2400</td><td>2j-2800</td><td>2j-3600</td><td>2j-2100</td><td>3j-1300</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV] ></td><td colspan="8">250</td></tr>
<tr><td>$p_{\mathrm{T}}(j_1)$ [GeV] ></td><td>250</td><td>300</td><td colspan="4">350</td><td>600</td><td>700</td></tr>
<tr><td>$p_{\mathrm{T}}(j_2)$ [GeV] ></td><td>250</td><td>300</td><td colspan="4">350</td><td colspan="2">50</td></tr>
<tr><td>$p_{\mathrm{T}}(j_3)$ [GeV] ></td><td colspan="7">–</td><td>50</td></tr>
<tr><td>$\lvert\eta(j_{1,2})\rvert$ <</td><td>0.8</td><td colspan="4">1.2</td><td colspan="3">–</td></tr>
<tr><td>$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ ></td><td colspan="6">0.8</td><td colspan="2">0.4</td></tr>
<tr><td>$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ ></td><td colspan="6">0.4</td><td colspan="2">0.2</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ [GeV$^{1/2}$] ></td><td>14</td><td colspan="5">18</td><td>26</td><td>16</td></tr>
<tr><td>m_{eff}(incl.) [GeV] ></td><td>1200</td><td>1600</td><td>2000</td><td>2400</td><td>2800</td><td>3600</td><td>2100</td><td>1300</td></tr>
</table>

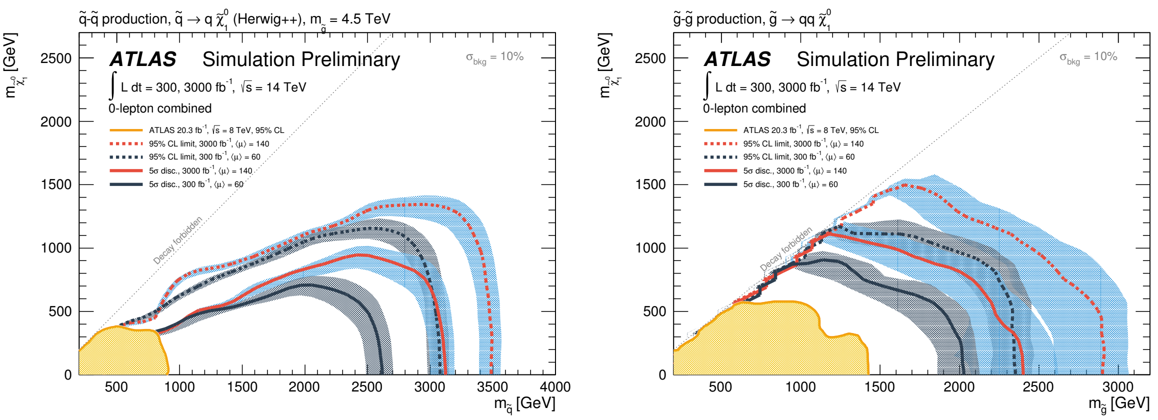
Table 5.1: Signal region definitions for the direct squark decay [32].

<table>
<tr><td>Targeted signal</td><td colspan="7">$\tilde{g}\tilde{g}, \tilde{g} \to q\bar{q}\tilde{\chi}_1^0$</td></tr>
<tr><td rowspan="2">Requirement</td><td colspan="7">Signal Region [Meff-]</td></tr>
<tr><td>4j-1000</td><td>4j-1400</td><td>4j-1800</td><td>4j-2200</td><td>4j-2600</td><td>4j-3000</td><td>5j-1700</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV] ></td><td colspan="7">250</td></tr>
<tr><td>$p_{\mathrm{T}}(j_1)$ [GeV] ></td><td colspan="6">200</td><td>700</td></tr>
<tr><td>$p_{\mathrm{T}}(j_4)$ [GeV] ></td><td colspan="4">100</td><td colspan="2">150</td><td>50</td></tr>
<tr><td>$p_{\mathrm{T}}(j_5)$ [GeV] ></td><td colspan="6">–</td><td>50</td></tr>
<tr><td>$\lvert\eta(j_{1,2,3,4})\rvert$ <</td><td>1.2</td><td colspan="5">2.0</td><td>–</td></tr>
<tr><td>$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ ></td><td colspan="7">0.4</td></tr>
<tr><td>$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ ></td><td colspan="6">0.4</td><td>0.2</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(N_j)$ [GeV$^{1/2}$] ></td><td>0.3</td><td colspan="3">0.25</td><td colspan="2">0.2</td><td>0.3</td></tr>
<tr><td>Aplanarity ></td><td colspan="6">0.04</td><td>–</td></tr>
<tr><td>m_{eff}(incl.) [GeV] ></td><td>1000</td><td>1400</td><td>1800</td><td>2200</td><td>2600</td><td>3000</td><td>1700</td></tr>
</table>

Table 5.2: Signal region definitions for direct gluino decay [32].

6 Multivariate Classification Techniques

Especially in searches for new physics in very rare processes it is essential to select the signal process with a high efficiency. Since the datasets are getting increasingly large, more sophisticated methods are necessary to analyse the data in an efficient way. One way to improve the signal identification are via so-called Multivariate Analysis Techniques (MVA).
In SUSY prospects studies for RUN III, the improvement was studied which can be achieved with respect to RUN I data using Cut & Count techniques shown in figure 6.1. It turned out that there is only a moderate gain and therefore these sophisticated methods are necessary.
In the previous chapter only Cut & Count techniques were used for the analysis presented by the ATLAS collaboration. The following chapter will give a brief introduction to more sophisticated methods which will be applied in the studies done for this thesis. The TMVA framework (Toolkit for Multivariate Data Analysis with ROOT) [46] provides such a workaround for several multivariate classification methods in particle physics which have been significantly improved in the last few years. It is integrated in the CERN analysis framework ROOT.



a) b)

Figure 6.1: SUSY prospect study for higher luminosities ($300\,\mathrm{fb}^{-1}$ and $3000\,\mathrm{fb}^{-1}$) and $\sqrt{s} = 14\,\mathrm{TeV}$ for a) squark pair production and b) gluino pair production [47].

6.1 Cut & Count

The simplest approach to classify data is the Cut & Count method. One picks a certain amount of variables and iterates over different cut combinations to optimise the signal yield. Only binary decisions are possible and no complicated cut functions. In order to use complex correlations between variables it is necessary to define new variables to cut on. At this point more advanced MVA methods can help to find new variables.

6.2 Multi-Bin Approach

Improvements over the pure Cut & Count method can be achieved by slicing the dataset into several orthogonal bins and performing a Cut & Count optimisation in each bin (multi-bin approach). This approach is especially promising if signal and background yields change differently as a function of given variables.
In addition, it is possible to replace signal regions with multi-bins in some cases, e.g. if some signal points are more sensitive on tight cuts and others not, described in more detail in the performed studies (see section 7.4).
To avoid large statistical uncertainties per bin one should stick to a reasonable amount of bins.

6.3 Boosted Decision Tree

One of the most common MVA methods used in particle physics analysis is the boosted decision tree method (BDT). A schematic overview of the operating principle of BDT is shown in figure 6.2. At each node/classifier one single variable is taken and a binary decision is made (yes/no) this is repeated until a stop criterion is fulfilled. With this structure several correlations between different variables can be realised and the phase-space can be splitted in many regions which can be classified as background or signal at the end.

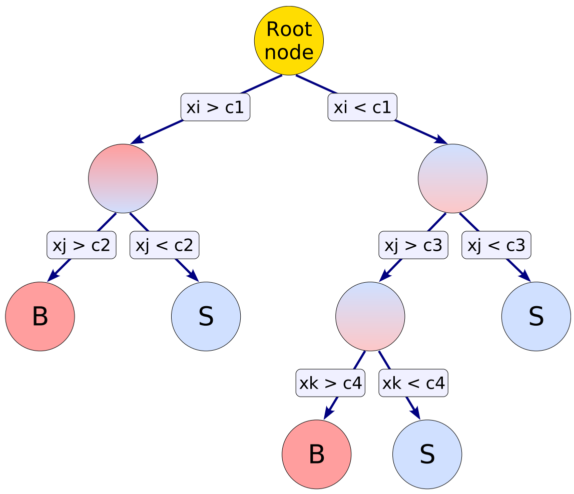


Figure 6.2: Schematic of BDT - binary decision tree [46].

In addition, a concept called *boosting* is applied which allows to move from one tree to several trees. Boosting allows to enhance the classification and regression performance of the method described above. It also increases the stability with respect to statistical fluctuations in the training sample. The idea is to sequentially apply the decision tree algorithm to reweighted (boosted) versions of the training data and afterwards to take a weighted majority vote of all the reweighted trained samples, they are also called a forest of trees. The two most commonly used boosting techniques are AdaBoost [48] and GradBoost [49], both will be applied in further studies.
Besides the boosting technique, there are other training parameters for the BDT, called hyper parameters. The most important hyper parameters are the following:

NTrees Number of trees in boosting.

MaxDepth Maximal allowed depth of the decision tree.

nCuts Number of grid points in variable range used in finding optimal cut in node splitting.

AdaBoostBeta Learning rate for AdaBoost algorithm and

Shrinkrate is the corresponding learning rate for GradBoost.

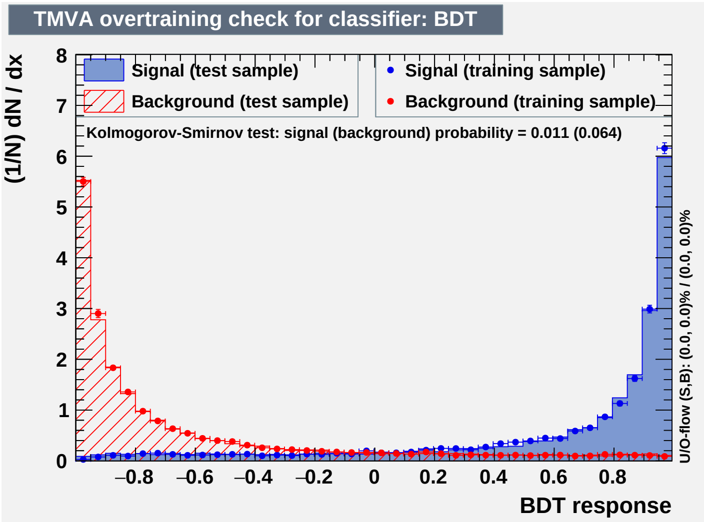


Figure 6.3: BDT score (response) distribution for an exemplary BDT training and testing produced with the TMVA GUI.

After the BDT is trained it can be applied it to an orthogonal testing sample and for each event a BDT response also called BDT score is obtained. This score indicates if the event is more background-like (small values) or more signal-like (large values). The BDT score distribution for one exemplary case is shown in figure 6.3. The better the classification, the more the background and signal distributions are separated. In addition, in this plot both the training and testing results are shown. Comparing the results from the testing and training can indicate a possible overfitting. Overfitting appears if the model contains too exact information about a particular data set or if there are large statistical fluctuations due to a too small training sample and hence the model may fail to fit additional data. It can be spotted if the shown BDT score distribution has a too large discrepancy between the training and application distribution. Another possibility to check for overfitting is to compare the ROC curves from training and application.
So, at the end the BDT combines for each event the input variables to a final score. Since smaller scores indicate more background like events it is feasible to select events with higher scores.

6.4 Training - Optimisation Strategy

To optimise the signal yield with the methods described above, the best parameters have to be determined which is done within a training procedure. The training and application procedure is illustrated in figure 6.4. First, the full MC sample is splitted into a training and a testing set with approximately the same amount of statistic. Next, the training of the different methods is performed with the training sample whereas each method has different training approaches. Finally, the optimised parameters from the training are applied to the testing sample to study the performance of the individual techniques.

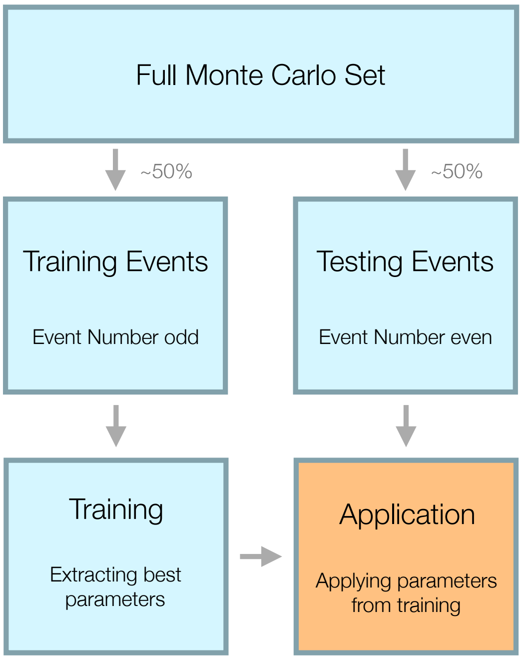


Figure 6.4: Optimisation strategy procedure using independent sets for training and testing.

This is a rather simple approach for the training and application. One could also interchange the testing and training set and perform a second optimisation to use the full statistic and moreover doing cross validation. The work in this thesis is meant to look for new techniques to improve the SR definitions. Afterwards more complex studies are necessary to implement a certain technique into the current analysis. Therefore, within this R&D study it is sufficient to use this simple approach.

If not indicated differently, all following results correspond to the testing sample.

6.5 Calculation of Significance

To quantify the outcome of a search in high energy physics, one uses the significance as reference, in this case the discovery significance. It is computed from the p-value which is the probability of finding data, under assumption of a model H, of equal or greater incompatibility with the predictions of H which corresponds to the Nullhypothesis

(Nullhypothesis has no signal).
Typically, the underlying statistics of the event count follows a Poisson distribution. The significance for the following studies is calculated via the RooStats function from ROOT `RooStats.NumberCountingUtils.BinominalExpZ` as it is used in the SUSY 0-lepton analysis (see sec. 5).

Assuming a Poisson statistic the p-value can be calculated using the normalised (regularised) incomplete beta function described in [50]:

$$I(z;a,b) = \frac{B(z;a,b)}{B(a,b)} = \frac{\int_0^z u^{a-1}(1-u)^{b-1}\mathrm{d}u}{B(a,b)}, \tag{6.1}$$

where

$$B(z;a,b) = \int_0^z u^{a-1}(1-u)^{b-1}\mathrm{d}u = z^a \sum_{n=0}^{\infty} \frac{(1-b)_n}{n!(a+n)} z^n, \tag{6.2}$$

with a, b, z parameters of the incomplete beta function and

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)}, \tag{6.3}$$

with Γ the gamma function and

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \tag{6.4}$$

The p-value is then given as

$$P(z;a,b) = 1 - I(z;a,b). \tag{6.5}$$

The parameters for the p-value calculation can be retrieved from the number of signal and background events and a correction factor τ

$$z = \frac{1}{1+\tau}, \tag{6.6}$$

$$a = s_{\mathrm{exp}} + b_{\mathrm{exp}}, \tag{6.7}$$

$$b = b_{\mathrm{exp}} \cdot \tau + 1. \tag{6.8}$$

The value for τ can be approximated via the relative background uncertainty and the number of expected background events

$$\tau \approx \frac{1}{b_{\mathrm{exp}}} \cdot \frac{1}{b_{\mathrm{unc}}^2}. \tag{6.9}$$

In order to estimate the systematic uncertainties for study purposes a fit was performed as shown in figure 6.5 done in [3]. This fit links the number of predicted background events with a systematic uncertainty via the following function:

$$f(N) = \frac{a}{N^b} + c, \tag{6.10}$$

where N is the number of SM background events in the signal region. The fitted parameter values are:

$$a = 0.40, \quad b = 0.53, \quad c = 0.077. \tag{6.11}$$

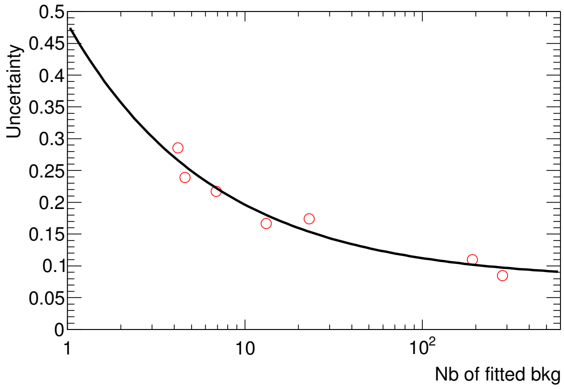


Figure 6.5: Systematic uncertainty of the 0-lepton analysis using $3.2\,\mathrm{fb}^{-1}$ of data as a function of the predicted number of background events. From private communication with Nikola Makovec.

In the following the notation is defined equivalent to [51] where the significance is defined as

$$Z = \Phi^{-1}(1-p), \tag{6.12}$$

with Φ^{-1} the inverse of the cumulative distribution of the standard Gaussian also called quantile and p the p-value.

Assuming the number of background events b is known, one gets the following approximated likelihood function

$$L(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)}, \tag{6.13}$$

with s the mean number of events from a signal model, μ the strength parameter and n the total number of events.
The test statistic for a discovery q_0 can be written as

$$q_0 = \begin{cases} -2 \ln \frac{L(0)}{L(\hat{\mu})} & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0, \end{cases} \tag{6.14}$$

where $\hat{\mu} = n - b$.
The bias of maximum likelihood estimators in general go to zero in the large sample limit. Thus the significance can be rewritten as

$$Z = \sqrt{q_0}. \tag{6.15}$$

For instance for a multi-bin analysis it is necessary to combine the significance of each bin to a combined significance. Each bin has its own likelihood function $L_i(\mu)$. The strength parameter μ is assumed to be the same for all bins. Moreover, the bins are assumed to be statistically independent. Thus the full likelihood function is given as

$$L(\mu) = \prod_i L_i(\mu). \tag{6.16}$$

Using the equations (6.14), (6.15) and (6.16) one gets the the following expression for the combined significance

$$Z_{\text{comb}} = \sqrt{q_{0,\text{comb}}} = \sqrt{-2\ln\left(\prod_i \frac{L_i(0)}{L_i(\hat{\mu})}\right)}. \tag{6.17}$$

Applying the logarithm property $\ln(a \cdot b) = \ln(a) + \ln(b)$, the previous expression simplifies to

$$Z_{\text{comb}} = \sqrt{-2\sum_i \ln\left(\frac{L_i(0)}{L_i(\hat{\mu})}\right)} = \sqrt{\sum_i q_{0,i}}. \tag{6.18}$$

From equation (6.14) follows that $q_0 = Z^2$ hence the combined significance Z_{comb} is equivalent to the quadratic sum of the single bin significances

$$Z_{\text{comb}} = \sqrt{\sum_i Z_i^2}. \tag{6.19}$$

7 Squark Pair Productions Studies

The analysis presented in chapter 5 was using a dataset with an integrated luminosity of $36.1\,\mathrm{fb}^{-1}$, whereas the expected luminosity for the end of RUN II is much higher. This allows the application of newer sophisticated techniques to optimise the signal regions for this large dataset.
One of these techniques are BDTs as described in section 6.3 which will be used in the following to optimise the signal region definitions.

At the time when this study was initiated, the luminosity at the end of RUN II was expected to be around $100\,\mathrm{fb}^{-1}$. Thus in the following, the squark studies are performed for this luminosity.

The target signal is the squark-pair production where the squarks decay directly into a neutralino and a quark, shown in figure 7.1.

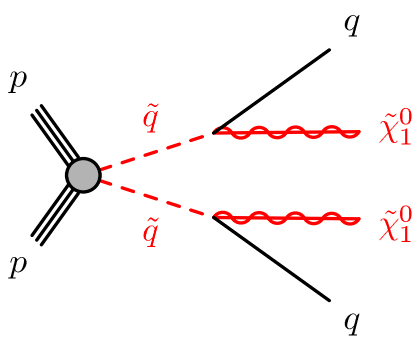


Figure 7.1: The decay topology of squark-pair production in the simplified models with direct squark decay [32].

The studies below are all fully based on Monte Carlo samples described in table 3.1. The W+jets, Z+jets, top and diboson backgrounds are taken into account. The multijet (QCD) background is negligible for this phase-space after applying a tight cut on $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ and $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and hence not considered for the following studies. Given the large size of multijet input samples, the analysis run-time speeds up significantly when skipping them.

In general, the luminosity of the simulated MC sample should be larger than the data luminosity. Therefore a reweighting of the MC samples to the data luminosity is necessary which will be done for this study.

7.1 Cut & Count Reoptimisation

In order to be able to compare the results from new techniques with the 0-lepton publication at a higher luminosity, it is necessary to reoptimise the cuts for a dataset with $100\,\mathrm{fb}^{-1}$.

The strategy is to fix the cuts on certain variables to the values from the 0-lepton publication (see table 5.1 & 5.2) and find the optimum for a set of chosen variables, preferably the most discriminating ones. This approach is chosen for simplification purposes since the Cut & Count results are meant to be a reference point to estimate the possible gain of more sophisticated techniques.

One examplary reoptimisation is shown in table 7.1 for the signal point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)\,\mathrm{GeV}$ corresponding to the signal region Meff-2j-1600 which was the best performing m_{eff} signal region in the 0-lepton publication for this signal point. In this case, the cuts on the variables $E_{\mathrm{T}}^{\mathrm{miss}}$, $p_{\mathrm{T}}(j_2)$, $\max(|\eta|_{i=1,2})$, $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ and $m_{\mathrm{eff}}(\mathrm{incl.})$ were reoptimised.
In general, the cuts on $|\eta|$ of the first two jets $|\eta(j_{1,2})|$ are the same and thus they can be simplified to $\max(|\eta|_{i=1,2})$.

Requirement	Original Cut & Count	Reoptimised Cut & Count
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV] >	250	250
$p_{\mathrm{T}}(j_1)$ [GeV] >	300	300
$p_{\mathrm{T}}(j_2)$ [GeV] >	300	300
$\max(\lvert\eta\rvert_{i=1,2})$ <	1.2	0.8
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ >	0.8	0.8
$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ >	0.4	0.4
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ [$\sqrt{\mathrm{GeV}}$] >	18	16
$m_{\mathrm{eff}}(\mathrm{incl.})$ [GeV] >	1600	1400
Significance	2.11	2.61

Table 7.1: Reoptimised cuts for training point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)\,\mathrm{GeV}$. All variables but $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $p_{\mathrm{T}}(j_2)$ were optimised. The original Cut & Count result corresponds to the signal region Meff-2j-1600 defined in tabble 5.1.

The reoptimisation, also called *training* in the following, is done by varying the cuts of each training variable. The cut intervals of each variable are extracted from the signal regions defined in the 0-lepton paper ranging from 0.5 to 1.5 times the signal region value[1]. The bin steps are further adapted for each single variable. Next, all possible cut combinations are

[1] If the best cut value corresponds to the minimum or maximum of the interval, it is checked if a lower or higher value does not give a better significance and changed if needed.

probed on the training dataset and the best selections are extracted.
Comparing the two results from table 7.1, the reoptimised selection shows an improvement of 20% on the significance. The $E_{\mathrm{T}}^{\mathrm{miss}}$, $p_{\mathrm{T}}(j_2)$ and $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ cuts stay the same while the cuts on $\max(|\eta|_{i=1,2})$, $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ and $m_{\mathrm{eff}}(\mathrm{incl.})$ were modified. The latter two variables were the most discriminating variables in the 0-lepton publication.
The $m_{\mathrm{eff}}(\mathrm{incl.})$ distribution after the application of the reoptimised cuts except the $m_{\mathrm{eff}}(\mathrm{incl.})$ cut itself is shown in figure 7.2 with the corresponding significance in the lower pad calculated using the total background and signal yields. In this plot the Multi-jet background is included only for demonstration purposes.

Due to the large number of signal points it is more feasible to draw contour lines similar to figure 5.3 to visualise a gain. While the 0-lepton publication shows exclusion contours (figure 5.3), the contour lines shown in the following correspond to a 3σ discovery significance.
The contour lines for the reoptimised cuts as well as for the original cuts from the 0-lepton publication are shown in figure 7.3. The reoptimised cut results show an improvement in the neutralino mass ($\sim 50\,\mathrm{GeV}$).

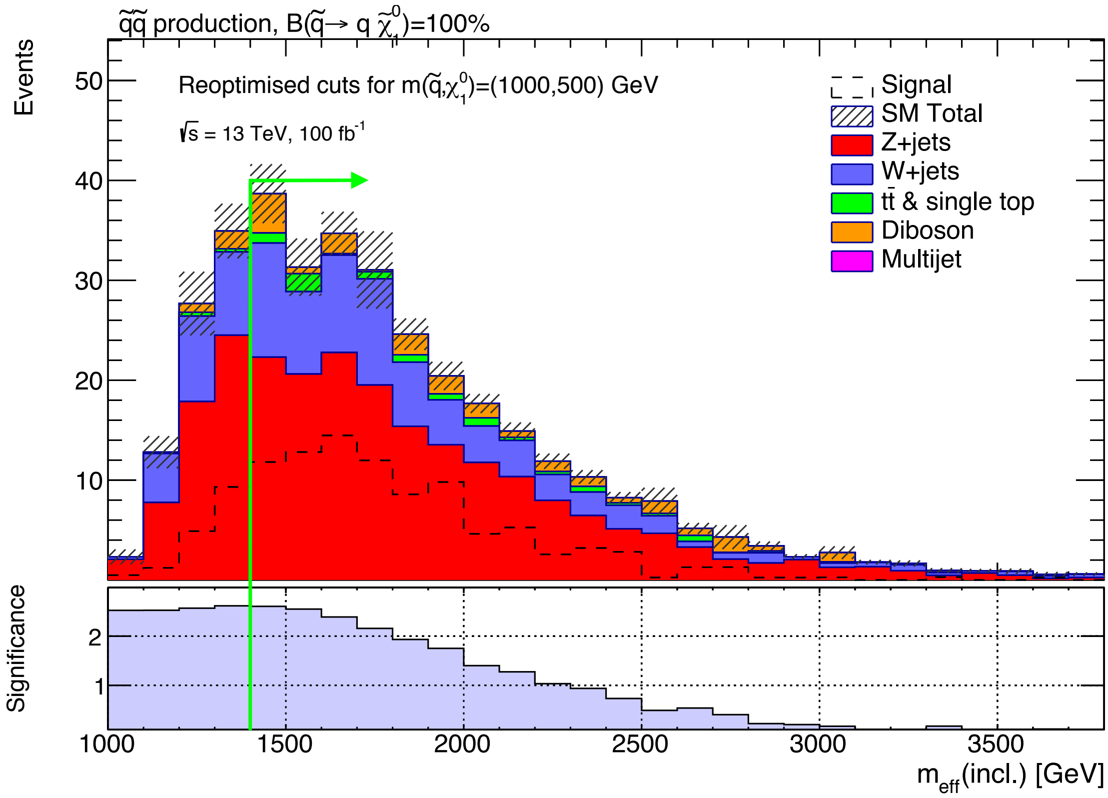


Figure 7.2: MC distribution of the $m_{\mathrm{eff}}(\mathrm{incl.})$ variable after applying the reoptimised cuts except the $m_{\mathrm{eff}}(\mathrm{incl.})$ cut, which is indicated via the green arrow-line. The lower pad shows the signal significance. The signal distribution for the mass point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)\,\mathrm{GeV}$ is shown as black dashed line.

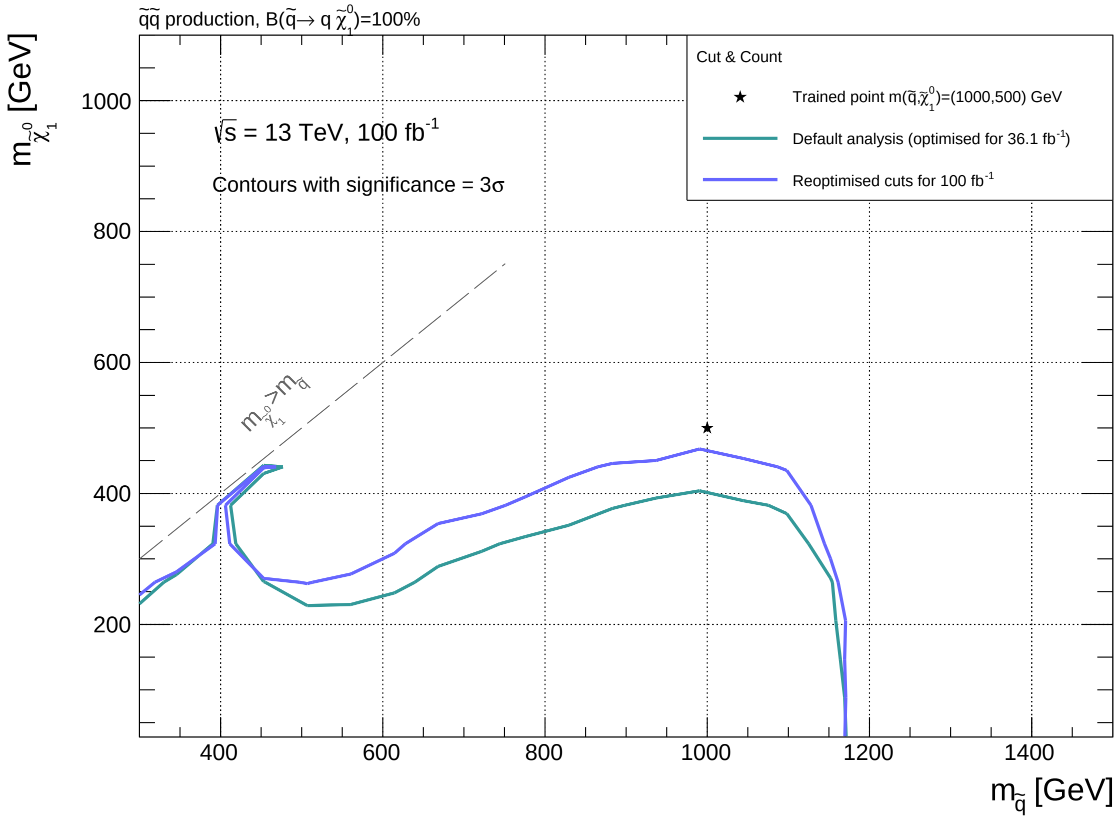


Figure 7.3: The 3σ discovery contours for the original signal region Meff-2j-1600 region (green) and the corresponding reoptimised signal region (blue). The star indicates the signal point which was reoptimised.

This exemplary optimisation nicely shows the need of a Cut & Count reoptimisation to be able to compare further results at higher luminosities. In the following Cut & Count always refers to the reoptimised Cut & Count method.

7.2 BDT Optimisation

The Boosted Decision Trees (BDT) as described in section 6.3 allow a more complex classification of background and signal events.
The procedure is similar to the Cut & Count reoptimisation. First, a certain number of signal points are picked. As a first guideline they are chosen such that they are close to the exclusion limits of the 0-lepton publication (see figure 5.3) and afterwards the choice of the signal points is further adapted to improve the 3σ discovery significance contour lines.

7.2.1 Variable Study

One of the most important parts of the BDT optimisation is the choice of the input variables. On the one hand one should use the most discriminating variables. On the other hand it is also not useful to overload the training with too many variables which could have a negative impact on the performance and would make the BDT unnecessarily more complex.

In order to achieve a good performance of the BDT, the individual input variables should have good separation of signal and background. The signal and background distributions are normalised to unity. Then the separation can be calculated as follows

$$D = \sum_{i=\mathrm{bin}_{\mathrm{min}}}^{\mathrm{bin}_{\mathrm{max}}} \frac{(N_{bkg}^{(i)} - N_{signal}^{(i)})^2}{N_{bkg}^{(i)} + N_{signal}^{(i)}}, \tag{7.1}$$

where $N_{bkg}^{(i)}$ and $N_{signal}^{(i)}$ are respectively the number of background and signal events in bin i. The expected separation for one bin is shown in figure 7.4. It is nicely visible that if the number of background and signal events are identical, the separation is zero and it is maximal if one distribution has zero events in this bin.

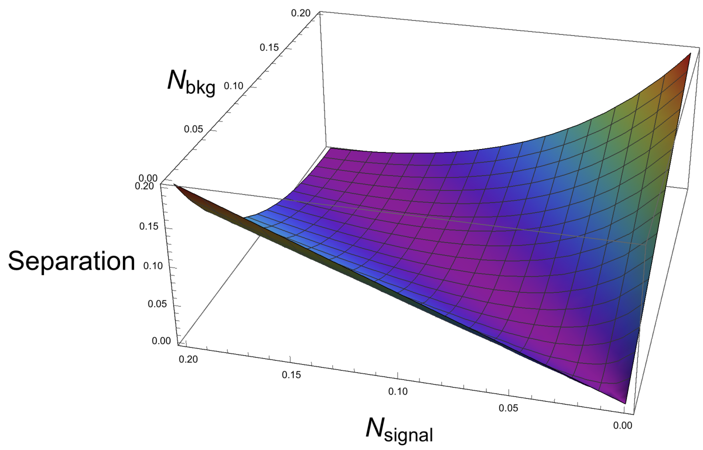


Figure 7.4: Separation for one bin as a function of number of background and signal events.

With the separation criterion from equation (7.1) it is thus possible to choose a first set of variables as input for the BDT.
The separation of the variables used in the 0-lepton publication for the squark signal regions (see table 5.1) are shown in table 7.2. The variables are sorted by their separation power. As expected, the m_{eff}(incl.) variable has the largest separation and is therefore a good choice as input variable. The normalised distributions of these variables are shown in figure 7.5 and 7.6 and the selection criteria are listed in A.1. The correlation matrix for the main variables is shown in figure 7.7.

Since the last two variables $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $p_{\mathrm{T}}(j_1)$ show such a low separation, they are for now not considered in the training of the BDT. Even though $E_{\mathrm{T}}^{\mathrm{miss}}$ is ranked as second best variable by its separation, it is put aside for the moment since m_{eff}(incl.) contains $E_{\mathrm{T}}^{\mathrm{miss}}$ (see chap. 5) and they are highly correlated with a correlation factor of 0.61 for both, signal and background. Also $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ depends on $E_{\mathrm{T}}^{\mathrm{miss}}$ and they show an even higher correlation of 0.87 for signal and 0.84 for the background. The first choice of input variables is then given as indicated in table 7.2.

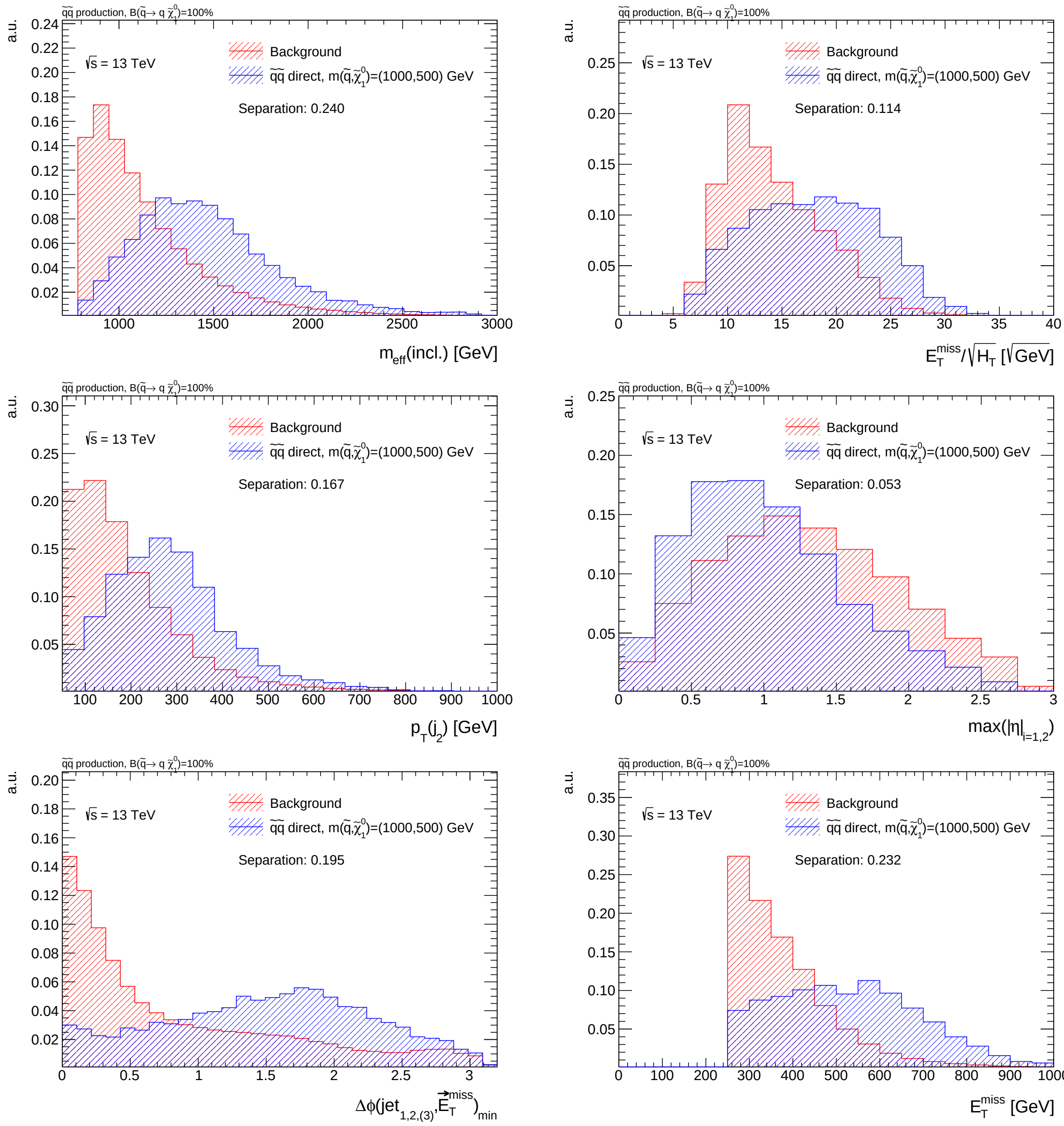


Figure 7.5: Normalised variable distributions: $m_{\text{eff}}(\text{incl.})$, $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$, $p_{\text{T}}(j_2)$, $\max(|\eta|_{i=1,2})$, $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ and $E_{\text{T}}^{\text{miss}}$. The background is shown in red and the signal with $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV in blue. In each plot the separation power for the background and signal distribution is given. The selection criteria are listed in A.1.

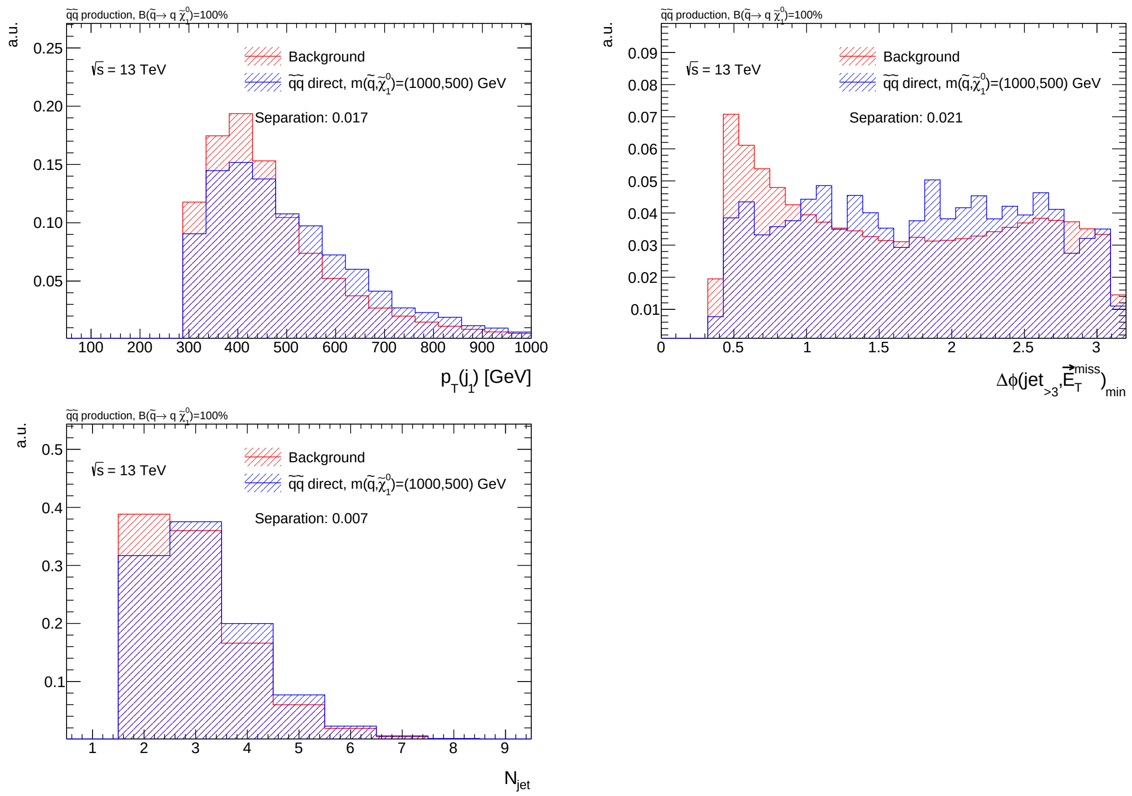


Figure 7.6: Normalised variable distributions: $p_{\mathrm{T}}(j_1)$, $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and N_{jet}. The background is shown in red and the signal with $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV in blue. In each plot the separation power for the background and signal distribution is given. The selection criteria are listed in A.1.

As next step, different input variable combinations are trained with the parameters taken from a default TMVA tutorial (hyper parameters described in sec 6.3):

- Boosting: AdaBoost,
- AdaBoostBeta: 0.15,
- NTrees: 200,
- MaxDepth: 4,

and the whole input string for the TMVA training can be found in appendix A.2.
Importantly, the preselection cuts on the SM background as well as on the signal sample $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV before the training are determined as for the Cut & Count optimisation, i.e. corresponding to the 0-lepton publication. The preselection cuts are based on variables which are not used as input variables for BDT[2].

The BDT score distribution for an exemplary BDT training with the signal point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV is shown in figure 7.8 with the input variable combination shown

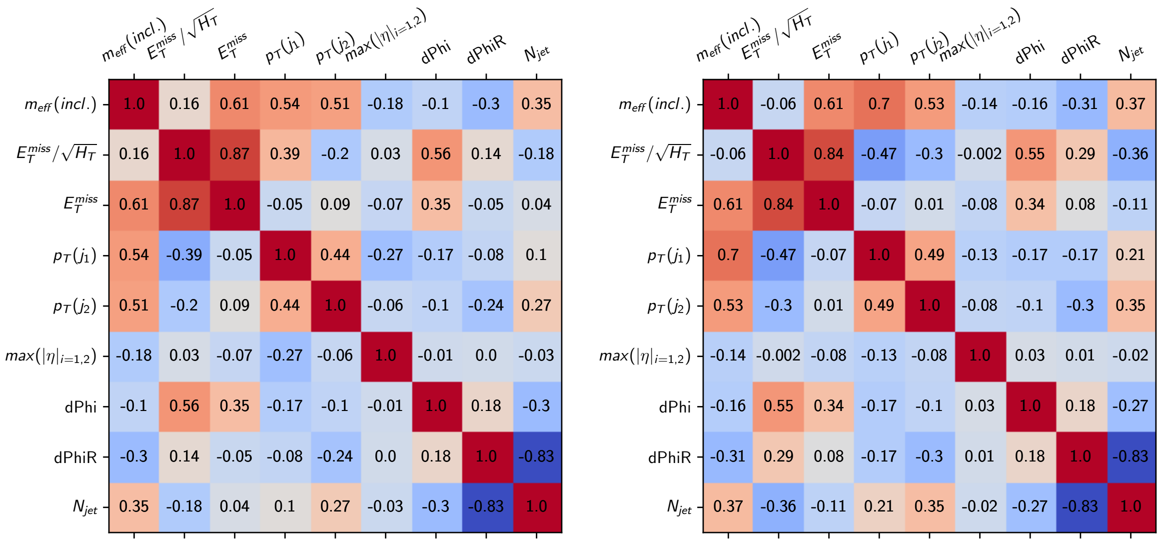
[2]For the $m_{\mathrm{eff}}(\mathrm{incl.})$ variable, there is always a pre-cut considered with $m_{\mathrm{eff}}(\mathrm{incl.}) > 800$ GeV.

Variable	Separation in %	Used in BDT
m_{eff}(incl.)	24.0	✓
$E_{\text{T}}^{\text{miss}}$	23.2	
$\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$	19.5	✓
$p_{\text{T}}(j_2)$	16.7	✓
$E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$	11.4	✓
$\max(\lvert\eta\rvert_{i=1,2})$	5.3	✓
$\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$	2.1	
$p_{\text{T}}(j_1)$	1.7	
N_{jet}	0.7	

Table 7.2: Separation of the variables used in the 0-lepton publication signal regions targeting squark pair production in simplified models. Input variables considered in BDT are indicated via check marks.

in table 7.2. The separation of the BDT score distribution of background and signal with a value of $\sim 46\%$ improved with respect to the best performing input variable (m_{eff}(incl.) $\sim 24\%$). The distribution of the test samples agree nicely with the distributions of the training samples except of some fluctuations which are due to low statistics in these areas. Disagreement would be an indication of overfitting.
The total statistic of one MC signal sample is around $10,000 - 16,000$ and is therefore a limiting factor. For the $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)\,\text{GeV}$ signal point ~ 6000 signal events and $1.5 \cdot 10^6$ background events are used for the BDT training and the same amount is also used for testing (but statistically independent).
Moreover, the ROC curve (signal efficiency versus background rejection) for training and testing shown in figure A.1 shows a very good agreement and does not indicate any overfitting. The large area under the ROC curve indicates that BDT provides a good separation of background and signal of $\sim 46\%$ for the testing sample.

In figure 7.9, four different BDT input variable combinations are shown. This plot is done by increasing the cut on the BDT score, calculating the number of signal and background events and drawing the corresponding values for the significance and number of signal events. The significance for the current and reoptimised Cut & Count are marked in the plot as dashed line and as star, respectively (taken from table 7.1).
According to this plot, the cut on the BDT score is chosen such that it corresponds to the best significance.
Apart from this variable study, several variable scans were performed to determine the best combination, e.g. p_{T}(jets) scans. As a result, the combination already indicated in table 7.2 stays the best choice and will be used in the following.



a) b)

Figure 7.7: Linear correlation matrix of a) signal $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV and b) background for the main variables considered for input to BDT. The names dPhi and dPhiR are abbreviations for $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\text{miss}})_{\text{min}}$ and $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\text{miss}})_{\text{min}}$, respectively.

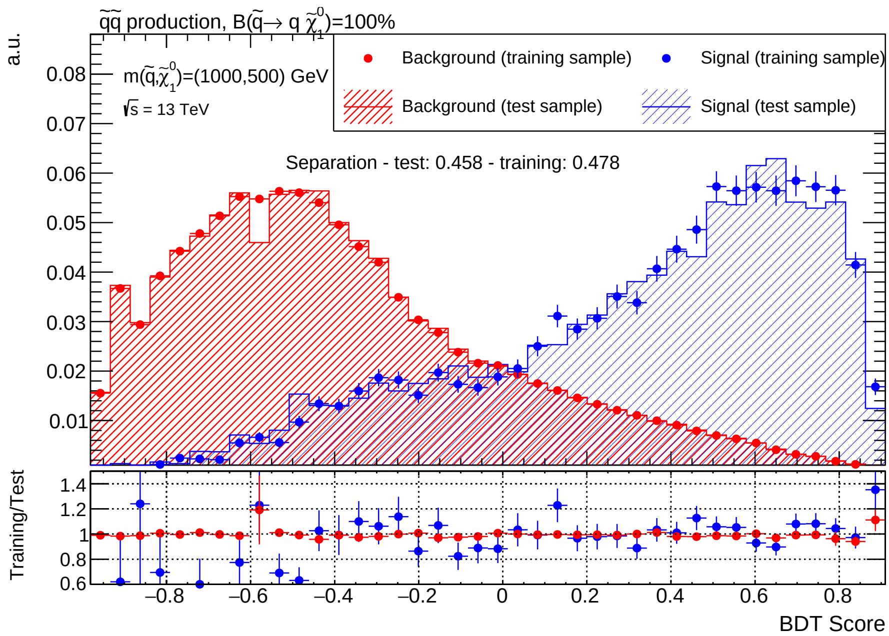


Figure 7.8: Normalised BDT score distribution of background (red) and signal (blue) for the training and testing sample using the signal training point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV. Used input variables: $m_{\text{eff}}(\text{incl.})$, $E_{\mathrm{T}}^{\text{miss}}/\sqrt{H_{\mathrm{T}}}$, $p_{\mathrm{T}}(j_2)$, $\max(|\eta|_{i=1,2})$ and $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\text{miss}})_{\text{min}}$. The bottom panel shows the ratio of training to testing distribution.

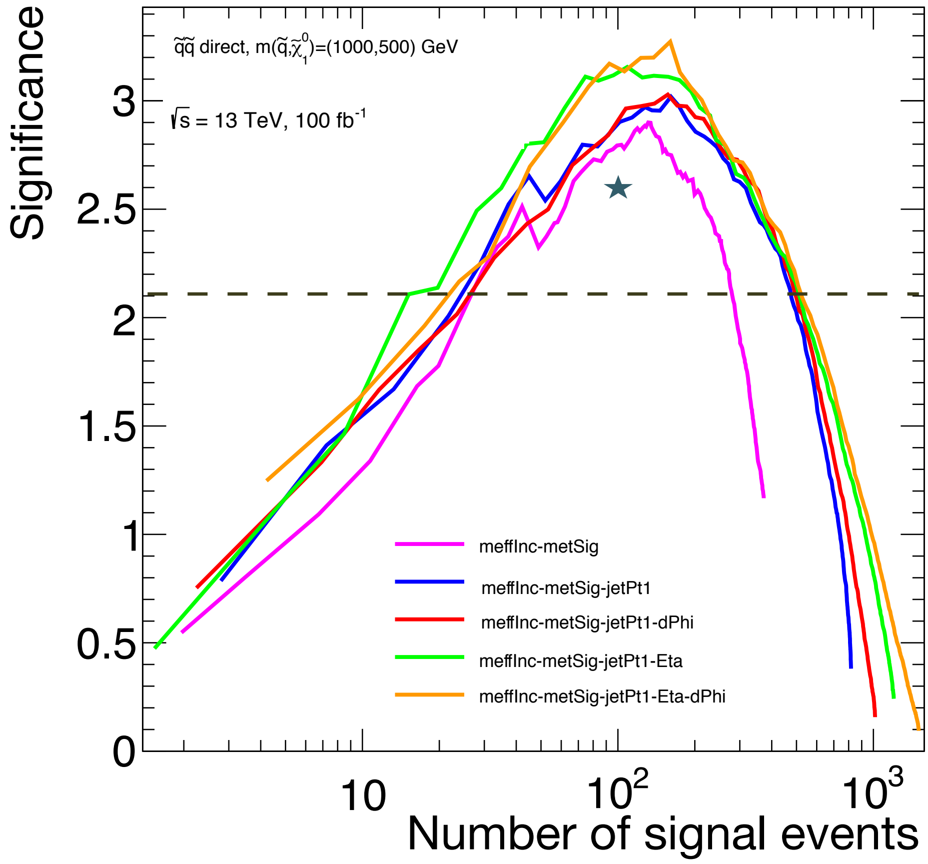


Figure 7.9: Significance plotted against number of signal events for different BDT input variables for the signal with $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV. The horizontal dashed line shows the significance of the default analysis and the star shows the value for the reoptimised Cut & Count method. The variables names in the legend are abbreviations for: meffInc for $m_{\text{eff}}(\text{incl.})$, metSig for $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$, jetPt1 for $p_{\text{T}}(j_2)$, Eta for $\max(|\eta|_{i=1,2})$ and dPhi for $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$.

7.2.2 Choice of Training Points

In the current 0-lepton publication the direct squark decay signal grid consists of 164 signal points. The kinematic properties of the signal points differ especially for different $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ values. This behaviour is for instance visible in the $m_{\text{eff}}(\text{incl.})$ distributions of two signal points with $\Delta m(\tilde{q}, \tilde{\chi}_1^0) = 1200$ GeV and $\Delta m(\tilde{q}, \tilde{\chi}_1^0) = 100$ GeV as shown in figure 7.10. Also the opposite is true, if the $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ is similar for two signal points so their kinematics are similar as well.

Consequently it is necessary to train a BDT for several signal points with different values of $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$. An exemplary result for four different points covering a wide range in the $m(\tilde{q}, \tilde{\chi}_1^0)$-plane is shown in figure 7.11. Each of these four BDT based signal regions is optimised for a specific $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$-region depending on the training point. In particular the BDT trained with the signal point $m(\tilde{q}, \tilde{\chi}_1^0) = (600, 500)$ GeV in the compressed region (low $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$) is highly optimised for this area and has almost no sensitivity for any other

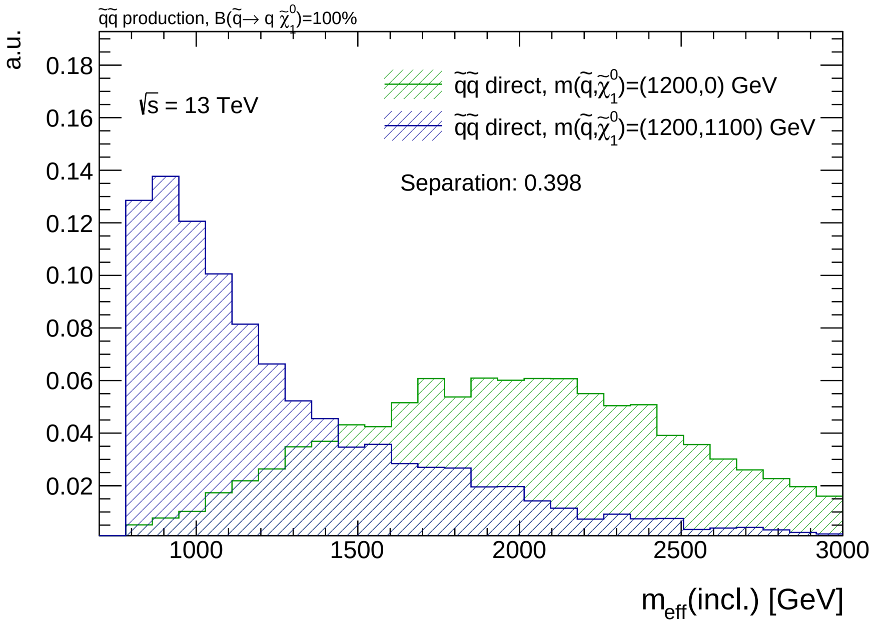


Figure 7.10: Normalised m_{eff}(incl.) distribution for two different signal points of direct squark decay. In green $m(\tilde{q}, \tilde{\chi}_1^0) = (1200, 0)$ GeV with large $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ and in blue $m(\tilde{q}, \tilde{\chi}_1^0) = (1200, 1100)$ GeV with low $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ value.

region. The same applies for the training point $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 0)$ GeV with high $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$.

The choice of the training points for the BDT is further optimised comparing their contour lines. For instance if the training points are close together with a similar mass difference the resulting contour lines are almost the same or one encloses the other one as shown in figure 7.12. In this case the training point $m(\tilde{q}, \tilde{\chi}_1^0) = (700, 600)$ GeV is discarded and the other one is kept.

This procedure is repeated until no improvement is visible anymore. As a result the following set of training points is retrieved:

1. $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 0)$ GeV,
2. $m(\tilde{q}, \tilde{\chi}_1^0) = (1200, 500)$ GeV,
3. $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV,
4. $m(\tilde{q}, \tilde{\chi}_1^0) = (800, 600)$ GeV,
5. $m(\tilde{q}, \tilde{\chi}_1^0) = (800, 500)$ GeV and
6. $m(\tilde{q}, \tilde{\chi}_1^0) = (600, 500)$ GeV.

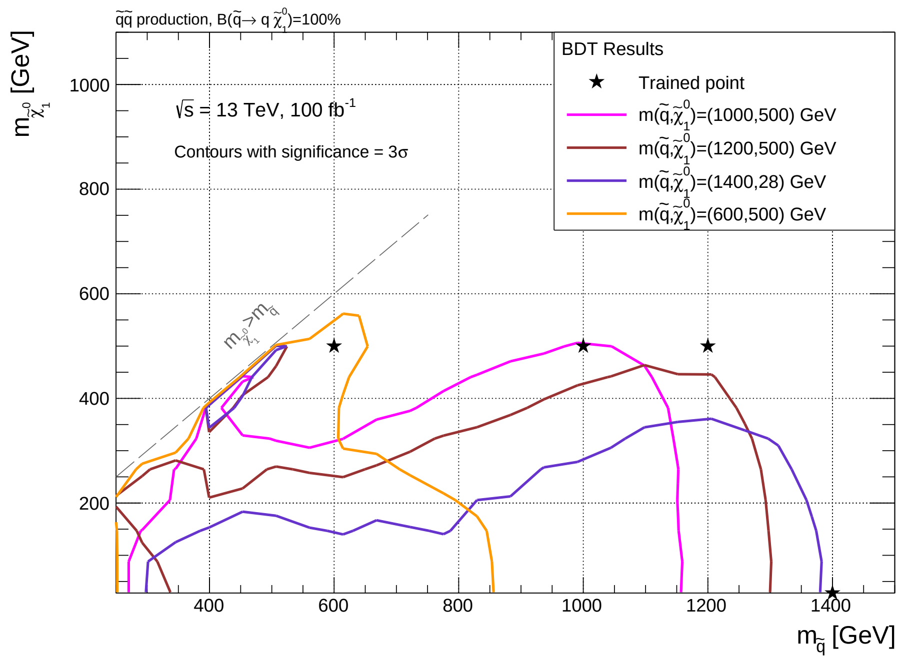


Figure 7.11: 3σ discovery significance contour lines corresponding to 4 BDT based signal regions. The signal points used in BDT training procedure are marked as stars.

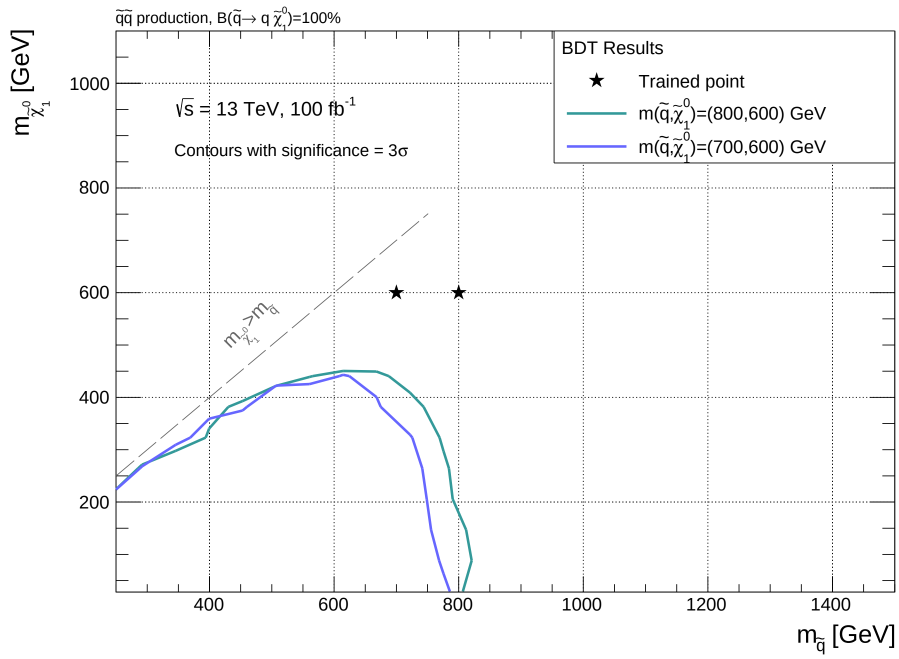


Figure 7.12: 3σ discovery significance contour lines for two very similar training points. The signal points used in BDT training procedure are marked as stars.

7.2.3 Oring Procedure

After the determination of the training points one gets six different contour lines. Drawing all these six lines in one plot is not well readable.
Therefore a method is needed to combine all these single results. The combination procedure is called *oring*. For each signal point, the BDT signal region yielding the highest significance is chosen. In this way one common contour line is obtained. The resulting combined contour line for the reoptimised Cut & Count analysis as well as for the BDT are shown in figure 7.13. Both, BDT and Cut & Count are trained with the same signal points listed above.

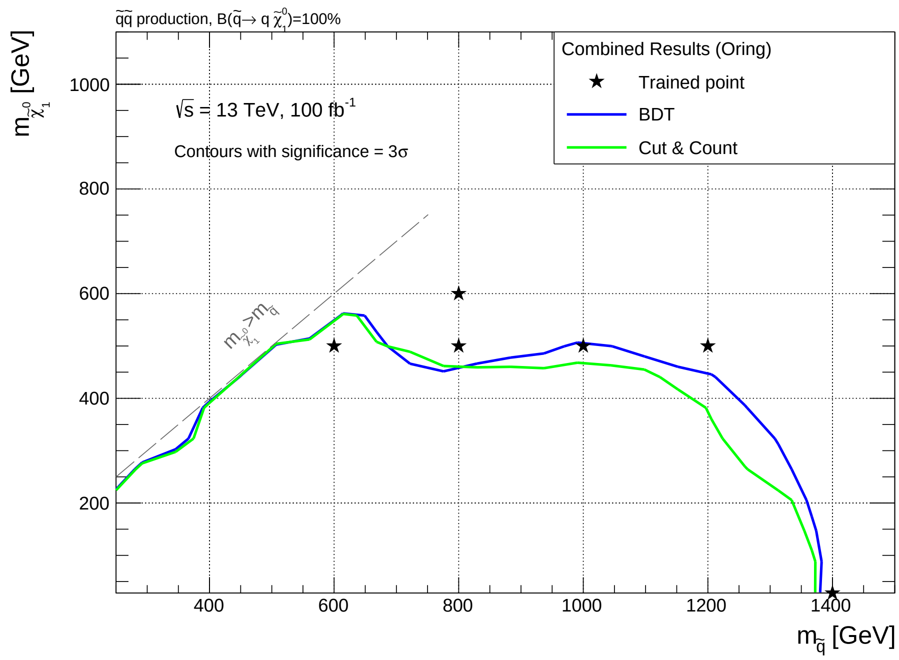


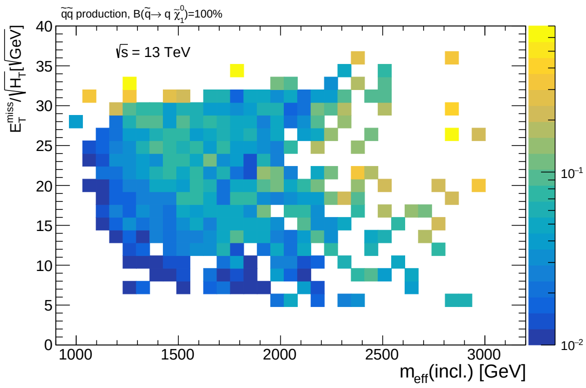
Figure 7.13: Contour lines with 3σ discovery significance for the combined (oring) BDT (blue) and Cut & Count (green) results.

The BDT method shows improvements in the region with $\Delta m(\tilde{q}, \tilde{\chi}_1^0) = 400 - 1000\,\text{GeV}$ but there is no gain in the more compressed region and for $\Delta m(\tilde{q}, \tilde{\chi}_1^0) > 1000\,\text{GeV}$.
Therefore the optimisation of the BDT hyperparameters was studied by iterating over a wide range of parameter definitions (mainly NTrees, MaxDepth, nCuts and AdaBoostBeta). However there was no significant gain visible which would justify the optimisation of all BDTs. Therefore the hyperparameters listed above will be used for all direct squark decay studies.

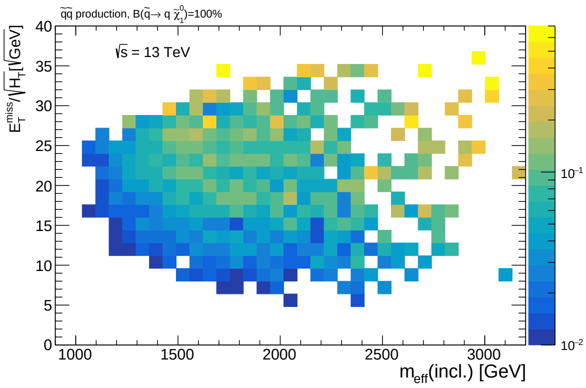
In conclusion, the pure BDT approach shows already some minor improvements in the moderate $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ region but for a possible implementation in the next round of the 0-lepton analysis it is not yet enough. Thus further studies are performed in the following using additional techniques.

7.3 N_{jet} Separation

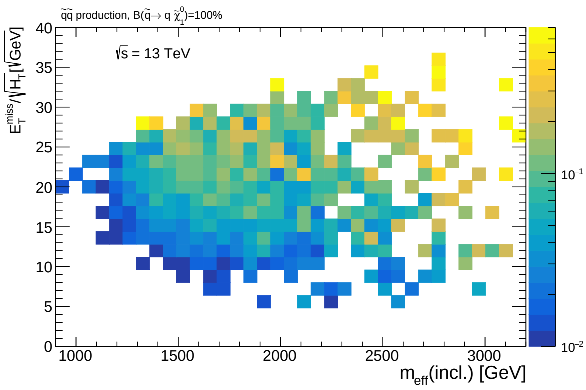
The signature in the detector of the direct squark decays consists of two quark jets and missing transverse momentum $E_{\text{T}}^{\text{miss}}$. Due to initial and final state radiation (ISR & FSR) the number of jets increases. Hence the additional jets originate mainly from gluon radiation. While the N_{jet} distribution as shown in figure 7.6 shows a very low separation, the m_{eff}(incl.) and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ distribution of signal over background changes with different jet multiplicities as shown in figure 7.14.



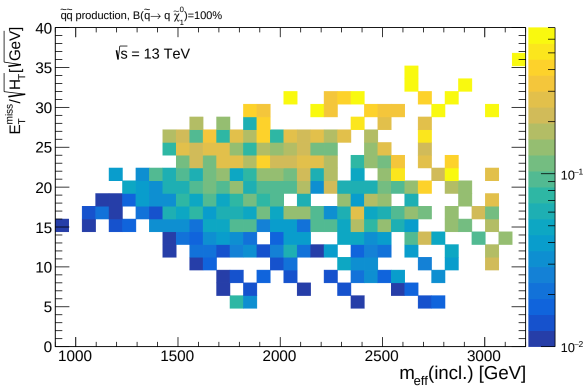
a) $N_{\text{jet}} == 2$



b) $N_{\text{jet}} == 3$



c) $N_{\text{jet}} == 4$



d) $N_{\text{jet}} \geq 5$

Figure 7.14: 2D variable distribution of $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ over m_{eff}(incl.) shown for different jet multiplicities. The signal yield is divided by the background yield indicated as colour code for the signal point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV.

This behaviour motivates a differentiation in the jet multiplicity for the signal region definitions. Due to the signal topology, the minimal required number of jets is two. Since almost all events are within the bins up to five jets (see figure 7.6) the signal regions are divided accordingly into four bins as shown in table 7.3.

The combined significance for several bins is calculated by adding the significances of the single bins quadratically as described in equation (6.19).

N° bin	N_{jet}
1	$== 2$
2	$== 3$
3	$== 4$
4	≥ 5

Table 7.3: Definition of the N_{jet} division for the signal regions.

Each bin is optimised separately for the Cut & Count method and also for the BDT. One exemplary result for the signal point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV is shown in table 7.4. While the results for the inclusive selection are better than the single results of the different jet multiplicities the combination of the four N_{jet} bins shows a major improvement, in the case of BDT 44% and for Cut & Count 54%.

N_{jet}	Significance		Quadr. sum	
	BDT	Cut & Count	BDT	Cut & Count
≥ 2	3.20	2.61	—	—
$== 2$	2.40	1.84	4.59	4.00
$== 3$	2.71	2.52		
$== 4$	2.18	1.79		
≥ 5	1.80	1.77		

Table 7.4: Results of Cut & Count and BDT for the signal point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV with an inclusive N_{jet} selection and also separated into different jet multiplicities.

This result for only one signal point is already promising and gives rise to a possible gain in the whole signal grid space. The combined results (oring) for the whole signal grid are shown in figure 7.15 for both, Cut & Count and BDT. In comparison to the unbinned case, the N_{jet} separation shows a significant improvement over the whole $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$-range in both methods except in the very compressed scenario ($\Delta m(\tilde{q}, \tilde{\chi}_1^0) < 50$ GeV).

The choice of the training points is also optimised for the N_{jet} separated training resulting in:

1. $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 400)$ GeV,
2. $m(\tilde{q}, \tilde{\chi}_1^0) = (1200, 500)$ GeV,
3. $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)$ GeV,
4. $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 600)$ GeV,

5. $m(\tilde{q}, \tilde{\chi}_1^0) = (800, 500)$ GeV,
6. $m(\tilde{q}, \tilde{\chi}_1^0) = (800, 600)$ GeV and
7. $m(\tilde{q}, \tilde{\chi}_1^0) = (600, 500)$ GeV.

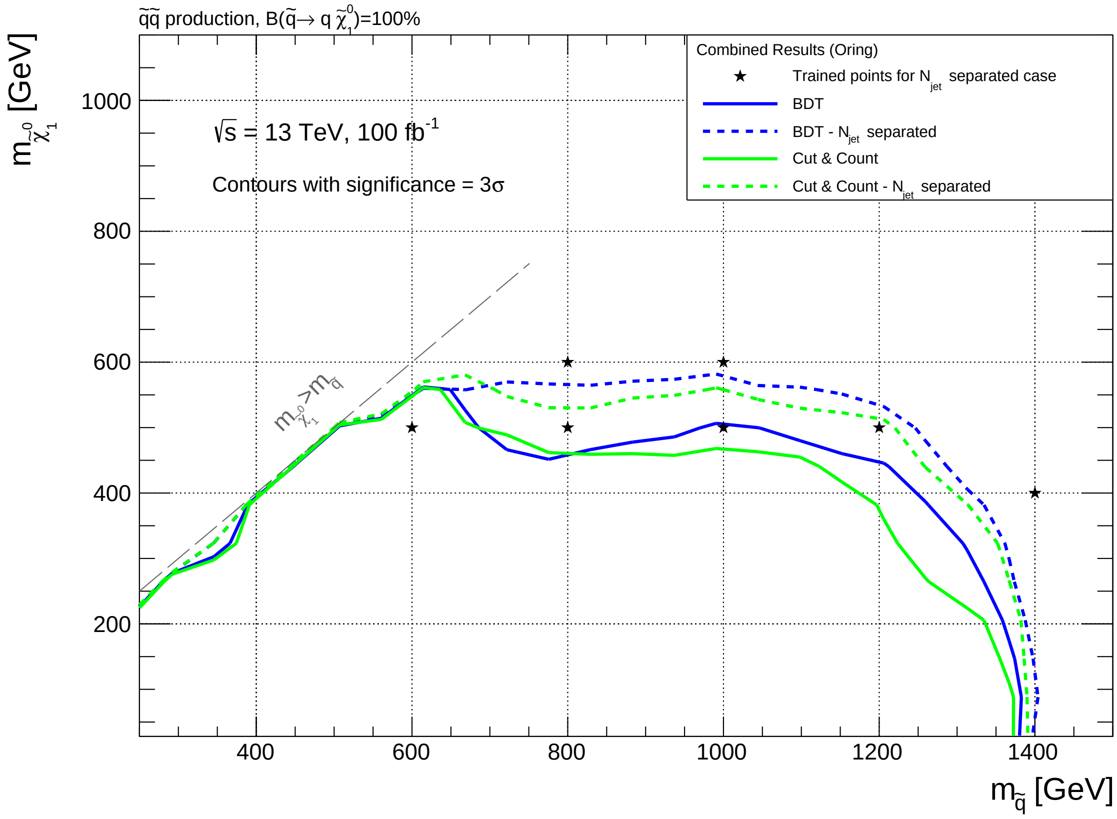


Figure 7.15: The 3σ contour lines for reoptimised Cut & Count, BDT and N_{jet} separated training (dashed contours). Training points for N_{jet} separated BDT and Cut & Count are marked as stars, the training points for the other methods are indicated in figure 7.13.

The results from the N_{jet} separated BDT are constantly better than those from the N_{jet} separated Cut & Count apart from the compressed region. However, the gain is minimal with constantly ~ 25 GeV for the neutralino mass and almost no gain in the squark mass. Considering the higher complexity of BDT to implement in the actual analysis makes the N_{jet} separated Cut & Count a more promising option since the major gain results from the binning into N_{jet} slices.

7.4 Multi-Bin Analysis

As shown in figure 7.15, already the binning into N_{jet} slices showed a great improvement. Based on this observation, the effect of binning into additional variables is studied, to test if further improvements are possible. Since the two main variables in [32] are $m_{\text{eff}}(\text{incl.})$ and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$, they will be used for the binning.

7.4.1 Bin Definitions

The definition of the bins are chosen similar to the signal regions of the 0-lepton publication. The N_{jet} bins stay the same as in the previous section. All these bin definitions are summarised in table 7.5.

	N_{jet}	$m_{\text{eff}}(\text{incl.})$ [GeV]	$E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ [$\sqrt{\text{GeV}}$]
1	$== 2$	$1200 - 1600$	$12 - 16$
2	$== 3$	$1600 - 2000$	$16 - 20$
3	$== 4$	$2000 - 2400$	$20 - 24$
4	≥ 5	$2400 - 2800$	$24 - 28$
5		$2800 - 3200$	$28 - 32$
6		≥ 3200	≥ 32

Table 7.5: Different categories for the Multi-Bin analysis. The combination of all single bins would give 144 bins in total.

By combining all three variables one ends up with $N_{N_{\text{jet}}}$ bins $\times$ $N_{m_{\text{eff}}(\text{incl.})}$ bins $\times$ $N_{E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}}$ bins $= 144$ bins whereas the combination of N_{jet} and $m_{\text{eff}}(\text{incl.})$ would give only 24 bins. Considering the different regions one will easily get a big amount of different categories which might consequently suffer from low statistics. Therefore it is feasible to also look for methods to reduce the number of bins and keep a balance between the complexity and gain.

7.4.2 Training Procedure

Since already the N_{jet} slicing decreased the statistics, a further binning will decrease it even more and it is not feasible to use a BDT (e.g. due to overfitting). Also, using a BDT would yield to an overwhelming complexity of the analysis. Thus the single bins are optimised only using the Cut & Count method with the following variables:

1. $\max(|\eta|_{i=1,2})$,
2. $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\min}$,
3. $p_{\text{T}}(j_2)$,

and if only N_{jet} and $m_{\text{eff}}(\text{incl.})$ are considered for binning also

4. $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$

is optimised. For all other variables the default selection from the current publication is applied.

7.4.3 Results

The Multi-Bin trainings are performed for five selected signal points covering a wide $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$-range, applied to the whole signal grid and shown in figure 7.16 as contour lines with 3σ discovery significance (more details about the training points are shown below).

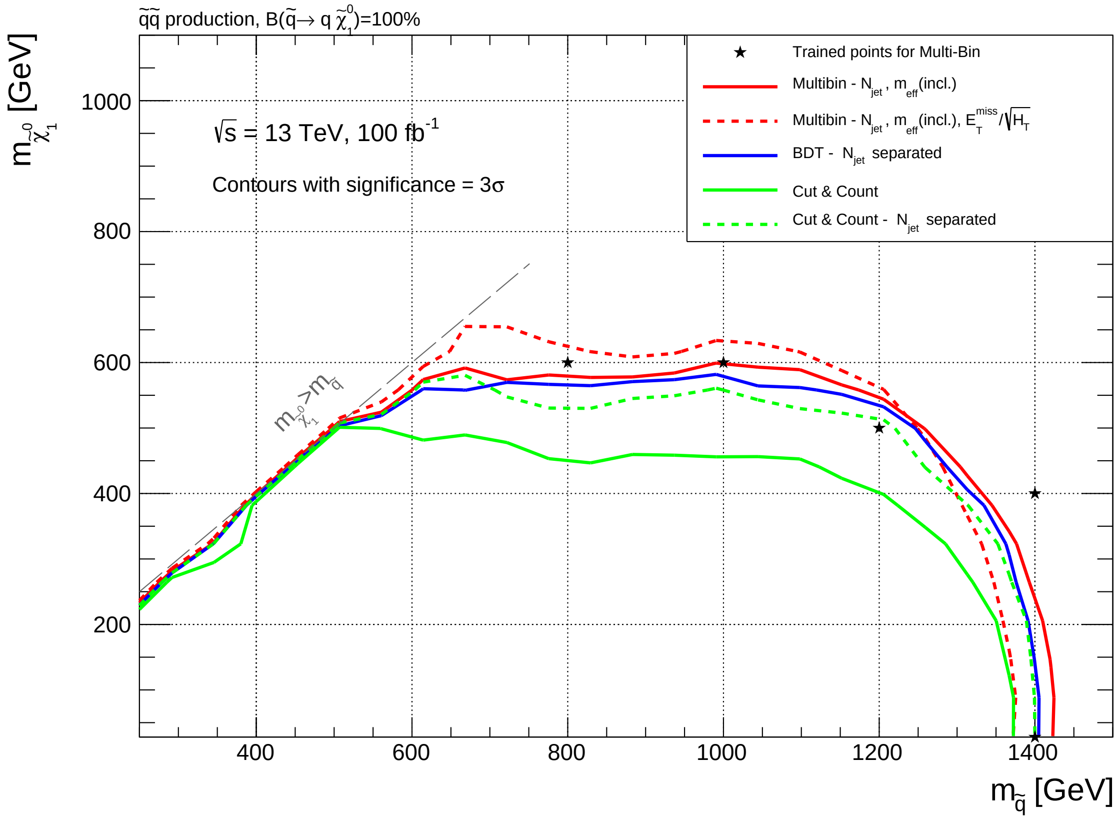


Figure 7.16: 3σ contour lines of the combined (oring) Multi-Bin results with different binning into N_{jet} denoted as "Cut & Count -N_{jet} separated" in the plot legend (green dashed contour), N_{jet} & m_{eff}(incl.) (red contour), N_{jet} & m_{eff}(incl.) & $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ (red dashed contour) and the N_{jet} separated BDT (blue contour) as well as the reoptimised Cut & Count method (green contour) as reference.

The Multi-Bin approach with N_{jet} and m_{eff}(incl.) bins (red contour) improves the contour line with respect to the N_{jet} separated Cut & Count method (green contour) over the whole $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$-range except in the very compressed scenario. The gain is constantly around 25-50 GeV for both squark and neutralino mass. Furthermore this Multi-Bin contour is also overall slightly better than the N_{jet} separated BDT contour, especially at high $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$.

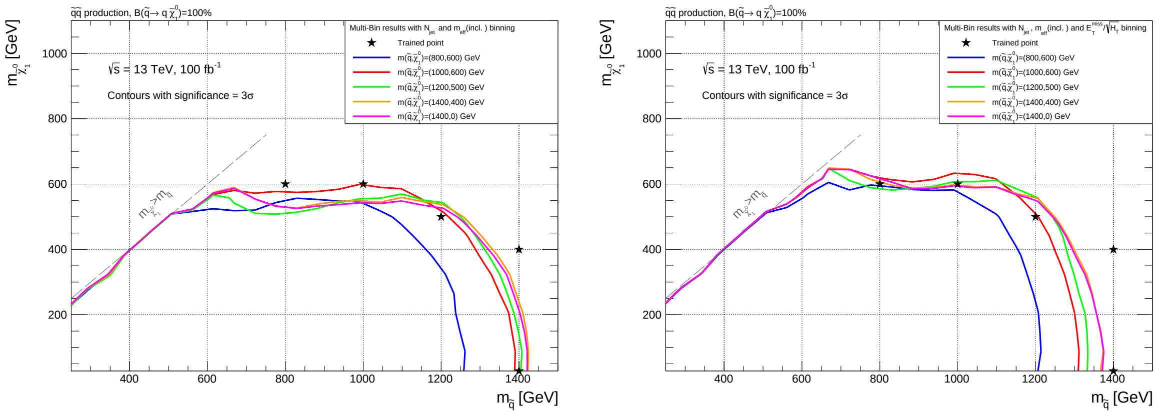
Moreover, the Multi-Bin approach with N_{jet}, $m_{\text{eff}}(\text{incl.})$ and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ bins gains in the regions with lower $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ values. However, in the region with higher $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ it is getting even less sensitive than the N_{jet} separated Cut & Count method. This behaviour is most likely due to the low statistics in each bin since the estimated systematic uncertainties depend on the number of background events (see eq. (6.10)).

7.4.4 Reduction Algorithm

Since the training was performed for five signal points, the number of bins for the Multi-Bin analysis increases by factor 5, i.e. 120 bins for the N_{jet} & $m_{\text{eff}}(\text{incl.})$ binning and in 720 bins by adding also $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ categories. This large amount of bins increases the complexity of the analysis in particular the number of control regions would increase significantly (in general for each signal region there are 4 control regions for top, W, Z and QCD backgrounds). Therefore the number of bins has to be reduced.
First, the five training points considered in figure 7.16:

1. $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 0)$ GeV,
2. $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 400)$ GeV,
3. $m(\tilde{q}, \tilde{\chi}_1^0) = (1200, 500)$ GeV,
4. $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 600)$ GeV and
5. $m(\tilde{q}, \tilde{\chi}_1^0) = (800, 600)$ GeV,

are shown as separated contour lines in figure 7.17 for a) N_{jet} and $m_{\text{eff}}(\text{incl.})$ binning and b) N_{jet}, $m_{\text{eff}}(\text{incl.})$ and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ binning in order to identify training points which do not contribute to the combined contour.



a) b)

Figure 7.17: Results for the Multi-Bin approach for five single training points using a) N_{jet} and $m_{\text{eff}}(\text{incl.})$ bins and b) N_{jet}, $m_{\text{eff}}(\text{incl.})$ and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ bins illustrated as 3σ contour lines.

For both Multi-Bin approaches only two training points mainly contribute to the outermost contour line:

1. $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 400)$ GeV and

2. $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 600)$ GeV.

By keeping only these two training points, the number of bins decreases to 48 and 288 for N_{jet} & m_{eff}(incl.) binning and additional $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ binning, respectively.

The next step is to take a look at the results of the single bins, how much they contribute to the quadratic sum at the end. These definitions are shown in table 7.6 for the training point $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 400)$ GeV and in table 7.7 for $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 600)$ GeV for the N_{jet} and m_{eff}(incl.) separated Multi-Bin.

Rank	N_{jet}	m_{eff}(incl.) [GeV]	Bin Nr.	N_{signal}	N_{bkg}	Z	Z_{comb}	ΔZ_{comb}	Retrieved Cuts $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ [GeV^{-1}]	$\Delta\phi$	$\max(\lvert\eta\rvert_{i=1,2})$	$p_{\text{T}}(j_2)$ [GeV]
1	==2	2400–2800	4	3.46	4.86	1.03	1.03	1.03	≥ 20	≥ 0.8	≤ 1.2	≥ 300
2	==3	2400–2800	10	5.23	13.52	0.97	1.41	0.38	≥ 20	≥ 0.4	≤ 1.6	≥ 300
3	==3	2000–2400	9	6.78	36.98	0.73	1.59	0.18	≥ 20	≥ 0.4	≤ 2.0	≥ 300
4	==4	2400–2800	16	4.29	15.58	0.73	1.75	0.16	≥ 16	≥ 0.8	≤ 2.0	≥ 250
5	==2	2000–2400	3	5.43	32.16	0.62	1.86	0.11	≥ 18	≥ 0.8	≤ 1.6	≥ 300
6	≥ 5	2800–3200	23	2.27	5.48	0.60	1.95	0.09	≥ 16	≥ 0.8	≤ 2.0	≥ 200
7	==3	2800–3200	11	2.27	5.71	0.59	2.04	0.09	≥ 18	≥ 0.8	≤ 1.6	≥ 300
8	≥ 5	2400–2800	22	3.52	16.66	0.56	2.12	0.08	≥ 14	≥ 0.8	≤ 1.6	≥ 200
9	==2	2800–3200	5	1.61	3.48	0.48	2.17	0.05	≥ 18	≥ 0.8	≤ 1.6	≥ 300
10	==4	2800–3200	17	1.92	6.53	0.44	2.21	0.04	≥ 14	≥ 0.8	≤ 2.0	≥ 200
11	==4	≥ 3200	18	1.39	4.27	0.35	2.24	0.03	≥ 16	≥ 0.4	≤ 1.2	≥ 150
12	≥ 5	≥ 3200	24	1.52	5.44	0.35	2.27	0.03	≥ 16	≥ 0.4	≤ 2.0	≥ 200
13	==4	2000–2400	15	3.09	28.85	0.34	2.29	0.02	≥ 18	≥ 0.8	≤ 1.6	≥ 250
14	==3	≥ 3200	12	1.51	6.16	0.32	2.32	0.03	≥ 16	≥ 0.4	≤ 2.0	≥ 150
15	≥ 5	2000–2400	21	3.09	44.17	0.26	2.33	0.01	≥ 16	≥ 0.8	≤ 2.0	≥ 150
16	==2	≥ 3200	6	1.05	3.71	0.23	2.34	0.01	≥ 14	≥ 0.8	≤ 2.0	≥ 100
17	==2	1600–2000	2	4.84	109.75	0.22	2.35	0.01	≥ 18	≥ 0.8	≤ 2.0	≥ 300
18	==3	1600–2000	8	3.54	87.34	0.18	2.36	0.01	≥ 18	≥ 0.8	≤ 2.0	≥ 300
19	==4	1600–2000	14	1.23	42.68	0.04	2.36	0.00	≥ 18	≥ 0.4	≤ 1.6	≥ 300
20	≥ 5	1600–2000	20	1.43	90.96	0.02	2.36	0.00	≥ 16	≥ 0.8	≤ 2.0	≥ 200
21	==2	1200–1600	1	1.45	334.20	0.00	2.36	0.00	≥ 14	≥ 0.8	≤ 2.0	≥ 300
22	==3	1200–1600	7	1.97	557.06	0.00	2.36	0.00	≥ 18	≥ 0.4	≤ 2.0	≥ 200
23	==4	1200–1600	13	0.52	254.40	0.00	2.36	0.00	≥ 18	≥ 0.4	≤ 2.0	≥ 200
24	≥ 5	1200–1600	19	0.86	686.81	0.00	2.36	0.00	≥ 14	≥ 0.4	≤ 2.0	≥ 50

Table 7.6: Event yields for the individual bins of the Multi-Bin approach with N_{jet} and m_{eff}(incl.) bins for the training point $m(\tilde{q}, \tilde{\chi}_1^0) = (1400, 400)$ GeV. Z denotes the significance of the single bin, Z_{comb} is the sequentially quadratically added significance up to the current line and ΔZ_{comb} is the additional significance the bin contributes to the quadratic sum. The bin numbering corresponds to the N_{jet} and m_{eff}(incl.) values. $\Delta\phi$ is the abbreviation for $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$.

Rank	N_{jet}	m_{eff}(incl.) [GeV]	Bin Nr.	N_{signal}	N_{bkg}	Z	Z_{comb}	ΔZ_{comb}	Retrieved Cuts $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ [GeV^{-1}]	$\Delta\phi$	$\max(\lvert\eta\rvert_{i=1,2})$	$p_{\text{T}}(j_2)$ [GeV]
1	==3	1200–1600	7	54.52	374.28	1.23	1.23	1.23	≥ 16	≥ 0.8	≤ 1.2	≥ 200
2	==4	1200–1600	13	30.85	189.88	1.17	1.70	0.47	≥ 16	≥ 0.8	≤ 1.2	≥ 200
3	==2	1200–1600	1	34.67	236.67	1.12	2.03	0.33	≥ 16	≥ 0.8	≤ 0.8	≥ 200
4	==4	1600–2000	14	22.12	130.37	1.10	2.31	0.28	≥ 14	≥ 0.8	≤ 2.0	≥ 250
5	==3	1600–2000	8	21.00	167.59	0.87	2.47	0.16	≥ 14	≥ 0.8	≤ 1.2	≥ 150
6	≥ 5	1600–2000	20	16.32	125.84	0.83	2.61	0.14	≥ 14	≥ 0.8	≤ 1.6	≥ 150
7	≥ 5	2000–2400	21	8.42	48.39	0.78	2.72	0.11	≥ 14	≥ 0.8	≤ 2.0	≥ 200
8	≥ 5	1200–1600	19	14.50	167.07	0.60	2.79	0.07	≥ 14	≥ 0.8	≤ 1.6	≥ 200
9	==4	2000–2400	15	4.77	43.03	0.44	2.82	0.03	≥ 14	≥ 0.4	≤ 1.2	≥ 250
10	==4	2800–3200	17	1.84	8.49	0.36	2.84	0.02	≥ 14	≥ 0.4	≤ 2.0	≥ 250
11	==3	2400–2800	10	3.16	30.71	0.33	2.86	0.02	≥ 14	≥ 0.4	≤ 1.6	≥ 200
12	==2	1600–2000	2	5.90	94.45	0.32	2.88	0.02	≥ 14	≥ 0.8	≤ 0.8	≥ 150
13	==3	2000–2400	9	2.61	26.27	0.29	2.89	0.01	≥ 14	≥ 0.8	≤ 0.8	≥ 200
14	==4	2400–2800	16	1.98	15.71	0.27	2.91	0.02	≥ 18	≥ 0.8	≤ 2.0	≥ 200
15	==4	≥ 3200	18	1.25	5.09	0.27	2.92	0.01	≥ 14	≥ 0.4	≤ 1.6	≥ 300
16	≥ 5	2800–3200	23	1.37	6.48	0.27	2.93	0.01	≥ 14	≥ 0.8	≤ 1.6	≥ 150
17	≥ 5	2400–2800	22	2.06	19.20	0.25	2.94	0.01	≥ 14	≥ 0.8	≤ 2.0	≥ 150
18	≥ 5	≥ 3200	24	1.05	4.57	0.20	2.95	0.01	≥ 18	≥ 0.4	≤ 1.6	≥ 200
19	==2	2000–2400	3	2.80	97.38	0.11	2.95	0.00	≥ 14	≥ 0.8	≤ 1.6	≥ 50
20	==3	2800–3200	11	0,66	8.03	0.03	2.95	0.00	≥ 14	≥ 0.8	≤ 2.0	≥ 200
21	==2	2400–2800	4	0.16	16.65	0.00	2.95	0.00	≥ 16	≥ 0.8	≤ 1.2	≥ 50
22	==2	2800–3200	5	0.00	11.26	0.00	2.95	0.00	≥ 14	≥ 0.4	≤ 2.0	≥ 50
23	==2	≥ 3200	6	0.00	5.94	0.00	2.95	0.00	≥ 14	≥ 0.4	≤ 2.0	≥ 50
24	==3	≥ 3200	12	0.33	3.48	0.00	2.95	0.00	≥ 20	≥ 0.4	≤ 1.2	≥ 300

Table 7.7: Event yields for the individual bins of the Multi-Bin approach with N_{jet} and m_{eff}(incl.) bins for the training point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 600)$ GeV. Z denotes the significance of the single bin, Z_{comb} is the sequentially quadratically added significance up to the current line and ΔZ_{comb} is the additional significance the bin contributes to the quadratic sum. The bin numbering corresponds to the N_{jet} and m_{eff}(incl.) values. $\Delta\phi$ is the abbreviation for $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$.

The results in these two tables are sorted and ranked according to their significance Z. The bin number is defined by the N_{jet} and m_{eff}(incl.) ranges also denoted in the table itself. However the bin numbering itself is not important while the fact that the bin numbers are the same for both tables is relevant. Although the N_{jet} and m_{eff}(incl.) cuts are the same for the individual bin numbers, the $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$, $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$, $\max(|\eta|_{i=1,2})$ and $p_{\text{T}}(j_2)$ cuts differ. These (optimised) cuts are also shown in the table.
Further, the Z_{comb} is the sequential quadratic sum of all previous significances up to this point in the table. It demonstrates how much one bin contributes to the final significance. In general, the first half of the bins contribute significantly whereas the second half causes only minor improvements. Additionally, the first nine highest ranked bins have no bin in common between the two tables.

This behaviour suggests to combine the two results into one Multi-Bin. For this purpose an algorithm is introduced following the procedure:

1. Rank the bins by their significance (as shown in the tables 7.6 and 7.7),
2. iterate over the bins (by their bin number) and choose the bin definition (they differ by their cuts) with the higher rank,
3. if they have the same rank, choose the one with the higher absolute significance value.

This approach can be applied to both Multi-Bin cases and the number of bins will be halved. However, considering the gain versus the complexity, the Multi-Bin with N_{jet} and m_{eff}(incl.) is the more promising method in order to be able to control systematic uncertainties and the amount of control regions later on.

The resulting contour line using this algorithm for the N_{jet} and m_{eff}(incl.) Multi-Bin is shown in figure 7.18. There is only a minor difference between the contour line of the combination of all points and the algorithm introduced before. For the proper implementation in the analysis, this difference would be covered by error bands.

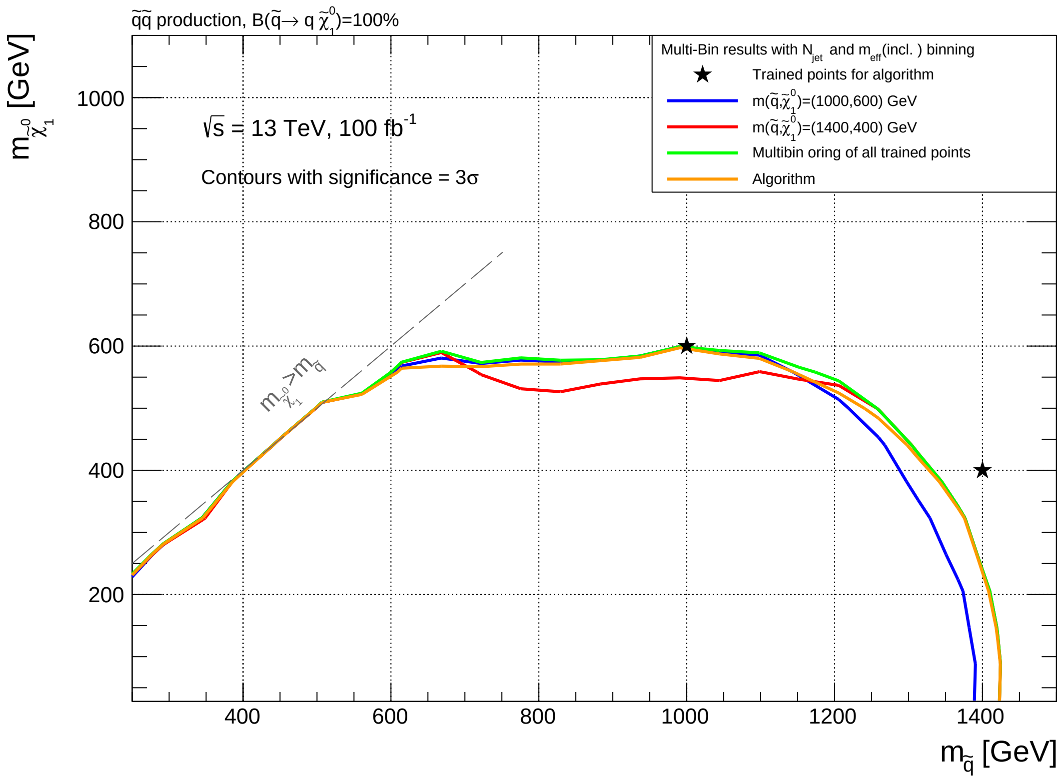


Figure 7.18: 3σ contour lines of the N_{jet} & m_{eff}(incl.) separated Multi-Bin results for two training points separately (blue and red contour), oring of all training points (green contour) and the results from the introduced algorithm (orange contour).

Since these 24 bins cover the whole $m(\tilde{q}, \tilde{\chi}_1^0)$-range at the end, there is no need for signal regions anymore, as used in the 0-lepton publication described in chapter 5.

The Multi-Bin results as well as the results presented before are summarised in figure 7.19. The BDT method showed first improvements with respect to the reoptimised Cut & Count method.
The slicing into N_{jet} improved the sensitivity for both, BDT and Cut & Count significantly. Furthermore, the division into more categories, $m_{\text{eff}}(\text{incl.})$ and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$, provided a further gain, at least in some regions.
In conclusion, the results of the algorithm with 24 bins (N_{jet} & $m_{\text{eff}}(\text{incl.})$ bins), which is a reasonable number considering the complexity, is a good choice for further investigations towards a possible implementation in the next round of the analysis.

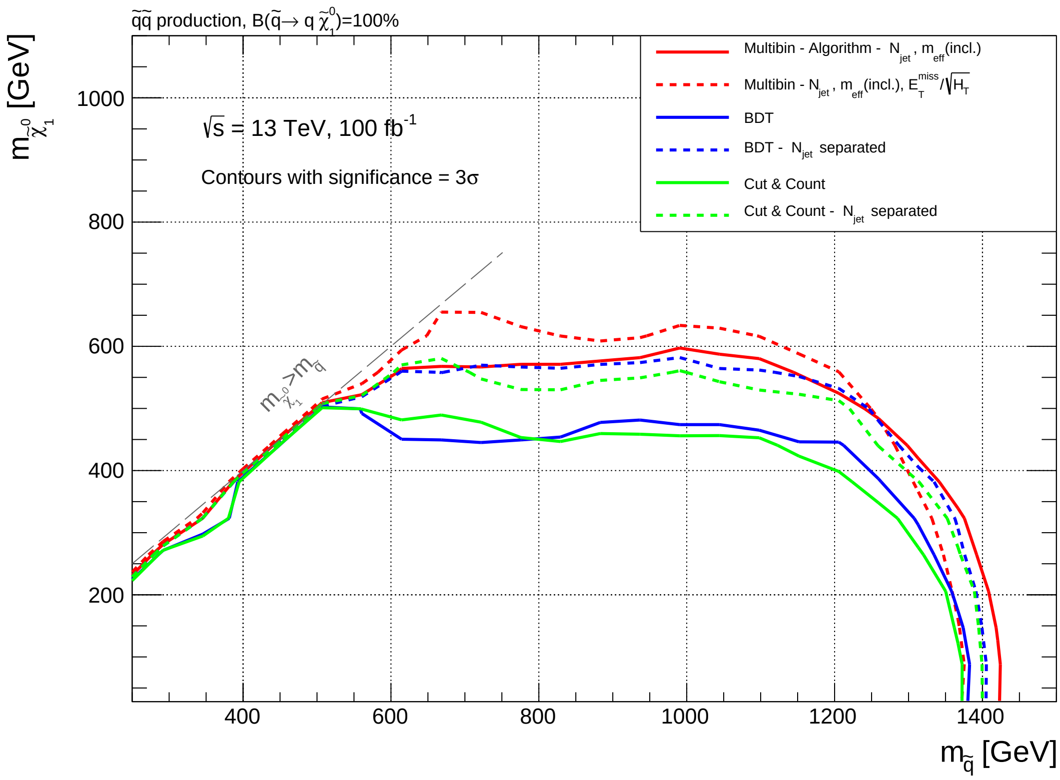


Figure 7.19: Summary plot with the 3σ contour lines of the combined (using oring) Multi-Bin (red contours), BDT (blue contours) and Cut & Count (green contours) results.

8 Gluino Pair Production Studies

After a significant gain in the redefinition of the squark signal regions as described in the previous chapter, a similar procedure will be applied to improve the gluino signal regions.

Due to the good performance of the LHC, the anticipated luminosity for the end of RUN II was corrected to be $\sim 120\,\text{fb}^{-1}$ at the beginning of this study. Since the squark and gluino studies are completely independent of each other, the further studies for the gluino signal region optimisation are done using $120\,\text{fb}^{-1}$.

The target signal is the gluino-pair production where the gluinos decay directly into a neutralino and two quarks, shown in figure 8.1.

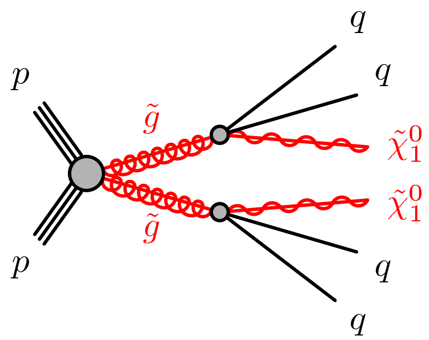


Figure 8.1: The decay topology of gluino-pair production in the simplified models with direct gluino decay [32].

The gluino studies are fully based on the Monte Carlo samples described in table 3.1. Just for completeness, in addition to the W+jets, Z+jets, top and diboson backgrounds also the Multi-jet (QCD) background is included in the following studies but it is still negligible.

8.1 Optimisation Strategy

In the case of the direct squark decays, the Multi-Bin approach was the most promising method. However, it is not clear which strategy is the best for the direct gluino decays. Therefore a brief scan of the different methods is done to decide the further optimisation strategy.

The following methods are tested:

- BDT
  1. single point training (as for squarks),
  2. combining several signal points for training (SRs),
  3. using different boosting techniques (AdaBoost, GradBoost),
  4. N_{jet} separated training,
- Cut & Count
  1. reoptimised for $120\,\text{fb}^{-1}$ and
  2. N_{jet} separated.

For the BDT training all variables listed in table 5.2 are used for a brief comparison. For the Cut & Count method, the following variables are used for optimisation: $m_{\text{eff}}(\text{incl.})$, $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4j)$, $p_{\text{T}}(j_{1,4})$, $\max(|\eta|_{i=1-4})$ and $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ and the cuts from the remaining variables are set to the values from the 0-lepton publication.

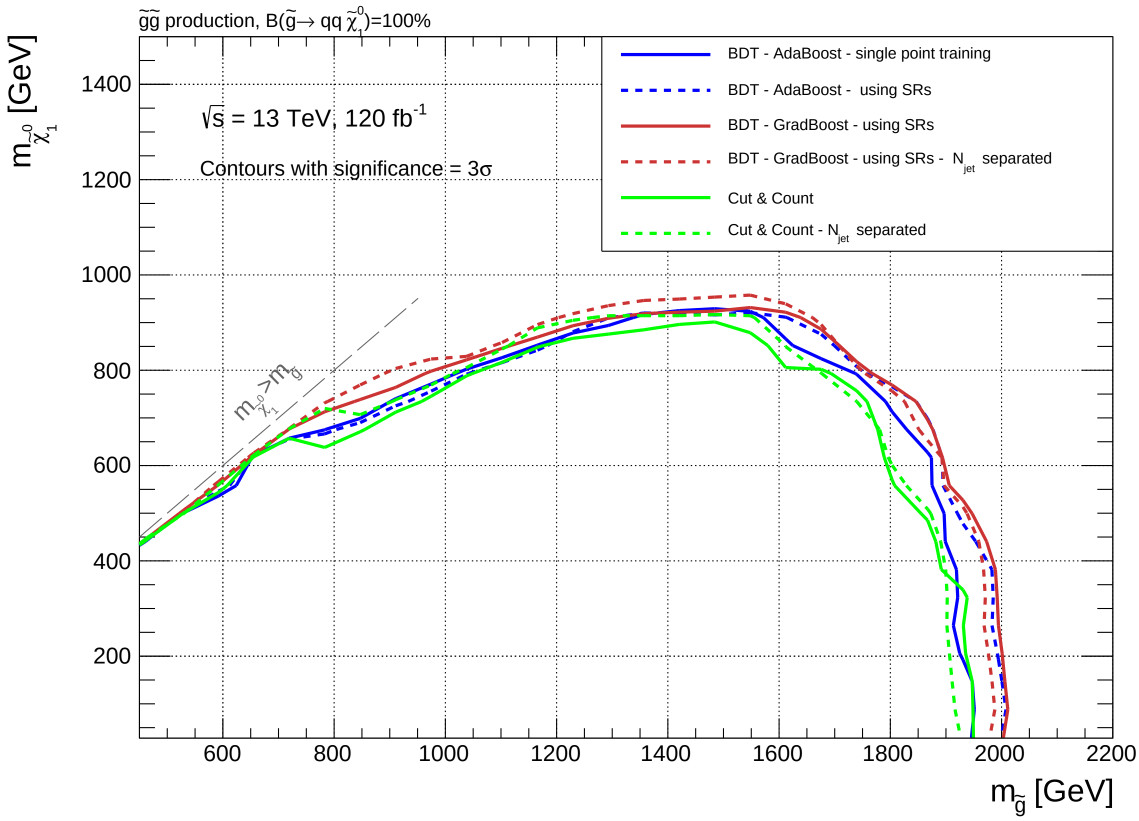


Figure 8.2: The 3σ contour lines for different BDT and Cut & Count setups with the purpose of a brief method scan. All contours are results of several training points (oring).

The results of the brief scanning are shown in figure 8.2. The N_{jet} slicing in general does not provide such a significant gain as achieved for the squarks, and is negligible considering the higher complexity. The BDTs improve especially the high $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$ regions. Also, the usage of signal regions, meaning the combinations of kinematically similar signal points for the BDT training (more detailed in sec. 8.2), results in better sensitivity (the signal regions are described in more detail below). The two boosting algorithms, AdaBoost and GradBoost, provide similar results whereas the GradBoost is slightly better.

In conclusion, the best technique from this rough scan is the BDT with GradBoost combined with the usage of signal regions. Therefore, in the following this combination is used as baseline for the studies. The focus of the studies is choosing input variables and finding a compromise between gain in significance and the complexity of the analysis.

8.2 Signal Region Definitions

After a rough estimation, the best approach was found to be to group several mass points to one signal region. The mass points for one signal region are chosen such that the event kinematics are similar for each point. This is the case for similar $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$ values.
The definitions of the BDT signal regions are listed in table 8.1 and are also illustrated in figure 8.3.

Training samples	BDT1	BDT2	BDT3	BDT4	BDT5
$m(\tilde{g}, \tilde{\chi}_1^0)$ [GeV]	$(2000, 600)$	$(1900, 900)$	$(1500, 900)$	$(1100, 900)$	$(1050, 950)$
	$(2100, 700)$	$(1950, 850)$	$(1550, 850)$	$(1150, 850)$	$(1150, 1050)$
	$(2100, 500)$	$(2000, 800)$	$(1600, 800)$	$(1200, 800)$	$(1250, 1150)$
	$(2300, 500)$	$(1900, 700)$	$(1600, 1000)$	$(1300, 900)$	$(1000, 800)$
	$(2200, 400)$	$(1800, 800)$	$(1700, 900)$	$(1350, 850)$	$(1100, 900)$
	$(2200, 600)$	$(1850, 750)$	$(1750, 850)$		$(1300, 1100)$
	$(2200, 800)$				$(1200, 1000)$
	$(2300, 700)$				
	$(2400, 600)$				
	$(2400, 800)$				
$\Delta m(\tilde{g}, \tilde{\chi}_1^0)$	$O(1.4 - 1.8)$ TeV	$O(1.0 - 1.2)$ TeV	$O(600 - 900)$ GeV	$O(200 - 500)$ GeV	$O(100 - 200)$ GeV

Table 8.1: BDT signal region definitions for gluino model with direct decay.

The re-grouped signal regions have the advantage that the statistics in the training increases and a possible overfitting gets less probable. One signal sample (one signal point) includes around $5,000 - 10,000$ events and the total signal statistic of the BDT2 signal region corresponds to $\sim 42,500$ events[1]. However, after the preselection cuts and the separation into training and testing samples, for the BDT2 signal region only $\sim 15,300$ signal events remain for each category and $\sim 108,000$ background events (before selection: $\sim 1.7 \cdot 10^6$). Therefore the combination of several signal points is useful to reduce statistical fluctuations.

[1] The number of events are here only the generated events, not the number of signal events corresponding to given luminosity.

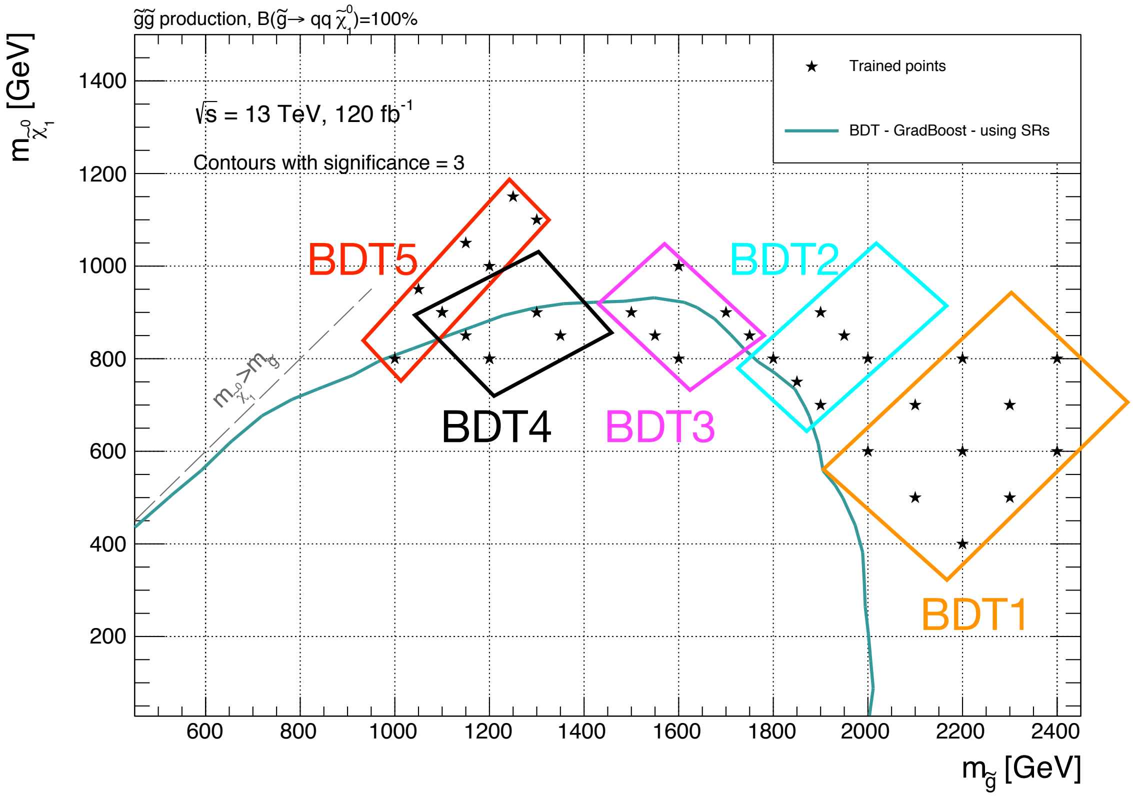


Figure 8.3: Definition of the five BDT based signal regions for the gluino model with direct decays. The contour line corresponds to the BDT with the new signal region definitions considered for the rough method estimation.

In addition, the training is not too much focused on one single signal point.

Since the BDT signal regions four and five are so close together, they share one point with $m(\tilde{g}, \tilde{\chi}_1^0) = (1100, 900)$ GeV. This mass point is included in the training of both signal regions. There is no bias or over-constraint through this double usage since the signal regions are complementary and never used together for the same mass point.

8.2.1 Preselection

Before the BDT training, loose preselection cuts are applied in each signal region to reduce background contributions and to speed up the BDT training. These preselection cuts are shown in table 8.2 for all the BDT based signal regions.

In the following these five signal regions are always used for the BDT training including the preselection cuts.

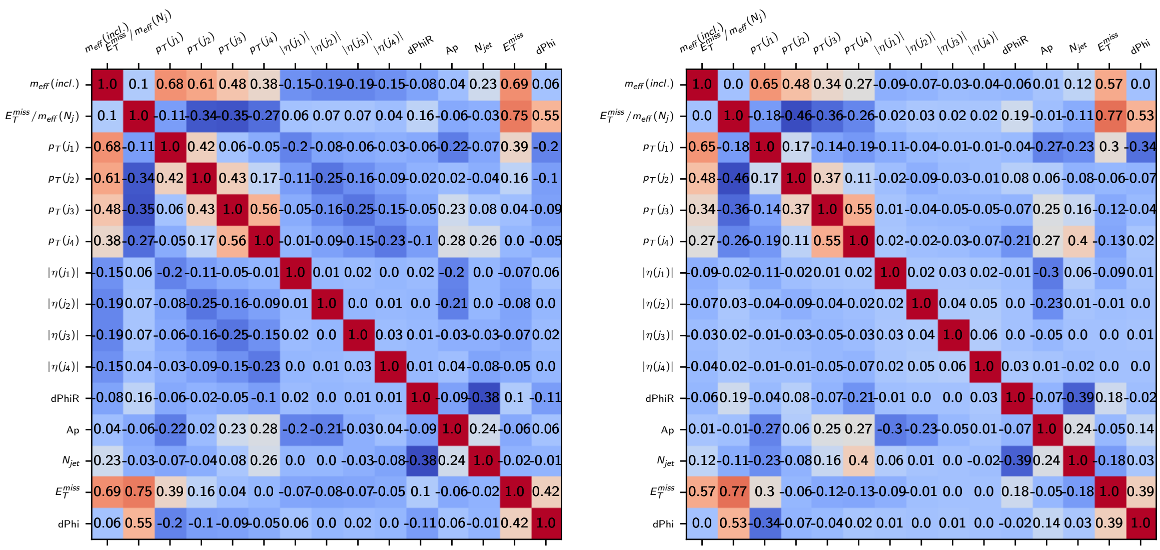
Targeted signal	$\tilde{g}\tilde{g}, \tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^0$				
Requirement	BDT based signal regions				
	BDT1	**BDT2**	**BDT3**	**BDT4**	**BDT5**
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 300				
$p_{\mathrm{T}}(j_1)$ [GeV]	> 200				
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	≥ 0.4				
N_{jet}	≥ 4				≥ 2
$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$	> 0.2				–
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ [GeV$^{1/2}$]	–				> 12
$m_{\mathrm{eff}}(\mathrm{incl.})$ [GeV]	> 1400		> 800		> 1400

Table 8.2: Preselection cuts for the five BDT based signal regions.

8.3 Variable Study for BDT Input

The focus of the gluino studies lies on optimising the BDT input variables and trying to reduce their number without loosing too much sensitivity.
In the following an exemplary variable study with the BDT based signal region BDT2 is performed. Afterwards, the same procedure will be applied to the remaining BDT based signal regions and presented in appendix B.3.
The variables which are used in the 0-lepton publication are described in section 5. In order to improve the BDT classification, different BDT input variable combinations are considered.

The correlation matrix for the background and the signal for these variables is shown in figure 8.4 as well as the single variable distributions in figure 8.5, B.1 and B.2.



a) b)

Figure 8.4: Linear correlation matrix of a) signal of BDT based signal region BDT2 and b) background for the main variables considered for input to BDT.

The separation of the single variables is shown in the corresponding variable plots and also listed in table 8.3. For the further studies the impact of these variables as BDT input variable is tested.

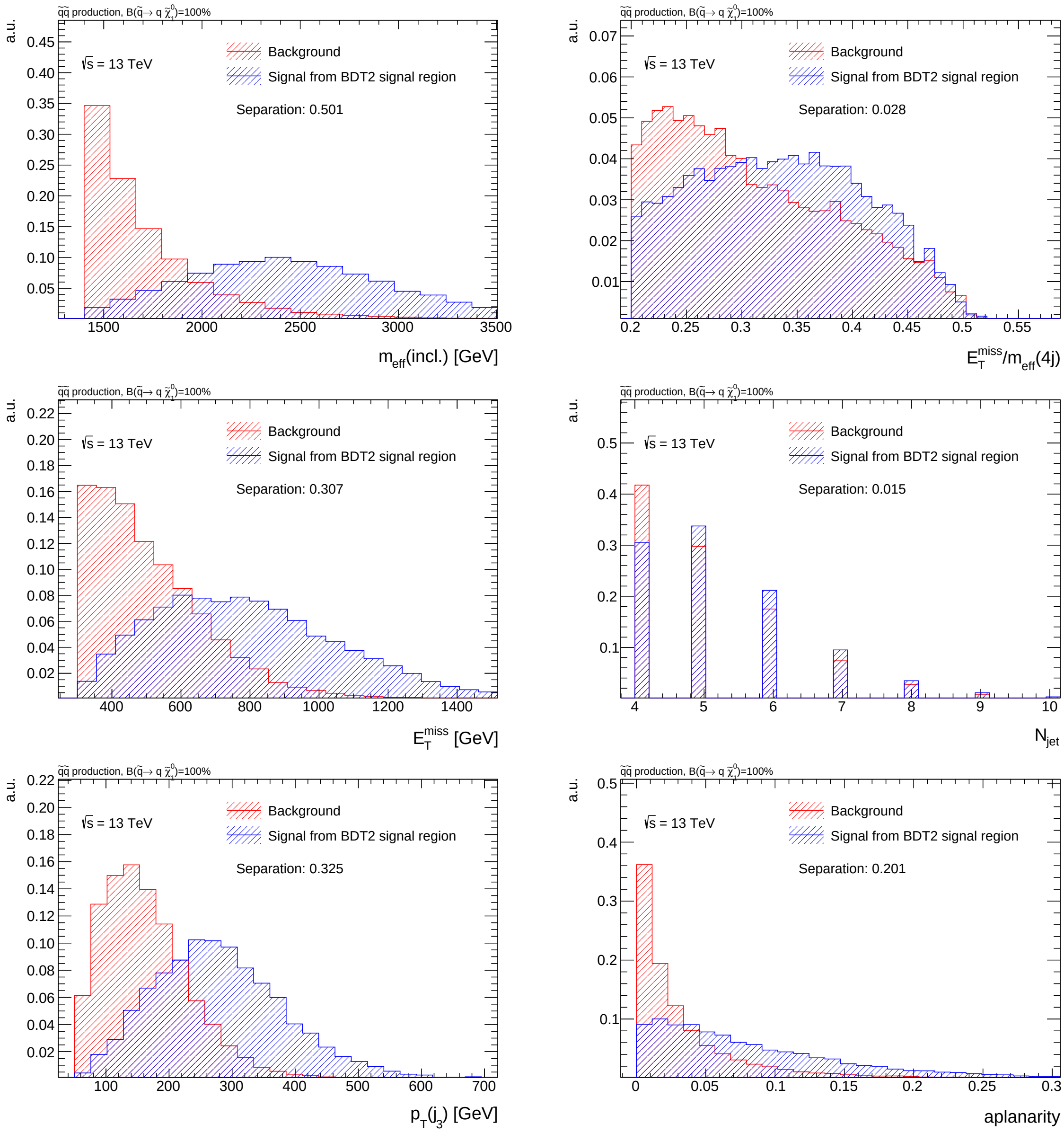


Figure 8.5: Normalised variable distributions: $m_{\text{eff}}(\text{incl.})$, $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(N_j)$, $E_{\text{T}}^{\text{miss}}$, N_{jet}, $p_{\text{T}}(j_3)$, and aplanarity. The background is shown in red and the signal of the BDT2 signal region in blue. In each plot the separation power of the background and signal distribution is given.

8.3.1 Variable Scan m_{eff}, $E_{\text{T}}^{\text{miss}}$, H_{T} and $|\eta(\text{jets})|$

First, a variable scan is performed with different combinations of m_{eff}, $E_{\text{T}}^{\text{miss}}$, H_{T} and $|\eta(\text{jets})|$ while keeping the BDT input variables: $p_{\text{T}}(j_{1-4})$, $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ and aplanarity. For the purpose of this study no preselection cut is applied on $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4j)$. The detailed information about the input variables for this study are listed in table 8.4. The corresponding contour plots with 3σ significance are shown in figure 8.6.

All contour lines are very similar to each other, especially in the region for which this BDT

Variable	Separation in %	Variable	Separation in %
$m_{\text{eff}}(\text{incl.})$	50.1	$\lvert\eta(j_2)\rvert$	6.7
$p_{\text{T}}(j_3)$	32.5	$\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$	6.1
$E_{\text{T}}^{\text{miss}}$	30.7	$\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$	5.4
$p_{\text{T}}(j_2)$	27.8	$\lvert\eta(j_3)\rvert$	3.7
$p_{\text{T}}(j_4)$	23.7	$E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(N_j)$	2.8
aplanarity	20.1	N_{jet}	1.5
$\lvert\eta(j_1)\rvert$	10.5	$\lvert\eta(j_4)\rvert$	0.7
$p_{\text{T}}(j_1)$	6.9		

Table 8.3: Separation of the variables used in the 0-lepton publication signal regions targeting gluino pair production in simplified models.

N°	$p_{\text{T}}(j_{1-4})$	$\lvert\eta\rvert(j_{1-4})$	$\max(\lvert\eta\rvert_{i=1-4})$	$\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$	Ap	$m_{\text{eff}}(\text{incl.})$	$E_{\text{T}}^{\text{miss}}$	Additional
1	✓	✓	–	✓	✓	✓	–	$E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4j)$, N_{jet}
2	✓	✓	–	✓	✓	✓	–	$E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4j)$
3	✓	–	✓	✓	✓	✓	–	$E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4j)$
4	✓	–	✓	✓	✓	✓	–	$E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(\text{incl.})$
5	✓	–	✓	✓	✓	✓	–	$E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}(\text{incl.})}$
6	✓	–	✓	✓	✓	✓	✓	–
7	✓	–	✓	✓	✓	–	✓	$H_{\text{T}}(\text{incl.})$
8	✓	–	✓	✓	✓	–	✓	$H_{\text{T}}(4j)$
9	✓	–	✓	✓	✓	–	✓	–

Table 8.4: BDT input variable combinations for the variable study with different combinations of m_{eff}, $E_{\text{T}}^{\text{miss}}$, H_{T} and $\lvert\eta(\text{jets})\rvert$. No cut on $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4j)$ is applied in the pre-selection. The preferred input variable choice is marked in gray.

signal region is designed. The region of interest is located near the trained points, indicated as stars. Since there are almost no differences visible one has the possibility to choose the simplest possible case to reduce the complexity. Therefore, this first variable scan yields to the nine variables (also marked gray in table 8.4):

1.-4. $p_{\text{T}}(j_{1-4})$,

5. aplanarity,
6. $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$,
7. $\max(\lvert\eta\rvert_{i=1-4})$,

8. m_{eff}(incl.) and
9. $E_{\text{T}}^{\text{miss}}$.

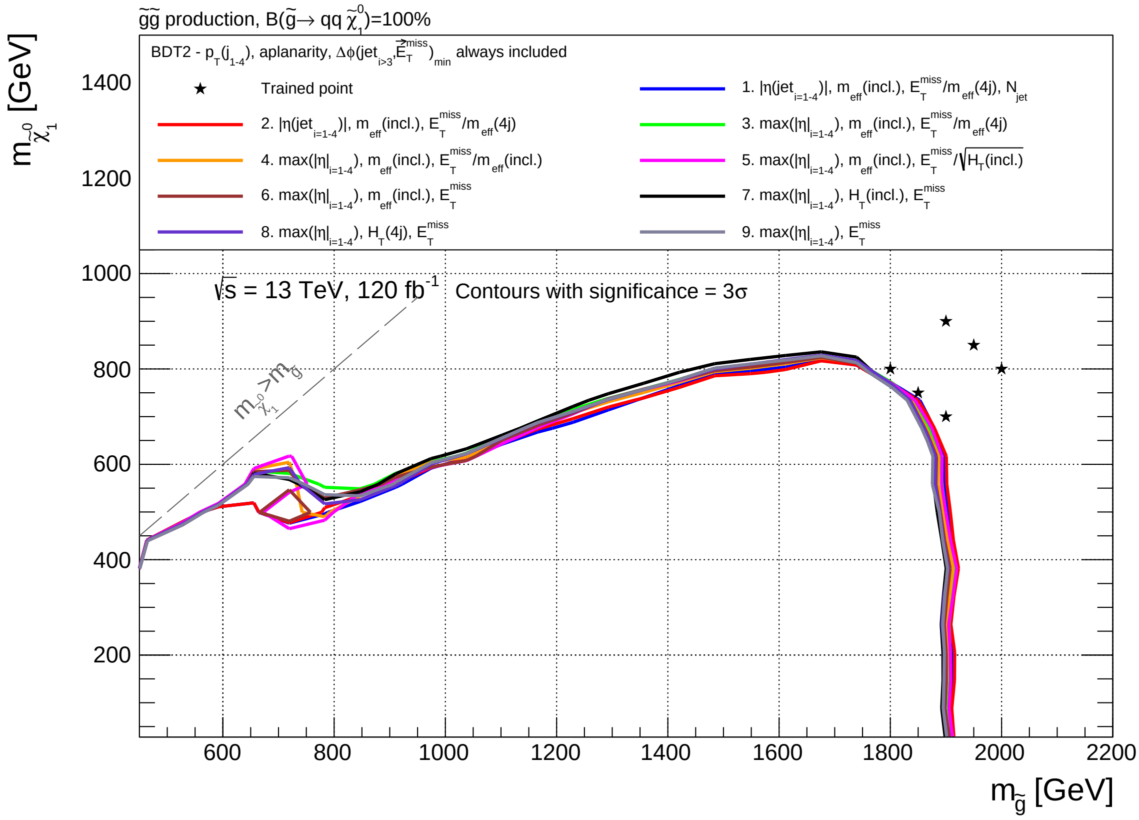


Figure 8.6: 3σ contour lines for different combinations of m_{eff}, $E_{\text{T}}^{\text{miss}}$, H_{T} and $|\eta(\text{jets})|$. No cut on $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4j)$ was applied in the pre-selection. The variable combination of study 6 shown in brown is the preferred choice.

In addition, the $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ and N_{jet} variables were tested as well, but there was no gain by including them as BDT input and hence these studies are not shown in detail.

8.3.2 Impact of $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ Variable on BDT Performance

Secondly, the impact of the $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ variable as BDT input is studied. Three different scenarios are considered for this variable: usage as BDT input variable, applying different preselection cuts without considering it as BDT input and completely discarding it as shown in detail in table 8.5. The resulting 3σ contour lines for these scenarios are illustrated in figure 8.7.
The scenario in which $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ is not considered at all provides the lowest sensitivity while the results of the other three attempts are very similar especially in the $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$ region the BDT2 is designed for.
Therefore also in this case it is possible to choose the less complex scenario using a preselection cut on $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ which reduces the number of input variables to eight. The looser

selection , meaning $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\min} \geq 0.2$ (highlighted in gray in table 8.5) is slightly better than the tighter cut and also allows to keep more statistics.

N°	$p_{\mathrm{T}}(j_{1-4})$	$\lvert\eta\rvert(j_{1-4})$	$\max(\lvert\eta\rvert_{i=1-4})$	$\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\min}$	Ap	$m_{\mathrm{eff}}(\text{incl.})$	$E_{\mathrm{T}}^{\mathrm{miss}}$
1	✓	✓	✓	✓	✓	✓	✓
2	✓	✓	✓	–	✓	✓	✓
3	✓	✓	✓	≥ 0.2	✓	✓	✓
4	✓	✓	✓	≥ 0.4	✓	✓	✓

Table 8.5: BDT input variable combinations for the variable study for $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\min}$. Preferred input variable choice marked in gray.

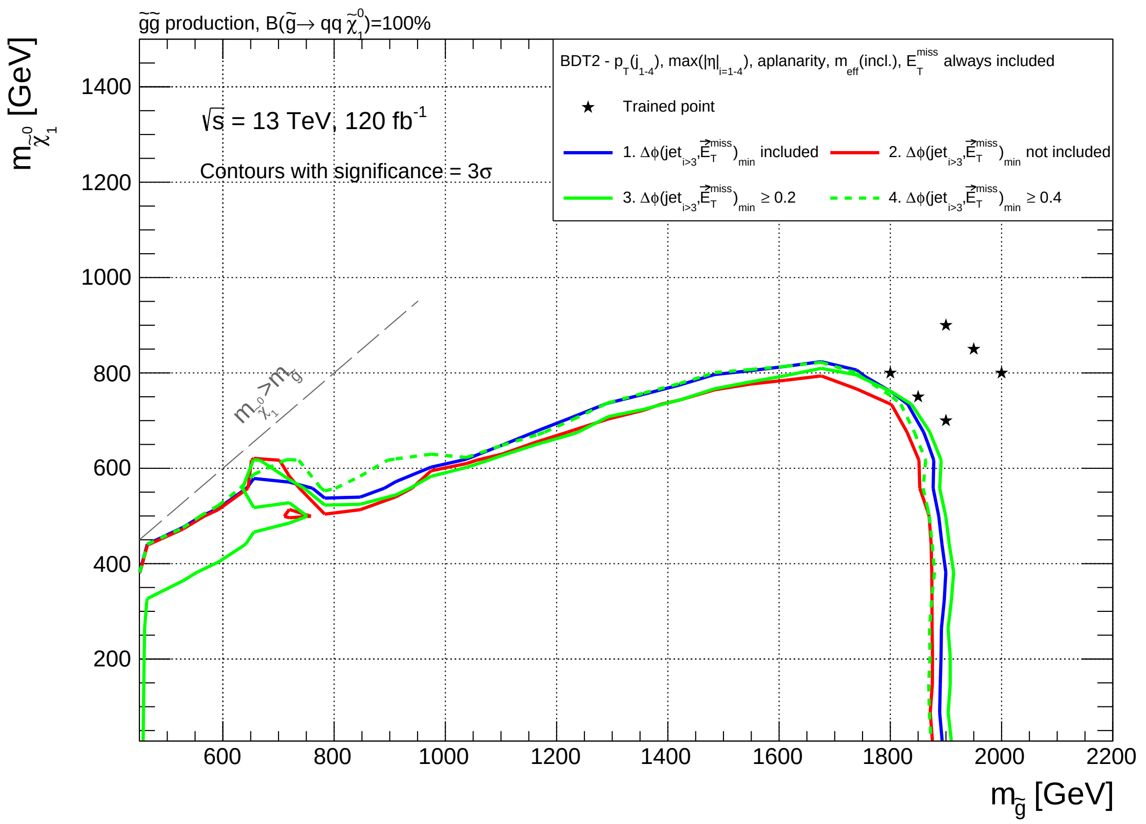


Figure 8.7: 3σ contour lines for different combinations of $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\min}$. The looser preselection (green solid contour) is the preferred choice.

8.3.3 Impact of Aplanarity Variable on BDT Performance

Moreover the impact of the aplanarity as BDT input variable is studied in like manner as for $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$. The different scenarios are listed in table 8.6 and the results are shown as 3σ contour lines in figure 8.8.
Both, discarding the aplanarity as well as applying a preselection cut provide less sensitive results than using the aplanarity as BDT input variable and hence the aplanarity will be further used as input variable for BDT also indicated in gray in table 8.6.

N°	$p_{\text{T}}(j_{1-4})$	$\lvert\eta\rvert(j_{1-4})$	$\max(\lvert\eta\rvert_{i=1-4})$	$\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$	Ap	$m_{\text{eff}}(\text{incl.})$	$E_{\text{T}}^{\text{miss}}$
1	✓	✓	✓	✓	✓	✓	✓
2	✓	✓	✓	✓	≥ 0.04	✓	✓
3	✓	✓	✓	✓	–	✓	✓

Table 8.6: BDT input variable combinations for the variable study for aplanarity. Preferred input variable choice marked in gray.

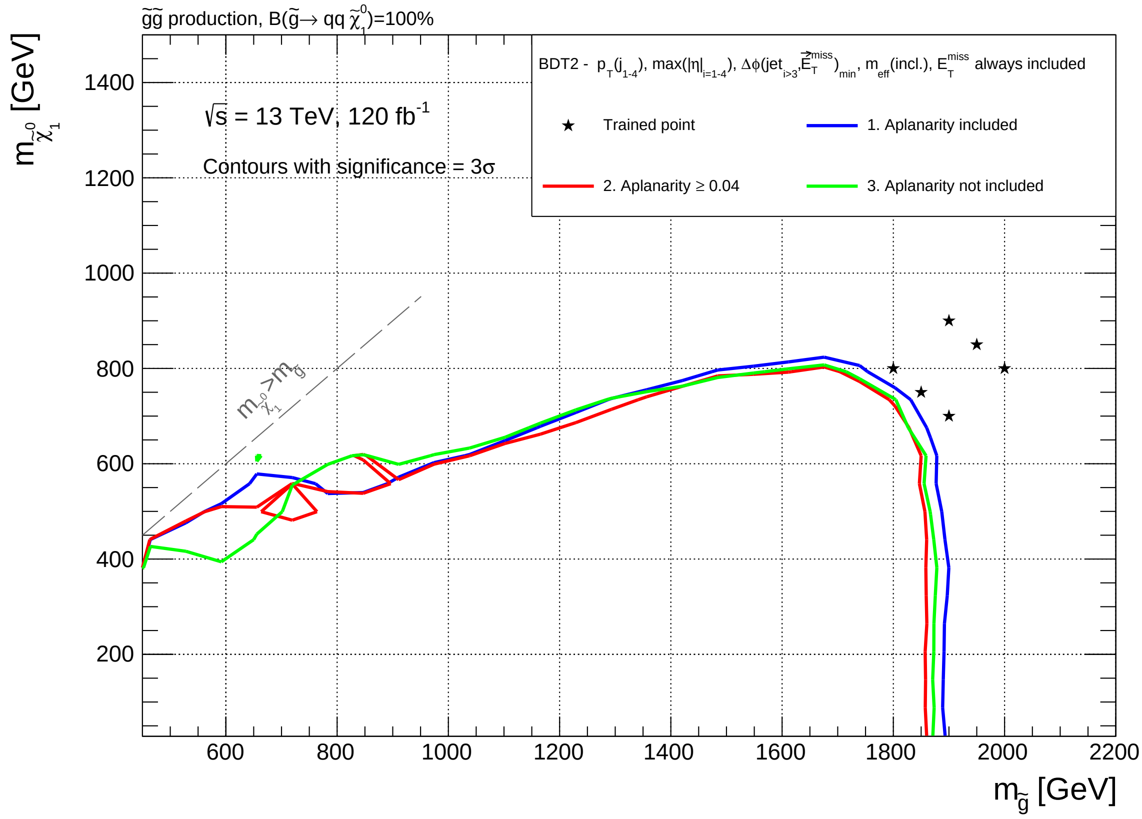


Figure 8.8: 3σ contour lines for different combinations of aplanarity. Including the aplanarity as BDT input shows the best results.

8.3.4 Impact of $\max(|\eta|_{i=1-4})$ Variable on BDT Performance

In the same way as above, the $\max(|\eta|_{i=1-4})$ variable is tested as BDT input using the scenarios shown in table 8.7. The resulting 3σ contour lines are displayed in figure 8.9. The gain using $\max(|\eta|_{i=1-4})$ as input variable for BDT is not really large but still does contribute and will be therefore kept as input variable.

N°	$p_\mathrm{T}(j_{1-4})$	$|\eta|(j_{1-4})$	$\max(|\eta|_{i=1-4})$	$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$	Ap	$m_\mathrm{eff}(\mathrm{incl.})$	$E_\mathrm{T}^\mathrm{miss}$
1	✓	✓	✓	✓	✓	✓	✓
2	✓	✓	≤ 2.0	✓	✓	✓	✓
3	✓	✓	–	✓	✓	✓	✓

Table 8.7: BDT input variable combinations for the variable study for $\max(|\eta|_{i=1-4})$. Preferred input variable choice marked in gray.

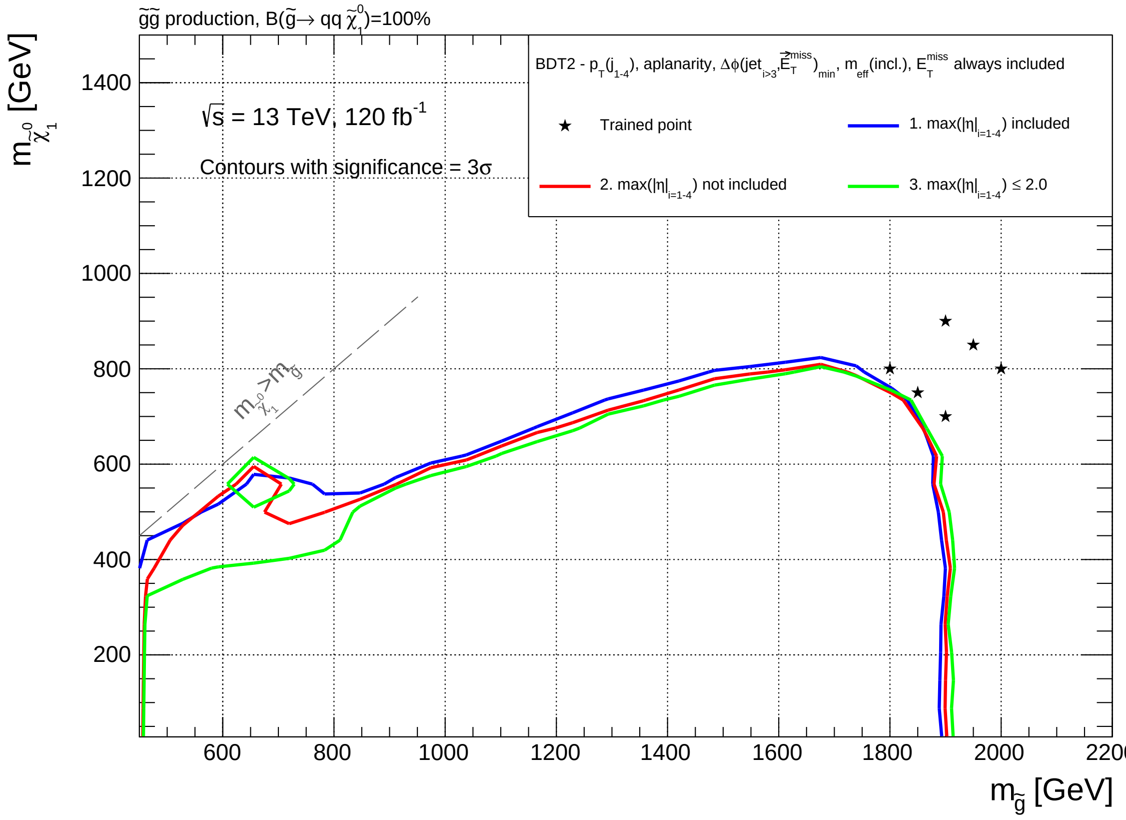


Figure 8.9: 3σ contour lines for different combinations of $\max(|\eta|_{i=1-4})$. The variable $\max(|\eta|_{i=1-4})$ will be further used as BDT input (blue contour).

8.3.5 Impact of $p_\mathrm{T}(j_{1-4})$ Variable on BDT Performance

In the previous variable scans all transverse momenta of the first four jets were included as input to the BDT. However it might not be necessary to train always on all of them. Therefore eleven different combinations of the p_T(jets) are investigated as shown in figure 8.10. The variables m_eff(incl.), $E_\mathrm{T}^\mathrm{miss}$ and aplanarity retain in the BDT training.
Without loosing much sensitivity it is possible to use only $p_\mathrm{T}(j_4)$ as BDT input variable.

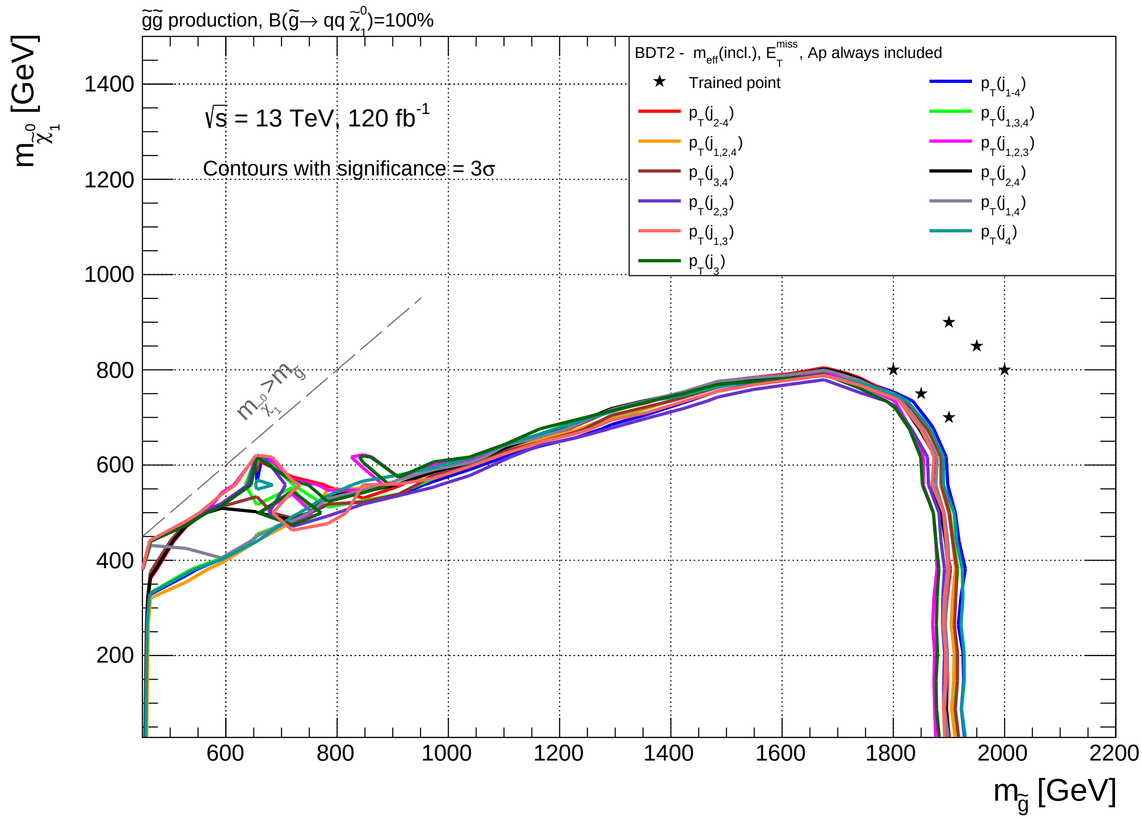


Figure 8.10: 3σ contour lines for different combinations of jet p_T. The preferred choice is the $p_\mathrm{T}(j_4)$ variable as BDT input.

8.3.6 Final Variable Combinations

Considering all the results from the previous studies the following variable combination is obtained for the BDT2 signal region:

1. m_eff(incl.),
2. $E_\mathrm{T}^\mathrm{miss}$,
3. $p_\mathrm{T}(j_4)$,
4. aplanarity,
5. $\max(|\eta|_{i=1-4})$ and

6. an additional preselection cut $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}} \geq 0.2$.

The 3σ contour line of this variable combination as well as for the extended variable set is shown in figure 8.11. There is only a minor difference visible while the number of BDT input variables was drastically reduced to five and an additional preselection cut was applied and hence the complexity is reduced.

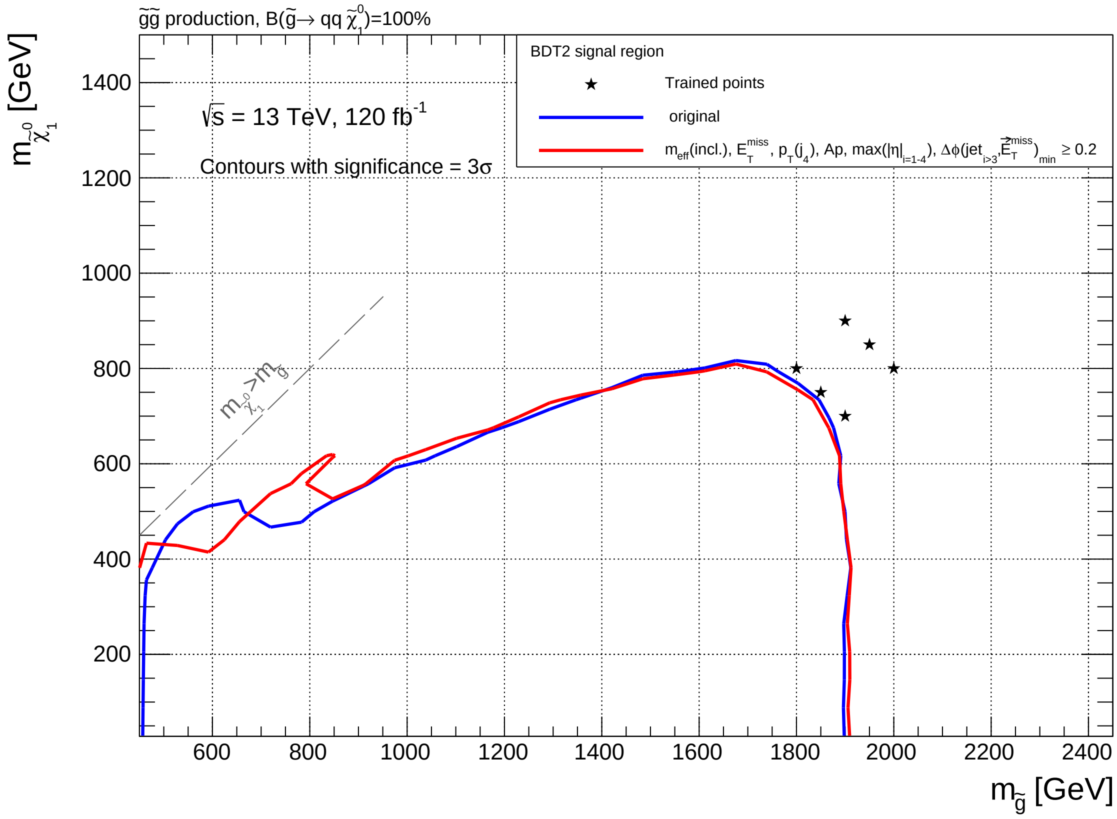
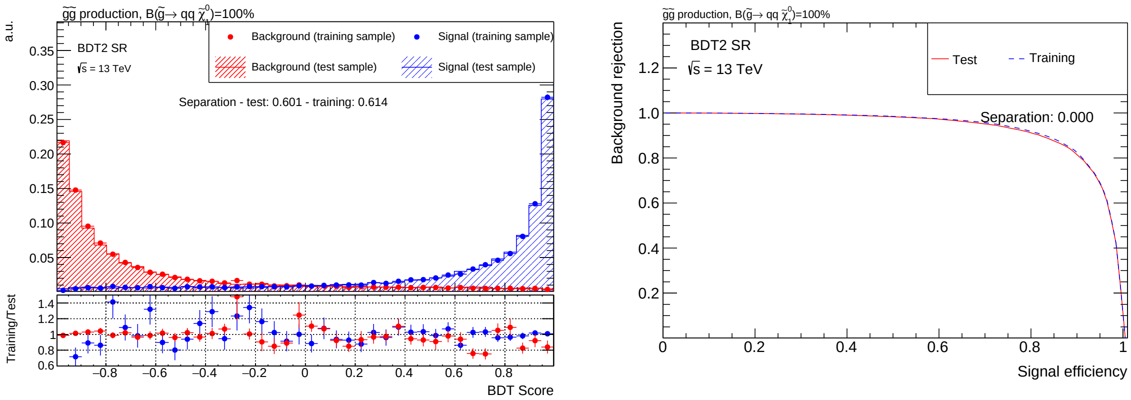


Figure 8.11: Comparison of the two 3σ contour lines for the reduced and extended set of input variables for the BDT2 signal region.

Overfitting Check

An overfitting check is performed for the BDT based signal region BDT2 with the new variable combination shown above. Figure 8.12 shows the BDT score distribution as well as the ROC curve for the BDT training and testing. The training over testing ratio of the BDT score distribution is consistent with one, except of few small fluctuations in the tails of the distributions and the ROC curves of training and testing agree nicely and hence no overfitting is visible.
The BDT method has a nice performance in this BDT signal region (BDT2) with a separation of $\sim 60\%$ of the background and signal distribution of the BDT score and a large area under the ROC curve.



a) b)

Figure 8.12: In a) the BDT score distribution for the BDT based signal region BDT2 is shown as well as the training over test ratio and in b) the corresponding ROC curve is displayed.

Combination of all BDT Signal Regions

The studies of the remaining BDT based signal regions are shown in appendix B.3. Importantly, the studies for the single BDT regions are not necessarily the best choice for the combination (oring) of all BDT regions. Therefore some minor adaptions are done compared to the studies in the appendix, taking into account the interplay between the regions.
The resulting BDT input variable combinations are summarised in table 8.8 and the new preselection cut definitions are shown in table 8.9.

SR	$m_{\text{eff}}(\text{incl.})$	$E_{\text{T}}^{\text{miss}}$	Ap	$\max(\lvert\eta\rvert_{i=1-4})$	$p_{\text{T}}(j_i)$	Additional variables
BDT1	✓	✓	✓	–	2–4	$H_{\text{T}}(\text{incl.})$
BDT2	✓	✓	✓	✓	4	–
BDT3	✓	✓	✓	✓	1,3,4	–
BDT4	✓	✓	–	✓	1–4	$\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$
BDT5	✓	✓	✓	–	1,2,4	N_{jet}, $\lvert\eta(j_{1-4})\rvert$

Table 8.8: BDT input variables for the different BDT based signal regions.

The comparison of the new variable definitions with the former ones and the Cut & Count method is shown in figure 8.13. The new variable combinations provide very similar results as the previous definition with less BDT input variables. In fact, in the compressed region there is a gain visible with respect to the former definition. Moreover, the BDT based signal regions show better sensitivity over the whole mass grid.

Targeted signal	$\tilde{g}\tilde{g}, \tilde{g} \to q\bar{q}\tilde{\chi}_1^0$				
Requirement	BDT based signal regions				
	BDT1	**BDT2**	**BDT3**	**BDT4**	**BDT5**
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 300				
$p_{\mathrm{T}}(j_1)$ [GeV]	> 200				
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	≥ 0.4				
N_{jet}	≥ 4				≥ 2
$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$	> 0.2				–
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ [GeV$^{1/2}$]	–				> 12
$m_{\mathrm{eff}}(\mathrm{incl.})$ [GeV]	> 1400		> 800		> 1400
$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	≥ 0.2			–	

Table 8.9: Adapted preselection cuts for the five BDT based signal regions.

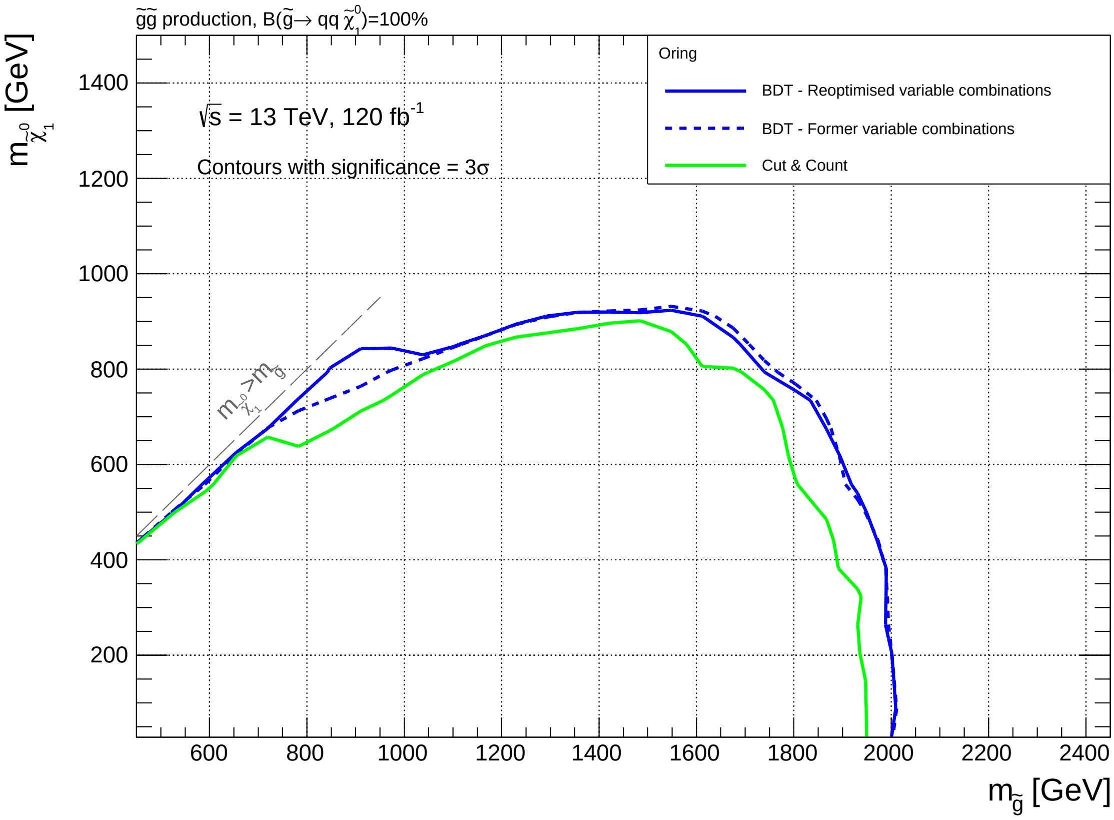


Figure 8.13: 3σ contour lines of the former (blue dashed contour) and reoptimised (blue solid contour) BDT input variable combinations. The results for the Cut & Count method are shown in green. All contours are results from the oring procedure.

8.4 Quark and Gluon Jet Separation

The SUSY signal of the directly decaying gluinos is quark-jet dominated (see the decay topology in figure 8.14) whereas the background is gluon-jet dominated. By finding a variable which is able to distinguish between these two signatures would help to better separate signal and background.

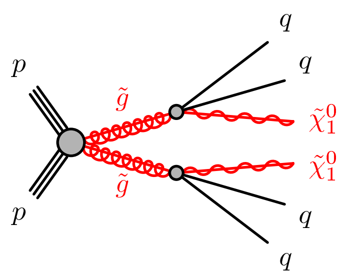


Figure 8.14: The decay topology of gluino-pair production in the simplified models with direct gluino decay [32].

The probability for gluon radiations is proportional to the colour factor C_F for the quark gluon interaction and C_A for the gluon gluon interaction. These colour factors as well as the Feynman diagrams for gluon radiations from quarks and gluons are shown in figure 8.15. Since the colour factors for the gluon radiation from a gluon is roughly twice as large as the colour factor for a gluon radiation from a quark, it is expected that the gluon jet contains more particles and therefore also more tracks[2].

Therefore the variable N_{trk} including the first four jets due to the signal topology

$$N_{\text{trk}} = \sum_{i=1}^{4} N_{\text{trk}}^{(i)}, \tag{8.1}$$

with $N_{\text{trk}}^{(i)}$ the number of tracks in jet i, is a good candidate to distinguish between quark and gluon jets. To test the impact of this variable it is added as BDT input variable on top of these described above. The resulting contour line is shown in figure 8.16. Especially in the medium $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$ region a gain is visible. In addition, the $N_{\text{trk}}^{(i)}$ for each jet separately was studied but the gain with respect to the simple sum was very small.

8.4.1 Validation Study for N_{trk}

The use of the N_{trk} variable shows a nice gain. However, in Run I of the LHC a disagreement of the $N_{\text{trk}}^{(i)}$ MC modelling for gluon initiated jets was observed as described in [52]. This observed discrepancy is most likely due to improper modelling of gluon jet hadronisation. Therefore it has to be checked if it is possible to use the N_{trk} variable for further studies. The extracted plot from this note is shown in figure 8.17. Since the p_T and $|\eta|$ range does not

[2] In principle one could also exploit jet shape variables, but since the calorimeter simulation has some problems to model properly showers, this study focuses on track quantities since they are better modelled.

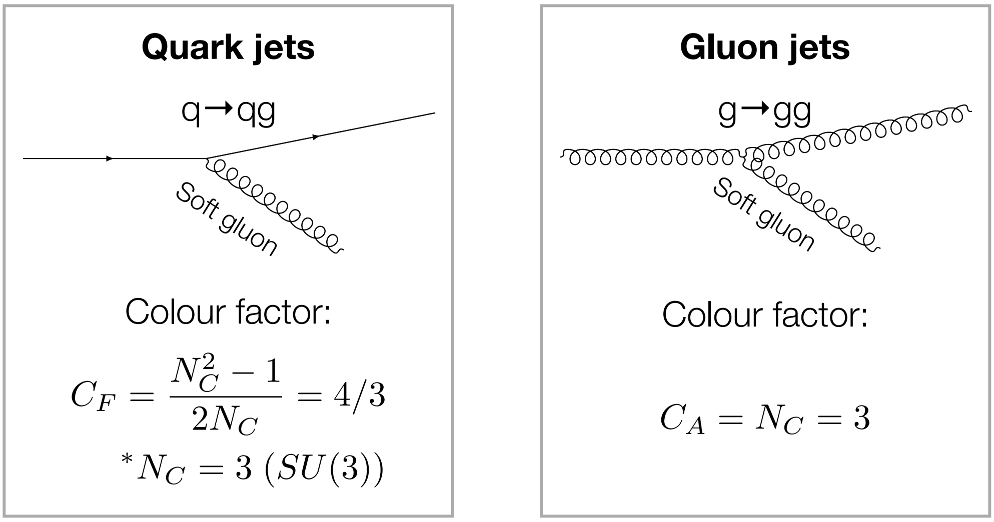


Figure 8.15: Difference between quark and gluon jets. The probability of gluon radiation is proportional to colour factor, indicated for both scenarios.

correspond to the regions of the 0-lepton publication, a direct conclusion for the given regions is not possible. Additionally, the MC samples for the SUSY 0-lepton analysis were produced with the Sherpa MC generator, in particular the W+jets, Z+jets and diboson samples which are the dominating processes in the SM background while for the study presented in [52] Herwig++ and Pythia8 are utilised. Therefore, in the following a new study is presented to estimate the impact of a possible mismodelling.

For this study the data corresponds to an integrated luminosity of 36.1 fb^{-1}, the same as taken for the 0-lepton publication [32].
For the purpose of this study, two W control regions are chosen since they are more gluon enriched than the top control regions:

- CRW of SR Meff-2j-1200,
- CRW of SR Meff-4j-2600.

In order to get more statistics the m_{eff}(incl.) cut for the 4 jet region is lowered to 1000 GeV and the region is further labeled as SR Meff-4j-2600*.
The preferred variable to add to the BDT studies is the sum of $N_{\mathrm{trk}}^{(i)}$ for the first four jets defined in equation (8.1). Thus, the sum and the single $N_{\mathrm{trk}}^{(i)}$ variables have to be studied. The cuts are extracted from the analysis described in [32]. The complete set of cut strings are also available in section B.1.

The observed distributions for the single $N_{\mathrm{trk}}^{(i)}$ distributions are shown in the figures 8.18 and 8.19.
The resulting distributions in the W control region for SR Meff-2j-1200 shown in figure 8.18 of MC and data are in rather good agreement with each other. Especially in the central part

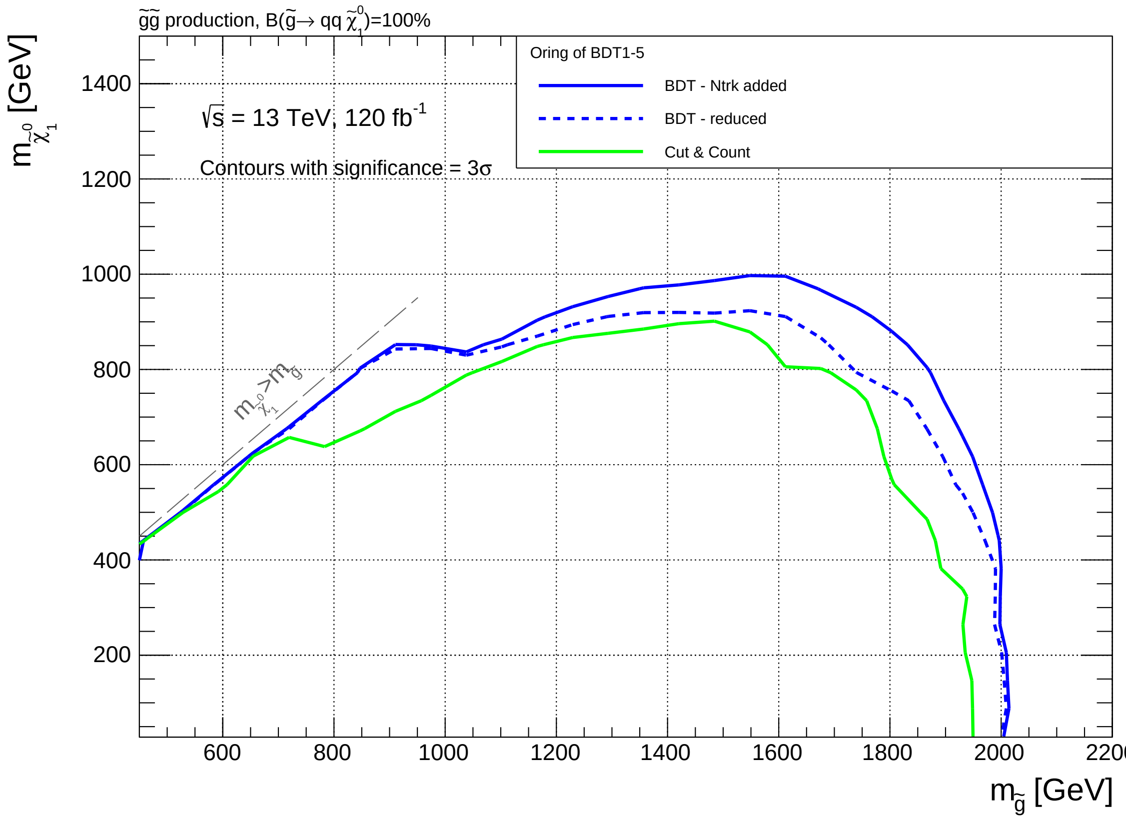


Figure 8.16: The 3σ contour lines with the BDT and Cut & Count results are presented above and compared to the BDT (blue solid contour) with the additional variable N_{trk} added in the training.

of the $N_{\text{trk}}^{(i)}$ distribution of the first jet (see figure 8.18 a)), the data/MC ratio is almost 1 and very flat. Similarly, the $N_{\text{trk}}^{(i)}$ distributions of the 2^{nd} to 4^{th} jet show a good agreement between data and MC although there are some fluctuations but they can be neglected. The peaks in the first bin in the figures 8.18 c) and d) are leptons which are considered as jets in the control region. Since leptons leave only one track in the detector and are mostly one of the jets with lower p_{T}, there is this peak at one.

Furthermore, the second set of plots in the W control region for SR Meff-4j-2600* shows some deviations. In particular the $N_{\text{trk}}^{(i)}$ distributions of the first two jets, as indicated in the figures 8.19 a) and b), show a general trend of MC overestimation. However, the data/MC ratio is relatively flat and can be compensated via a fit. Consequently, this behaviour is not a problem for further studies.
Moreover, the distributions of the 3^{rd} and 4^{th} jet may indicate a slight mismodelling but there are no major deviations in the data/MC ratio visible.

Finally, the sum of $N_{\text{trk}}^{(i)}$ which will be used in the BDT studies, is shown in figure 8.20 for both control regions. In fact, the N_{trk} variable is quite well described in the CRW for SR Meff-2j-1200 (figure 8.20 a)), whereas, slight deviations are visible in the CRW for SR

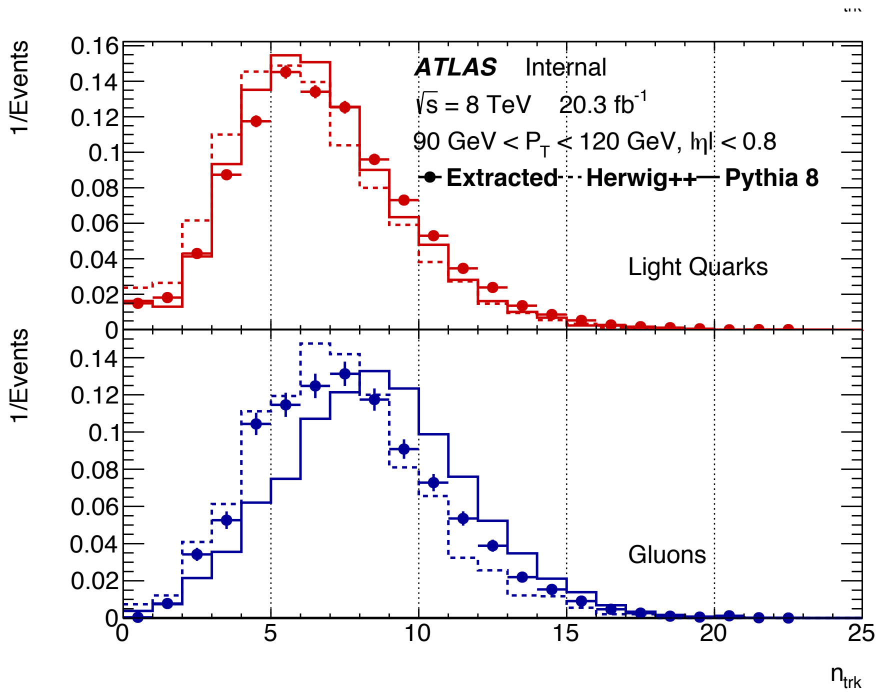


Figure 8.17: Extracted templates for $N_{\text{trk}}^{(i)}$ comparing data, Pythia8 and Herwig++ for an intermediate p_{T} bin and $|\eta| < 0.8$ [52].

Meff-4j-2600* (figure 8.20 b)).

To summarise, the studied distributions of the single $N_{\text{trk}}^{(i)}$ and the sum of the first four $N_{\text{trk}}^{(i)}$ are mostly in good agreement between data and MC. The small deviations which are observed stay within a reasonable range. Also, the showed distributions do not contain systematic uncertainties and thus the discrepancies may well be covered by the full uncertainty band. Hence, the N_{trk} variable can be used in further studies to improve the signal region definitions.

Figure 8.21 shows different scenarios with additional systematic uncertainties applied. If the additional systematic uncertainties stay within 10% the N_{trk} variable still improves the sensitivity. However if they exceed 20% there is no gain and it would not be useful to include the N_{trk} variable.

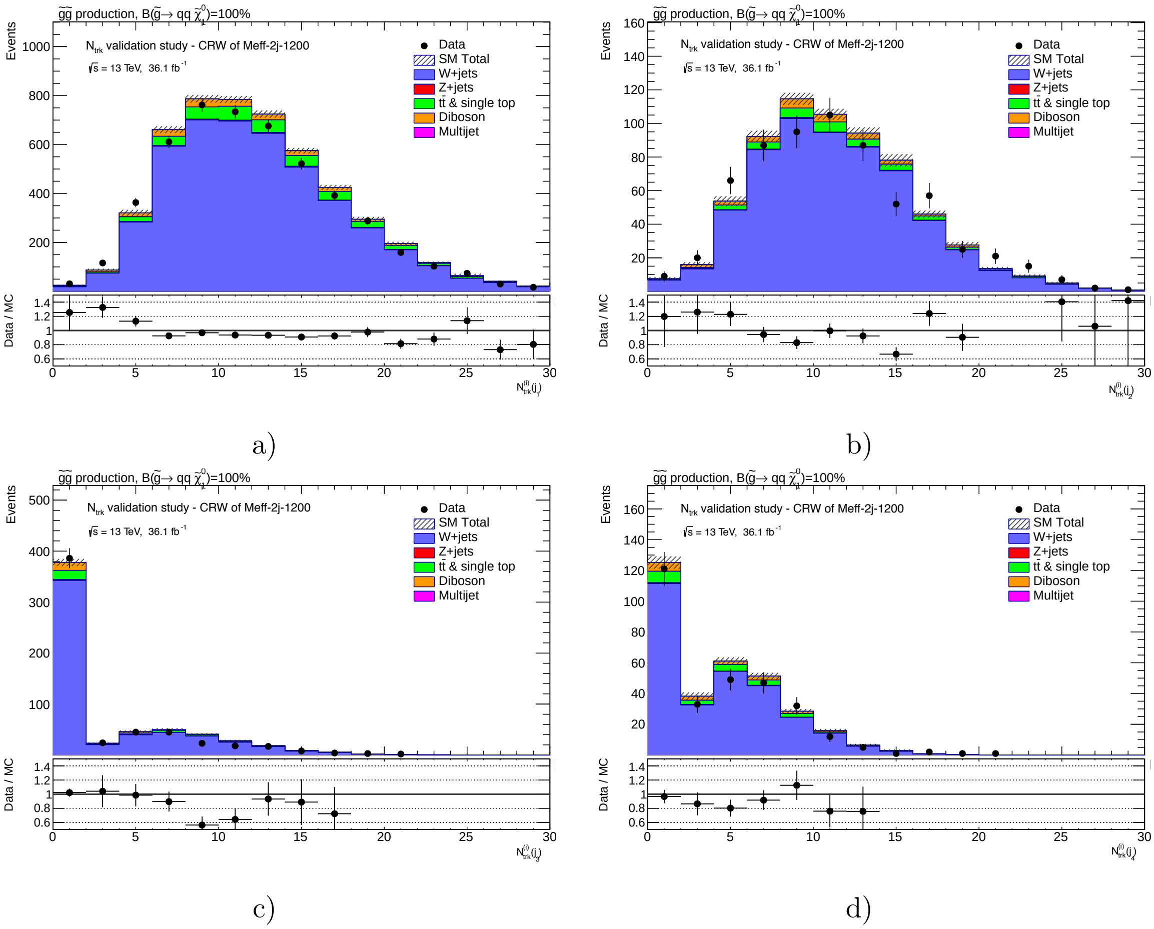


Figure 8.18: Observed $N^{(i)}_{\mathrm{trk}}$ distributions in the control region CRW for SR Meff-2j-1200 with the default cut $p_{\mathrm{T}}(\text{jets}) \geq 50\,\mathrm{GeV}$. a) $N^{(i)}_{\mathrm{trk}}$ distribution of the first jet with $N_{\mathrm{jet}} \geq 1$, b) $N^{(i)}_{\mathrm{trk}}$ distribution of second jet with $N_{\mathrm{jet}} \geq 2$ and $p_{\mathrm{T}}(j_2) \geq 250\,\mathrm{GeV}$, c) $N^{(i)}_{\mathrm{trk}}$ distribution of third jet with $N_{\mathrm{jet}} \geq 3$ and d) $N^{(i)}_{\mathrm{trk}}$ distribution of fourth jet with $N_{\mathrm{jet}} \geq 4$. Only statistical uncertainties considered.

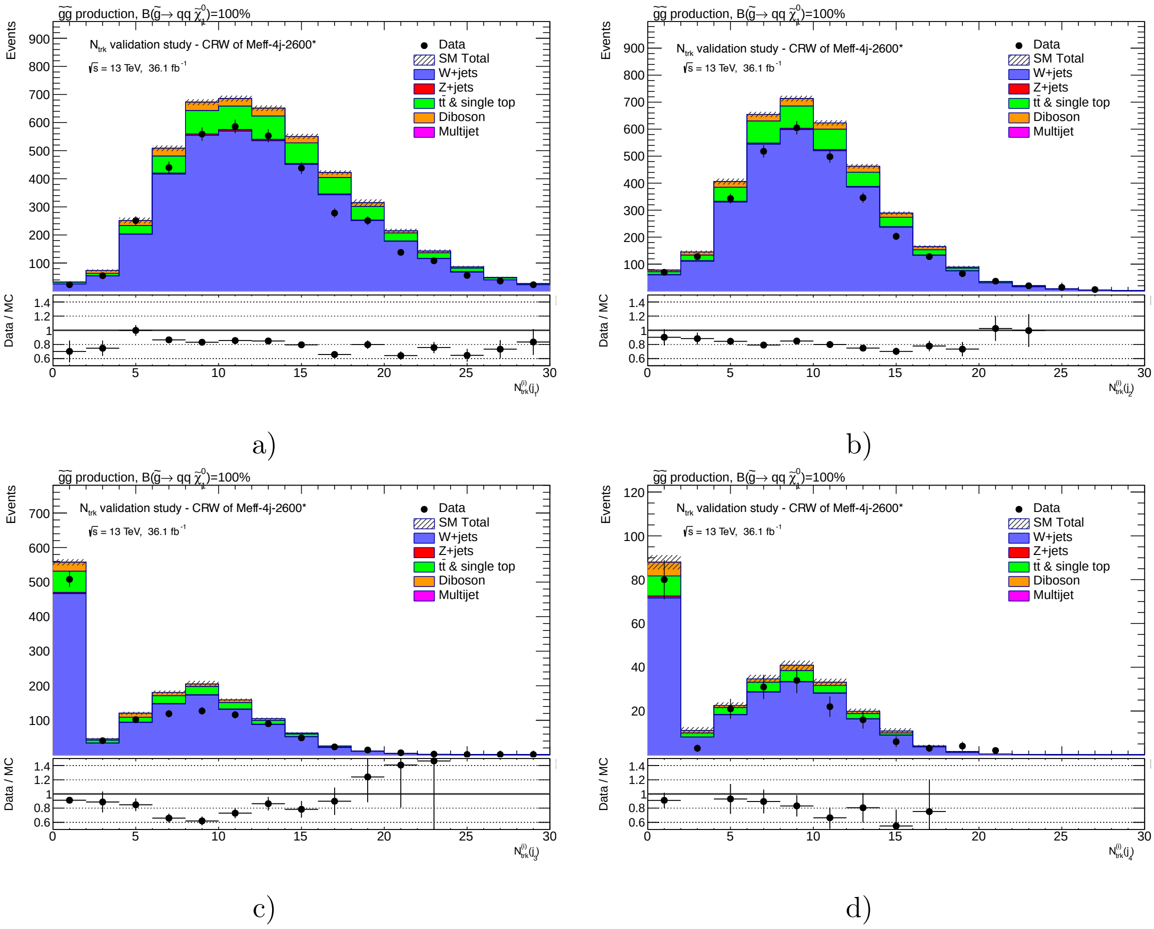
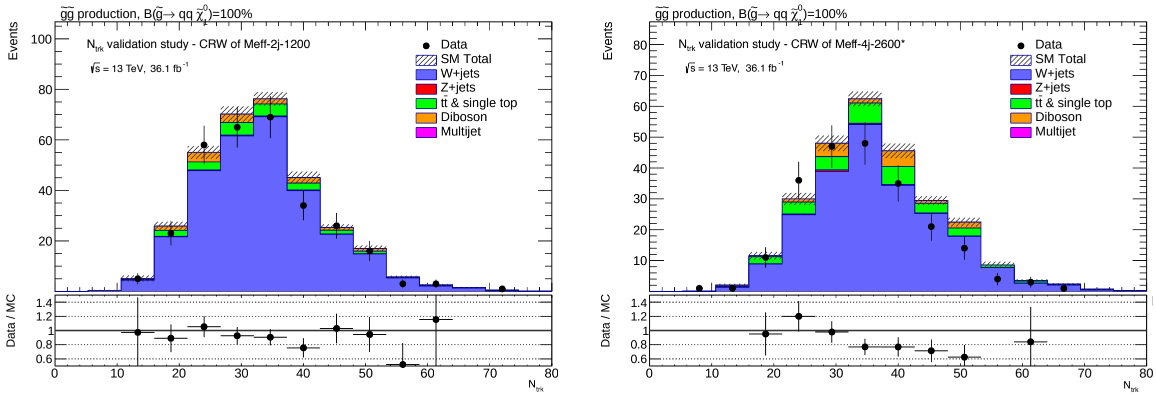


Figure 8.19: Observed $N_{\mathrm{trk}}^{(i)}$ distributions in the control region CRW for SR Meff-4j-2600* (but $m_{\mathrm{eff}}(\mathrm{incl.}) \geq 1000\,\mathrm{GeV}$ due to statistics). a) $N_{\mathrm{trk}}^{(i)}$ distribution of the first jet with $N_{\mathrm{jet}} \geq 1$ and no cut on higher jet p$_T$s, b) $N_{\mathrm{trk}}^{(i)}$ distribution of second jet with $N_{\mathrm{jet}} \geq 2$ and $p_{\mathrm{T}}(j_2) \geq 150\,\mathrm{GeV}$ and no cut on higher jet p_{T}s, c) $N_{\mathrm{trk}}^{(i)}$ distribution of third jet with $N_{\mathrm{jet}} \geq 3$ and $p_{\mathrm{T}}(j_3) \geq 150\,\mathrm{GeV}$ and no cut on higher jet p$_T$s and d) $N_{\mathrm{trk}}^{(i)}$ distribution of fourth jet with $N_{\mathrm{jet}} \geq 4$ and $p_{\mathrm{T}}(j_4) \geq 150\,\mathrm{GeV}$. Only statistical uncertainties considered.



a) b)

Figure 8.20: Observed N_{trk} (sum of first four $N_{\mathrm{trk}}^{(i)}$) distributions in the control regions CRW for a) Meff-2j-1200 and b) Meff-4j-2600*. Only statistical uncertainties considered.

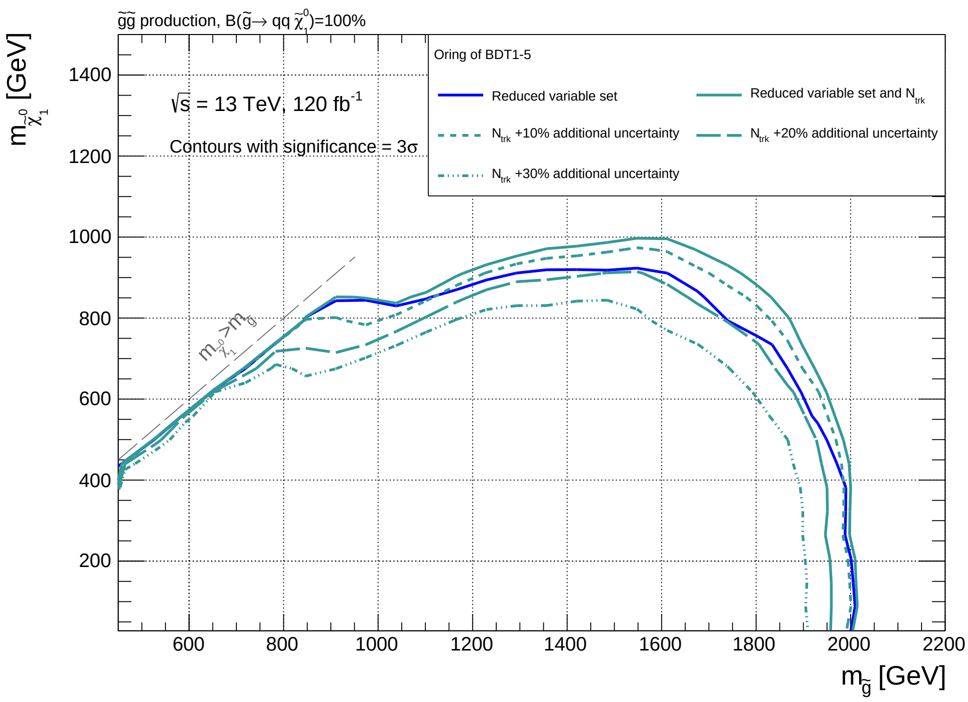


Figure 8.21: 3σ contour lines for different additional systematic uncertainties (10%, 20% and 30%) for the BDT including the N_{trk} variable. As reference point the BDT with the reduced variable set is shown in blue.

9 Conclusion

Modern particle physics requires large experiments like the ATLAS detector at the LHC. Furthermore, complex data processing tools and optimisation studies are needed to improve the measurements and searches performed at such large experiments.
Due to the limitations of the current Standard Model of particle physics, searches for new physics is one of the main tasks at CERN experiments. Supersymmetry (SUSY) represents an elegant model to describe and solve several problems which cannot be explained by the Standard Model.
The SUSY analysis of the ATLAS experiment presented in section 5 searches for squark and gluino decays into hadronic final states in $36.1\,\mathrm{fb}^{-1}$ of pp collision data at $\sqrt{s} = 13\,\mathrm{TeV}$. Two selection strategies are used to search for these SUSY particles with the typical detector signature of several high-p_T jets and large missing transverse momentum. One of these selection methods is focusing on the m_{eff}(incl.) variable. The results are further interpreted in terms of simplified SUSY benchmark models. Since no significant excess above the SM expectation was found, limits on the mass of the squarks, gluinos and neutralinos were placed, representing an improvement in the exclusion of the SUSY parameter space compared to former analyses.
However, the utilisation of only cut-based analysis techniques, as done in the most recent 0-lepton publication, is not sufficient anymore to achieve significant improvements. Therefore, more sophisticated analysis tools are necessary to optimise the signal efficiency and discovery significance in particular analyses. One of these tools are boosted decision trees (BDT), implemented in the TMVA framework operating with the ROOT software. This method allows for a better discrimination of signal and background events using more complex selection criteria due to correlations between two or more variables.
Another method is the Multi-Bin approach where the data set is sliced into several orthogonal bins and each bin is optimised via the Cut & Count method.
These two methods are investigated in this thesis to improve the signal regions for the pair production of squarks and gluinos in direct decays. The studies are exclusively performed with simulated events.

To be able to compare the results of the new methods at an integrated luminosity of $100\,\mathrm{fb}^{-1}$ for the squark pair production, a Cut & Count reoptimisation is performed for this luminosity. The first BDT study is done using the input variables with the best background to signal separation from the 0-lepton publication. After optimising the choice of the signal training points, the BDT shows first improvements compared to the reoptimised Cut & Count.
Furthermore, the impact of slicing the simulated samples into orthogonal subsets depending on the jet multiplicity is studied. The different bins of the N_{jet} separated simulated samples are optimised using the Cut & Count method and BDT. After the optimisation the significances of the single bins are combined. For both methods, there is a significant enhancement visible compared to the simple reoptimised Cut & Count and BDT over the whole $m(\tilde{q}, \tilde{\chi}_1^0)$-range except in the compressed mass region (low $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$).

Given these results, the binning into two additional variables m_{eff}(incl.) and $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ is investigated. The single Multi-Bin categories are optimised using the Cut & Count method, since BDT-based signal regions would largely increase the complexity of the analysis. While the Multi-Bin analysis with N_{jet} and m_{eff}(incl.) yields an improvement over the whole $m(\tilde{q}, \tilde{\chi}_1^0)$-range, the additional $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ binning improves only for lower $\Delta m(\tilde{q}, \tilde{\chi}_1^0)$ values and looses sensitivity for high values.
In order to reduce the number of bins and thus the complexity, the signal training points are reduced to two, which is possible without loosing sensitivity and an algorithm based on the significance ranking is introduced. It is shown that the Multi-Bin approach with N_{jet} and m_{eff}(incl.) categories consisting of 24 bins covers the whole $m(\tilde{q}, \tilde{\chi}_1^0)$-range. Therefore the classical signal regions as used in the 0-lepton publication are not needed anymore.
In summary, the final amount of bins from N_{jet} and m_{eff}(incl.) categories is a reasonable number considering the complexity of the analysis. A gain up to approximately 150 GeV in the neutralino mass and an enhancement of about 200 GeV in the squark mass is achieved compared to the reoptimised Cut & Count method. Concluding the performed studies, the Multi-bin approach together with the significance ranking algorithm is a good choice for further investigations towards a possible implementation in an updated version of the analysis.

The second target signal is the gluino pair production studied at an integrated luminosity of 120 fb^{-1}. In order to determine the strategy of the study, a rough method scan is performed. The best method after this brief scan is to use a BDT with GradBoost and the combination of several signal points with similar kinematics to signal regions for the BDT training.
This method is further improved by scanning different BDT input variable combinations to optimise the gain and reduce the number of input variables as much as possible without loosing sensitivity. The results from the reduced set of input variables show similar improvements as the inclusion of the full variable set from the 0-lepton analysis. In fact, there is also an improvement visible in the compressed region. Overall, the BDT increases the sensitivity over the whole mass grid compared to the reoptimised Cut & Count method.
Since the SUSY signal is quark-jet dominated and the background contains mostly gluon-jets, they differ in their number of tracks associated with the jets. Therefore, a new variable, called N_{trk} (sum of tracks in the first four jets) is introduced. Considering this variable as BDT input, the sensitivity improves significantly. However, indications of MC mismodelling of the $N_{\text{trk}}^{(i)}$ variable has been seen in RUN I and thus a validation in the W control region is performed using pp collision data with $\sqrt{s} = 13$ TeV and an integrated luminosity of 36 fb^{-1}. In general, simulation and data are in rather good agreement and only small systematic uncertainties have to be introduced to cover some small effects. If the additional systematic uncertainties are less than 10%, the inclusion of the N_{trk} variable can still improve the discovery potential of the analysis.
In conclusion, the most promising method is the BDT including the N_{trk} variable for the gluino pair production. Compared to the reoptimised Cut & Count method, a gain of about 200 GeV in the neutralino and gluino mass is reached.

In summary, the results of the studies performed in this thesis show improvements for the squark and gluino pair production using different sophisticated techniques beyond the simple cut-based methods. Since the studies are under development, further studies as well as a full implementation of all systematic uncertainties and control regions are necessary to include

these techniques into the next publication of the 0-lepton analysis.
Further improvements are possible using neural networks to extract new variables for the Cut & Count method, especially in the more compressed regions. Also the utilisation of more variables which are able to distinguish between quark-jets and gluino-jets will be promising.

A Additional Material for Squark Studies

A.1 Input variables

Cut string for $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)\,\text{GeV}$ and the plots shown in 7.5 & 7.6:

```
"(((veto==0||veto==410500) && ( abs(jetEta[0])>2.4 || jet1Chf/jetFracSamplingMax[0]>0.1)
&& (dPhiBadTile>0.3 || RunNumber>284484)) && (abs(timing)<4 || timing==-99999)
&& ((cleaning & 0x30F)==0) && met >= 250.000000 && nJet>=2 && jetPt[0]>=300.000000
&& (dPhiR >=0.400000 || nJet<=3 )) && EventNumber%2==0"
```

A.2 BDT Parameters

Training string BDT:

```
"!H:!V:NTrees=200:MaxDepth=4:BoostType=AdaBoost:AdaBoostBeta=0.15:
SeparationType=GiniIndex:nCuts=100:PruneMethod=NoPruning"
```

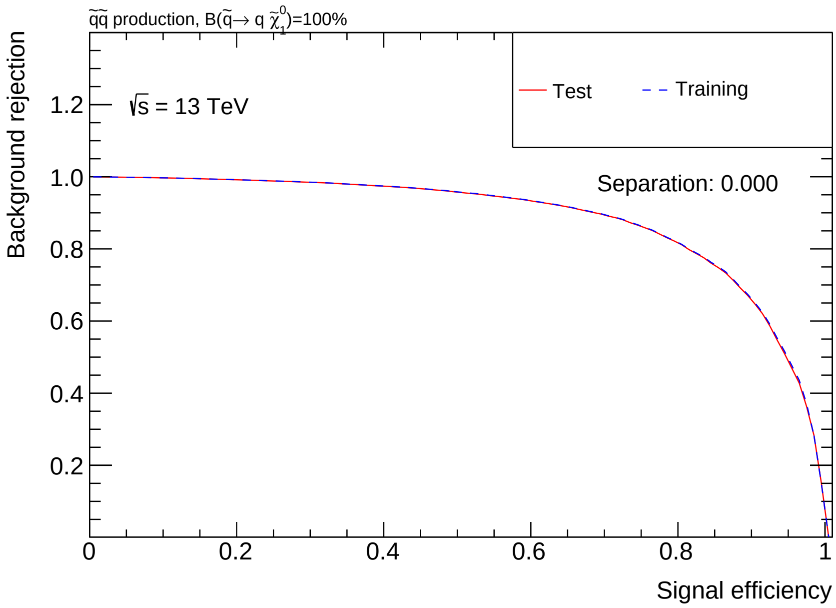


Figure A.1: ROC curve for the training (blue dashed line) and testing (red line) sample using the signal training point $m(\tilde{q}, \tilde{\chi}_1^0) = (1000, 500)\,\text{GeV}$. Used input variables: $m_{\text{eff}}(\text{incl.})$, $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$, $p_{\text{T}}(j_2)$, $\max(|\eta|_{i=1,2})$ and $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$.

B Additional Material for Gluino Studies

B.1 Ntrk Validation Study

Cut string for CRW for SR2j-1200:

"((veto==0||veto==410500) && met >= 250.000000 && nJet >= 2 && jetPt[0] >= 250.000000 && jetPt[1] >= 250.000000 && ((cleaning & 0x30F)==0) && met/sqrt(meffInc-met) >= 14.000000 && (lep1Triggered & 0x03)!=0 && lep1Signal==1 && nBJet==0 && lep1Pt>27 && 30<mt && mt<100)"

Cut string for CRW for SR4j-2600*:

"((veto==0||veto==410500) && meffInc >= 1000.000000 && met >= 250.000000 && nJet >= 4 && jetPt[0] >= 200.000000 && jetPt[1] >= 150.000000 && jetPt[2] >= 150.000000 && jetPt[3] >= 150.000000 && ((cleaning & 0x30F)==0) && met/(met + jetPt[0] + jetPt[1] + jetPt[2] + jetPt[3]) >= 0.200000 && (lep1Triggered & 0x03)!=0 && lep1Signal==1 && nBJet==0 && lep1Pt>27 && 30<mt && mt<100)"

Additional cleaning cuts for both:
"&&(abs(jetEta[0])>2.4 || jet1Chf/jetFracSamplingMax[0]>0.1) && (abs(jetEta[1])>2.4 || jet2Chf/jetFracSamplingMax[1]>0.1) && (dPhiBadTile>0.3 || RunNumber>284484)"

B.2 Remaining Variable Distributions for BDT2

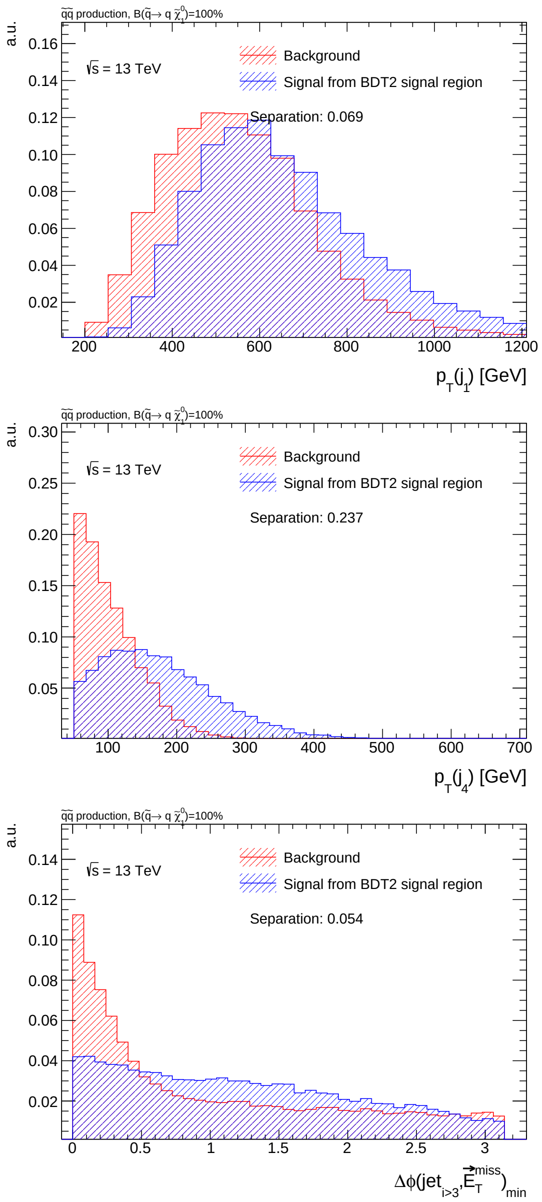


Figure B.1: Normalised variable distributions: $p_{\mathrm{T}}(j_{1,2,4})$, $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$. The background is shown in red and the signal of the BDT2 signal region in blue. In each plot the separation power for the background and signal distribution is given.

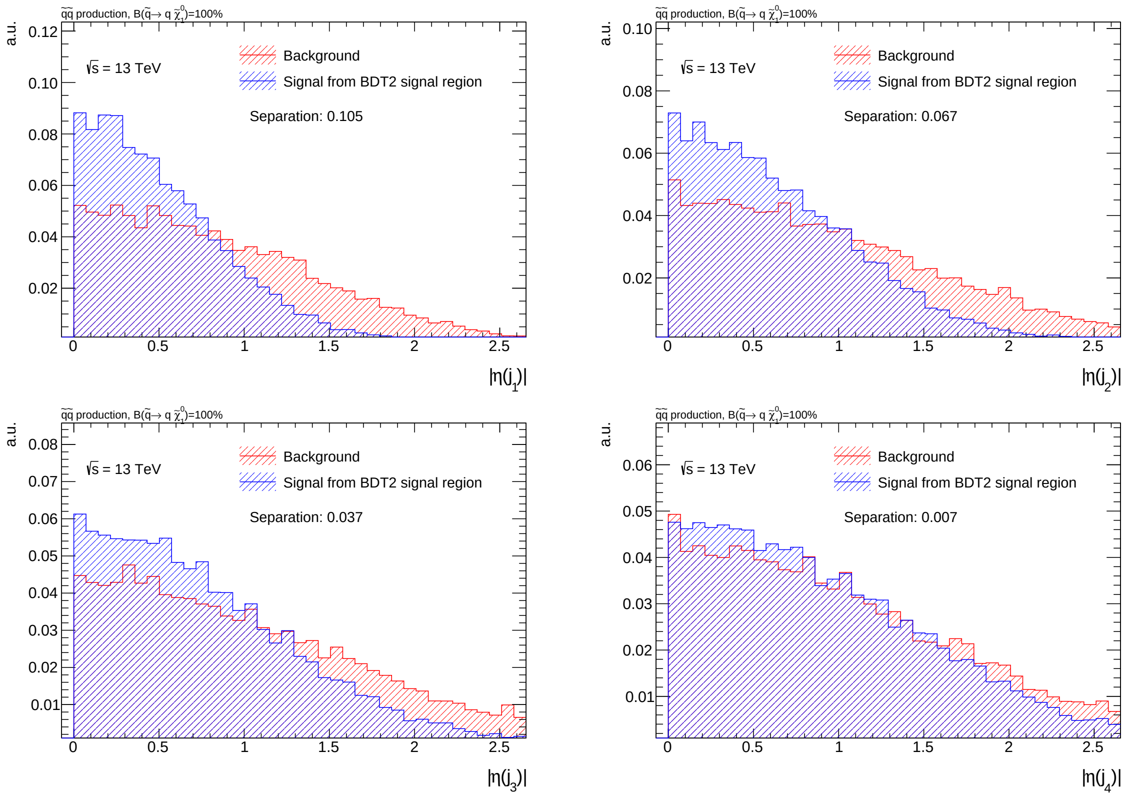


Figure B.2: Normalised variable distributions $|\eta(j_{1-4})|$. The background is shown in red and the signal of the BDT2 signal region in blue. In each plot the separation power for the background and signal distribution is given.

B.3 Gluino Variable Study for Remaining Regions

In the following the variable studies are presented for the four remaining BDT based signal regions: BDT1, BDT3, BDT4 and BDT5. Since the examplary study was already performed in section 8.3 the results are only presented as tables and plots and the final variable combinations are separately indicated.

B.3.1 BDT1

N°	$p_T(j_{1-4})$	$\|\eta\|(j_{1-4})$	$\max(\|\eta\|_{i=1-4})$	$\Delta\phi(\text{jet}_{i>3}, \vec{E}_T^{\text{miss}})_{\min}$	Ap	$m_{\text{eff}}(\text{incl.})$	E_T^{miss}	Additional
1	✓	✓	–	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(4j)$, N_{jet}
2	✓	✓	–	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(4j)$
3	✓	–	✓	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(4j)$
4	✓	–	✓	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(\text{incl.})$
5	✓	–	✓	✓	✓	✓	–	$E_T^{\text{miss}}/\sqrt{H_T(\text{incl.})}$
6	✓	–	✓	✓	✓	✓	✓	–
7	✓	–	✓	✓	✓	–	✓	$H_T(\text{incl.})$
8	✓	–	✓	✓	✓	–	✓	$H_T(4j)$
9	✓	–	✓	✓	✓	–	✓	–
10	✓	✓	✓	✓	✓	✓	✓	–
11	✓	✓	✓	–	✓	✓	✓	–
12	✓	✓	✓	≥ 0.2	✓	✓	✓	–
13	✓	✓	✓	≥ 0.4	✓	✓	✓	–
14	✓	✓	✓	✓	✓	✓	✓	–
15	✓	✓	✓	✓	≥ 0.04	✓	✓	–
16	✓	✓	✓	✓	–	✓	✓	–
17	✓	✓	✓	✓	✓	✓	✓	–
18	✓	✓	≤ 2.0	✓	✓	✓	✓	–
19	✓	✓	–	✓	✓	✓	✓	–

Table B.1: BDT input variable combinations for the variable studies for the BDT based signal region BDT1. The single studies are separated by two horizontal lines. Preferred input variable choices are marked in gray for the single studies. Lines 6 and 7 are both marked since the combination is taken.

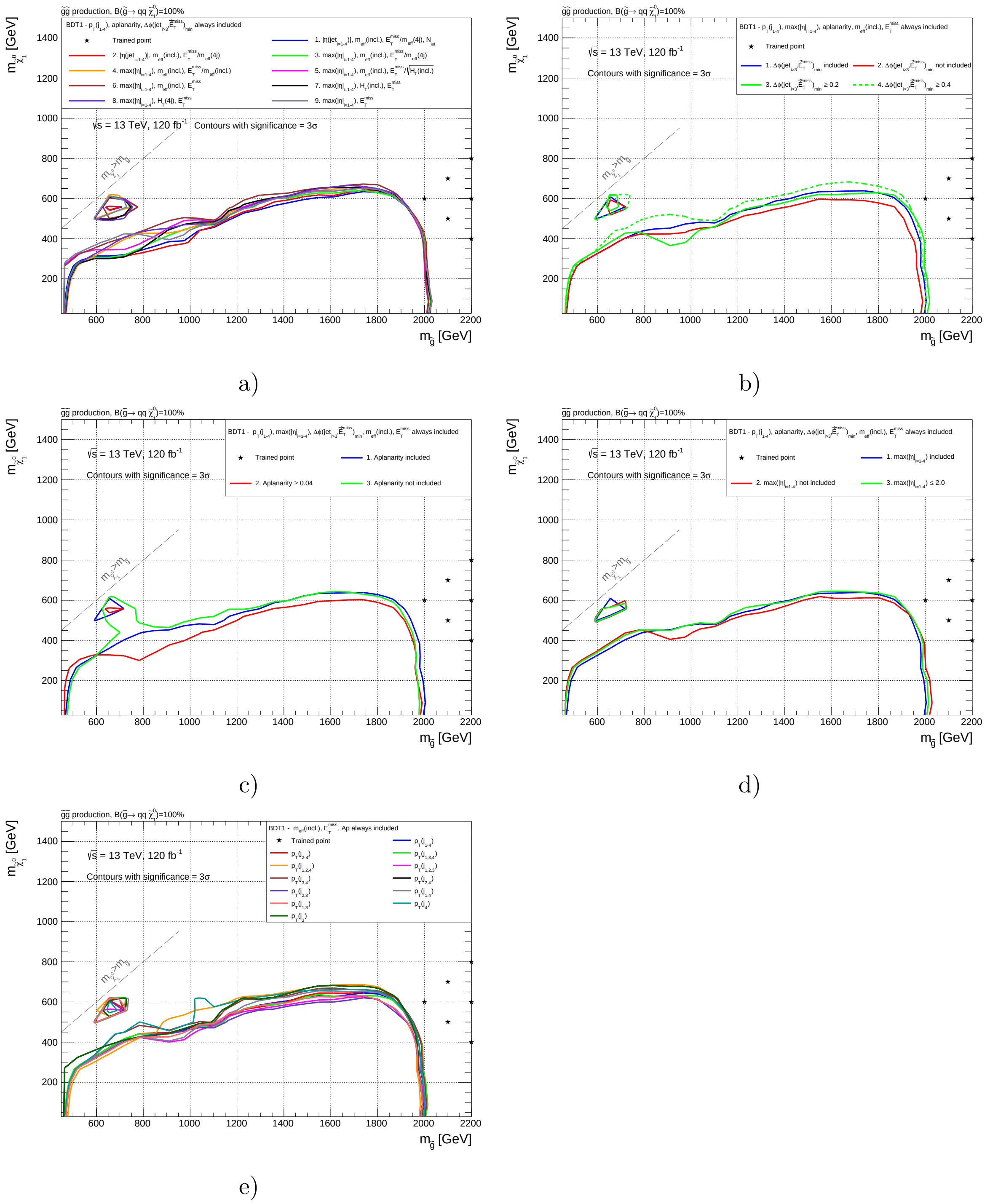


Figure B.3: Contour plots with significance of 3 with different variable combinations of a) m_{eff}, $E_{\text{T}}^{\text{miss}}$ and H_{T}, b) $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$, c) aplanarity, d) $\max(|\eta|_{i=1-4})$ and e) $p_{\text{T}}(j_i)$ for the BDT based signal region BDT1.

B.3.2 BDT3

N°	$p_T(j_{1-4})$	$\lvert\eta\rvert(j_{1-4})$	$\max(\lvert\eta\rvert_{i=1-4})$	$\Delta\phi(\text{jet}_{i>3}, \vec{E}_T^{\text{miss}})_{\text{min}}$	Ap	$m_{\text{eff}}(\text{incl.})$	E_T^{miss}	Additional
1	✓	✓	–	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(4j)$, N_{jet}
2	✓	✓	–	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(4j)$
3	✓	–	✓	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(4j)$
4	✓	–	✓	✓	✓	✓	–	$E_T^{\text{miss}}/m_{\text{eff}}(\text{incl.})$
5	✓	–	✓	✓	✓	✓	–	$E_T^{\text{miss}}/\sqrt{H_T(\text{incl.})}$
6	✓	–	✓	✓	✓	✓	✓	–
7	✓	–	✓	✓	✓	–	✓	$H_T(\text{incl.})$
8	✓	–	✓	✓	✓	–	✓	$H_T(4j)$
9	✓	–	✓	✓	✓	–	✓	–
10	✓	✓	✓	✓	✓	✓	✓	–
11	✓	✓	✓	–	✓	✓	✓	–
12	✓	✓	✓	≥ 0.2	✓	✓	✓	–
13	✓	✓	✓	≥ 0.4	✓	✓	✓	–
14	✓	✓	✓	✓	✓	✓	✓	–
15	✓	✓	✓	✓	≥ 0.04	✓	✓	–
16	✓	✓	✓	✓	–	✓	✓	–
17	✓	✓	✓	✓	✓	✓	✓	–
18	✓	✓	≤ 2.0	✓	✓	✓	✓	–
19	✓	✓	–	✓	✓	✓	✓	–

Table B.2: BDT input variable combinations for the variable studies for the BDT based signal region BDT3. The single studies are separated by two horizontal lines. Preferred input variable choices are marked in gray for the single studies.

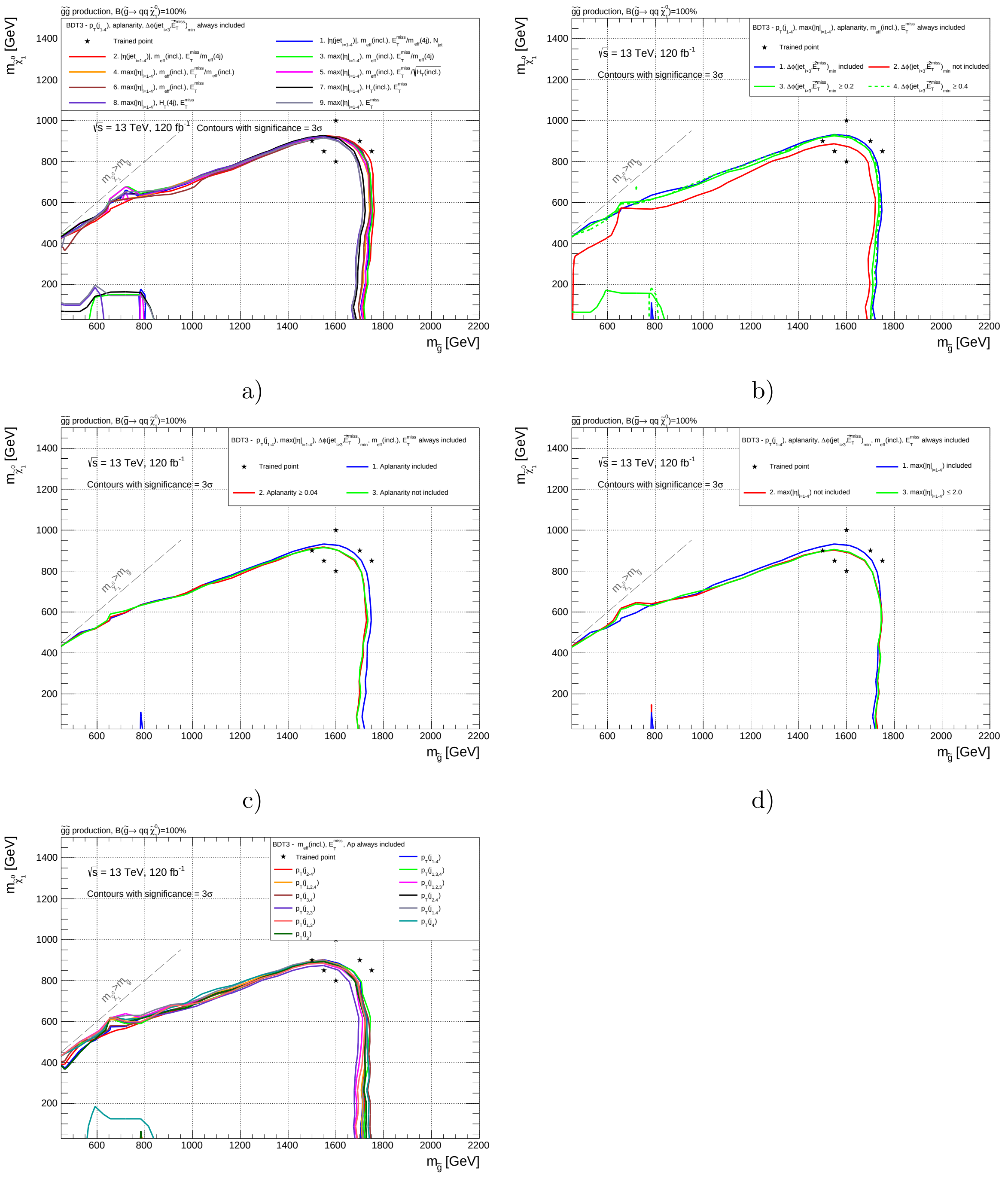


Figure B.4: Contour plots with significance of 3 with different variable combinations of a) m_{eff}, $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}, b) $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, c) aplanarity , d) $\max(|\eta|_{i=1-4})$ and e) $p_{\mathrm{T}}(j_i)$ for the BDT based signal region BDT3.

B.3.3 BDT4

N°	$p_{\mathrm{T}}(j_{1-4})$	$\lvert\eta\rvert(j_{1-4})$	$\max(\lvert\eta\rvert_{i=1-4})$	$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	Ap	$m_{\mathrm{eff}}(\mathrm{incl.})$	$E_{\mathrm{T}}^{\mathrm{miss}}$	Additional
1	✓	✓	–	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$, N_{jet}
2	✓	✓	–	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$
3	✓	–	✓	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$
4	✓	–	✓	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(\mathrm{incl.})$
5	✓	–	✓	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}(\mathrm{incl.})}$
6	✓	–	✓	✓	✓	✓	✓	–
7	✓	–	✓	✓	✓	–	✓	$H_{\mathrm{T}}(\mathrm{incl.})$
8	✓	–	✓	✓	✓	–	✓	$H_{\mathrm{T}}(4j)$
9	✓	–	✓	✓	✓	–	✓	–
10	✓	✓	✓	✓	✓	✓	✓	–
11	✓	✓	✓	–	✓	✓	✓	–
12	✓	✓	✓	≥ 0.2	✓	✓	✓	–
13	✓	✓	✓	≥ 0.4	✓	✓	✓	–
14	✓	✓	✓	✓	✓	✓	✓	–
15	✓	✓	✓	✓	≥ 0.04	✓	✓	–
16	✓	✓	✓	✓	–	✓	✓	–
17	✓	✓	✓	✓	✓	✓	✓	–
18	✓	✓	≤ 2.0	✓	✓	✓	✓	–
19	✓	✓	–	✓	✓	✓	✓	–

Table B.3: BDT input variable combinations for the variable studies for the BDT based signal region BDT4. The single studies are separated by two horizontal lines. Preferred input variable choices are marked in gray for the single studies.

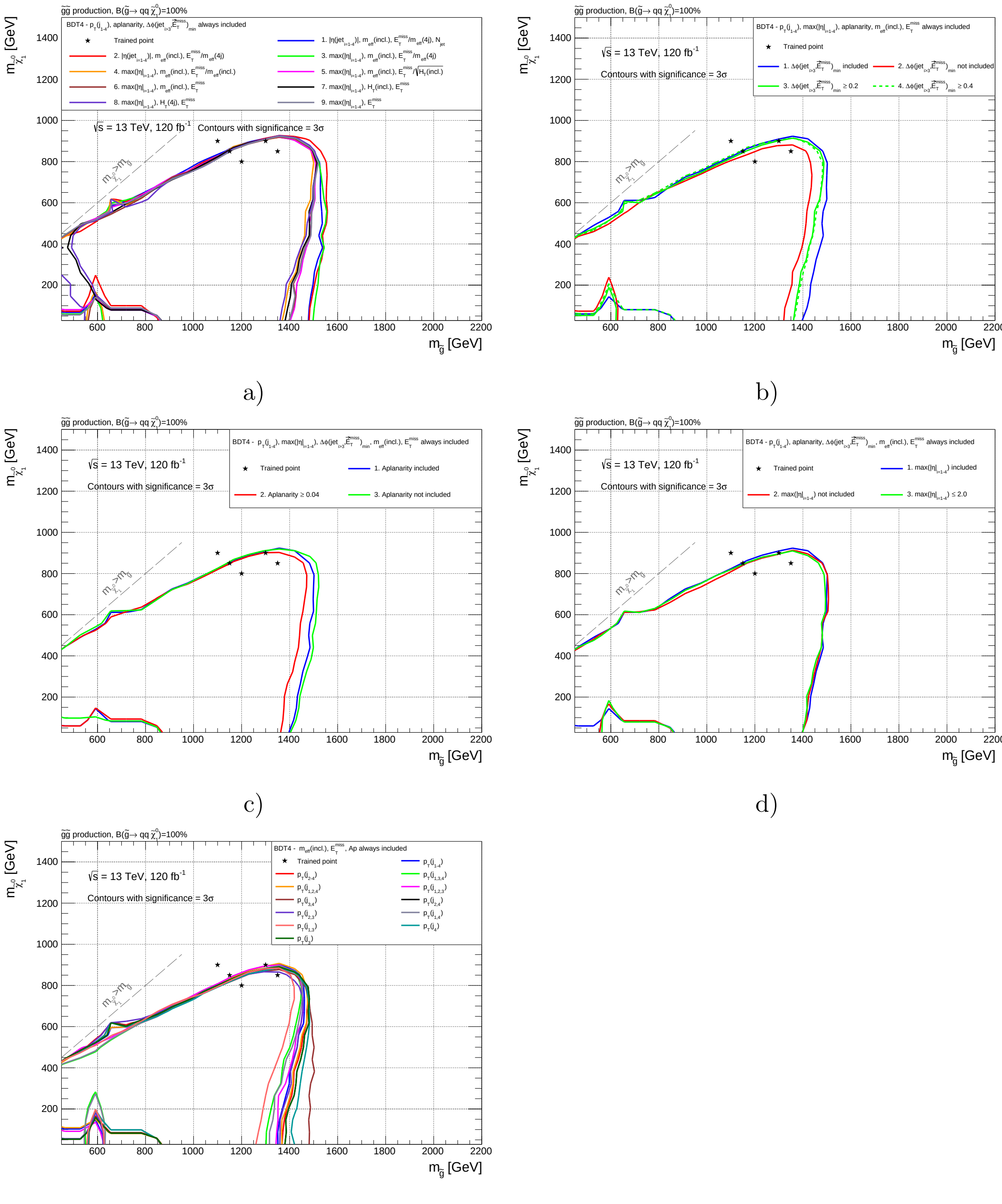


Figure B.5: Contour plots with significance of 3 with different variable combinations of a) m_{eff}, $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}, b) $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, c) aplanarity, d) $\max(|\eta|_{i=1-4})$ and e) $p_{\mathrm{T}}(j_i)$ for the BDT based signal region BDT4.

B.3.4 BDT5

N°	$p_{\mathrm{T}}(j_{1-4})$	$\|\eta\|(j_{1-4})$	$\max(\|\eta\|_{i=1-4})$	$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	Ap	$m_{\mathrm{eff}}(\mathrm{incl.})$	$E_{\mathrm{T}}^{\mathrm{miss}}$	Additional
1	✓	✓	–	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$, N_{jet}
2	✓	✓	–	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$
3	✓	–	✓	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4j)$
4	✓	–	✓	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(\mathrm{incl.})$
5	✓	–	✓	✓	✓	✓	–	$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}(\mathrm{incl.})}$
6	✓	–	✓	✓	✓	✓	✓	–
7	✓	–	✓	✓	✓	–	✓	$H_{\mathrm{T}}(\mathrm{incl.})$
8	✓	–	✓	✓	✓	–	✓	$H_{\mathrm{T}}(4j)$
9	✓	–	✓	✓	✓	–	✓	–
10	✓	✓	✓	✓	✓	✓	✓	–
11	✓	✓	✓	–	✓	✓	✓	–
12	✓	✓	✓	≥ 0.2	✓	✓	✓	–
13	✓	✓	✓	≥ 0.4	✓	✓	✓	–
14	✓	✓	✓	✓	✓	✓	✓	–
15	✓	✓	✓	✓	≥ 0.04	✓	✓	–
16	✓	✓	✓	✓	–	✓	✓	–
17	✓	✓	✓	✓	✓	✓	✓	–
18	✓	✓	≤ 2.0	✓	✓	✓	✓	–
19	✓	✓	–	✓	✓	✓	✓	–

Table B.4: BDT input variable combinations for the variable studies for the BDT based signal region BDT5. The single studies are separated by two horizontal lines. Preferred input variable choices are marked in gray for the single studies.

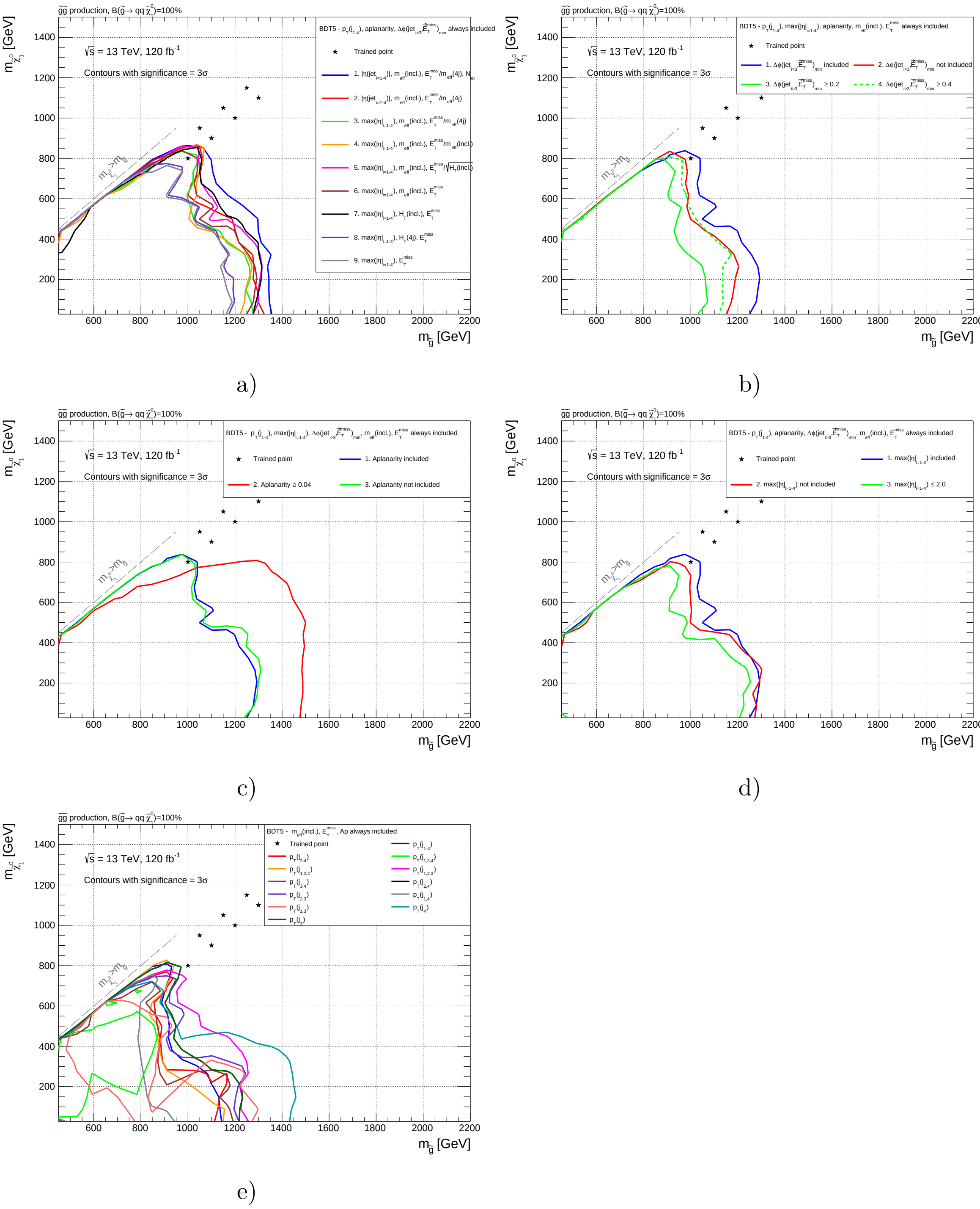


Figure B.6: Contour plots with significance of 3 with different variable combinations of a) m_{eff}, $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}, b) $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, c) aplanarity, d) $\max(|\eta|_{i=1-4})$ and e) $p_{\mathrm{T}}(j_i)$ for the BDT based signal region BDT5.

C Bibliography

[1] ATLAS Collaboration.
Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC.
Phys.Lett., B716:1–29, 2012.

[2] CMS Collaboration.
Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC.
Phys. Lett., B716:30–61, 2012.

[3] ATLAS Collaboration.
Search for squarks and gluinos in final states with jets and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector. .
(ATLAS-CONF-2017-022), 2017.
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2017-022/.

[4] Michael E. Peskin and Daniel V.Schroeder.
An introduction to quantum field theory.
Perseus Books Publishing, 1995.

[5] Particle data group.
Official website: http://pdg.lbl.gov/.

[6] Murray Gell-Mann.
Symmetries of baryons and mesons.
Phys. Rev., 125:1067–1084, Feb 1962.

[7] Sheldon L. Glashow.
Partial-symmetries of weak interactions.
Nuclear Physics, 22(4):579 – 588, 1961.

[8] Abdus Salam.
Gauge unification of fundamental forces.
Rev. Mod. Phys., 52:525–538, Jul 1980.

[9] Steven Weinberg.
A model of leptons.
Phys. Rev. Lett., 19:1264–1266, Nov 1967.

[10] Tadao Nakano and Kazuhiko Nishijima.
Charge independence for v-particles*.
Progress of Theoretical Physics, 10(5):581–582, 1953.

[11] F. Englert and R. Brout.
Broken symmetry and the mass of gauge vector mesons.
Phys. Rev. Lett., 13:321–323, Aug 1964.

[12] Peter W. Higgs.
Broken symmetries and the masses of gauge bosons.
Phys. Rev. Lett., 13:508–509, Oct 1964.

[13] Wim de Boer.
The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology.
In *Time and Matter 2013 (TAM2013) Venice, Italy*, 2013.

[14] ATLAS Collaboration.
Measurement of the higgs boson mass from the $h \to \gamma\gamma$ and $h \to zz^* \to 4\ell$ channels with the atlas detector using 25 fb^{-1} of pp collision data. measurement of the higgs boson mass from the $h \to \gamma\gamma$ and $h \to zz^* \to 4\ell$ channels with the atlas detector using 25 fb^{-1} of pp collision data.
Phys. Rev. D, 90(CERN-PH-EP-2014-122. CERN-PH-EP-2014-122):052004. 35 p, Jun 2014.

[15] Y. Fukuda et al.
Evidence for oscillation of atmospheric neutrinos.
Phys. Rev. Lett., 81:1562–1567, 1998.

[16] Adam G. Riess et al.
Observational evidence from supernovae for an accelerating universe and a cosmological constant.
Astron. J., 116:1009–1038, 1998.

[17] S. Perlmutter et al.
Measurements of Omega and Lambda from 42 high redshift supernovae.
Astrophys. J., 517:565–586, 1999.

[18] Stephen P. Martin.
A Supersymmetry primer.
1997.
[Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[19] NNPDF - NNLO PDF of proton.
https://nnpdf.hepforge.org/images/nnpdf23_nnlo_allpdfs.png.

[20] Chris Blanks.
Strangeness at lhcb.
http://imperialhep.blogspot.de/2011/08/strangeness-at-lhcb.html, August 2011.

[21] http://www.isgtw.org/sites/default/files/img_2014/sherpa_sim_0.jpeg.

[22] J. Alwall et al.
The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations.
JHEP 07 (2014) 079, arXiv: 1405.0301 [hep-ph].

[23] R. D. Ball et al.
Parton distributions with lhc data.
Nucl. Phys. B 867 (2013) 244, arXiv: 1207.1303 [hep-ph].

[24] T. Sjöstrand et al.

An introduction to pythia 8.
Comput. Phys. Commun. 191 (2015) 159, arXiv: 1410.3012 [hep-ph].

[25] ATLAS Collaboration.
Summary of atlas pythia 8 tunes.
ATL-PHYS-PUB-2012-003 (2012).

[26] T. Gleisberg et al.
Event generation with sherpa 1.1.
JHEP 02 (2009) 007, arXiv: 0811.4622 [hep-ph].

[27] C. Oleari S. Alioli, P. Nason and E. Re.
A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box.
JHEP 06 (2010) 043, arXiv: 1002.2581 [hep-ph].

[28] S. Mrenna T. Sjöstrand and P. Skands.
Pythia 6.4 physics and manual.
JHEP 05 (2006) 026, arXiv: hep-ph/0603175.

[29] P. Z. Skands.
Tuning monte carlo generators: The perugia tunes.
Phys. Rev. D 82 (2010) 074018, arXiv: 1005.3457 [hep-ph].

[30] H.-L. Lai et al.
New parton distributions for collider physics.
Phys. Rev. D 82 (2010) 074024, arXiv: 1007.2241 [hep-ph].

[31] J. Pumplin et al.
New generation of parton distributions with uncertainties from global qcd analysis.
JHEP 07 (2002) 012, arXiv: hep-ph/0201195.

[32] ATLAS Collaboration.
Search for squarks and gluinos in final states with jets and missing transverse momentum using 36 fb^{-1} of $\sqrt{s}$=13 TeV pp collision data with the ATLAS detector.
2017, ARXIV:1712.02332.

[33] ATLAS Collaboration.
The atlas experiment at the cern large hadron collider.
JINST 3 (2008) S08003.

[34] Julie Haffner.
The CERN accelerator complex. Complexe des accélérateurs du CERN.
https://cds.cern.ch/record/1621894, Oct 2013.
General Photo.

[35] CMS Collaboration.
The compact muon solenoid technical proposal.
CERN-LHCC-94-38.

[36] ALICE Collaboration.
The alice experiment at the cern lhc.
JINST 3 S08002.

[37] LHCb collaboration.
A large hadron collider beauty experiment.

CERN-LHCC-98-004.

[38] ATLAS Public Results - Luminosity Public Results Run2.
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[39] Joao Pequenao.
Computer generated image of the whole ATLAS detector.
https://cds.cern.ch/record/1095924, Mar 2008.

[40] Karolos Potamianos.
The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider.
PoS, EPS-HEP2015:261, 2015.

[41] ATLAS Collaboration.
Alignment of the ATLAS Inner Detector with the initial LHC data at $\sqrt{s} = 13$ TeV.
(ATL-PHYS-PUB-2015-031), Jul 2015.

[42] ATLAS - Tile Calorimeter - University of Chicago.
https://hep.uchicago.edu/images/Calorimeters_labels.png.

[43] ATLAS - muon spectrometer.
http://inspirehep.net/record/834063/files/MuonSystem_d3.png.

[44] Aranzazu Ruiz-Martinez and ATLAS Collaboration.
The Run-2 ATLAS Trigger System.
Technical Report ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb 2016.

[45] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez.
The Anti-k(t) jet clustering algorithm.
JHEP, 04:063, 2008.

[46] A. Hoecker et al.
TMVA Users Guide.
http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf.

[47] ATLAS Collaboration.
Search for Supersymmetry at the high luminosity LHC with the ATLAS experiment.
(ATL-PHYS-PUB-2014-010), Jul 2014.

[48] Yoav Freund and Robert E. Schapire.
Experiments with a new boosting algorithm.
In *IN PROCEEDINGS OF THE THIRTEENTH INTERNATIONAL CONFERENCE ON MACHINE LEARNING*, pages 148–156. Morgan Kaufmann, 1996.

[49] Jerome H. Friedman.
Greedy function approximation: Agradient boosting machine.
Annals of Statistics 29 (2000) 1189–1232.

[50] G. Choudalakis and D. Casadei.
Plotting the differences between data and expectation.
The European Physical Journal Plus, 127(2):25, Feb 2012.

[51] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells.
Asymptotic formulae for likelihood-based tests of new physics.

Eur. Phys. J., C71:1554, 2011.
[Erratum: Eur. Phys. J.C73,2501(2013)].

[52] ATLAS Collaboration.
Discrimination of Light Quark and Gluon Jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector.
(ATL-COM-PHYS-2015-350), Jul 2016.

List of Figures

List of Tables

Erklärung

Ich versichere hiermit, dass die vorliegende Masterarbeit selbstständig verfasst und keine weiteren als die angegebenen Hilfsmittel benutzt sowie die Stellen der Arbeit, die in anderen Werken dem Wortlaut oder dem Sinn nach entnommen sind, durch Angaben der Quellen sichtbar gemacht wurden. Außerdem versichere ich, dass die vorgelegte Masterarbeit nicht Bestandteil eines anderen Prüfungsverfahrens ist und war.

Freiburg, 16.03.2018

(Ort, Datum) (Unterschrift)