Standard Model Background Studies for SUSY Searches with the CMS Detector at the LHC

Muammer Altan Çakir

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik des
KARLSRUHER INSTITUT FÜR TECHNOLOGIE

genehmigte

Dissertation

von

M.Sc. Muammer Altan Çakır
aus Ankara

Tag der mündlichen Prüfung: 02.07.2010

Referent: Prof. Dr. W. de Boer, Institut für Experimentelle Kernphysik,
Karlsruher Institut für Technologie

Korreferent: Prof. Dr.G. Quast, Institut für Experimentelle Kernphysik,
Karlsruher Institut für Technologie

"Prediction is very difficult, especially about the future."
Niels Bohr

Contents

Chapter 1

Introduction

High energy and particle physics enter a new era with the start-up of the Large Hadron Collider (LHC) at CERN. This is a proton-proton accelerator located at the Swiss-French border with a circumference of 27 km. The LHC allows for very precise test of the Standard Model (SM), which provides a detail description of the strong, weak and electromagnetic interactions. The SM describes the most observations extremely well and as such has been discussed in Nobel prize lectures and standard textbooks [1–7]. Although these interactions have been tested in previous experiments, the SM contains some inconsistencies leading to the belief that it is really part of some larger theory. Extensions to the SM try to fix the inconsistencies and add some simplicity and beauty to the model. One such extension is Supersymmetry (SUSY)(see reviews [8–13]) which provides an attractive way towards a unification of the electroweak and strong interactions. In addition it provides a candidate for Dark Matter (DM) and a natural mechanism for electroweak symmetry breaking.

The Compact Muon Solenoid (CMS) is a new experiment designed to observe and record these interactions. Physical theories, the SM and extensions like SUSY, predict what could happen. From all theories it is crucial to compare these predictions, which are simulated by Monte-Carlo (MC) generators, with the experimental observations [14–18]. In particular, it is crucial to understand theoretical and experimental uncertainties in these comparisons.

Physics beyond the SM is expected to consist of new particles, else these particles would have been discovered already at present accelerators [19–22]. Heavy particles decay to lighter particles, which will have higher transverse momentum (P_T) with respect to to the beam axis than the decay of light particles simply because of kinematics. In addition, SUSY is expected to lead to missing transverse energy (E_T^{Miss}) from the escaping neutralinos, the candidates for dark matter. So in calculating the SM background for the prediction of SUSY particles it is important to have a good description of P_T and E_T^{Miss} especially for the production of the heaviest SM particles, like the heavy gauge $W^{\mp}$, Z bosons and the top quark, which are naturally the main SM background for the production of even heavier particles.

In general the popular parton shower generators, like Pythia and Herwig, yield a good description of the event structure in hadronic interactions, except for the large P_T - tails [23–27]. This is expected, since the parton shower models are based on the leading-log resummation of all leading logarithms occurring in the cascade of quark and gluon emission (called parton shower) starting after the interaction of the initial partons. These emissions are dominated by the divergences of the soft- and collinear parton radiation, so the hard process are not taken into account correctly in the leading-log approximation (LLA). This can be remedied by calculating the hard interactions with a matrix-element (ME) calculation and let the emerging partons hadronize via the parton shower model. However, to combine the two approaches, ME and LLA, one has to be careful not to double count parton production in the overlapping regions of phase space. This is usually achieved by *matching schemes*. Since different generators use different shower algorithms, different ME and different matching schemes, one expects differences in the observables.

This has indeed been observed at the Tevatron, the proton-antiproton collider at the Fermi National accelerator Laboratory in Batavia, Illinois, which has a center of mass energy of 1.96 TeV (to be compared with 14 TeV at the LHC) [28–31]. It was found that some generators describe well the shape of the P_T spectrum, but not to magnitude, others describe the magnitude better, but not the P_T. The main question for the LHC is: how large become the differences between different generators if extrapolated to LHC energies, given the tuned input parameters sets for Tevatron energies?

The large center of mass energy of the LHC could allow for the production of SUSY particles, which leave signatures characterized by missing transverse energy (E_T^{Miss}) from escaping neutrinos in addition to the multi-jets and/or multi-lepton final states from decaying SUSY-particles. The main subject of the thesis is the calculation of the discovery reach for SUSY particles in these various topologies with early CMS-SUSY searches at $\sqrt{s}$=10 TeV and an integrated luminosity of $\mathcal{L}$=100pb^{-1} with special emphasis on the systematic uncertainties estimated from MC-comparisons. 100pb^{-1} can be accumulated within one year of running. The most popular simulation programs for hadron collisions have been compared: MadGraph, Alpgen, Sherpa and MC@NLO. The major characteristics of physics observables are studied with and without CMS detector reconstruction.

The thesis is organized as follows: After describing the SM and SUSY with the corresponding experimental constraints in Chapter 2, the basic technical details on the LHC and the CMS detector follow with the reconstruction of physics objects and the CMS software analysis framework in Chapter 3. The Monte-Carlo methods and approaches are discussed in Chapter 4. Chapter 5 is devoted to the various MC generator comparisons, followed by the discovery reach for the various SUSY searches in Chapter 6.

Chapter 2

Standard Model and Supersymmetry

The matter of the universe we live in is made up of tiny building blocks held together by appropriate forces. The known forces are coded in the so-called Standard Model (SM) of particle physics, which describes simultaneously the strong nuclear force, the electromagnetic force and the weak force. These forces mediate the interactions between the matter particles consisting of 6 quarks and 6 leptons by the exchange of gauge bosons, the gluon for the strong force, the photon for the electromagnetic force and the heavy $W^{\mp}$, Z^0 bosons for the weak force.

Although, the SM has been confirmed in numerous high energy experiments with extreme good precision in the past decades, it cannot explain why the different forces have such different strengths, the large hierarchy between the electroweak and gravity scale, the existence dark matter and so on. The SM can be divided into three parts: the fundamental matter particles, which are spin 1/2 fermions; the carriers of the fundamental forces, which are spin 1 bosons: and the Higgs mechanism, which allows particles to obtain mass and predicts a hitherly undiscovered spin 0 Higgs boson. The chapter will finish with presenting the problems of the SM, giving reasons for the need to expand it by new physics, e.g. *Supersymmetry*.

2.1 Standard Model of Elementary Particle Physics

The Standard Model describes the universe in terms of 12 fundamental matter particles, which are spin 1/2 fermions. These particles interact through four fundamental forces, the exchange of integer spin gauge boson particles. These four fundamental fields are the electromagnetic field, the weak field, the strong field, and the gravitational field. The gravitational field is not included in the Standard Model, as its relative strength is insignificant for the small masses of elementary particles.

The matter particles are classified as either quarks or leptons. The leptons and quarks both interact though the electromagnetic and weak forces, but only the

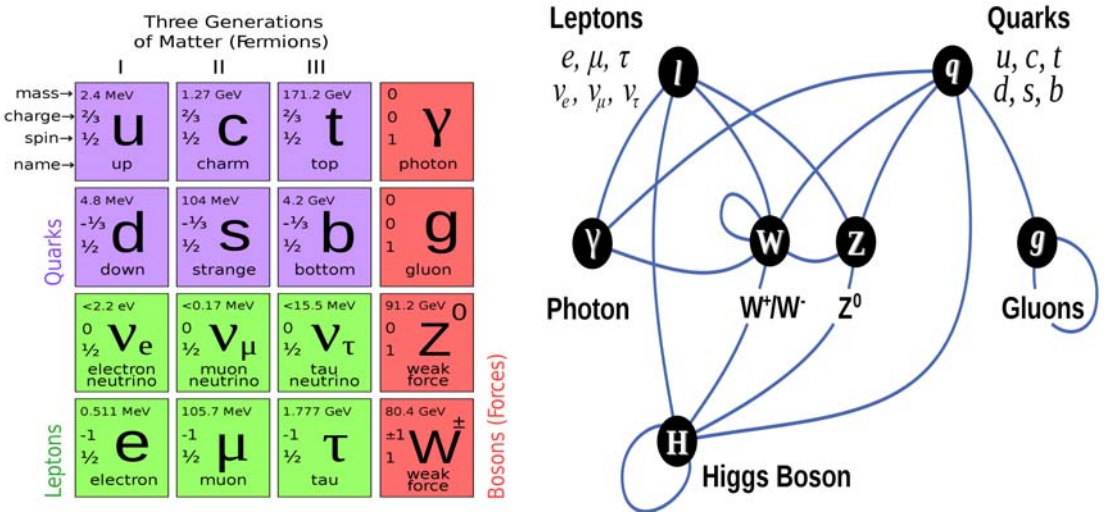


Figure 2.1: Left: The Standard Model of elementary particles, with the gauge bosons in the rightmost column. Right: Summary of interactions between particles described by the Standard Model [32].

quarks interact through the strong force. The properties of the leptons and quarks are summarized in Figure 2.1. The force particles (the photon, the Z and $W^{\pm}$, the gluons) mediate interactions between matter particles. The Standard Model is a quantum field theory based on a combination of gauge symmetry groups. Each particle is fully specified by quantum numbers and its mass. The electric and weak charges (see Section 2.1.3) specify the interaction strength of the electroweak force, which unifies the electromagnetic force (described by Quantum-Electro-Dynamics (QED)) with the weak force; similarly, the color charge defines the interaction strength of the strong force (described by Quantum-Chromo-Dynamics (QCD)). The mass of a particle is given by its interaction with the Higgs field. The main idea of all symmetry groups will be reviewed and discussed with the physical observables in the next subsections.

2.1.1 Local Gauge Symmetry

In the absence of interactions, massless fermions can be described by the Lagrangian[1]:

[1]The position of an object as a function of time, is determined by the lagrangian, which is defined as:

$$L(q_i, \dot{q}_i, t) = T - U \tag{2.1}$$

where the T and U denote kinetic and potential energy of a particle with the coordinates q_i and their derivatives $\dot{q}$. The equation of motion (Euler-Lagrange equations) can be defined via the Lagrangian:

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) = (\frac{\partial L}{\partial q_i}) \quad (i = 1, 2, 3) \tag{2.2}$$

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu \partial_\mu \Psi), \tag{2.4}$$

with γ_μ being Dirac matrices, Ψ a fermionic field and $\bar{\Psi}$ its adjoint field. The corresponding equation of motion is the Dirac equation which describes relativistic spin-1/2 fermions. Symmetry groups and gauge invariance are the guiding principles for introducing additional terms in the Lagrangian to describe interactions are those of symmetry groups and gauge invariance. These principles are known to give the correct interaction terms for electromagnetism and have been successfully applied in understanding the electroweak force and the strong force [33]. Unitary symmetry group transformations of a group SU, where S means special and implies a determinant equal | 1 |, can be written as:

$$\Psi \rightarrow \Psi' = U\Psi \quad ; \quad U = e^{i\tau_j \theta_j(x)}$$

where τ_j is the j^{th} generator of the group and $\theta(x)$ is an arbitrary function of the 4-coordinate x. For the group U(1) there is one such generator, a scalar constant, whereas for SU(N), N = 2, 3, there are $N^2 - 1$ generators which can be represented as traceless NxN matrices. The requirement that the Lagrangian (Eq. 2.4) has local gauge invariance, i.e. allowing $\theta_j(x)$ in Eq 2.5 to be a function of space-time, leads to the introduction of new field(s), A_μ, through a so-called minimal substitution:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - ig\frac{\tau^i}{2}A_{i\mu} \tag{2.5}$$

where τ_i are the generator(s) of the symmetry group in question, g is a coupling constant, and $A_{i\mu}$ are spin-1 gauge boson fields. The gauge transformation properties of the gauge fields are defined to cancel the term arising from the partial derivative in Eq. 2.4 acting on $\theta_j(x)$, defined in Eq. 2.5. The substitution defined by Eq. 2.5 gives one new vector field, with an associated vector boson, per generator of the symmetry group. The vector field(s) per construction interact with the fermionic field of Eq. 2.4, and can be interpreted as the messenger particle of the force arising from the gauge symmetry. For SU(2)⊗U(1) the gauge fields describe the electroweak force, whereas the SU(3) gauge fields describe the strong force.

2.1.2 Lagrangian

The Standard Model is a quantum field theory based on the principle of gauge invariance. It is specified in terms of a Lagrangian density, $\mathcal{L}_{\mathcal{SM}}$, which determines the equations of motion through a minimization of the action, $\mathcal{S}=\int d^4x\mathcal{L}_{\mathcal{SM}}$. The

Hence, if the coordinates of L (from classical mechanic expression) are replaced by the field density of ϕ, the corresponding Lagrangian becomes $\mathcal{L}(\phi_i, \partial\phi_i, x_\mu)$ via $\partial_\mu\phi_i \equiv \frac{\partial\phi_i}{\partial x^\mu}$ and Euler-Lagrange equation for the Lagrange density:

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)} = \frac{\partial\mathcal{L}}{\partial\phi_i} \quad (i = 1, 2, 3, 4).\right. \tag{2.3}$$

Lagrangian is required to be invariant under the Lorentz transformation, so as to give a relativistic theory, and it contains terms describing free fields as well as interactions between different fields. The SM Lagrangian is defined as follows:

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{\mathcal{QCD}} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \tag{2.6}$$

The two first terms, $\mathcal{L}_{EW}$ and $\mathcal{L}_{\mathcal{QCD}}$, describe free fermions, free bosons (gauge bosons) associated with the SU(2)⊗U(1) and SU(3) gauge symmetries, the interaction between fermions and gauge bosons, and the interactions among gauge bosons themselves. The terms $\mathcal{L}_{Higgs}$ and $\mathcal{L}_{Yukawa}$ introduce the Higgs particle and non-zero gauge boson and fermion masses.

2.1.3 Weak Interactions

The SU(2)⊗U(1) symmetry of the SM gives the unification of the electromagnetic and weak forces through the Glashow-Weinberg-Salam model [34]. The $SU(2)_L$ symmetry distinguishes between *Left handed*(L) and *Right handed*(R)[2] fermion states,

$$\Psi_L = P_L\Psi \quad and \quad \Psi_R = P_R\Psi \tag{2.7}$$

where P_L and P_R projection operators which satisfy P_LP_R=$P_RP_L = 0$ and P_L+P_R = 1, implying $P_{L,R} = P^2_{L,R}$. The gauge bosons of the $SU(2)_L$ group couple only to left-handed fermions, and the Lagrangian has the form

$$\mathcal{L}_{EW} = i\bar{L}\gamma^\mu D_\mu L + i\bar{R}\gamma^\mu D_\mu R - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}W^{\mu\nu} \tag{2.8}$$

where the L represents left-handed fermion doublets and R the corresponding right-handed singlets. The covariant derivative is given in analogy to Eq. 2.5,

$$D_\mu = \partial_\mu + ig_1\frac{Y}{2}B_\mu + ig_2\tau_i W^i_\mu \tag{2.9}$$

with B_μ being the $U(1)_Y$ gauge field, g_1 the $U(1)_Y$ coupling constant and Y is the group generator. Similarly, for $SU(2)_L$, g_2 is coupling constant,(τ_i) are the group generators and W^i_μ (i=1, 2, 3) represent the gauge fields. Combining the left- and right-handed components of these fields into a single Lagrangian leads to *electroweak unification*. Two of the $SU(2)_L$ fields can be combined linearly into the physical, charged fields:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}[W^1_\mu \mp iW^2_\mu], \tag{2.10}$$

[2]The helicity of a particle is Right-handed if the direction of its spin is the same as the direction of its motion. It is Left-handed if the direction of spin and motion are opposite. Mathematically, helicity is the sign of the projection of the spin vector onto momentum vector: Left is negative, Right is positive.

which have been observed as $W^{\pm}$ *bosons*. The remaining neutral $SU(2)_L$ field mixes with the $U(1)_Y$ field through the weak mixing angle, θ_W, giving the fields A_μ and Z_μ, represented as

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W, \quad \textit{and} \quad Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W. \tag{2.11}$$

The weak mixing angle is defined by requiring the A_μ field to correspond to *photons*. This is fulfilled by the choice $g_1 \sin\theta_W = g_2 \cos\theta_W = \mathrm{e}$, with e is the absolute value of the electric charge of an electron. The field of Z_μ is observed as the *Z boson*. The charge corresponding to the $U(1)_Y$ symmetry is called hyper-charge, Y, whereas for $SU(2)_L$ it is referred to as weak isospin, I_3. These charges are related to the electrical charge, Q, through

$$Q = \frac{1}{2}Y + I_3 \tag{2.12}$$

2.1.4 Strong interactions and QCD

The $SU(3)$ symmetry of the Standard Model describes the strong force between quarks. The theory of $SU(3)$ is referred to as QCD that emerges when the naive quark model is combined with SU(3) local gauge symmetry [35–37]. The quark model classifies the large number of hadrons in terms of a few, more fundamental constituents. Baryons consist of three quarks, while mesons are made of a quark and an anti-quark. In this sense, the quark model requires an additional quantum state because the Pauli exclusion principle would not allow for a particle with three same quark states with spin 3/2. The only way to construct a symmetric wave-function is the postulating of an additional quantum number, which is called *color*. Quarks can exist in three different color states, called *red, green and blue*. Remember from local gauge invariance

$$\Psi \rightarrow \Psi' = U\Psi \quad ; \quad U = e^{i\theta_j(x)T^i}$$

In electrodynamics, there is only one electric charge, and a gauge transformation involves a single phase factor, $\mathrm{U}=e^{i\theta_j(x)}$. In QCD one has three different colors and U becomes a (complex) unitary 3x3 matrix. These matrices form the fundamental representation of the group SU(3), where 3 equals the number of colors. The matrix U has 8 independent elements and can therefore be parameterized in terms of 8 generators T^i above. The Lagrangian of QCD is given as

$$\mathcal{L}_{\mathcal{QCD}} = i\bar{q}_f \gamma^\mu D_\mu q_f - \frac{1}{4} G^i_{\mu\nu} G^{\mu\nu}_i \tag{2.13}$$

with q_f being a color triplet of quarks of flavor f. The covariant derivative, D_μ is given by

$$D_\mu = \partial_\mu - ig_3 \frac{\lambda_i}{2} G^i_\mu \tag{2.14}$$

Here, λ_i (i = 1...8) are the eight $SU(3)$ generators and G_i the corresponding gauge fields with g_3 being the coupling constant. The quanta of the $SU(3)$ fields are called *gluons*.

2.1.5 The Higgs Mechanism

Introducing mass terms -$m^2\bar{\Psi}\Psi$, where Ψ represents any SM field, would break $SU(2)_L$ invariance of the SM (see Eq. 2.4) . This follows from $\bar{\Psi}\Psi{=}\bar{\Psi}(P_L{+}P_R)\Psi{=}\bar{\Psi}(P_L^2{+}P_R^2)\Psi{=}\bar{\Psi}_R\Psi_L + \Psi_L\Psi_R$ and since left- and right-handed fields transform differently under the $SU(2)_L$ gauge transformation defined by Eq 2.5

$$\begin{aligned} \Psi_L &\rightarrow \Psi'_L = U\Psi_L \\ \Psi_R &\rightarrow \Psi'_R = \Psi_R \end{aligned} \tag{2.15}$$

Instead of direct mass terms masses are generated in the SM by postulating the interactions with a scalar bosonic field with a non-zero vacuum expectation value, referred to as the Higgs field [38]. The Higgs field is a scalar, complex $SU(2)_L$ doublet with Y=1 of one charged and one neutral Higgs particle:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.16}$$

described by the Lagrangian

$$\mathcal{L}_{Higgs} = |D^\mu \Phi|^2 - V(\Psi) \quad ; \quad V(\Psi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \tag{2.17}$$

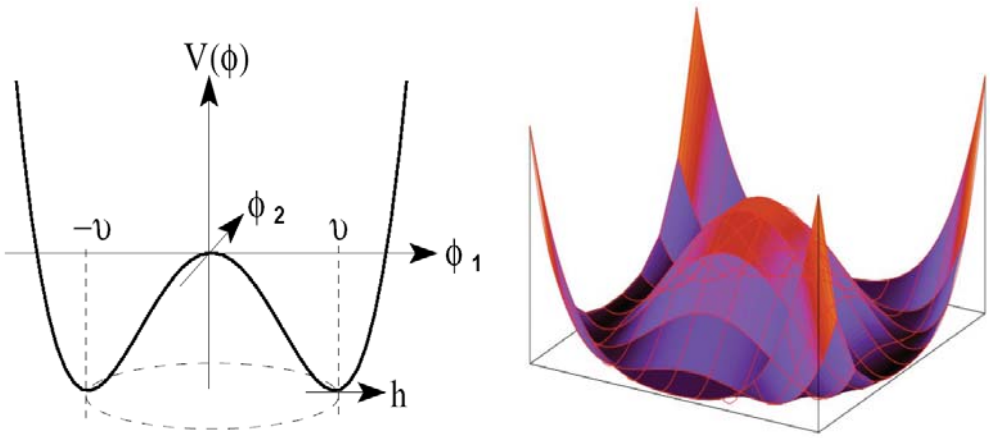


Figure 2.2: The Standard Model Higgs potential for a complex field Φ [39].

with $\lambda > 0$ for the scalar potential $V(\Psi)$ to be bounded from below. The covariant derivative, defined in Eq. 2.9, introduces coupling terms between the Higgs field and the gauge bosons. For $\mu^2 < 0$, the ground state, Φ_0, can be written as

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon \end{pmatrix} \quad with \quad \upsilon = \sqrt{\frac{-\mu^2}{\lambda}} \tag{2.18}$$

corresponding to a non-vanishing vacuum expectation value

$$< 0|\Phi^2|0 >= \frac{\upsilon^2}{2} \tag{2.19}$$

Despite the $SU(2)_L \mathrm{x} U(1)_Y$ symmetry of the Lagrangian the ground state is not symmetric. This is called Spontaneous Symmetry Breaking (SSB). The field Φ can be expanded around the ground state as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta_1 + i\eta_2 \\ \upsilon + h + i\eta_3 \end{pmatrix} = \begin{pmatrix} 0 \\ \upsilon + h \end{pmatrix} \tag{2.20}$$

The last equality follows from transforming from an arbitrary to the unitary gauge. The excitations h around the minimum corresponds to the quanta of the field, the Higgs boson. In the unitary gauge the three η_i fields appear as longitudinal degrees of freedom for the Z and $W^{\pm}$ gauge bosons [40]. The three gauge bosons thus acquire masses, given as

$$m_W = \frac{1}{2}\upsilon g_2 \quad , \quad m_Z = \frac{m_W}{\cos\theta_W} \tag{2.21}$$

In contrast, the A_μ field, defined in Eq. 2.11, remain massless. The masses m_W and m_Z can be expressed in terms of well-measured gauge couplings, as required by the Higgs mechanism: masses are generated by interactions instead of ad hoc mass terms in the Lagrangian. Therefore the mass ratios are predicted as ratios of gauge couplings, in agreement with current experimental data. This Higgs mass given by:

$$m_H = \sqrt{-2\mu^2} \tag{2.22}$$

and is not predicted by the SM since μ is essentially a free parameter, only required to satisfy $\mu^2 < 0$.

Since the fermions also have a mass, additional terms have to be added explicitly to the Lagrangian. These terms, called Yukawa terms, are defined for electrons as follows,

$$\mathcal{L}_{Yukawa} = h_e \left[(\bar{\nu}_e, \bar{e}_L) \begin{pmatrix} 0 \\ \upsilon + h \end{pmatrix} e_R + \bar{e}_R (0, \upsilon + h) \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \right]$$

$$\mathcal{L}_{Yukawa} = h_e \upsilon (\bar{e}_L e_R + \bar{e}_R e_L) + (\bar{e}_L e_R + \bar{e}_R e_L) h \tag{2.23}$$

The second part describes the interaction of an e^+e^--pair with a Higgs boson h. The first term has the form of a Dirac mass term for a spin 1/2 particle with mass m_e=$h_e v$. Similar Yukawa terms are introduced for the other charged leptons and for the quarks and the couplings of the Higgs boson to fermions are always proportional to the fermion masses. The general $\mathcal{L}_{Yukawa}$ in Eq. 2.6 is defined as follows

$$\mathcal{L}_{Yukawa} = h_e^i \bar{L}^i \Phi e_R^i + h_d^{ij} \bar{Q}^i \Phi d_R^{ij} + h_u^{ij} \bar{Q}^i \Phi^C u_R^{ij} + h.c \tag{2.24}$$

Here i and j are the indices of the three generations, L^i is a left-handed lepton doublet, Q^i is a left handed quark doublet and e_R, d_R and u_R are the right-handed partners of the isospin doublets. The mass matrix of the quarks is not diagonal, i.e. interaction eigenstates are not mass eigenstates. The diagonalization is done with a 3x3 matrix, called the Cabibbo-Kobayashi-Maskawa(CKM) matrix [41].

2.1.6 The problems in the Standard Model and what lies beyond?

Although the SM is one of the best verified theories in physics, the SM is nevertheless incomplete. **Experimental** arguments in support of this are [42–44]:

- The solar and atmospheric neutrino data, interpreted as neutrino oscillations, require **neutrinos to have mass**.
- Cosmological observations have established the existence of **cold dark matter** in the Universe for which there is no candidate in the SM.
- Observations of supernovae at large red shift as well as the cosmic microwave background radiation suggest that the bulk of the energy the Universe resides in a novel form, called *dark energy*. This could be the cosmological constant first introduced by Einstein, or something else.
- Gravity exists.

There are also **theoretical** considerations that suggest that the SM cannot be the complete picture.

- understanding of particle masses and mixing patterns requires a **large number of underlying parameters in the SM**[3].

[3]The SM has $18 + 3$ free parameters: three parameters α, α_s and $\sin\theta_W$ corresponding to the coupling, 2 parameters of the Higgs potential corresponding to the mass of the Z- and Higgs boson, $9 + 3$ parameters Yukawa couplings for the 6 quarks and 3 leptons. Since neutrinos are not massless, there are 3 further couplings. And finally 4 parameters from CKM mixing matrix in the quark sector and 4 parameters for the MNS mixing matrix in the lepton sector.

- The **choice of the symmetry group of SM** and particle representations in $SU(3)_C$x$SU(2)_L$x$U(1)_Y$. A higher symmetry group, such as $SU(5)$ or $SO(10)$, in which all quarks and leptons can be arranged in larger multiplets, is more appealing [45].

- Origin of the spontaneous **electroweak symmetry breaking**. The new scalar fields are introduced "by hand " via an arbitrary scalar potential; there is no understanding of why the squared mass parameter for the Higgs field is negative.

- **Hierarchy problem**: why is the electroweak scale so small compared with a possible unification scale at the order of $M_W \approx 10^{-17} M_{Planck}$.

- The radiative corrections to M_H include quadratically divergent terms. To keep these corrections small compared with typical mass scales one has to introduce a cut-off scale, above which the SM is not valid anymore. For the SM this means new physics in the TeV region.

- The idea of **gauge coupling unification** for all symmetries is to have the same gauge coupling at high energies $M_{GUT} \approx 10^{16}$. At low energies according to the renormalization group evolution. The gauge unification is the basic motivation of the gauge unification theory (GUT) and the string theories which attempt to incorporate all fundamental interactions including gravity. However the experimental results of the values of the low energy gauge couplings show that the SM can not unify the gauge couplings accurately.

Bosons and fermions provide corrections to the Higgs mass with a different sign. If one postulates *a partner for each particle, which differs by* $1/2$ *unit in spin*, but otherwise has the same quantum numbers, this performs an automatic cancellation of the form

$$\delta m_H^2 = -\frac{\lambda_F^2}{16\pi^2}(\Lambda^2 + m_F^2) + \frac{\lambda_F^2}{16\pi^2}(\Lambda^2 + m_B^2) + ... = \mathcal{O}(\frac{\alpha}{4\pi})|m_B^2 - m_F^2|, \qquad (2.25)$$

Assuming the Yukawa couplings for fermions and bosons to be equal ($\lambda_F = \lambda_B$). This reduces the corrections to an acceptable level, as long as the masses do not differ much more than about a TeV. The fine tuning and Hierarchy problem, explained before, can therefore be handled in an elegant way. According to this argument Supersymmetric partners of the SM particles should not to be too heavy and can be found at LHC energies.

Introduction of particles at the TeV scale changes the slope of the energy dependence of the couplings in such a way that perfect unification is possible. In Figure 2.3 the running in the SM and the MSSM are compared [46].

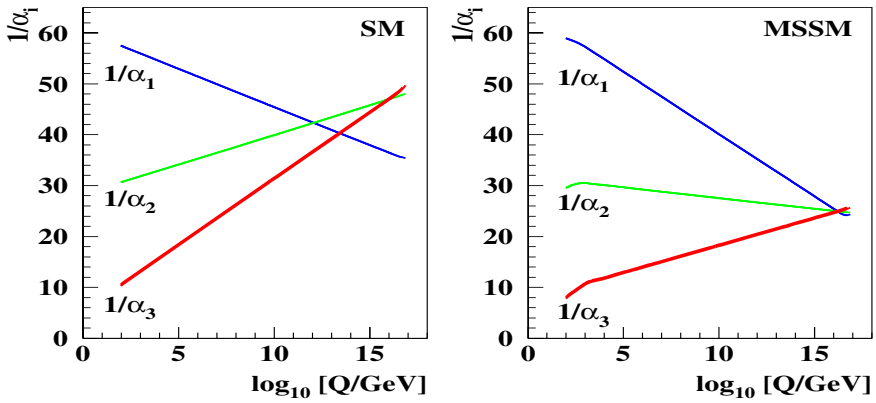


Figure 2.3: Unification of the coupling constants in the SM (Left) and Minimal Supersymmetric SM(Right) [46].

2.2 Supersymmetry

It is easily seen from the examples that a new symmetry is needed for stabilizing a scalar (the Higgs) mass against quantum corrections. Therefore one searches for a theory including such a symmetry that can protect the Higgs mass from quadratically divergent corrections. This symmetry must connect fermions and bosons. There must be a generator of this symmetry that turns a bosonic state into a fermionic one, and vice versa. If this were possible, it would imply that bosons and fermions are merely different manifestations of the same state and in some sense would correspond to an ultimate form of unification. For a long time, it was believed that such a symmetry transformation was not possible to implement into physical theories. At present, however, such transformations can be defined, and in fact there exist theories that are invariant under such transformations. These transformations are known as Supersymmetry (SUSY) transformations. This new symmetry, which connects bosons and fermions, is called *Supersymmetry*.

$$Q|Boson\rangle \propto |Fermion\rangle \quad and \quad Q|Fermion\rangle \propto |Boson\rangle \tag{2.26}$$

Obviously, also the hermitian conjugate $Q^\dagger$ has to be a symmetry generator and both carry a spin of 1/2. It can also be shown that the supersymmetry generators commute with the space-time momentum operator [47].

2.2.1 The Minimal Supersymmetric Standard Model(MSSM)

There is a large number of possible supersymmetric theories. All models considered in this analysis are based on the Minimal Supersymmetric Model(MSSM), which is a direct supersymmetrization of the SM, except for the fact that a second Higgs doublet field has to be introduced and R-parity conservation is assumed.

Superfield	Bosons	Fermions	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Gauge					
$\mathbf{G^a}$	g^a	$\tilde{g}^a$	8	0	0
$\mathbf{V^k}$	W^μ $(W^\pm, Z_0)$	$\tilde{W}^\mu$ $(\tilde{W}^\pm, \tilde{Z}_0)$	1	3	0
$\mathbf{V'}$	B/γ	$\tilde{B}$ / $\tilde{\gamma}$	1	1	0
Matter					
$\mathbf{L_i}$	Sleptons $\tilde{L}_i = (\tilde{\nu}, \tilde{e})_L$	Leptons $L_i = (\nu, e)_L$	1	2	-1
$\mathbf{E_i}$	Sleptons $\tilde{E}_i = \tilde{e}_R$	Leptons $E_i = e_R$	1	1	2
$\mathbf{Q_i}$	Squarks $\tilde{Q}_i = (\tilde{u}, \tilde{d})_L$	Quarks $Q_i = (u, d)_L$	3	2	1/3
$\mathbf{U_i}$	Squarks $\tilde{U}_i = \tilde{u}_R$	Quarks $U_i = u_R$	3	1	$-4/3$
$\mathbf{D_i}$	Squarks $\tilde{D}_i = \tilde{d}_R$	Quarks $D_i = d_R$	3	1	2/3
Higgs					
$\mathbf{H_d}$	Higgs H_d	Higgsino $\tilde{H}_d$	1	2	-1
$\mathbf{H_u}$	Higgs H_u	Higgsino $\tilde{H}_u$	1	2	1

Table 2.1: Particle spectrum of the MSSM [12]. Gauge and Chiral multiplets are represented in the MSSM.

The *minimal* is identified with the minimal number of new supersymmetric particles. Supersymmetry associates known bosons with new fermions and known fermions with new bosons. Table 2.1 shows the particle spectrum of the MSSM.

As left- and right-handed fermions have different gauge transformations (see Section 2.1.5), there are two scalars for each fermion called sfermions. These sfermions form together with the SM fermions *chiral supermultiplets*. The notation left or right refers to the SUSY-partner of a left- or right-handed fermion. The superpartner of the SM gauge bosons are spin 1/2 gauginos that also have two helicity states. These gauge-bosons and gauginos form a *gauge* or *vector supermultiplet*.

In the SM, baryon- and lepton-numbers are conserved because of gauge invariance. In supersymmetric theories it is possible to construct renormalizable operators that do not conserve these numbers, but are still consistent with SM gauge symmetries and supersymmetry. As the proton has a lifetime of more than 10^{33} years, terms that violate both baryon and lepton numbers have to be small. With the introduction of R-parity conservation these terms are excluded. R-parity is defined by:

$$R = (-1)^{3(B-L)+2S} \tag{2.27}$$

where B and L are baryon- and lepton-numbers, respectively, and S is the spin. All SM particles have even R-parity, while their superpartners are R-odd. Therefore, there can be no mixing between SM particles and sparticles. Furthermore, the lightest supersymmetric particle (LSP) has to be stable, if R is conserved [13]. This is assumed for this analysis.

2.2.2 Lagrangian of the MSSM

To build the MSSM Lagrangian it must be invariant under the gauge group $SU(3)_C$x $SU(2)_L$x$U(1)_Y$, it must also include the superfields with the particle content of the Table 2.1. In addition it must contain the terms that break supersymmetry since, as no scalar particles have been found at the electroweak scale, we may directly infer that, even if SUSY exists, it must be broken. The MSSM lagrangian can be composed of two main parts:

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{SUSY-Breaking}} \quad . \tag{2.28}$$

The first term, $\mathcal{L}_{\text{SUSY}}$, can be divided into two terms: the first one is generally represented as the *Gauge-terms* part that include the supersymmetric generalization of the terms containing the kinetic energy of the gauge fields and the interaction of fermions and gauge fields. The second one is the supersymmetric version of the Higgs fields, defined as *Yukawa-terms*.

$$\mathcal{L}_{\text{SUSY}} = \mathcal{L}_{\text{Gauge}} + \mathcal{L}_{\text{Yukawa}} \quad . \tag{2.29}$$

The superpotential[5] W_R of the Yukawa term $\mathcal{L}_{\text{Yukawa}}$ can be written as:

$$W_R = \varepsilon_{ij}(h_U^{ab} Q_a^j U_b^C H_u^i + h_D^{ab} Q_a^j D_b^C H_d^i + h_L^{ab} L_a^j E_b^C H_d^i + \mu H_u^i H_d^j) \quad , \tag{2.30}$$

with the $SU(2)$ indices $i, j = 1, 2$, the generation labels $a, b = 1, 2, 3$, the Yukawa couplings $h_{U,D,L}$ and the antisymmetric tensor ε_{ij}. R refers to the R-parity defined in Eq. 2.27. Charge conjugated fields are denoted by the label C. This part of the Lagrangian is similar to that of the SM, but in comparison, the superpotential contains only superfields rather than the ordinary fields of the SM. Additional the last term describes the Higgs mixing, which is absent in the SM, since here only one Higgs field appears.

From the Lagrangian (Eq. 2.29) we can obtain the full MSSM particle spectrum, as well as the interactions, which contain the usual gauge interactions, the fermion-Higgs interactions that correspond to the two Higgs doublet model (2HDM) [48], and the pure SUSY interactions. A detailed treatment of this Lagrangian, and the derivation of the forthcoming results can be found in [49].

Soft SUSY Breaking

As discussed before, SUSY must be a broken symmetry. So far, there is no understanding of how this breaking takes place. The common approach is to assume that SUSY is broken in a *hidden sector*, that is essentially decoupled from the physics at the *low* energy scale ($\mathcal{O}$(1 TeV), where it can be reachable by the collider experiments. The effects of SUSY breaking is only transferred to our world by messenger

[5] A given set of chiral superfields and arbitrary functions of them.

interactions. As the physics of this messenger sector is still unknown, it is only possible to introduce *soft* terms containing all possible soft SUSY breaking operators in the Lagrangian by hand. *Soft* means here that these terms are consistent with all desired symmetries and do not lead to the reappearance of the quadratic divergences in the one-loop corrections to the Higgs mass. The soft terms in the Lagrangian can be written as:

$$\begin{aligned} L_{\text{Soft}} &= m_{H_d}^2 H_d^\dagger H_d + m_{H_u}^2 H_u^\dagger H_u + B\mu \left(H_u^T i\tau_2 H_d + h.c.\right) \\ &+ \sum_i \left(\tilde{m}_{Q_i}^2 \tilde{Q}_i^\dagger \tilde{Q}_i + \tilde{m}_{L_i}^2 \tilde{L}_i^\dagger \tilde{L}_i + \tilde{m}_{U_i}^2 \tilde{U}_i^\dagger \tilde{U}_i + \tilde{m}_{D_i}^2 \tilde{D}_i^\dagger \tilde{D}_i + \tilde{m}_{E_i}^2 \tilde{E}_i^\dagger \tilde{E}_i + h.c.\right) \\ &+ \sum_{i,j} \left(A_u^{ij} h_u^{ij} \tilde{\bar{U}}_i H_u \tilde{Q}_j + A_d^{ij} h_d^{ij} \tilde{\bar{D}}_i H_d \tilde{Q}_j + A_e^{ij} h_e^{ij} \tilde{\bar{L}}_i H_d \tilde{L}_j + h.c.\right) \\ &+ \frac{1}{2} \sum_{l=1}^{3} M_l \tilde{\lambda}_l \tilde{\lambda}_l + h.c. \quad . \end{aligned} \tag{2.31}$$

Here $H_{u,d}$ are the Higgsdoublets, μ is the mass parameter of the superpotential (Eq. 2.30), B is the bilinear coupling, $A_{u,d,e}^{ij}$ is the trilinear coupling of interactions of Higgs fields and sfermions, the indices $i, j = 1, 2, 3$ run over the three generations and $\tilde{\lambda}_l$ are the gauginos, the superpartners of the gauge bosons.

Including *all* these terms leads to a total of 178 free parameters [42]. It is very hard to make a phenomenological analysis with such a huge parameter space. There are some models that reduce the amount of those free parameters severely. The minimal supergravity (mSUGRA) is one of them and will be described in the next sections. In the presented analysis the mSUGRA model is used [50], here supersymmetry is broken by a coupling to a yet unknown supergravity theory, assuming the unification at M_{GUT} of:

$$M_1(M_{\text{GUT}}) = M_2(M_{\text{GUT}}) = M_3(M_{\text{GUT}}) \equiv m_{1/2} \tag{2.32}$$

$$\tilde{m}_{E,L,U^i,D^i,Q^i}(M_{\text{GUT}}) = m_{H_{u,d}}(M_{\text{GUT}}) \equiv m_0 \tag{2.33}$$

$$A_t(M_{\text{GUT}}) = A_b(M_{\text{GUT}}) = A_\tau(M_{\text{GUT}}) \equiv A_0 \quad . \tag{2.34}$$

At the GUT scale m_0 is the common mass for all sleptons and squarks. The mass of all gauginos at the GUT scale is $m_{1/2}$. The renormalization group equations (RGE)[6] are used to calculate the values at other energy scales. The remaining free parameters are

[6] Irrespective of the model the quantities depend on the scale at which the theory is renormalized (when calculating all higher order diagrams all integrals must remain finite). The main reason is that Green functions are truncated at a specific order and thus there is an explicit dependence on the scale of renormalization. For collider processes, for instance, it is necessary to compute all masses and couplings at the scale relevant for the collider.

- $m_{1/2}$, the unified gaugino mass $m_{1/2}$ at the GUT scale
- m_0, the unified scalar mass m_0 at the GUT scale
- B, the bilinear coupling or equivalent $tan\beta \equiv \frac{v_2}{v_1}$, the ratio of the vacuum expectation values of the two Higgs fields
- A_0, the unified trilinear coupling A_0 at the GUT scale
- μ, the Higgs field mixing parameter.

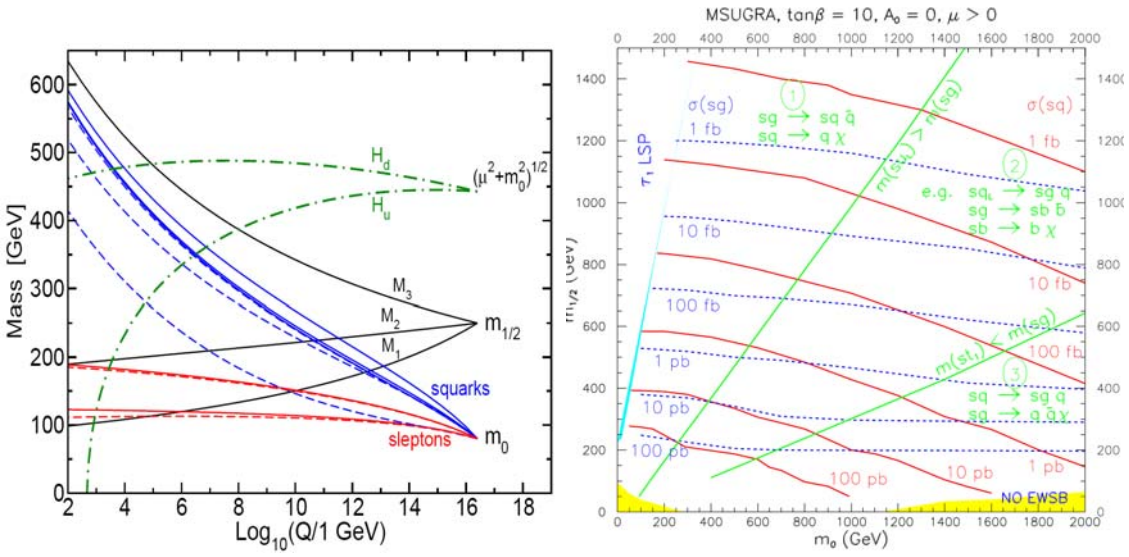


Figure 2.4: Left: RG evolution of scalar and gaugino mass parameters in the MSSM with typical minimal supergravity-inspired boundary conditions imposed at GUT scale. The mass parameter H_u runs negative, provoking electroweak symmetry breaking [12]. Right: Regions of the m_0 versus $m_{1/2}$ plane showing the production cross-sections and with the main squark and gluino decays. [51]

2.2.3 SUSY Mass Spectrum

In the MSSM, the masses of the SUSY particles can be calculated via the renormalization group equations (RGE), which are derived from the Lagrangian. With a given initial condition at the GUT scale, the solution of the RGE link the values at the GUT scale with the electroweak scale and thus determine the mass matrices of gauginos, squarks and leptons.

Neutralinos and Charginos

Neutralinos and Charginos are the mass eigenstates of the neutral and charged fields, respectively. Their mass eigenstates are mixed states of gauginos and higgsinos:

$$\chi = \begin{pmatrix} \tilde{B} \\ \tilde{W}^3 \\ \tilde{H}_d^0 \\ \tilde{H}_u^0 \end{pmatrix} \quad , \quad \Psi = \begin{pmatrix} \tilde{W}^+ \\ \tilde{H}^+ \end{pmatrix} \quad , \tag{2.35}$$

where χ and Ψ are the Majorana neutralino and Dirac chargino fields, respectively.

The corresponding neutralino mass matrix can be written as:

$$M^{(0)} = \begin{pmatrix} M_1 & 0 & -M_Z \cos\beta \sin\theta_W & M_Z \sin\beta \sin\theta_W \\ 0 & M_2 & M_Z \cos\beta \cos\theta_W & -M_Z \sin\beta \cos\theta_W \\ -M_Z \cos\beta \sin\theta_W & M_Z \cos\beta \cos\theta_W & 0 & -\mu \\ M_Z \sin\beta \sin\theta_W & -M_Z \sin\beta \cos\theta_W & -\mu & 0 \end{pmatrix} \tag{2.36}$$

with the gaugino masses M_1, M_2, the weak mixing angle θ_W and $\tan\beta$, the ratio of two Higgs vacuum expectation values. The physical masses of the neutralinos are given by the eigenvalues of this matrix. The neutralino mass eigenstates are denoted as $\chi_1^0, \chi_2^0, \chi_3^0, \chi_4^0$ with $m_{\chi_1^0} \leq m_{\chi_2^0} \leq m_{\chi_3^0} \leq m_{\chi_4^0}$.

The mass matrix for the charginos is given by:

$$M^{(c)} = \begin{pmatrix} M_2 & \sqrt{2} M_W \sin\beta \\ \sqrt{2} M_W \cos\beta & \mu \end{pmatrix}, \tag{2.37}$$

which leads to two chargino eigenstates $m_{\chi_{1,2}^\pm}$ with the mass eigenvalues

$$\begin{aligned} m^2_{\chi^\pm_{1,2}} &= \frac{1}{2} M_2^2 + \frac{1}{2}\mu^2 + M_W^2 \\ &\quad \mp \frac{1}{2}\sqrt{(M_2^2 - \mu^2)^2 + 4M_W^4 \cos^2 2\beta + 4M_W^2 (M_2^2 + \mu^2 + 2M_2 \mu \sin 2\beta)} \quad . \end{aligned} \tag{2.38}$$

The gluino is the only color octet fermion. Since SU(3) is unbroken, the gluino does not mix with other MSSM particles. The mass of the physical particle is defined by the gaugino mass parameter $m_{\tilde{g}} \equiv M_3$.
Approximately, the gaugino mass parameters at the electroweak scale are:

$$M_3 \simeq 2.7 m_{1/2} \tag{2.39}$$

$$M_2 \simeq 0.8 m_{1/2} \tag{2.40}$$

$$M_1 \simeq 0.4 m_{1/2} \quad . \tag{2.41}$$

The physical masses of the neutralinos are obtained by diagonalizing the mass matrix Eq. 2.36. In the mSUGRA model, the lightest neutralino is dominantly bino-like

and the next-to-lightest neutralino is mostly wino-like, with masses close to M_1 and M_2, respectively. The mass of the lightest chargino is approximately given by M_2. Hence the masses of the next-to-lightest neutralino and the lightest chargino are similar, and approximately two times the mass of the lightest neutralino.

Sleptons and Squarks

The masses of left-handed and right-handed fermions are equal. But their superpartners are bosons and the masses of left-handed and right-handed sfermions can be different:

$$\tilde{m}^2_{e_L} = \tilde{m}^2_{L_i} + m^2_{E_i} + M_Z^2 \cos(2\beta)\left(-\frac{1}{2} + \sin^2\theta_W\right) \tag{2.42}$$

$$\tilde{m}^2_{\nu_L} = \tilde{m}^2_{L_i} + M_Z^2 \cos(2\beta)\left(\frac{1}{2}\right) \tag{2.43}$$

$$\tilde{m}^2_{e_R} = \tilde{m}^2_{E_i} + m^2_{E_i} - M_Z^2 \cos(2\beta)\left(\sin^2\theta_W\right) \tag{2.44}$$

$$\tilde{m}^2_{u_L} = \tilde{m}^2_{Q_i} + m^2_{U_i} + M_Z^2 \cos(2\beta)\left(+\frac{1}{2} + \sin^2\theta_W\right) \tag{2.45}$$

$$\tilde{m}^2_{d_L} = \tilde{m}^2_{Q_i} + m^2_{D_i} + M_Z^2 \cos(2\beta)\left(-\frac{1}{2} + \frac{1}{3}\sin^2\theta_W\right) \tag{2.46}$$

$$\tilde{m}^2_{u_R} = \tilde{m}^2_{U_i} + m^2_{U_i} + M_Z^2 \cos(2\beta)\left(\frac{2}{3}\sin^2\theta_W\right) \tag{2.47}$$

$$\tilde{m}^2_{d_R} = \tilde{m}^2_{D_i} + m^2_{D_i} - M_Z^2 \cos(2\beta)\left(\frac{1}{3}\sin^2\theta_W\right) \quad . \tag{2.48}$$

On the right side of the equations, the terms denoted as $\tilde{m}$ are calculated with the RGE, the mass terms m are the fermion masses. The index i denotes the three generations.

Furthermore non-negligible Yukawa couplings lead to a mixing between the electroweak eigenstates and the mass eigenstates of the third generation sleptons and squarks. Due to small Yukawa couplings the mixing is negligible for the first and second generation. Therefore the mass eigenstates corresponds to the interaction eigenstates, which have been introduced above. The mass matrices for the third generation can be written as:

$$\mathcal{M}^{\tilde{t}} = \begin{pmatrix} \tilde{m}^2_{t_L} & m_t(A_t - \mu\cot\beta) \\ m_t(A_t - \mu\cot\beta) & \tilde{m}^2_{t_R} \end{pmatrix} \tag{2.49}$$

$$\mathcal{M}^{\tilde{b}} = \begin{pmatrix} \tilde{m}^2_{b_L} & m_b(A_b - \mu\tan\beta) \\ m_b(A_b - \mu\tan\beta) & \tilde{m}^2_{b_R} \end{pmatrix} \tag{2.50}$$

$$\mathcal{M}^{\tilde{\tau}} = \begin{pmatrix} \tilde{m}^2_{\tau_L} & m_\tau(A_\tau - \mu\tan\beta) \\ m_\tau(A_\tau - \mu\tan\beta) & \tilde{m}^2_{\tau_R} \end{pmatrix} \quad . \tag{2.51}$$

The mass eigenstates of the third generation are:

$$m^2_{\tilde{t}_{1,2}} = \frac{1}{2}\left(\tilde{m}^2_{t_L} + \tilde{m}^2_{t_R}\right) \pm \sqrt{\frac{1}{4}\left(\tilde{m}^2_{t_L} - \tilde{m}^2_{t_R}\right)^2 + m_t^2\left(A_t - \mu\cot\beta\right)^2} \quad (2.52)$$

$$m^2_{\tilde{b}_{1,2}} = \frac{1}{2}\left(\tilde{m}^2_{b_L} + \tilde{m}^2_{b_R}\right) \pm \sqrt{\frac{1}{4}\left(\tilde{m}^2_{b_L} - \tilde{m}^2_{b_R}\right)^2 + m_b^2\left(A_b - \mu\tan\beta\right)^2} \quad (2.53)$$

$$m^2_{\tilde{\tau}_{1,2}} = \frac{1}{2}\left(\tilde{m}^2_{\tau_L} + \tilde{m}^2_{\tau_R}\right) \pm \sqrt{\frac{1}{4}\left(\tilde{m}^2_{\tau_L} - \tilde{m}^2_{\tau_R}\right)^2 + m_\tau^2\left(A_\tau - \mu\tan\beta\right)^2} \quad . (2.54)$$

Electroweak Symmetry Breaking

The Higgs sector of the MSSM is that of a 2HDM, with some SUSY restrictions. After expanding (Eq. 2.28) the Higgs potential reads[7]

$$\begin{aligned} V &= m_1^2\,|H_d|^2 + m_2^2\,|H_u|^2 - m_3^2\,\left(\epsilon_{ij}\,H_d^i\,H_u^j + \text{h.c.}\right) \\ &+ \frac{1}{8}(g^2+g'^2)\,\left(|H_d|^2 - |H_u|^2\right)^2 + \frac{1}{2}\,g^2\,|H_d^\dagger\,H_u|^2\,. \end{aligned} \quad (2.55)$$

In the SM the non-trivial minumum of the Higgs potential is introduced the coefficient μ^2 in the super-potential 2.30. The scalar potential should developed a well defined local minimum in which electroweak symmetry is spontaneously broken if the following conditions are met [52]: $(B\mu)^2(m^2_{H_u}+\mu^2)(m^2_{H_d}+\mu^2)$ and $m^2_{H_u} + m^2_{H_d} + 2\mu^2 > 2|B\mu|$. The neutral Higgs bosons fields acquire a vacuum expectation value (VEV), with the notations :$\langle H_d\rangle \equiv \upsilon_d.\cos\beta$,$\langle H_u\rangle \equiv \upsilon_u.\sin\beta$ and $\upsilon^2 \equiv \upsilon_d^2 + \upsilon_u^2$, $\tan\beta \equiv \upsilon_u/\upsilon_d$

The VEVs must satisfy:

$$M_W^2 = \frac{1}{2}g^2(v_u^2+v_d^2) \equiv g^2\frac{v^2}{2} \quad (2.56)$$

$$M_Z^2 = \frac{1}{2}(g^2+g'^2)v^2 \equiv M_W^2\cos^2\theta_W \quad (2.57)$$

$$\tan\beta = \frac{v_u}{v_d}\ , \quad 0<\beta<\frac{\pi}{2} \quad (2.58)$$

$$\tan\theta_W = \frac{g'}{g} \quad (2.59)$$

where θ_W and gauge boson masses have already been measured. Here, the additional parameter $\tan\beta$ is an unknown of the model, and it signals the presence of more than one single Higgs doublet.

These VEV's make the Higgs fields to mix. There are five physical Higgs fields: a couple of charged Higgs bosons ($H^\pm$); a *pseudoscalar* Higgs ($CP=-1$) A^0; and two scalar Higgs bosons ($CP=1$) H^0 (the heaviest) and h^0 (the lightest) [48].

[7] g is the coupling constant of the corresponding gauge group as follows : $U(1)_y \to g = g'$, $SU(2)_L \to g = g$, $SU(3)_C \to g = g_s$.

All the masses of the Higgs sector of the MSSM can be obtained with only two usual parameters, the first one is $tan\beta$, and the second one is a mass; usually this second parameter is taken to be either the charged Higgs mass $m_{H^\pm}$ or the pseudoscalar Higgs mass m_{A^0}. We will take the last option. From (2.55) one can obtain the tree-level mass relations between the different Higgs particles,

$$\begin{aligned} m^2_{H^\pm} &= m^2_{A^0} + M^2_W \ , \\ m^2_{H^0,h^0} &= \frac{1}{2}\left(m^2_{A^0} + M^2_Z \pm \sqrt{\left(m^2_{A^0} + M^2_Z\right)^2 - 4\, m^2_{A^0}\, M^2_Z \cos^2 2\beta} \right) \end{aligned} \tag{2.60}$$

The immediate consequence of such a constrained Higgs sector, is the existence of absolute bounds (at tree level) for the Higgs masses:

$$0 < m_{h^0} < m_Z < m_{H^0}, \quad m_W < m_{H^\pm} \tag{2.61}$$

where experiments have already bounded m_h, the lightest Higgs mass, to be larger than 114 GeV. Therefore, these tree-level relations are far from representing the reality; one needs radiative effects to be incorporated into the Higgs potential (radiative electroweak symmetry breaking).

2.2.4 Constraints on the Experimental Parameter Space and the LHC

As discussed in section 2.1.6, there are many reasons to expect some new physics at the TeV scale, which should be accessible to the LHC. These include the Higgs boson and further SUSY particles, which stabilize the Higgs mass and hence the electroweak scale. There are also general arguments that the lightest stable weakly-interacting particle, the candidate for dark matter, should weigh a TeV or less. There is also an abundance of lower energy experiments that may be sensitive indirectly to TeV-scale physics. On the other hand, there are several examples of new physics that may well lie beyond the LHCs reach. Physics at these scales could only be explored indirectly by the LHC and other low-energy experiments. The main constraints in the SUSY searches are:

- Gauge coupling unification, which fixes the scale of SUSY soft breaking of be of the order of 1 TeV.
- M_Z value: to get the right value of M_Z requires proper adjustment of parameters that determine the value of the μ parameter for a given value of m_0 and $m_{1/2}$.
- Yukawa coupling constant unification: the masses of the top, bottom quark and τ can be obtained from the low energy values of the running couplings

$$m_t = y_t \upsilon \sin\beta, \quad m_b = y_b \upsilon \cos\beta, \quad m_\tau = y_\tau \upsilon \cos\beta \tag{2.62}$$

These are written for the running masses. They can be rewritten as a pole mass without any radiative corrections.

- Branching ratio of the b $\rightarrow s\gamma$: The agreement between the observed rate for the decay and SM expectation yields significant constraints on off-diagonal squark mass squared matrix elements. The SM contribution to this process comes from the W-t loop and provides a prediction which is close to the experimental value, thus leaving little room for SUSY contributions.

- Dark matter constraint: The lightest supersymmetric particle (LSP) must be neutral and stable, candidate for the dark matter.

- The anomalous magnetic moment measurement of the muon, if it is assumed to be the denoted by a_μ, in many extension of the SM the new physics contributions to the lepton magnetic moment are proportional to m_l^2, hence a_μ is to more new physics sensitive than the better measured magnetic muon of the electron. The $E821$ experiment at Brookhaven National Laboratory has measured $a_\mu = (g-2)_\mu/2$ [53]. The SM and SUSY contribution to this term shows differences on the theoretical value which is mainly taken into account for the SUSY searches.

- Proton life time: There are two sources of proton decay in SUSY grand unification models. The first one is related to the s-channel exchange of heavy gauge boson which is the same as in non-SUSY models. To agree with experiment, the unification scale has to be above 10^{15} GeV which is satisfied in SUSY models. The second one depends on the specific SUSY model that the proton decay takes place due to the loop diagrams with the exchange of heavy higgsino triplets [43].

- Cosmological constraints: Since R-parity is assumed to be conserved in the MSSM, the lightest SUSY particle is absolutely stable. This has profound implications for cosmology and, in particular, may imply that the relic LSPs left over from the Big-bang could account for the bulk matter in the universe. Further, the requirement that relic density of LSPs should be in accord with astrophysical measurements of the dark matter density of the Universe leads to important constraints on SUSY model parameters.

- **Collider experiments:** Experimental lower limits for SUSY masses : the lower limits of SUSY masses have been determined in LEP and Tevatron experiments with the corresponding center of mass energies. Furthermore the Higgs mass limit at LEP of 114 GeV requires rather heavy stop masses.

In collider experiments, our main interest, $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$ and $\tilde{q}\tilde{q}$ processes expected to be the dominant processes for sparticle production in the mSUGRA model (see details in chapter 5). The cascade decay signatures will generally be complex and give rise to events with jets and isolated leptons together with missing transverse energy (E_T^{Miss}) at the LHC experiment. Jets from primary decays of squarks or gluinos can

be energetic and have large transverse momenta, as expected for massive particles. Leptons that originate further down the cascade chain are typically softer than the primary jets in these events. The event topologies can be classified as before by the number of identified isolated leptons in these events:

1. E_T^{Miss} channel : an inclusive channel requiring Large E_T^{Miss} plus ≥ 2 jets plus any number of identified leptons,
2. Zero lepton channel: a subset of the E_T^{Miss} channel which in addition *vetoes* any isolated leptons
3. Single lepton channel : a subset of E_T^{Miss} channel containing a single isolated lepton,
4. OS channel: a subset of E_T^{Miss} containing two-opposite sign isolated leptons
5. SS channel: a subset of E_T^{Miss} containing same-sign isolated leptons
6. Tri-Lepton channel: a subset of E_T^{Miss} containing three isolated leptons

The experiments at LEP2 have already placed stringent bounds on MSSM searches, and Tevatron experiments may well find evidence for the Higgs masses during the LHC run. Nevertheless, it will be an important for the CMS (and ATLAS) experiment to establish a proper search for SUSY and Higgs bosons content of the MSSM, to determine as much as possible about their properties (see Chapter 6).

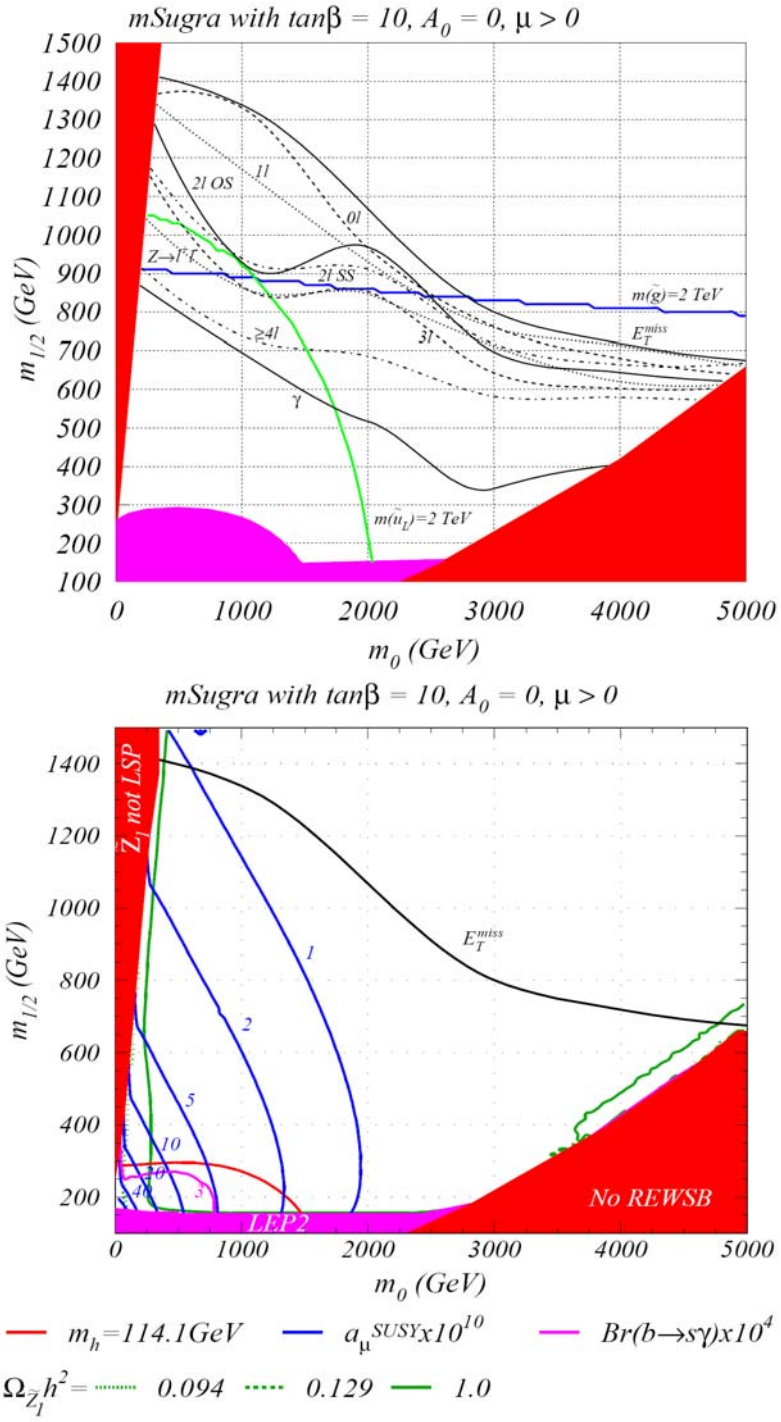


Figure 2.5: The reach of LHC in the m_0 and $m_{1/2}$ parameter plane of the mSUGRA model, with $\tan\beta$=10, A_0=0 and $\mu > 0$, assuming $100fb^{-1}$ integrated luminosity. Upper: The red (magenta) regions are excluded by the constraints discussed in the text. The various SUSY searches has been presented with E_T^{Miss}. Lower: Contours of several low energy observables in the m_0 and $m_{1/2}$ plane of the mSUGRA model. The relic density contours (green) has been presented for Cold Dark Matter (CDM), together with a contour of m_h=114.1 GeV (red), contours of muon anomalous magnetic moment a_μ (x10^{10})(blue) and contours of $b \to \gamma s$ branching fraction(x10^4) (magenta) [52].

Chapter 3

The CMS Experiment at the Large Hadron Collider

The physics program at present and future colliders is aimed at a truly comprehensive exploration of the TeV scale. In addition to the precision physics of the Standard Model, recent years have seen the emergence of an impressive variety of proposals for what physics may be uncovered by these machines in just a few years. The ideas range from hypotheses of new fundamental matter or forces(additional family or Z'), to new space-time symmetries (Supersymmetry), or even new spatial dimensions with the possible production of microscopic black holes.

The Large Hadron Collider (LHC) has started operation this March at CERN near Geneva and is expected to answer several of the questions raised above. In this chapter, the basic properties of the colliders are discussed and the associated experiments are introduced focusing on CMS.

3.1 Large Hadron Collider

The LHC at CERN is the world's largest particle accelerator. It is twenty-seven kilometers in circumference in a tunnel from the former Large Electron Positron Collider (LEP) spanning the French-Swiss border outside of Geneva. Between 1989 and 2000 LEP was used for electron-positron collisions with a center-of-mass energy of up to 209 GeV. The experiments precisely measured the properties of W and Z boson and confirmed the predictions of the Standard Model. However, the LEP energy was restricted by the energy loss per turn due to synchrotron radiation:

$$-\Delta E = \frac{4\pi\alpha}{3R}\beta^3\gamma^4 \quad with \quad \beta = \frac{\nu}{c} \approx 1 \quad and \quad \gamma = \frac{E}{mc^2} \tag{3.1}$$

According to this equation the electron radiates 2.3% of its energy per turn at LEP experiment, which has to be compensated by the accelerator facility. There are two

general approaches for controlling the loss of energy and to reach higher center-of-mass energies. Either the radius of the facility can be enlarged or eliminated (Linear collider project) or heavier particles with a higher mass than the electron are accelerated. The acceleration of protons is the concrete approach at the moment, since these particles are roughly 2000 times heavier than electrons and the mass enters the energy loss with the fourth power. Hence synchrotron radiation becomes negligible but it has other disadvantages for physics analyses. The proton is a composite object of quarks and gluons, each carrying only a fraction x of the proton's momentum and the parton density decreases rapidly with increasing x. Therefore, the beam energy at hadron colliders has to be foreseen to be well above the energy scale of the desired interactions. The Tevatron is a collider that accelerates protons and anti-protons in a 6.28 km ring to energies of up to 1 TeV per beam.

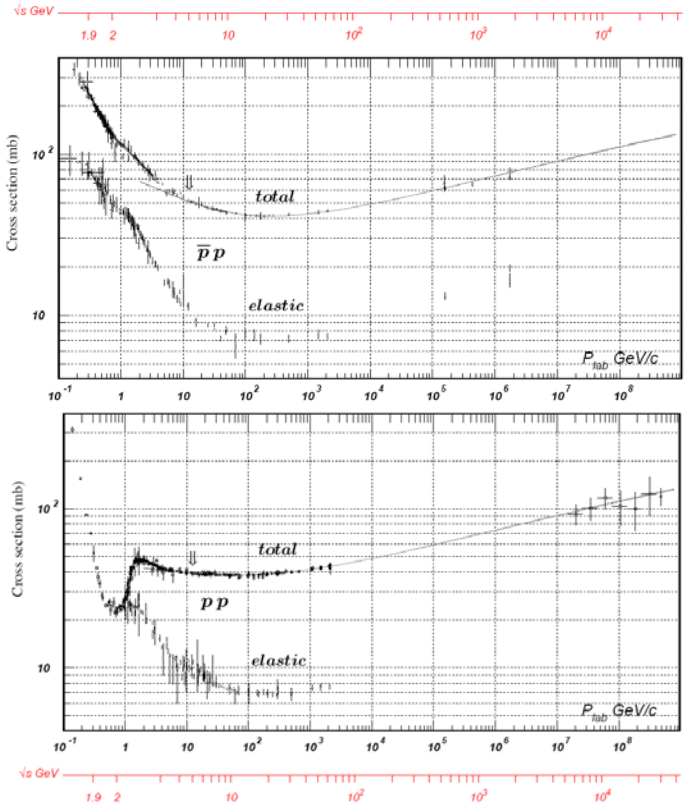


Figure 3.1: Total and elastic cross sections for pp and $p\bar{p}$ collisions as a function of laboratory beam momentum and total center-of-mass energy(red scale). Data points on the scale are a combined effort of both accelerator and cosmic shower physics experiments [54].

The LHC at CERN is designed to collide proton beams of center-of-mass energies of 14 TeV and a luminosity of $10^{34} cm^{-2} s^{-1}$. It is a two ring superconducting hadron accelerator and collider in a 27 km long tunnel of eight straight sections and eight arcs between 45m and 170 m below the surface. As the LHC is a proton-proton collider, both beams cannot share the same phase space in a single ring, but two rings with counter-rotating beams are necessary. The LHC can also collide heavy ions with an energy of 2.8 TeV and a peak luminosity of $10^{27} cm^{-2} s^{-1}$.

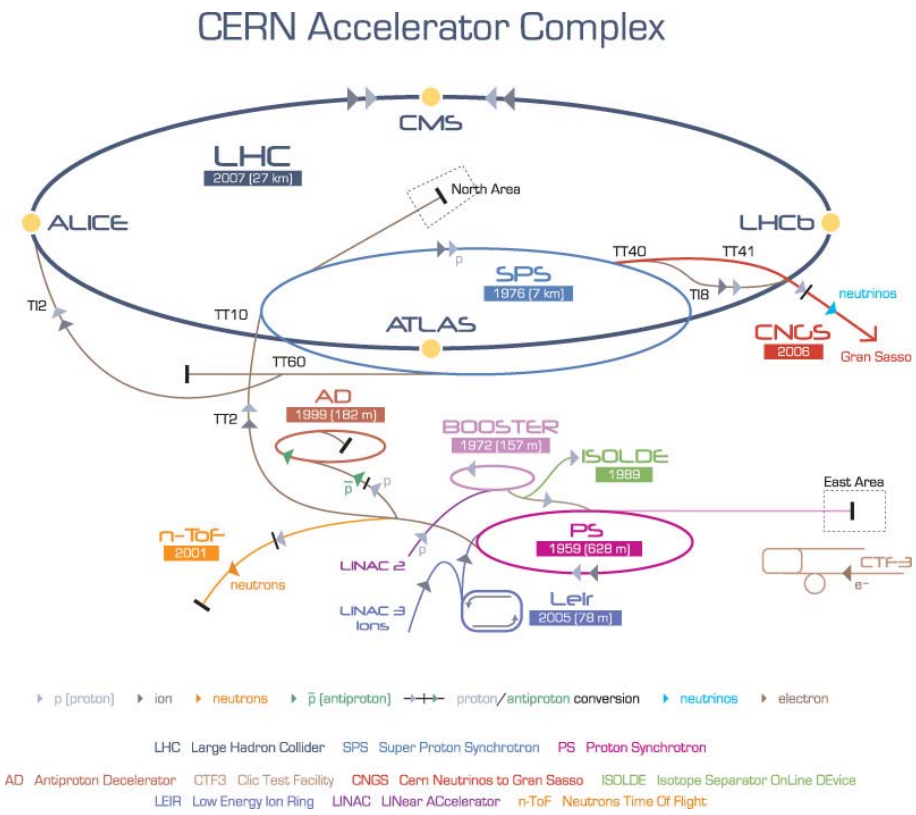


Figure 3.2: Schematic layout of the LHC accelerator complex. The four main experiments are LHCb, which explores b-physics in proton-proton collisions, ALICE, which investigates heavy-ion physics in lead-lead collisions, and the two general purpose experiments ATLAS and CMS [55].

In December 2009, the LHC started operation successfully and provided two proton beams, circulating in opposite directions with energies of 900 GeV and 2.36 TeV for a short time. In March 2010, the LHC has started higher luminosity operation again and provided two proton beams at 3.5 TeV each. This time, the beam energy is limited by the maximal strength of the magnetic dipole field compensating the centrifugal force of the charged particles due to the curvature of the ring. The LHC magnets provide a maximum magnetic field of 8.33 T, which corresponds to 7 TeV beam. To reach the high magnetic field required high currents are needed. To avoid excessive resistive losses, the magnets are superconducting. The magnet coils are made of niobium-titanium (NbTi) cables which become superconducting below a temperature of 10 K. They are cooled down with super-fluid liquid helium at a temperature of about 1.9 K for operations. With this setup a center-of-mass energy of the proton-proton collisions of 14 TeV can be reached at the LHC.

The number of events $\mathcal{N}$ (event) per second generated in collisions with a cross section σ and machine luminosity $\mathcal{L}$ is

$$\mathcal{N}(events)/s = \mathcal{L} \cdot \sigma(events) \tag{3.2}$$

The total proton-proton cross section expected at the LHC is 110 mb with 60 mb from inelastic collisions, 40 mb from elastic collisions and 10 mb from diffractive events. The machine luminosity is a function of the beam parameters. The luminosity for head-on collisions of two such bunches consisting out of n particles with a collision frequency f is given by

$$\mathcal{L} = f_{rev} \frac{n^2}{4\pi\sigma_x\sigma_y}, \tag{3.3}$$

where σ_x, σ_y denote the gaussian transverse profiles of the beam. For Gaussian-distributed beam with N_b particles per bunch, n_b bunches per beam, revolution frequency f_{rev}, relativistic gamma factor γ_r, normalized transverse beam emittance ϵ_n and the amplitude function(beta) at the collision point β^*:

$$\mathcal{L} = F \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n\beta^*}, \tag{3.4}$$

A non-zero crossing angle at the interaction point reduces the luminosity, reflected by the geometric luminosity reduction factor F.

In order to reach the designed luminosity $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1} of the LHC, the bunch crossing frequency, the number of proton bunches as well as the number of protons per bunch have to be chosen. At the nominal intensity each beam will consist of around 2800 bunches with up to 1.15×10^{11} protons per bunch and the bunch spacing is 25 ns corresponding to bunch crossing frequency of 40 MHz. With the high number of protons per bunch, the average number of inelastic collisions between protons per bunch-crossing is large ($\sim 10 - 20$), which are called pile-up events. This complicates the search for rare interesting physic processes, since they will be contaminated by pile-up events. The designed luminosity leads to around 1 billion proton-proton interactions per second.

The LHC initially runs at an energy of 3.5 TeV per beam after start up in March 2010 until a significant data sample has been collected. Thereafter the energy will be taken towards 7 TeV per beam. At the end of 2010, the LHC runs with lead ions for the first time. Afterwards the LHC will shut down and work will begin on moving the machine towards 7 TeV per beam [56].

A layout of the LHC accelerator complex is schematically presented in Figure 3.2, it consists of several facilities for the subsequent acceleration of the protons. The LINAC2 generates protons with 50 MeV. Then the Proton Synchrotron Booster

(PSB) and in turn the Proton Synchrotron (PS) accelerates the protons up to a total energy of 26 GeV. Then the protons are accelerated in 24 cycles to an energy of 450 GeV in the SPS before inserting them into the main LHC ring. The filling of each of the two LHC rings takes about four and a half minutes. After the filling procedure the protons are accelerated for 20 minutes to their nominal energy of 7 TeV. Beside the protons runs, the LHC is designed to collide heavy ions like lead(Pb) with a center-of-mass energy of 1148 TeV.

Four particle detectors are designed to observe corresponding collisions. Two of them are installed as special-purpose detector: the ALICE [57, 58] experiment for heavy-ion physics and LHC*b* experiment for b-physics. The ATLAS [59] experiment and the Compact Muon Solenoid (CMS) are multi-purpose detectors, which are designed to cover the various possible range of physics processes, among them to discover the Higgs particle and to search for supersymmetry or other physics beyond the Standard Model. In the following the CMS experiment, in which IEKP-Institute is involved with a large group of more than 40 people, will be described in more detail.

3.2 CMS Experiment

The CMS detector, which has a length of 24 m, a diameter of 14.6 m and a weight of about 14500 tons, is one of the two general purpose experiments that takes data at the LHC. The CMS structure is a typical one for the experiments at colliders: a cylindrical central section (*the barrel*) closed at its end by two caps (*the endcap*), as sketched in Figure 3.3. The main distinguishing features of CMS are a high-field solenoid, a full silicon-based inner tracking system and a fully active scintillating crystals-based electromagnetic calorimeter, based on scintillating crystals.

The broad LHC physics program requires four main design choices [61]:

- Identify *muons* with high efficiency and good momentum resolution over a wide momentum range and with a large angular acceptance. Resolve dimuon masses at a level of 1% at $M_{\mu\mu}$= 100 GeV/c^2 and determine muon charges up to momenta of 1 TeV/c.
- Reconstruct *charged particles* in the inner tracker with high efficiency and good momentum resolution. Trigger efficiently on τ *leptons and b jets.*
- Resolve *electromagnetic energy, diphoton and electron masses* to excellent levels (1% at $M_{\mu\mu}$= 100 GeV/c^2). Cover a large pseudo-rapidity range, efficiently reject π^0 mesons and isolate photons and leptons at high luminosities.
- Resolve *missing transverse energy and di-jet mass to good levels.*

As shown pictorially in Fig. 3.4 particles emerging from the interaction point first traverse the tracking system, made entirely of high resolution silicon detectors which

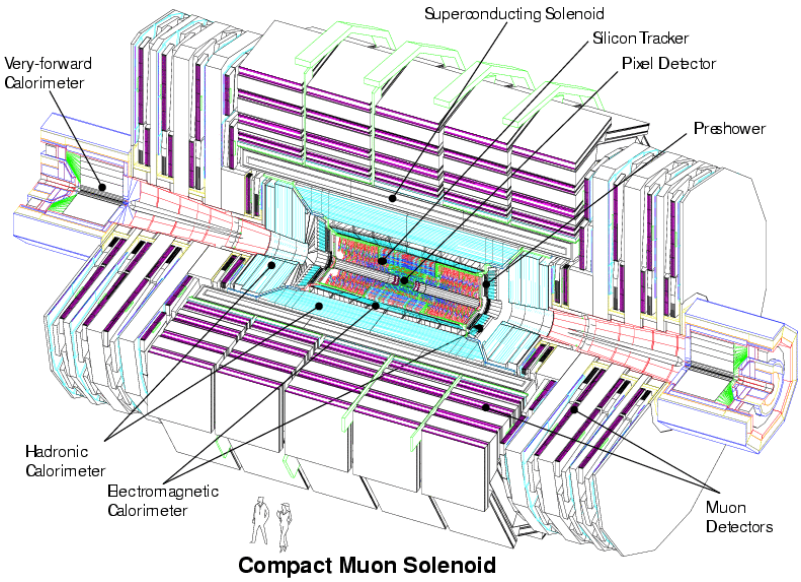


Figure 3.3: Compact Muon Solenoid experiment [60].

allow to measure their momentum. Outside the tracking system are calorimeters that measure the energy of particles. The tracking system should interfere with the particles as little as possible, whereas the calorimeters are designed to stop the particles. The electromagnetic calorimeter (ECAL) measures the energy of photons and electrons whereas the hadron calorimeter (HCAL) is designed principally to detect any particle made up of quarks. The size of the magnet allows the tracker and calorimeters to be placed inside its coil, resulting in an overall compact detector. In the outer part of the detector, the iron magnet return yoke confines and guides the magnetic field. All particles, expect muons and weakly interacting particles, such as neutrinos, should be stopped within the calorimeters and the iron return yoke. Hence only muons will be measured in the muon system, so muons provide a clean signature, whereas electrons and photons measured in the ECAL require sophisticated techniques for discrimination and identification.

Since muon final states form a very clean signature for many of the interesting processes that will be searched for at the LHC, the detection of muons is among the highest priorities for CMS. Concerning the measurement of the muon momentum the choice of the magnetic field configuration is an important issue. The CMS collaboration decided to use a solenoid magnet which produces a strong magnetic field along the beam axis bending the charged particles in the transverse plane. The muon momentum reconstruction depends on the strength of the magnetic field B and on its length L as:

$$\frac{\sigma_{P_T}}{P_T} = \frac{\sigma_s}{s} = \sigma_s \frac{8P_T}{0.3BL^2}, \tag{3.5}$$

with the transverse momentum of the particle P_T and the sagitta s, determining the amount of bending, of the particle track inside the magnetic field. Hence the

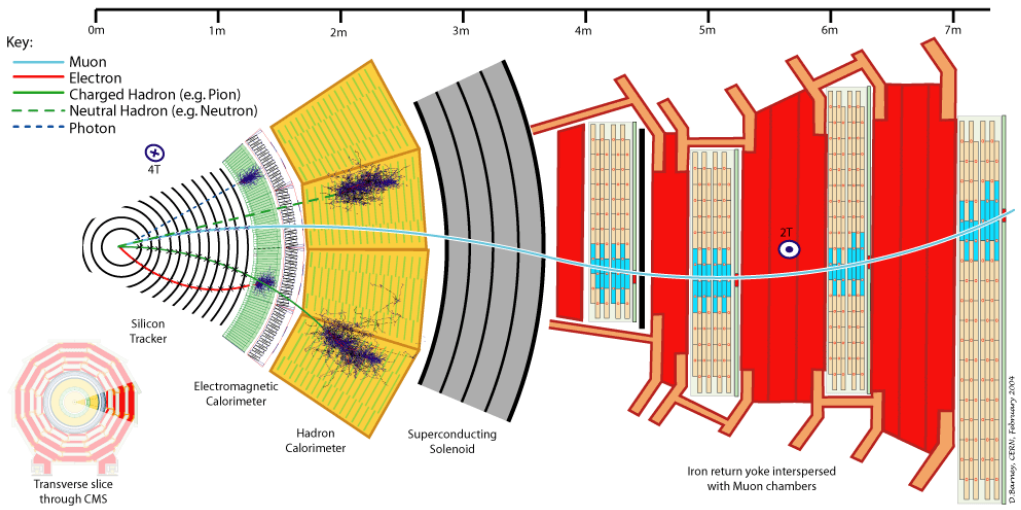


Figure 3.4: Slice of the CMS detector and pictorial presentation of the measurement of electrons, photons, hadrons (e.g. pions) and muons in the different subsystems. Electrons and photons deposit their whole energy in the electromagnetic calorimeter, whereas hadrons are measured in the hadron calorimeter. All particles, except for muons and only weakly interacting particles will be stopped in the inner region of the detector. Muons pass the detector material including the superconducting solenoid and are detected in the muon system. Therefore muons provide a very clean signature. On the contrary electrons and photons have to be distinguished with sophisticated techniques [62].

precision of the momentum measurement is proportional to $1/(BL^2)$. Whereas the ATLAS experiment uses a rather moderate magnetic field, the CMS experiment uses a high field strength of 3.8 Tesla and a rather compact detector.

3.2.1 Coordinate System

CMS has adopted a coordinate system with the origin at the nominal interaction point, the *y-axis* pointing vertically upward and the *x-axis* pointing toward the LHC center. Azimuthal angles are measured with respect to the *x-axis* in the *xy-plane*, with radial coordinate r. Polar angles are measured with respect to the *z-axis* and often represented by the pseudo-rapidity η= -ln tan $\theta/2$. The transverse momentum P_T and transverse energy E_T are computed as:

$$P_T = \sqrt{p_x^2 + p_y^2} \quad \text{and} \quad E_T = \sqrt{E_x^2 + E_y^2}. \tag{3.6}$$

In a reconstructed event the imbalance of energy measured in the transverse plane is called missing transverse energy.

3.2.2 Inner Tracking System

Starting from the beam pipe, the first detector component is the tracking system, which provides a precise and efficient measurement of the trajectories of charged particles with a high momentum resolution. The tracking system offers high efficiencies also for low momentum particles ($p = 1 - 5$ GeV) which is essential e.g. for studies with muon final states, since muons from SM bosons and their SUSY partners can be identified by the low particle activity around the muon track. Moreover it provides a precise reconstruction of their vertex position used to identify decays of heavy hadrons.

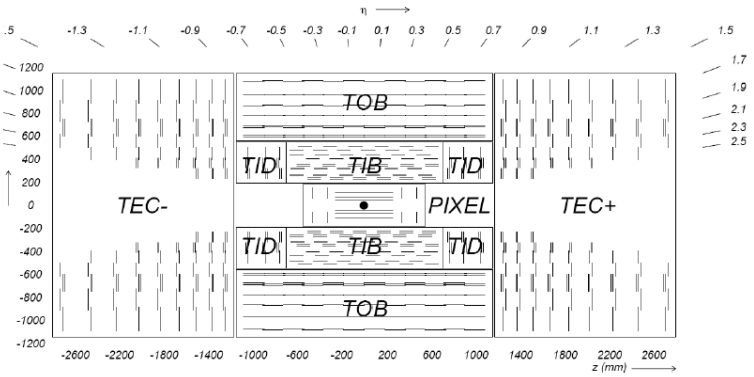


Figure 3.5: Layout of the CMS tracking system [63]. Each line represents a detector module, double lines indicate stereo modules.

A layout of the CMS tracker is shown in Fig. 3.5. The outer radius of the CMS tracker extends up to nearly 110 cm, while its total length is approximately 540 cm. The pseudo-rapidity coverage corresponds to $|\eta| < 2.5$. The design of the tracking system is strongly related to the challenging experimental conditions at the LHC, where the large number of pile up events at the designed luminosity leads to a huge amount ($\sim$ 1000) of charged particles produced per bunch crossing. In order to identify the particle tracks and assign them to the corresponding bunch crossing, a high granularity and fast response is mandatory. Additionally the tracking system has to stay unharmed by the high radiation environment and the material budget has to be minimized in order to limit secondary phenomenas like multiple scattering, bremsstrahlung, photon conversions and nuclear interactions.

To meet the mentioned requirements, the tracking system is composed of silicon pixel detectors placed in the inner regions up to $r < 10$ cm, where the occupancy is highest, followed by silicon microstrip detectors.

Pixel Detector

The silicon pixel detector is composed of a total of 66 million pixels with a size of 100 μm$\times$ 150 μm, which are arranged in the barrel in three concentric cylinders at

mean radii of 4.4 cm, 7.3 cm and 10.2 cm around the beam axis and in the end-caps in two layers perpendicular to the beam axis in the z direction at ±34.5 cm and ±46.5 cm. Charged particles passing through the silicon pixel create electron-hole pairs affected both by the reverse bias which is applied to the pixels and the magnetic field of the CMS solenoid. Consequently the drifting electron-hole pairs experience a Lorentz force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{3.7}$$

with the charge of the electron q, the magnitude of the electric field E, the electron drift velocity v and the magnitude of the magnetic field B. Since in the barrel the magnetic field is perpendicular to the electric field, the charge carriers are deflected at an angle to the electric field lines known as the Lorentz angle, resulting in a better resolution than the width of each individual pixel because the charge is shared between several pixels. However, in order to benefit from the same effect in the end-caps, the pixels in the end-caps are not mounted perpendicular to the beam pipe in the disks, but rotated by 20° about their radial symmetry axis, known as the turbine-blade geometry. The position resolution is ~ 10 μm in the r-ϕ and ~ 20 μm in the z direction [64].

Strip Detector

The silicon strip detector has sensor area of over 200 m^2 contains a total of 9.6 million silicon strips. As shown in Figure 3.5, the silicon strip detector is subdivided in the inner barrel (TIB), the tracker outer barrel (TOB), the tracker inner disks (TID) and the tracker end-caps (TEC). The TIB consists of 4 layers of silicon sensors with a thickness of 320 μm and a strip pitch varying from 80 to 120 μm. Since the strips are parallel to the beam axis, the z coordinate cannot be determined by a single strip detector. Therefore the first 2 layers are made with stereo modules i.e. two detectors are placed back to back with the strips at an angle of 0.1 rad to each other. In the TIB the resolution is between 23 and 34 μm in the r-ϕ direction and 230 μm in z. Due to smaller radiation levels in the outer region, the thickness of the silicon sensors is chosen as 500 μm together with a wider strip pitch of 120 to 180 μm. Also in the TOB the first two layers are made with stereo modules. The resolution is between 35-52 μm in the r-ϕ direction and 530 μm in z. Stereo modules are attached in the first two layers of the TID and in the first two as well as the fifth layer of the TOB. The strips in the TID and TEC are perpendicular to the beam axis with a thickness of 320 μm in the TID and the first three layers of the TOB and 530 μm for the rest of the TOB [65].

3.2.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is designed to measure electrons and photons with high accuracy and contributes to the jet energy measurement by determination of the electromagnetic component of jets. The ECAL consists of 75848 lead tungstate (PbWO4) crystals, a material with a high density ~ 8.3 g/cm^3, short

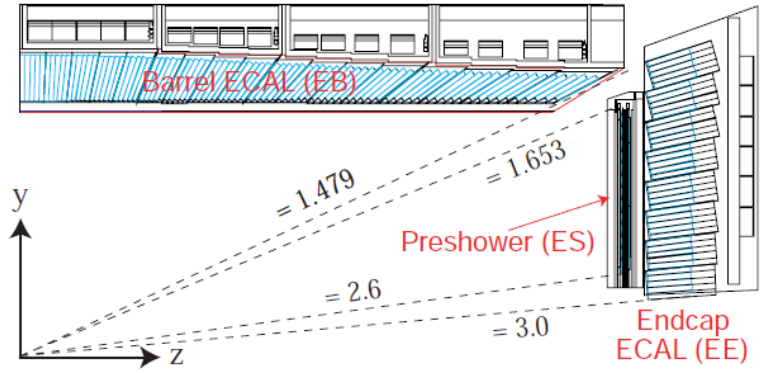


Figure 3.6: Layout of one quarter of the electromagnetic calorimeter (ECAL) [60].

radiation length $X_0 = 0.89$ cm and small Moliere radius 2.2 cm, giving the scale of the transverse dimension of the contained electromagnetic showers. Since a smaller Moliere radius means better shower position resolution and better shower separation due to a smaller degree of shower overlaps, the choice of lead tungstate enabled a compact design for the ECAL without suffering from a worse shower resolution. Furthermore, the material has a short scintillation light decay time, since in 25 ns $\sim 80\%$ of the light is collected.

Fig. 3.6 shows the electromagnetic calorimeter subdivided in the barrel part (EB) made of 61200 lead tungstate crystals and two end-caps (EE) each made of 7324 crystals. The EB with an inner radius 129 cm covers the region $|\eta| < 1.479$. The crystals in the EB have a front face cross-section of 22×22 mm^2, each crystal covers a range of $\Delta\eta \times \Delta\phi = 0.00174 \times 0.00174$. The length of 230 mm corresponds to a radiation length of $\sim 26\ X_0$.

The electromagnetic calorimeter in the end-caps covers the pseudorapitidy range of $1.479 < |\eta| < 3.0$ with crystals each having a front face cross-section of 28.6×28.6 mm^2 and a length of 220 mm which corresponds to a radiation length of $\sim 24.7\ X_0$. The preshower detector (ES) covers the range $1.653 < |\eta| < 2.6$ and is placed in front of the crystal calorimeter in order to identify neutral pions and to increase the position measurement of electrons and photons. The ES is a two layer sampling calorimeter, the first layer is composed of lead radiators followed by a second layer of silicon strip sensors.

The energy resolution of the ECAL can be parameterized as a function of energy:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{3.8}$$

where S is the stochastic term that includes the effects of the fluctuations in the photon statistics and the shower containment, N is the noise term that comes from electronics and pile-up, and C is the constant term that arises due to calibration

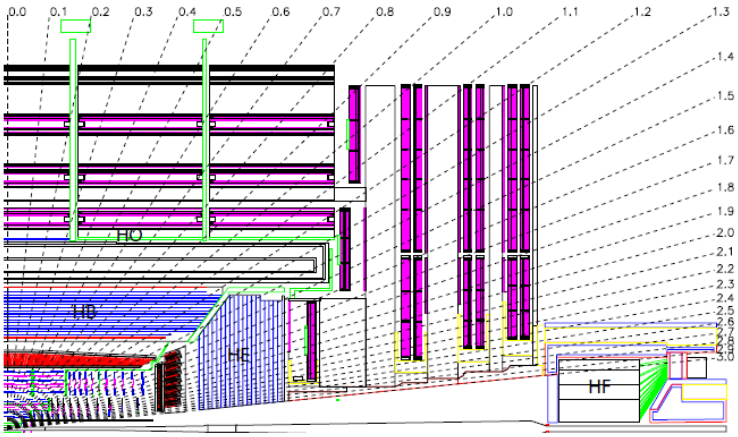


Figure 3.7: Layout of the hadron calorimeter (HCAL) [60].

errors and various systematic errors. The constant term dominates the resolution at high energies, its target value is 0.5% for both the barrel and the end-caps. The energy resolution of the ECAL barrel super-modules was studied with test beam electrons with energies between 20 and 250 GeV. For electrons entering at the center of the studied 3×3 arrays of crystals the energy resolution parameters have been determined as $S = 2.8\%, N = 41.5$ MeV and $C = 0.3\%$ [63, 66, 67].

3.2.4 Hadron Calorimeter

The hadron calorimeter (HCAL) is designed to measure the hadronic component of jets and the missing transverse energy of events together with the ECAL, which requires a good overall coverage. The HCAL is a sampling calorimeter consisting of scintillator tiles with interposed absorber plates made of steel and brass. The hadrons entering the calorimeter interact with the nuclei of the detector material, which creates a hadronic shower measured by the scintillators. The optical signal is detected with hybrid photo diodes (HPD) mounted at the ends of the barrel. Since most of the shower energy is stored in the absorber material, the energy resolution of the HCAL is worse compared to the ECAL resolution.

Figure 3.7 shows the HCAL subdivided in several subsystems. The hadron barrel (HB) calorimeter covers a pseudo-rapidity range $|\eta| < 1.3$ and is placed between the EB and the solenoid magnet. The HB consists of two barrels of 18 identical brass alloy absorber plates with wavelength shifting fiber readout (WLS) arranged parallel to the beam axis. The hadron endcap (HE) calorimeter covers the range $1.305 < |\eta| < 3.0$ and consists also of brass and scintillator. In both parts of the calorimeter the segmentation is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ except near $|\eta| \sim 3.0$, where the segments is doubled. The hadron forward (HF) calorimeter with coverage $3.0 < |\eta| < 5.0$ consists of quarts fibers embedded in iron, since in these regions the high rate of hadrons require the use of radiation hard material. Cerenkov light

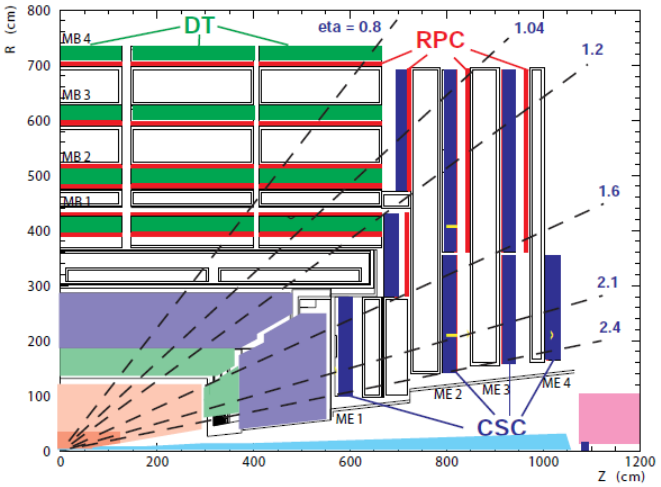


Figure 3.8: Longitudinal view of the CMS muon system [60].

generated in the fibers is transmitted to the photo multiplier tubes (PMT). In the barrel region outside the magnet, the hadron outer (HO) calorimeter covers the pseudo-rapidity range $|\eta| < 1.26$ extending the total depth of the HB in interaction lengths to a minimum of 11 λ_I [68, 69].

3.2.5 Magnet

The superconducting solenoid is 13 m long, has an inner diameter of 6 m and provides a strong magnetic field of 3.8 T. This high field is required for a momentum resolution of $\Delta p/p \sim 10\%$ for 1 TeV muons. The magnetic flux is returned by a 1.5 m thick saturated iron yoke, which weighs about 10000 tons hosting several layers of muon detectors. At full current on energy of ~ 2.6 GJ is stored in the magnet. The tracking system, the electromagnetic calorimeter and the hadronic calorimeter, except for the hadron outer calorimeter are situated inside the superconducting solenoid [70].

3.2.6 Muon System

Since various interesting processes in physics beyond the Standard Model as well as electroweak-, Higgs- and B-physics lead to muon final states, the muons are considered as a crucial tool for discovery and precision measurements. Hence the reconstruction, identification, correct charge assignment and precise measurement of the momentum of muons along with triggering of events using muons is of highest priority. Since muons are heavy ($m_\mu \sim 105.65$ MeV), they emit less bremsstrahlung radiation compared to electrons. Muons pass the calorimeter system depositing only little amount of ionizing energy, whereas other particles like electrons, photons and hadrons will loose their whole energy and will be stopped in the calorimeter. Only a small amount of hadrons pass the calorimeter (so called punch-through).

Fig. 3.8 shows the muon system situated outside the magnetic coil as the outermost detector component of CMS. The central barrel part (MB) covering the pseudo-rapidity range $|\eta| < 1.2$ and the two end-caps (ME) covering the range $0.9 < |\eta| < 2.4$ form the entire muon system, which consists of three types of gaseous detectors. In the MB drift tube (Δr) chambers are used. Additionally, in both regions resistive plate chambers (RPC) are used, which provide a fast response with good time resolution in order to identify the correct bunch crossing but with a coarser position resolution than the DT's and CSC's.
In the barrel region in total 250 DT chambers organized in four layers are arranged parallel to the beam axis in cylinders interleaved with the iron return yoke. The barrel consists of 5 wheels, arranged along the beam pipe of different z-position. Each of the 3 innermost layers have of 12 chambers with each covering a 30° azimuthal angle, whereas the outer section holds 14 chambers. Depending on the station, each chamber has 1 or 2 RPCs attached. DTs consist of 1.2 mm diameter and 9.6 mm long aluminum tubes with stainless steel anode wires at their center and the cells are filled with a gas mixture of Ar and CO_2. The electrons generated in the DTs by charged particles move to the anode wire in the center and the high electric field close to the wire amplifies the signal. The track position is measured by the traveling time of the electrons, which have a maximum drift time of 400 ns. The time resolution is 5 ns leading to a single point resolution of ~ 200 μm. The precision of the measurement of the muon vector is ~ 100 μm in ϕ position or ~ 1 mrad in the direction for each station. In the central region high P_T muons can be reconstructed from up to 44 measured points, when it passes 4 DT chambers.
The end-caps hold 4 disks of CSCs and RPCs attached perpendicular to the beam axis consisting in total of 468 CSCs each measuring up to 6 space coordinates in r, ϕ and z. The spatial resolution is typically about ~ 200 μm with an angular resolution in ϕ in the order of 10 mrad. The CSCs have a faster response and a finer segmentation than the DTs. The CSCs are 1×2 m^2 trapezoidal chambers consisting of six gas gaps, all having a plane of radial cathode strips and a plane of gold-plated anode wires running almost perpendicular to the strips in the middle of the chamber. They are filled with a mixture of Ar-CO_2-CF_4 gas. A charged particle entering the chamber ionizes the gas, leading to an avalanche at the anode which then induces a charge on a group of cathode strips. The signal on the wires is fast and is used for the Level-1 trigger [60, 71].

3.2.7 Trigger and Data Acquisition

The collision of proton bunches at the LHC with a frequency of 40 MHz creates an enormous amount of experimental data. Since the rate of collisions is too high to store each event on tape, the CMS trigger aims to select only interesting events and reduces the event rate to a manageable amount. The entire CMS trigger system consists of the Level-1 (L1) trigger based on custom electronics and the High Level Trigger (HLT), a software trigger relying on commercial processors.

For a trigger decision, the L1 trigger uses coarsely segmented data from calorimeter

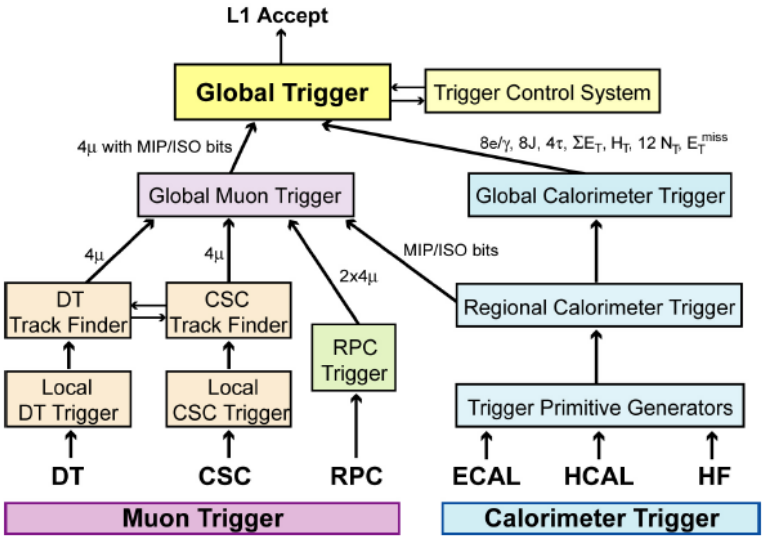


Figure 3.9: Design of the CMS L1-Trigger [63].

and muon detectors, while holding all the high-resolution data in pipeline memories in the front-end electronics. If an event is accepted by the L1 trigger decision, the high-resolution data is further analyzed by the HLT. Fig. 3.9 shows a sketch of the L1 trigger design. The global muon trigger combines and evaluates the information of the three muon systems components, where track segments identified by DTs and CSCs are used for a rough track reconstruction in the regional muon trigger, while in parallel tracks are reconstructed by using only information of RPCs. The global calorimeter trigger builds electron, photon, jets, sum of E_T and MET candidates using the information of the regional calorimeter trigger. Finally the global trigger uses the information of both global muon and global calorimeter trigger to decide whether a event is dropped or accepted to be further processed by the HLT. The HLT software system processes the event on a filter farm, reducing the event rate further to 150 Hz. Assuming an event size of 1.5 MB, a data stream of 225 MB/s has to be stored for latter processing. Since the HLT is a software based trigger, the algorithms used for HLT selection are flexible and adaptable [63, 72].

3.2.8 CMS Computing Model

In order to manage the enormous amount of data produced at the experiments, the LHC Computing Grid Project (LCG) has been developed, see e.g. [73, 74]. The CMS experiment uses decentralized mass storage and computing resources. At the Tier 0 center at CERN the raw data passing the HLT trigger and emerging from the data acquisition system is accepted and a first reconstruction of the physics objects takes place. The raw and reconstructed data is stored in the mass storage system of the Tier 0 and copied to the associated Tier 1 centers for further processing, where the size of the data is compressed (AOD format) and dedicated filters are applied

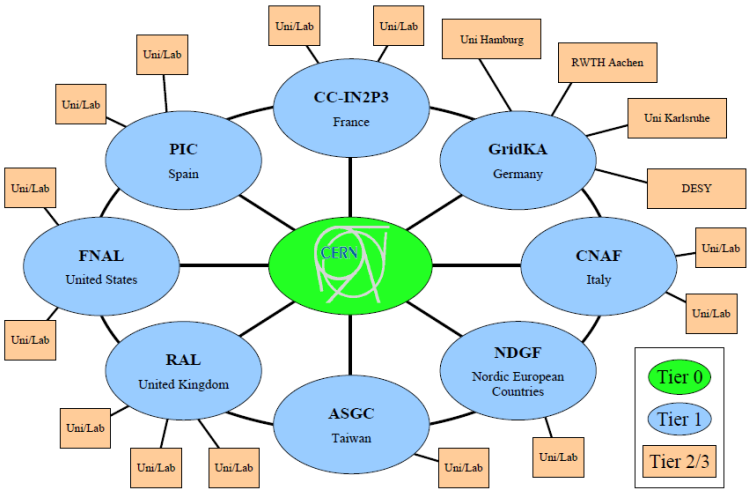


Figure 3.10: Illustration of the Tier structure of the LHC Computing Grid [73].

(skims). The skimmed datasets are copied to the Tier 2 centers, which offer capacity for analysis, calibration activities and Monte Carlo simulation. The last category, the Tier 3 centers are designed for interactive analysis by local groups. Fig. 3.10 illustrates the Tier structure schematically.

3.3 Analysis Framework and Event Reconstruction

This section is meant to provide an overview of the actual software implementation in CMS of the experimental analysis used for the detection of physics objects. In order to perform this analysis, the different detector components inside the detector have to be taken into account according to the physics process reconstruction. A large amount of Monte-Carlo simulation has to be processed and analyzed with respect to the current knowledge of particle physics as well as the design and expected response of the detector. A modular collection of software forms the CMS software framework(CMSSW), which is built around the concept of an *event*, has been developed for such a purpose by the CMS collaboration.

3.4 The simulation of events from LHC collisions

Samples of simulated events are used for the interpretation of the data collected by high energy experiments to determine the expected distributions of the final state particles. The event simulation requires:

- the generation of the four-momenta of the final state particles;

- the simulation of the interaction of the generated particles with the detector
- the simulation of the `digitization` phase (analogue electronic signals resulting from the particle-material interaction, get shaped, discriminated and read out by dedicated electronic devices)
- running on the simulated digitized signals from different level of trigger;
- running the programs for the reconstruction of the event both at the local sub-detector and to build the higher level analysis objects used for the final physics analysis on the simulated digitized signals.

As a result, if the collected raw data has the same format than the output of the simulation after digitization step, the same reconstruction software as used on the real data can be applied on the simulated data.

3.5 The CMS Software Framework

The CMSSW framework is a modular software built around the `Event Data Model` (EDM). The steps explained above have been processed with a clear data model, modular testing procedures and a simple data structure. It runs with a single executable `cmsRun` and many plug-in modules which run the different algorithms. This allows to use the same architecture to process collision and Monte-Carlo data.

The CMS event data model, represented in Figure 3.11, is based on the concept of an event as a C++ object container for all raw and reconstructed data of a physics event. The main steps of the CMSSW framework have been defined as follows:

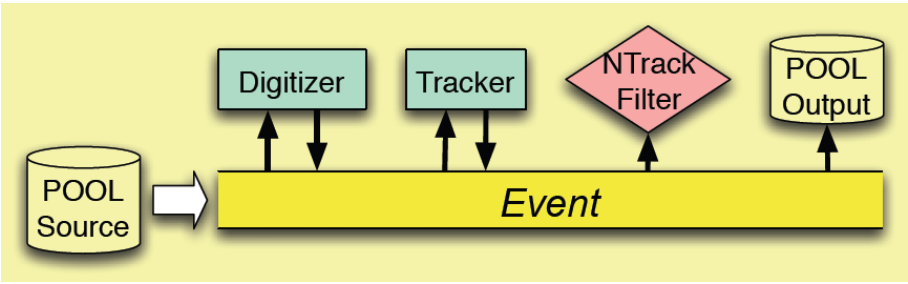


Figure 3.11: The CMS software `Event Data Model` [75]. First a source creates the *Event*, which is passed to execution paths. The *Producers* add data to the *Event* and once all paths have been executed, the *OutputModule* stores the output to external media.

- Pool `Source:` Read events and gives access the corresponding information via ROOT files [76]. It provides the *Events* to be processed.
- `Producer:` Processes the events and add this new information

- `Filter:` Read processed events to determine for further processing
- `Analyzer:` Analyze the properties of the event to produce histograms
- Pool `Output:` writes all processed events and stores the data information to a file.

3.6 Monte-Carlo Event Simulation

An event generator is a computer software that helps calculating the complex particle interactions as studied in high energy physics. These interactions are defined as scattering events produced in collider experiments. The different types of events and processes require the full sequence of calculations in an automatic way. Thus, event generators that simulate the events from the Lagrangian structure of the physics model up to cross section calculations are crucial to understand high-energy particle physics studies. In this sense Monte-Carlo techniques, extensively discussed in the next chapter, are used to determine the different variables according to the desired probability distributions and to generate events with all details. In recent years several packages have been developed for a wide range of collider experiments including a tuning of the generators to provide the best possible description of available collision data. The most important Monte-Carlo event generators are below.

- `PYTHIA` [77] is a general purpose Monte Carlo event generator, written in Fortran by T. Sjostrand and others and maintained by the Lund University theory division (The new Pythia 8 version written and developed in C++ [78]). It puts particular emphasis on the detailed simulation of QCD parton showers using the famous Lund string model for producing soft particles.`PYTHIA` provides full simulation of the hard hadron-hadron scattering of the LHC that is successfully implemented in CMSSW framework [79].
- `HERWIG` [80] is a general purpose Monte Carlo event generator, written in Fortran by G. Corcella and others (The new Herwig++ version written and developed in C++ [81]). It has particular emphasis on the detailed simulation of QCD parton shower. `HERWIG` provides a full simulation of the hard hadron-hadron scattering at LHC that is successfully implemented in the CMSSW framework [82].
- `MC@NLO` [83] is a generator to implement the scheme which is proposed for combining a Monte Carlo event generator with Next-to-Leading-Order calculations [27] of rates for QCD processes. The processes available are those of Higgs boson, single vector boson, vector boson pair, heavy quark pair, single top (with and without associated W), lepton pair and associated Higgs+W/Z production in hadron collisions. it is used Herwig for collinear approximation and fragmentation models in CMSSW framework.

- MADGRAPH/MADEVENT [84] MadEvent is a multi-purpose tree-level event generator, written in Fortran by Fabio Maltoni et. al. It uses the matrix elements produced by MadGraph. From these matrix elements, MadEvent calculates the cross sections and produces unweighted events. The amplitudes calculated by MadGraph are integrated, using a method called Single- Diagram-Enhanced, to give the cross sections. This decomposes the integration into a number of independent integrations. The 4-vectors of the particles produced in the event are calculated from these results. These events can then be passed to a shower Monte Carlo program which is Pythia in the CMSSW framework. To consistently combine multi-parton matrix elements with the QCD parton cascades is the matching approach of MLM [85] is employed.

- ALPGEN [86] is a Monte Carlo event generator specializing on multi-jet events og hadronic collisions. Alpgen was written in Fortran by Michelangelo L. Mangano et. al. Alpgen performs leading order calculations of the matrix elements of these events at parton level. Alpgen does not include the production of sparticles. A series of integrations is then performed in order to find an optimized integration grid, which contains the matrix elements. This grid is then used to calculate the kinematics and generate the events. These events will be weighted so that the contribution of each subprocess to the cross section can be evaluated. These events are then unweighted, by comparing the weight of each event to the maximum weight in the sample. These unweighted events are stored in a event file [87], which is then passed to PYTHIA for parton showering in the CMSSW framework. To consistently combine the multi-parton matrix elements with the QCD parton cascades the approach of MLM is employed.

- SHERPA [88] is a general purpose Monte Carlo event generator, written in C++ by F. Krauss and others. It contains a very flexible tree-level matrix-element generator, AMEGIC [89], for the calculation of hard scattering processes within the Standard Model and various new physics models. The emission of additional QCD partons off the initial and final states is described through a parton-shower model, APACIC [90]. To consistently combine multi-parton matrix elements with the QCD parton cascades the approach of Catani, Krauss, Kuhn and Webber(CKKW) is employed [91]. A simple model of multiple interactions, the old Pythia Model, is used to account for underlying events in hadron–hadron collisions. The fragmentation of partons into primary hadrons is described using a phenomenological cluster-hadronization model [92]. SHERPA is successfully implemented in the CMSSW framework [93]. Physics validation and software implementation of Sherpa in the CMS software are mainly done in Karlsruhe CMS-SUSY group.

- SUSYHIT [94] is a program package for the computation of supersymmetric particle decays within the theoretical framework of the MSSM. The code is based on two existing programs HDECAY [95] and SDECAY [96] for the calculation of the decay widths and branching ratios of the MSSM Higgs bosons and the SUSY particles, respectively. In the presented analysis the SUSY particle spectrum is calculated with SOFTSUSY [97] and passed to SUSYHIT.

3.7 Detector Simulation

In the following subsections the way how the main components of the CMS detector are taken into account in the detector simulations will be outlined. The detailed CMS detector and physics simulation are currently based on GEANT4 [98] simulation toolkit, which provides a rich set of physics processes describing electromagnetic and hadronic interactions in the detector in detail. GEANT provides a rich of set physics processes to model the electromagnetic an hadronic interactions in the presence of a magnetic field based on the detailed geometry of the detector and properties of the different particles. This includes energy loss through bremsstrahlung, ionization and multiple scattering and the processes of electromagnetic and hadronic showering. The result of energy loss in a dedicated volume of a sub-detector is stored in the form of *simulated hits*. Thus, the response of readout electronics of the different detector components, which is called *digitization*, is simulated.

The simulation has been done with full and fast simulations of the CMS detector. The fast simulation skips the time consuming simulation of the interaction of particles with the detecting material (GEANT4 - see Section 3.7) and digitization step. Simulated hits in the tracking detectors and shower evolution in the calorimetry are obtained using a parametrized approximation that aims at reproducing the full GEANT4 result. Concerning the simulation of the hits in the tracker detector, a simplified geometry is used, which is made of cylinders and disks that are assigned a depth in radiation and interaction lengths that approximate the full geometry result. Propagation of particles between layers is performed analytically; when the particle traverses a layer, the effect of the interaction with the material is calculated according to parametrized function. The hit position and error are then assigned with a Gaussian smearing with respect to the analytically calculated impact point. Track reconstruction starts with the emulation of the seeding efficiency. The pattern recognition step is skipped, the hit-to-track associations being taken directly from the MC information: the track fit is performed using the standard algorithms.

Calorimetry energy deposits are obtained in two steps. First the shower is simulated as if the calorimeter were a homogeneous medium; then detector effects, such as crystal granularity, inefficiency, magnetic field influence, are simulated. Energy deposits are then turned into reconstructed signals simulating noise and zero suppression. Muon tracks in the muon detector are simulated using a parametrization of resolution and efficiencies, but without the simulation of hits in the muon chambers. The matching with tracker tracks is done using standard algorithms.

3.7.1 Reconstruction of Physical Objects

Physics at hadron-hadron collisions is characterized by the presence of a variety of particles, each of them with its intrinsic properties. Understanding these particles drives the reconstruction, identification algorithms, which yield the *physical objects* as results.

3.7.2 Jets

Due to the color symmetry in QCD, only colorless final states are allowed as final state. Thus, a single quark or gluon state cannot be observed directly. What is instead seen in the detector is the result of the *hadronization* process, in which the parton is fragmented into colorless hadrons. An accurate measure of the quark or gluon energy is achieved by clustering the energy deposits in the calorimeters, defined as *hadronic jet* or simply *jet*.

A jet algorithm collects neighboring objects (calorimetric towers, particles). One basic technique of a clustering algorithm is to add up all the energy inside a cone around some specified direction. Besides giving a precise prediction of the initial parton energy and direction, an algorithm should also be *collinear safe*, so that the obtained result is unchanged if the energy carried by one particle is distributed among two close particles, and *infrared safe*, which means that the algorithm is stable against soft particles. The following jet reconstruction algorithms exist in the CMSSW framework:

- `Iterative Cone`: This is a simple cone-based algorithm. Input objects with $E_T > 1$ GeV sorted by descending order are used as seeds for the iterative search for stable cones associating all the inputs with R=$\sqrt{\Delta\phi^2 + \Delta\eta^2} < R_C$ from the cone axis, with R_C the cone amplitude. This algorithm has a short and well predictable computation time, but it is neither collinear nor infrared safe [99].

- `Midpoint Cone`: This algorithm is also based on fixed angular magnitude R_C cones. An improvement compared to the IC algorithm is given by considering as seeds also each pair of jet candidates closer than $2R_C$ [100].

- `SISCone`: the **S**eedless **I**nfrared **S**afe **Cone** jet algorithm is a cone cluster algorithm, designed to be infrared safe. As `Iterative Cone` algorithm is collinear and infrared unsafe, it is used by using a dedicated number of seeds for the jet clustering. However, the `SiSCone` algorithm avoids this by searching for all stable cones that follows a more advanced approach with an additional computing time [101].

- `Fast` k_T: this is an implementation of the k_T algorithm that dramatically reduces the computational time. It merges two inputs if they are near in the metric $d_{ij}=\frac{2min(P^2_{T,i},P^2_{T,j})\Delta R^2_{ij}}{D}$. The parameter D plays the role similar to the cone amplitude R_C [102].

The resulting jets are called also *raw* jets, as no correction has been applied to them. The major corrections to be applied: the dependence on the calorimeter response (**Offset-L1**), corrections for variations in the jet response with pseudo-rapidity relative to a control region (**Relative-L2**) and corrections of the transverse energy of a calorimeter jet to a particle jets in the control region (**Absolute-L3**). The main

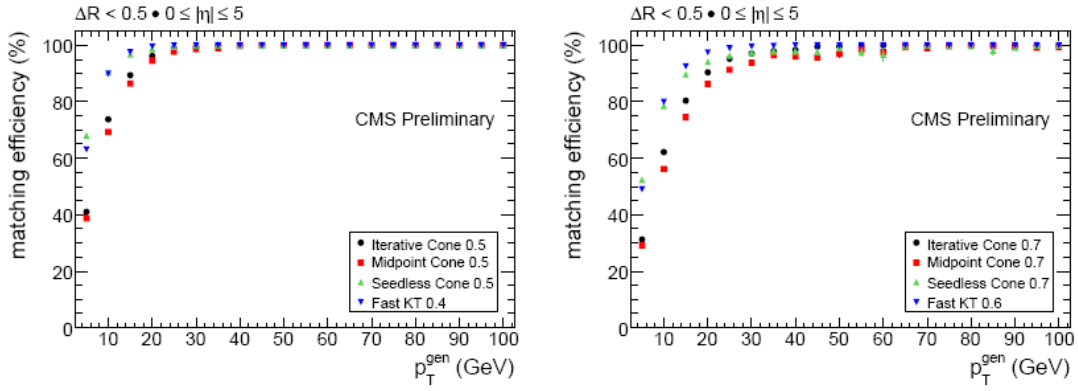


Figure 3.12: Matching efficiency vs P_T^{gen} for R = 0.5 / D = 0.4(left) and R = 0.7 / D = 0.6(right). Matching efficiency is defined as the ratio of number of particle jets matched to the calorimeter jet within the the corresponding ΔR and the total number of particle jets. The efficiencies of jets reconstructed with the Fast k_T and SiSCone algorithms indicate better performance [103].

goal of Offset-L1 correction is the electronic noise and physics noise of the system which are appeared in the calorimeter readouts and pile-up events. Additional complication is the energy thresholds applied to reduce data size from selective readout (SR) in electromagnetic calorimeter (ECAL) and zero suppression (ZS) in hadronic calorimeter (HCAL). With SR-ZS, noise effect depends on energy deposit which needs to properly take into account before subtracting noise. The main goal of the L2 correction is the extract the relative jet response with respect to barrel with the following conditions: larger statistics with better absolute scale and small dependence versus pseudo-rapidity of the jet. The main goal of the L3 correction is balance on transverse plane via two methods: γ +jet and Z+jets. The response of a calorimeter is not equal with respect to hadronic and electromagnetic interactions. To a large degree that is because in hadronic interactions some fraction of energy is lost to low-energy nuclear effects which do not result in an ionization or scintillation energy deposit.

3.7.3 Muons

Muons are reconstructed in CMS using three subsequent steps, namely local (muon chamber), standalone (muon system) and global (inner tracking and muon system) reconstruction.

The local reconstruction uses the position of hits in the muon chambers and forms segments in each chamber which are further used to generate state vectors (track segments), consisting of position, direction and a rough estimate of the muon transverse momentum. Both tracking detectors (DT and CSC) and the RPCs are used, where the RPCs have a worse spatial resolution, but complement the DTs and CSCs especially in the regions with problematic geometrical coverage, mostly in the barrel-endcap overlap region.

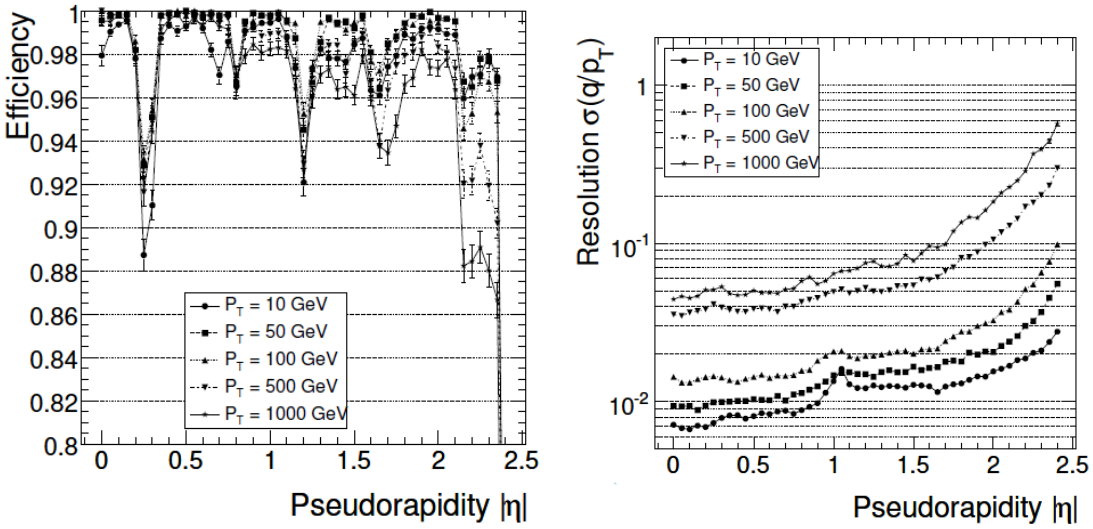


Figure 3.13: Global muon reconstruction efficiency (Left) and resolution (Right) as a function of the pseudo-rapidity [60].

The segments, of which the muon trajectories are built, starting from the innermost chambers, subsequently including the outer chambers, are used as inputs for a Kalman Filter fit [104]. The result is then propagated till a nominal interaction point. After a χ^2 (to reject bad hits, mostly from showering, delta rays and pair production) check on the track goodness: one obtains object the called *Stand Alone Muon* (STA) yield a *Global Muon*.

The global muon reconstruction is obtained from the muon trajectories by including the information of the inner tracking system. Using the standalone reconstructed muon as input, the muon trajectory is extrapolated from the innermost muon station to the outermost region of the inner tracking system. The regions of interest in the tracker are determined by the extrapolation of the muon track taking into account muon energy loss in the material, multiple scattering effects and the uncertainties of the extrapolation. Inside the regions of interest in the inner tracking system, regional seeds are build using two hits from different layers. Then the track reconstruction algorithm runs, transforming each seed into a set of trajectories working from inside-out. Based on the hit multiplicity and χ^2 tests, the trajectory cleaner resolves ambiguities between multiple trajectories arising from the same seed. In the last step, the reconstructed tracks are fitted using the hits in the muon chambers from the original standalone reconstruction together. The final muon candidates are selected on the basis of a χ^2 cut [44]. The muon reconstruction efficiency and the momentum resolution for the global reconstructed muons is shown in Fig. 3.13. The reconstruction efficiency is typically 95-99%, except for pseudo-rapidity regions between 2 DT wheels ($|\eta| = 0.25$ and $|\eta| = 0.8$) and in the transition region between the DT and CSC systems ($|\eta| = 1.2$). For low-momentum muons, the muon resolution around $\sim$ 2% is dominantly obtained by the resolution in the silicon tracker [60].

3.7.4 Electrons

Electron reconstruction in CMS requires at the most basic level a cluster energy in the ECAL that is matched to a track in the inner tracker. Energy clustering is designed to collect all the energy from an electron incident on the ECAL. Clusters are defined as regions of energy deposit that are centered upon a high energy seed crystal and which extend from that seed crystal in the η and ϕ directions until energy deposition either falls below a present threshold or begins to rise again. The electron energy is taken from the energy of the cluster[1]. The track is used to deduce the charge of the electron, and for background rejection through quality cuts on track momentum and cluster energy matching and track and cluster position matching. The CMS reconstruction code [105] provides different levels of electron identification, namely *Robust*, *Loose* and *Tight*. Robust identification is designed to be simple, easy to understand, and insensitive to expected uncertainties in detector alignment and calibration. It does not use electron classification which is a cut on E_{seed}/p_{in}. Loose and Tight identification involves tighter cuts, cuts on E_{seed}/p_{in}, and electron classification. Electrons passing Tight identification are a subset of those passing Loose identification, which are in turn a subset of those passing Robust identification.

3.7.5 Missing Transverse Energy (E_T^{Miss})

The summed momenta of all particles involved in a collision must be the same after the collision as before. Since there is no momentum in the transverse plane before the collision, there is also no net momentum in the transverse plane after the collision. Neutrinos and weakly interacting particles are not directly detected by CMS, but they do carry momentum that contributes to the balance. This weakly-interacting particles are detected by an imbalance of measured momentum. This inbalance of momentum is called missing transverse energy. The Missing transverse momentum ($\vec{E_T^{Miss}}$) is defined as the sum over energy deposits in uncorrected, projective calorimeter towers:

$$\vec{E_T^{Miss}} = -\sum (E_n \sin\theta_n \cos\phi_n \hat{\mathbf{i}}) + (E_n \sin\theta_n \sin\phi_n \hat{\mathbf{j}}) \tag{3.9}$$

The missing transverse energy (E_T^{Miss}) is defined as $\vec{E_T^{Miss}}$. Despite the fact that this quantity has a very simple definition, it is sensitive to detector malfunctions (e.g hot or dead channels) or particles hitting poorly instrumented regions of the detector. To the measured E_T^{Miss}, called als raw E_T^{Miss}, corrections from jet energy response and energy scale, muons (which have negligible energy deposits in the calorimeters) and also particle-based algorithms can be applied to improve the E_T^{Miss} resolution and its central value.

[1] In certain rare case when the cluster energy is believed to be badly mis-measured, the electron energy is taken from the measured momentum of the track.

The current implementation (used in the analysis) of E_T^{Miss} uses Monte-Carlo corrected jets, taking into account the detector response and energy scale:

$$\vec{E}_{T,Corr.}^{Miss} = \vec{E}_T^{Miss} - \sum_{i=1}^{N_{jets}} [\vec{p}_{T,i}^{corr.} - \vec{p}_{T,j}^{corr.}]. \tag{3.10}$$

The corrected missing transverse energy definition is used in this thesis.

3.8 Software Tools

- ROOT [76, 106] is an object-oriented framework designed for the challenges of data analysis in high-energy physics. It provides facilities for visualization of physics results with a large collection of classes which includes four-vectors, geometry packages and statistical tool packages. In addition it provides a high-performance input/output system as well as functionality for histogramming and fitting routines. In the following chapters, those properties are used for the physics analysis and figures of this thesis.

Chapter 4

Simulations of the LHC processes

The collisions of particle producing accelerators and colliders have different forms of interaction that depend on the primary particles (such as electron-positron (LEP) and proton-(anti)proton LHC(Tevatron)) with their properties. In these interactions, the colliding particles are considered as the initial state of a physical system. During the collision these particles can annihilate or scatter, thus producing different particles which are defined as the final state. The relations between initial and final state interactions of a physical event are related through the scattering matrix which is called S-Matrix.

The number of subprocesses in pp collisions describing a given process is so large that automatic tools have been developed to reduce the calculations since interactions at higher energies open a large spectrum of possible final states and consequently increase the number of processes to compute. It is necessary to impose the higher order calculations one can divide them into two categories, virtual(or loop) and real gluon radiation as shown in Figure 4.1. The final-state particles generated by event generators can be fed into the detector simulation, allowing a precise prediction and verification for the entire system of experimental setup.

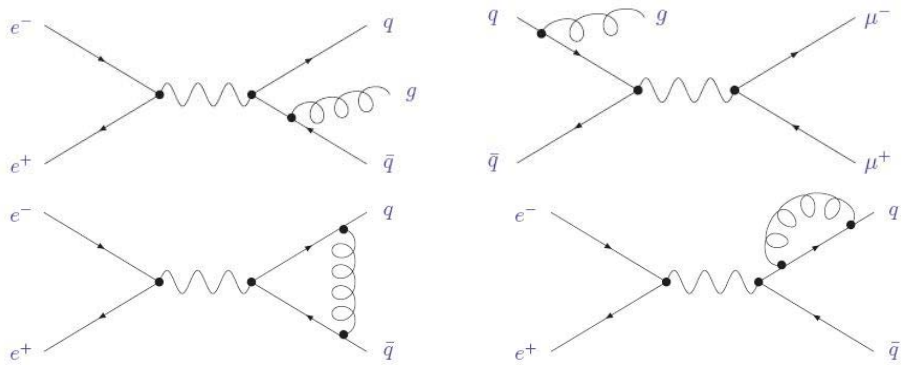


Figure 4.1: A schematic view of the real (upper) and virtual (lower) emission.

4.1 The Hard Processes Formalism and Factorization Theorems

In hadron colliders scattering process can be classified in two way: hard or soft. QCD is the underlying theory for all such processes but the approach and level of understanding is very different for those two cases. For hard process, e.g. high P_T jet production, the structure of the event properties can be predicted with good precision using the perturbation theory. For soft processes, the cross section calculation and the rates of the event are dominated by the non-pertubative QCD effects, which are not well understood so far. An understanding of the rates and the characteristics of predictions for hard processes, for the high P_T events, using pertubative QCD is crucial for LHC. The basic approximation for cross-section calculations in hadronic collision is defined as:

$$\sigma(AB) = \int \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a) f_{b/B}(x_b) \cdot \hat{\sigma}_{AB\to X}. \tag{4.1}$$

where for the Drell-Yan [107] process. X is defined as two opposite sign (OS) leptons and $ab = q\bar{q}, g\bar{g}$ with $M_X{=}M_{l\bar{l}}^2$ when the center of mass energy, s$\to\infty$, goes to infinity, $\tau{=}M_{l^+l^-}^2/s$ fixed. For the measured cross sections, the good agreement provided confirmation of the parton model. The production of hadrons and photons with large transverse momentum was also successful for the hard processes. However, the problems appeared when perturbative corrections from real and virtual gluon emission were calculated (see Figure 4.1). The calculations of gluons emitted from incoming quarks conceive large logarithms that spoiled the convergence of the perturbative expansion. In the mean time it was subsequently realized that these logarithms were the same as those that arise in Deep Inelastic Scattering (DIS) structure function calculations, and could therefore be absorbed some sets of equations, called as DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) equation. Figure 4.2 presents multiple small-angle parton emission in DIS for higher-order contributions.

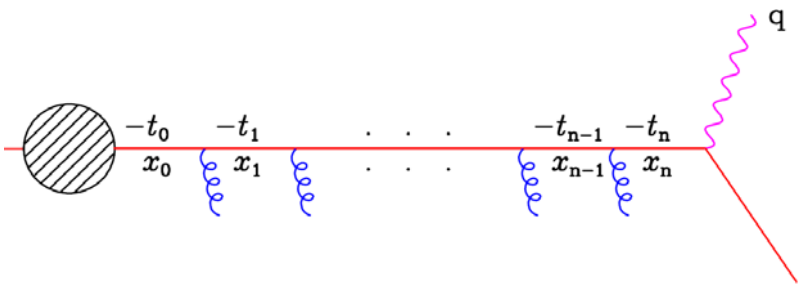


Figure 4.2: Incoming quark from target hadron, initially with low virtual mass-squared -t_0 and carrying a fraction x_0 of hadrons momentum. It is finally struck by virtual mass squared q^2=-Q^2.

Cross section depends on Q^2 and on momentum fraction distribution of partons seen by f(x, Q^2). Thus, all logarithms appearing in Drell-Yan corrections could be

factored into renormalized parton distributions in this way, and the factorization theorem shows that this is a general attribute of the hard scattering process [108]. The Q^2 appears in the parton distribution functions (PDFs) and is a large momentum scale that characterizes the hard scattering e.g $M^2_{l^+l^-}, P_T$. The cross-section calculation simply becomes

$$\sigma(AB) = \int \mathrm{d}x_a \mathrm{d}x_b f_{i,a}(x_a, Q^2) f_{j,b}(x_b, Q^2) \cdot \hat{\sigma}_{AB\to X}. \tag{4.2}$$

In this sense the leading logarithm approximation is equivalent $Q^2 \approx 2M^2_{l^+l^-}$. Naturally the final step is the finite corrections which are considered after logarithms.

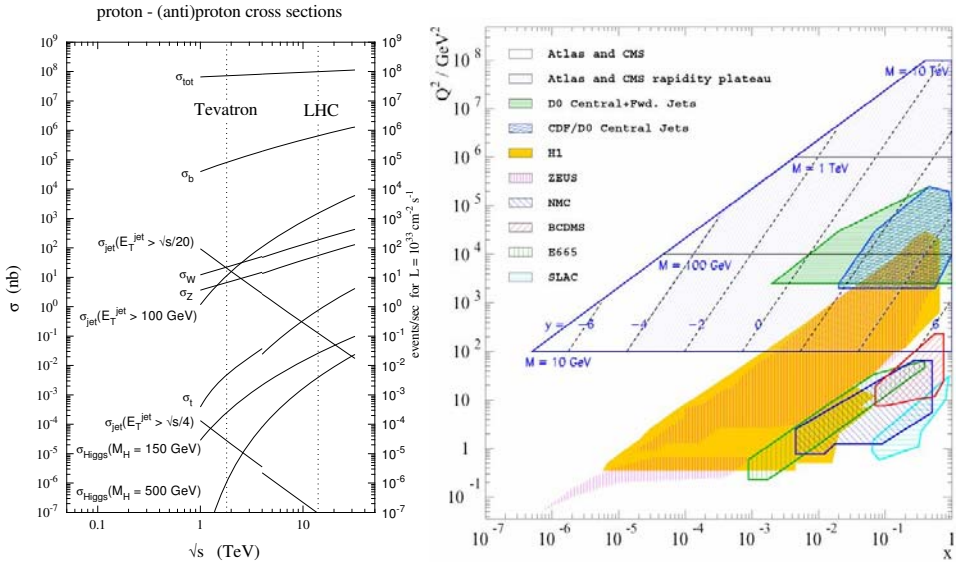


Figure 4.3: Left: The SM cross sections as function of center of mass energy. Right: Graphical representation of the relationship between parton (x, Q^2) variables and the kinematics variables corresponding to a final state of mass M produced with rapidity y at the collider [109].

Therefore they have to be calculated separately for each process of order n in the strong coupling constant (α_s^n):

$$\sigma(AB) = \int \mathrm{d}x_a \mathrm{d}x_b f_{i,a}(x_a, \mu_F^2) f_{j,b}(x_b, \mu_F^2) \cdot [\hat{\sigma_0} + \alpha_s(\mu_R^2)\hat{\sigma_1} + \alpha_s^2(\mu_R^2)\hat{\sigma_2} + ...)]_{AB\to X} \tag{4.3}$$

μ_F^2 is the **factorization scale** which can be considered as the scale that separates the long and short distance physics, μ_R^2 is the **renormalization scale** for the QCD running coupling. The corresponding cross-section in Eq. 4.3 is invariant under changing these parameters if calculated to all orders of perturbation theory. However, the individual $\hat{\sigma_1}$ depend on the explicit scale choice, μ_F^2 and μ_R^2, on the parton distributions and the coupling constant in these equations. Furthermore, it is necessary to include more terms in the perturbation series.

In the absence of higher order corrections one can try to make an educated guess of these scales by avoiding large logarithmic corrections, which are typically of the form $\log(Q^2/M^2)$, where M is the typical mass scale of the process. So if one chooses Q^2 to be of order M^2 the large logarithmic terms become small. To check the uncertainty of the choice of these scales one typically varies them by a factor two up and down. The cross section for various processes are shown in Figure 4.3 (left) together with allowed range in x needed for the production of a particle with a certain mass M^2=Q^2 (right). The calculation of a typical cross section for a given inclusive final state X + anything at leading order (LO) process can be expressed as follows: check the leading order process that contributes to X and calculate the corresponding $\hat{\sigma}_0$. The appropriate combination of parton distribution functions for initial partons is determined and a specific choice for the factorization(μ_F) and renormalization(μ_R) scales has to be choosen. Finally the numerical integration over the x_a and x_b is performed. The Monte-Carlo modeling of physics process in pp collision is shown pictorially in Figure 4.4. Modern event generators typically factorize the entire process in steps according to the different kinematics regimes.

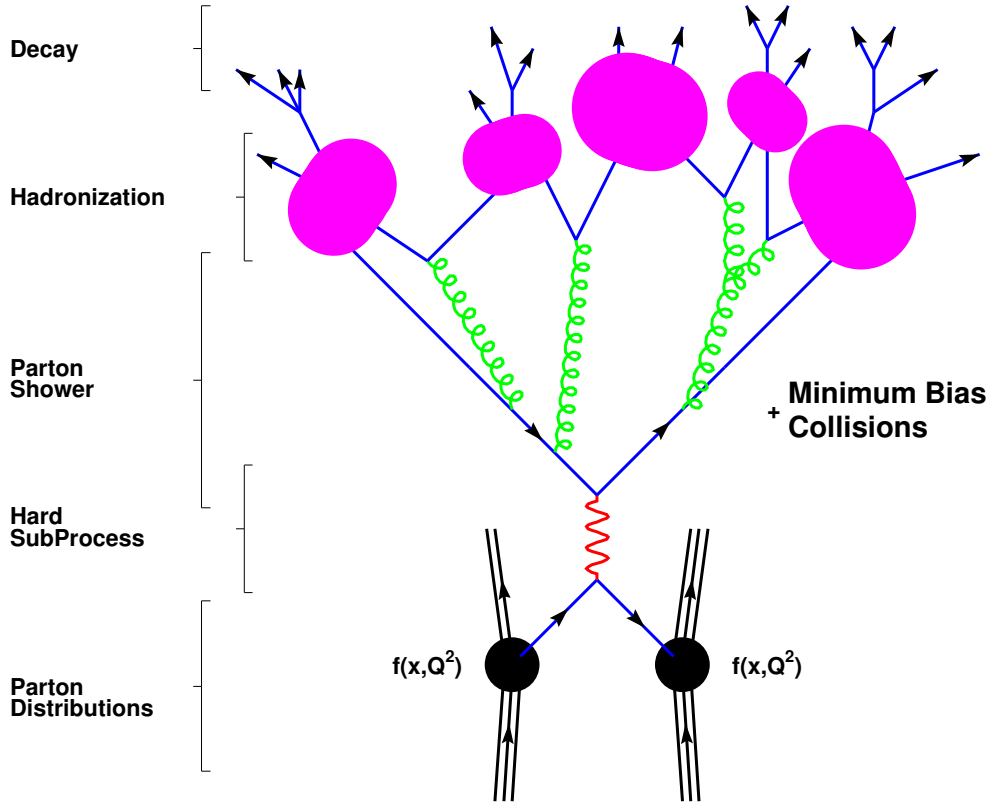


Figure 4.4: The basic structure of the showering and hadronization event generators for particle physics experiments [110].

The time evolution of the event goes from bottom to top in Figure 4.4. Two **protons** (each indicated by three solid lines to denote their valence quark content) collide and a parton is resolved at scale **Q** and **momentum fraction x** in each one. The parton density as function of x and Q^2 is encoded in the **parton distribution function** which is labeled by $f(x, Q^2)$. The quark and anti-quark annihilate into an s-channel resonance denoted by a wavy line. The resonance then decays into a fermion anti-fermion pair. This part of the event is called the **hard subprocess**. The generator incorporates higher order QCD effects by allowing the (anti)quarks to branch into qg pairs, while the gluons may branch into $q\bar{q}$ or gg pairs. The resultant partons may also branch, resulting in a cascade of branchings. This part of the event is labeled **parton shower** in the figure (showering of or cascade of partons). The event now consists of a number of elementary particles, including quarks, anti-quarks, and gluons which are not allowed to exist in isolation, as dictated by colour confinement. Next, the program groups the colored partons into colour-singlet composite hadrons using a phenomenological model referred to as **hadronization**. The hadronization scale is in the non-perturbative regime and the programs use fairly crude phenomenological models, which contain several non-physical parameters that have to be tuned using experimental data. After hadronization, many **short-lived resonances** will be present and are their decays simulated by the program. The generators also add features of the **underlying event**. The **beam remnants** are the colored remains of the proton which are left behind when the parton which participates in the hard subprocess is pulled out. The motion of the partons inside the proton results in a small (≈ 1 GeV) primordial transverse momentum, which yield a recoil energy of the beam remnants. The beam remnants are colour connected to the hard subprocess and so should be included in the same hadronization system. **Multiple parton-parton interactions**, wherein more than one pair of partons from the beam protons interact, are also accounted for. In a final step, **pile-up** from other proton-proton collisions in the same bunch crossing are added to the event [110].

Although parton shower (PS) generators, like Pythia and Herwig, provide an excellent description in regions which are dominated by soft and collinear gluon emission, it is important to consider a good description of processes where the partons are energetic and widely separated. In addition to effects of interference between amplitudes, matrix element (ME) calculations provide a good description of these processes. However, the ME calculations do not taken into account the interference effects in soft and collinear gluon emissions which cannot be resolved, as shown in Figure 4.5 (see details in Section 4.2.3).

It is essential not to trust blindly the results of any single event generator, but always to have several cross-checks. Sound judgment and some understanding of the generator are necessary prerequisites for a successful use. There exist many generators in recent times that can make calculations up to several jets in the final state. These predictions are tested well with many experiments and are tuned for further predictions.

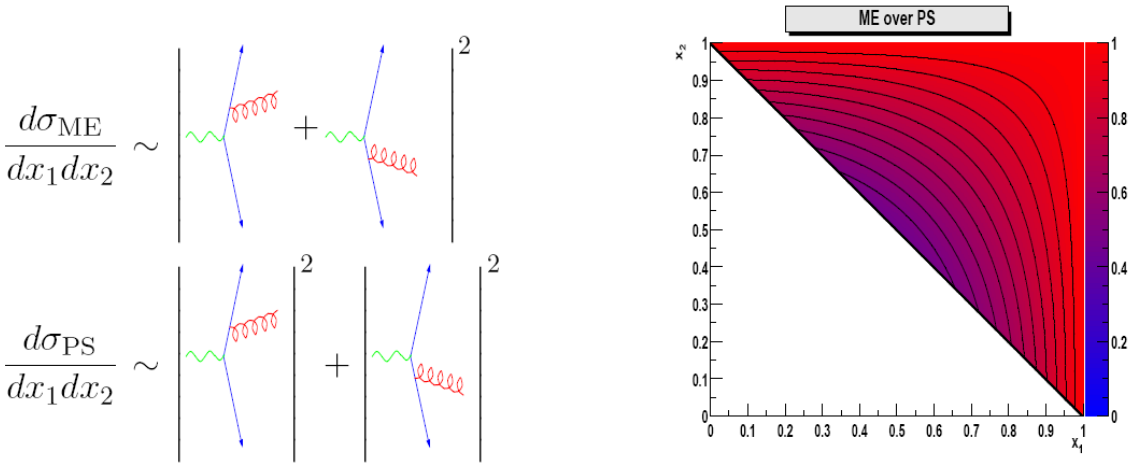


Figure 4.5: Pictorial representation of contributions to single gluon emission in $e^+e^- \to q\bar{q}g$ as given by the ME and the PS. As indicated, the PS does not take into account the interference contributions present in the ME. This leads to the ratio ME/PS as depicted in the contour plot. At the soft and collinear boundaries of the phase space in the x1-x2 plane the PS correctly reproduces the ME, whereas in the region of hard gluon emission the PS omits the interference contribution [111].

Among the general purpose and the hard process ones there are many tree-level (LO) generators: `PYTHIA`, `HERWIG`, `MADGRAPH/MADEVENT`, `ALPGEN` and `SHERPA`, which are able to calculate matrix element predictions for a number of processes, and to match the matrix element informations with parton showers. A few generators, which can perform the full Next-to-leading-order (NLO) calculation with all virtual corrections, are also available for a limited number of processes; an example is the program called `MC@NLO`. To begin with, the simulation prescriptions as implemented in the generators will be reviewed in detail.

4.2 Generator Tools and Approaches

In recent years, the theoretical and experimental particle physics have seen a very intense activity in improving existing Monte-Carlo tools and developing new ones, in view of the ongoing Tevatron Run II analysis and of the LHC start up. There are two important ranges of the LHC: the increase of a factor of ten in the available available center of mass energy and the increase of a factor 10^3 in luminosity. In view of these facts the uncertainties in theoretical predictions could become a limiting factor in extracting information from data analysis. In this chapter, the most widely used QCD Monte-Carlo programs, are considered by reviewing the features of their theoretical implementations in the next subsections.

4.2.1 Matrix Element

The central part of any generated event is the signal process, described by a corresponding matrix element (ME). There are major difficulties in order to get this formidable calculational task at hadron colliders:

1. The number of Feynman diagrams, i.e quantum mechanical transition amplitudes, grows roughly factorially with the number of external particles.
2. The integration over the phase space of the outgoing particles becomes a severe task. To achieve convergence of the Monte-Carlo procedure process- and cut-dependent phase space are required.
3. The partonic initial state is not fixed to a specific flavour pair (former LEP time), but can be any combination of two quarks, anti-quarks and gluons. All in all, the number of sub-processes that has to be considered increases dramatically with the number of final state jets.

In recent years, parton level generators, have been constructed that deal with the problems outlined above.

ALPHA Algorithm

The Alpgen MC relies on the `ALPHA` algorithm [112] which is based on the Schwinger-Dyson method to recursively define one-particle off-shell Green's functions, which are then numerically evaluated through a specific representation of their ingredients without Feynman diagrams. Strictly speaking, scattering amplitudes are connected with the Legendre transform of the Lagrangian at tree-level. This is valid for tree-level expressions. This approach significantly reduces the factorial growth of the number of terms to be calculated with the number of final state particles.

Helicity Amplitudes

The basic idea behind helicity amplitude is to calculate the set of particles momenta and helicities for a given Feynman diagram. The amplitude includes complex numbers that can be calculated by sampling the phase space and then they are squared and summed. Sherpa[1] and MadGraph/MadEvent use the helicity amplitudes method in different ways to make use of the Feynman diagrams. Madgraph/Madevent generates all Feynman diagrams for a process under consideration and employs the method of helicity amplitudes through the HELAS library [113]. In Sherpa, full sets of Feynman diagrams are constructed automatically and are translated by the program into helicity amplitudes [114].

[1]The new extension has been implemented to the new versions-1.2.X [www.sherpa-mc.de].

This approach relies on decomposition of the amplitude into scalar products of four-momenta and into spinor products. Knowing the fourmomenta, the scalar products can be calculated easily, and for fixed helicities, the spinor products can be computed as well.

4.2.2 Parton Shower

In the framework of a Monte-Carlo event generator the parton shower provides the link between perturbatively calculable differential cross sections at the parton level and models for their transition to observable hadrons with phenomenological parameters, which need to be tuned to data. For the calculation of an observable distribution in perturbation theory, the observable should be factorized with respect to the emission of soft and collinear gluons. In particular, if $\vec{p_a}$ is any momentum occurring in its definition, it must be invariant under the branching

$$\vec{p_a} \rightarrow \vec{p_b} + \vec{p_c} \tag{4.4}$$

whenever $\vec{p_b}$ and $\vec{p_c}$ are parallel, called collinear, or one of them is small, called soft or infrared by definition. A matrix element for a $2 \rightarrow$ n process is divergent ($q \rightarrow qg$, $\bar{q} \rightarrow \bar{q}g$, $g \rightarrow gg$) in these soft and collinear regions. These divergences can be removed by including virtual corrections, but such calculations require much more complex calculations and are available only for a limited set of processes. Therefore, parton showers provide an alternative way to handle this complexity and to remove divergences by cutting these phase-space regions at parton level and replace them by phenomenological hadronization model.

Final State Showers

In the shower formulation, the kinematics of each branching is given in terms of two variables, Q^2 and z. The main differences between generators on the market appear with the different interpretation of these variables. Q^2 has the dimension of a squared mass, and is related to the mass of the transverse momentum scale of the branching. z is defined as the energy fraction carried by the one of the two emerging partons (see Eq. 4.6). Strictly speaking, it gives the sharing of the energy of a particle a and momentum between the two daughters in equation 4.4, with parton b taking a fraction z and parton c a fraction $1\text{-}z$ via an ordering variable t. Different choices are possible for the ordering variable. A common choice, used in Sherpa[2], is the virtuality of the parton that is going to split $Q^2{=}p_a^2$. Other choices to treat this variable are also possible in case of treating this variable. If the transverse momentum is used as an ordering variable, the parton shower is called p_T-ordered showering which is implemented in new versions of the Pythia generator. Another choice is the energy weighted opening angle of each emission, resulting in a so-called angular ordered shower, which is used in Herwig.

[2]The new extension has been implemented to the new versions-1.2.X [www.sherpa-mc.de].

The branching of a parton a into two partons b+c shown in Figure 4.6 is considered. Assuming for the momenta of the partons

$$t = p_a^2 \gg p_b^2, p_c^2 \tag{4.5}$$

and the energy fraction z is

$$z = \frac{E_b}{E_a} \quad and \quad 1 - z = \frac{E_c}{E_a} \tag{4.6}$$

and the angle between the outgoing partons as $\theta = \theta_b + \theta_c$, one obtains for small angles

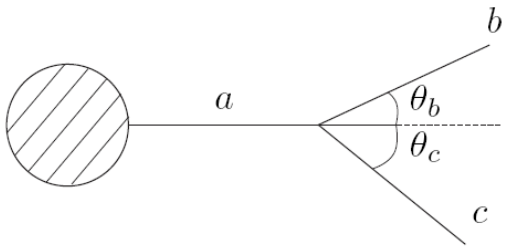


Figure 4.6: Branching of a parton a in two partons b and c

$$t = (p_b + p_c)^2 = 2E_bE_c(1 - \cos\theta) \approx z(1 - z)E_a\theta^2. \tag{4.7}$$

Using transverse momentum conservation $|p_{b_T}|^2 = |p_{c_T}|^2$ and making the same approximation for small angles,

$$\frac{|p_{b_T}|^2}{E_a^2} = \frac{|p_b|^2}{E_a^2}\sin^2\theta_b \approx z^2\theta_b^2 \tag{4.8}$$

$$\frac{|p_{c_T}|^2}{E_a^2} = \frac{|p_c|^2}{E_a^2}\sin^2\theta_b \approx (1 - z)^2\theta_c^2 \tag{4.9}$$

it is therefore possible to conclude that for small angles

$$\theta = \theta_b + \theta_c = \theta_b(1 + \frac{z}{1 - z}) = \frac{\theta_b}{1 - z} = \frac{\theta_c}{z} \tag{4.10}$$

From equations 4.4, 4.6 and 4.10, the differences between generators can be defined as follows:

- **Virtuality Ordering:** t $= (p_b + p_c)^2 \approx E_a^2\theta^2 z(1 - z)$

- **Transverse Momentum (p_T) Ordering:** $\mathrm{t} = p_{b_T}^2 = p_{c_T}^2 \approx E_a^2\theta^2 z^2(1-z)^2$
- **Angular Ordering:** $E_a^2\theta^2$

The splitting(or branching) probability $P_{a\to bc}$ is expressed in terms of `z` and `t`. It gives the probability for a gluon to split into two quarks with the energy fraction z. Similarly, one can look at the branching in which a light quark emits a gluon, or decays into a gluon emitting a quark, and at one in which a gluon decays into two other gluons [77]

$$dP_{a\to bc}(t,z) = \sum_{bc} \frac{\alpha_{abc}}{2\pi} P_{a\to bc}(z)dtdz \tag{4.11}$$

In the limit for transverse momentum and virtuality ordered showers, the exact matrix element in the collinear emission has to handle all divergences in 4.11[3] and the soft divergence arises from the splitting functions for z=1.

In order to handle these divergences, one needs to impose the conservation of the total probability. From integration equation 4.11, the total probability of branching between t and t+dt is obtained as

$$dP_a(t) = \left(\sum_{bc} \int_{z_{mix(t)}}^{z_{max(t)}} \frac{\alpha_{abc}}{2\pi} P_{a\to bc}(z)dz \right) dt \tag{4.12}$$

This is the integral of branching probability over all allowed z values and the probability that no branching occurred between the starting shower t_0 and t. Hence the expression of the branching probability

$$dP_a(t) = dP_a(t) exp\left(-\sum_{bc} \int_{t_0}^{t} dt' \int_{z_{mix(t')}}^{z_{max(t')}} \frac{\alpha_{abc}}{2\pi} P_{a\to bc}(z')dz' \right). \tag{4.13}$$

It means that there can be no longer branch at t if a parton has already branched at a time $t' < t$. The exponential terms above is defined as `Sudakov form factor` representing the non-emission probability between scale t_0 and t for final state shower. In addition to that the Sudakov form factor has also an interpretation in terms of Feynman diagrams while the branching probability given in equation 4.11 can be regarded as an approximation of the exact matrix element for real emission, the Sudakov form factor is an approximation of the complete virtual corrections.

Initial State Showers

Initial state showers are considerably more difficult. The shower is initiated by a parton selected from the parton distribution structure functions at small Q^2. This parton may now branch, but in the branching one daughter is timelike[4], whereas

[3]The detail extension of the splitting functions can be found in [77].

[4]The emission of the harder gluons in the hard scattering.

the other is space-like[5]. The time-like parton may develop a shower, very much like the final state shower case, although typically with less allowed phase space and therefore less extensive. The space-like parton may branch again, yielding a new pair of one time-like and one space-like daughter etc.

Following with the same approach like in the final state shower, the splitting probability can be written down with DGLAP equations [115]. The parton distribution function for `a` parton of type `b` evolves:

$$\frac{df_a(x,t)}{dt} = \sum_{a,c} \int \frac{dx'}{x'} f_a(x',t) \frac{\alpha_{abc}}{2\pi} P_{a\to bc}\left(\frac{x}{x'}\right) \tag{4.14}$$

The initial state shower Sudakov form factor is quite different than the one from final state since it contains the PDF's by definition[6]. As a result, the branching probability for initial state showers can be written via equations 4.13 and 4.14. The hard event is produced first and one reconstructs the initial state showering backwards in time. It is chosen from t to t-dt scale in order to recombine parton `b` with parton `a`.

$$dP_b(t) = \frac{df_b(x,t)}{f_b(x,t)} = |dt| \sum_{a,c} \int \frac{dx'}{x'} \frac{f_a(x',t)}{f_b(x,t)} \frac{\alpha_{abc}}{2\pi} P_{a\to bc}\left(\frac{x}{x'}\right) \tag{4.15}$$

Remembering the Sudakov form factor for FSR, the branching probability between scale t_{max} and t$< t_{max}$ for initial ISR is

$$S_b(x, t_{max}, t) = exp\left(- \int_t^{t_{max}} dt \sum_{a,c} \int \frac{dx'}{x'} \frac{f_a(x',t)}{f_b(x,t)} \frac{\alpha_{abc}}{2\pi} P_{a\to bc}\left(\frac{x}{x'}\right) \right). \tag{4.16}$$

Thus, the probability for parton `b` to recombine into a between t and t-dt is:

$$dP_b(t) = -\frac{S_b(x, t_{max}, t)}{dt} dt \tag{4.17}$$

and

$$dP_b(t) = \left[\sum_{a,c} \int \frac{dx'}{x'} \frac{f_a(x',t)}{f_b(x,t} \frac{\alpha_{abc}}{2\pi} P_{a\to bc}(\frac{x}{x'}) \right] S_b(x, t_{max}, t) dt \tag{4.18}$$

To summarize: the three quantities in equation 4.18 should be determined for each branching during the initial state radiation: the scale or ordering value t, the flavour a, the energy fraction $z = \frac{x}{x'}$. In hadron collisions, the effect from the flavour

[5]The emission of the softer gluons in the hard scattering.

[6]However, ISR backward radiation provides the correct configuration at the ME.

composition of the fraction is especially unknown for the correct prediction. They are determined with specific Monte-Carlo methods which are out of our scope. Further details can be obtained from [77].

Parton Showers and resummation

As shown in Eq. 4.3 before, the expansion in terms of α_s contains terms of the form $\alpha_s^n(\mu_R^n)\hat{\sigma_n}$ or $\alpha_s^n L^m$ (m$\leq$ 2n), where L being the cutoff scale for resolved emission. This scale includes logarithmic terms depending on the cutoff scale over squared center of mass energy s, defined as L= ln $\frac{Q_{cut}}{s}$. The evolution of these logarithms can spoil the convergence of the series. The large values of the logarithms need to be treated since the order n in the perturbative expansion is meaningful if successive terms in the series are small. The treatment of these large logarithms is called `resummation`. They can be ordered as:

- $\alpha_s^{2n} L^{2n}$ are the leading-log terms **LL**
- $\alpha_s^{2n} L^{2n-1}$ are the next-to-leading-log terms **NLL**

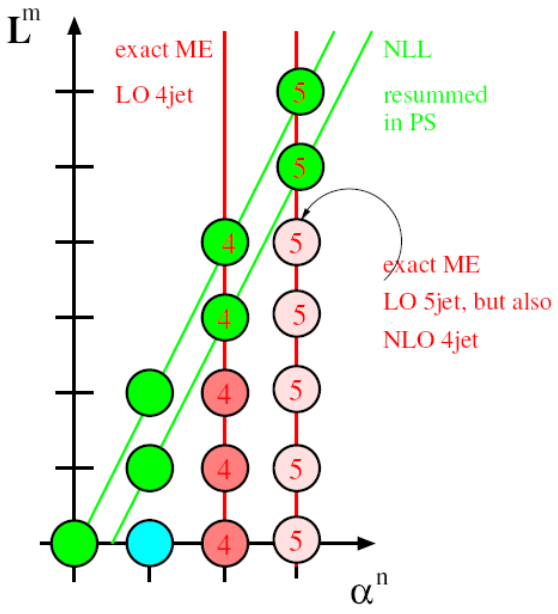


Figure 4.7: A schematic description of α_s orders and large logarithm powers that enter a multi-parton calculation [111].

4.2.3 Matrix Element and Parton Shower Matching

The analysis of multi-particle final states at LHC becomes increasingly important in the search for the production of SM processes and decay of new, heavy particles such as squarks and gluinos in SUSY. Therefore, in order to perform such analysis, their simulation in Monte-Carlo event generators should be as correct as possible. There are two complementary approaches to model the production of multi-particle final states. Matrix element and Parton shower structures of the MC methods can be summarized with their main facts:

Matrix elements are important for

- tree-level or beyond to describe particle production in specific process through Feynman diagrams up to several parton in the final state
- ME take into account all quantum interferences describing well separated parton configurations.

However, there are some problems as follows:

- ME cannot describe the internal structure of the jet because of the divergences in the soft and collinear regions.
- ME do not produce enough partons in order to combine hadron level information at final states.

Parton showers are important for

- their universality and give realistic parton configurations.
- In order to handle divergences, the Sudakov form factor is used in the collinear limit which is suitable for describing the evolution of jets.
- It can be used to evolve partons to a common scale at which no perturbative description is possible any more.

However, there is a problem: to fill efficiently the phase space for the well separated parton configurations since they are derived via the collinear approximation.

The main idea of a matching scheme is to avoid double counting, which prevents some events to appear twice, once for each path, and it ensures that each configuration is generated by at least one of the allowed paths.

Apparently, a combination of both approaches explained above allows a better description of a particle over the full phase-space in Figure 4.7. A way of merging multi-particle matrix elements with subsequent parton showers consistently at leading logarithmic accuracy (LL) and take into account important parts of the next-to-leading logarithms (NLL) was first formulated for $e^+e^- \rightarrow$ hadrons in [91].

There exist two approaches to the ME+PS methods. One based on CKKW, that has been outlined in [116] and implemented in the Sherpa generator and the MLM scheme, based on [117], that has been implemented in Alpgen, Madgraph/Madevent. The main structure of the methods have been represented in the next sections with their experimental observations at Tevatron experiment.

CKKW Matching

The CKKW matching algorithm is based on one or more *resolution parameters* which split the phase space into two regions, one of hard and large angle emission described by the ME and the other of soft/collinear emissions to be described by the PS, compared in Figure 4.5. These resolution parameters are important to put a cutoff scale between soft/collinear and fixed order calculations. In this sense the dependence on the resolution parameter is shifted beyond NLL. This algorithm is implemented in Sherpa and reviewed below:

The CKKW matching depends on the measure of the parton-parton separation as used in the $k_\perp$-jet algorithm [118, 119]. Thus the distance among two final state partons is defined as

$$y_{ij} = \frac{2min\{E_i^2, E_j^2\}(1-\cos\theta)}{s} \tag{4.19}$$

s being the center of mass squared energy, E_{ij} the parton energies and θ_{ij} their relative angles. The angular separation between a parton and the incoming partons (the beam) is defined as

$$y_i = \frac{p_{\perp i}^2}{s} \tag{4.20}$$

The separation between ME and PS partons is achieved by introducing a resolution parameter Y_{sep} and

- requiring that ME partons are resolved $y_{i,j}$, $y_i > Y_{sep}$ as well as
- vetoing PS emissions at a scale harder than Y_{sep}.

The CKKW approach can be summarized with the main steps for hadron-hadron collisions:

- The calculation of the cross sections, σ_{n_i}, is considered for each parton multiplicity n and for each different combination i of partons in the ME. The separation Y_{sep} is applied to avoid divergences with fixed α_s^{ME}.
- With all possible configurations n, the selected probability can be determined as $P_{n,i} = \sigma_{n_i} / \sum_{m,j} \sigma_{m,j}$

- and the momentum of the partons can be obtained by the corresponding matrix element squared.

- The splitting scale is reconsidered through a $k_\perp$-clustering of the partons emerging from the ME. The clustering proceeds up the core $2 \to 2$ hard process. This leads to a series with associated values of $k_\perp$ distance $y_1, y_2, ..., y_n$. In this case two factors appear: For each clustering i an α_s correction $\alpha_s^{cutoff}/\alpha_s^{ME}$ is applied with the Sudakov form factor correction [111].

- Events are accepted or rejected according to their weight.

- The accepted events are showered with a veto on the emission above Y_{sep}.

MLM Matching

An alternative prescription has been proposed by M. Mangano. The purpose of the algorithm is summarized:

- The first step is the generation of parton-level configurations for all final-state parton multiplicities n up to a given N. The kinematic cuts are defined following:

$$p_\perp^{parton} > p_\perp^{min}, \quad |\eta_{parton}| < |\eta_{max}|, \quad \Delta R_{jj} > R_{min} \tag{4.21}$$

  where p^{parton} and η_{parton} are the transverse momentum and the pseudo-rapidity of the final state partons, and ΔR_{jj} is their minimal separation in the (η, ϕ)-plane. The parameters p^{min}, η_{min} and R_{min} are called generation parameters, and are the same for all n=1,...,N.

- The necessary tree branching structure is defined for each event, allowing however only branchings, which are consistent with the colour structure of the events, which in Alpgen is extracted from the ME calculation [85]. For a pair of final state partons, the $k_\perp$-measure defined by

$$d_{ij} = \Delta R_{ij}^2 min(p_{\perp_i}^2, p_{\perp_j}^2) \tag{4.22}$$

  is used, where $\Delta R_{ij}^2 = \Delta\eta_{ij}^2 + \Delta\phi_{ij}^2$

- The $k_\perp$-value at each vertex is used as a scale for the relative power of α_s. The factorization scale for the parton densities is given by the hard scale of the process, such as $Q_0^2 = m_W^2 + p_{\perp_W}^2$. The hard scale Q_0 is adopted for all powers of α_s corresponding to the non-merged clusters.

- Events then are showered and the evolution for each parton starts at the scale determined by the default showering algorithms on the basis of the kinematics and colour connections of the event. After evolution, a jet cone algorithm is applied to the partons produced in the pertubative phase of the shower. Jet are

defined by a cone size R_{clus}, a minimum transverse energy $E_{\perp}^{clus}$ and a maximum pseudo-rapidity η_{max}^{clus}. These parameters are called matching parameters, and should be kept the same for all samples n=1,...,N.

- Starting from the hardest parton, the jet which is closest to it in the (η, ϕ)-plane is selected. Strictly speaking, the parton and jet match if the distance between the parton and jet is smaller than $1.5R_{clus}$. The matched jet is removed from the list of jets, and the matching test for subsequent partons is performed. The events is fully matched if each parton matches to a jet. Events, which do not match, are rejected. A typical example is when two partons are so close that they cannot generate independent jets, and therefore cannot match.

The approach used in Madgraph/Madevent is based on the MLM prescription, but uses a different jet algorithm for defining jet matching. The phase space separation between the different multi-jet processes is achieved using the $k_{\perp}$-measure as in Sherpa, while the Sudakov reweighting is performed by rejecting showered events that do not match to the parton-level jets, as in Alpgen [120, 121].

NLO Matching

The scheme implemented in the MC@NLO generator describes the hard emission just like a NLO order calculation, including the NLO normalization. In addition it simulates additional collinear particle emissions using the Sudakov factor. This is precisely what the parton shower does. It describes the entire $p_{\perp}$ range of jet emission for the **first** and **hardest** radiated jet consistently. Additional jets, which do not appear in the NLO calculation are simply added by the parton shower in the collinear approximation [122]. To summarize: To combine a fixed order NLO calculation with a parton shower one can think of the parton shower as a contribution which cancels a properly defined subtraction term which it can be included as part of the real emission contribution.

4.2.4 Fragmentation and Decays

Fragmentation is a non-perturbative phenomenon, and such can not be understood from first principle. Fragmentation is the process of the formation of hadrons out of quarks and gluons. As with the time-like parton shower (see Section 4.2.2), experience from e^+e^- annihilation helps constrain models significantly. Due to the colour confinement, these cannot exist individually. In the Standard Model they combine with quarks and anti-quarks spontaneously created from the vacuum to form of hadrons. Two main fragmentation methods exist: the **String model** (implemented in `Pythia`) and the **Cluster model** (implemented in `Herwig`).

The String Model

In this model the color field form strings, as expected from the self-interaction of the gluons. The strings are depicted in Figure 4.8. As they separate from each other the color field lines tighten and acquire a string shape configuration, the energy stored per unit length in the colour field tends to be uniform. If the energy density is high enough, the string can break and produce a quark-antiquark pair. In the simplest case, the string is stretched between a quark and an antiquark. The hadronic system is generated by pair creation inside the string. One goes to a frame where two string ends have opposite momenta, and, starting from each string end one has a fragmentation function to describe the probability to generate a hadron carrying away a given fraction of the longitudinal momentum of the string [77].

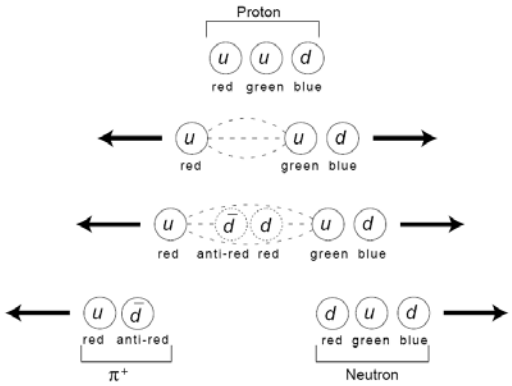


Figure 4.8: Schematic representation of the color field as the string form. Quark-antiquark are connected by "strings", gluons form "kinks" on these string, the harder the gluon, the stronger kink, string is treated as 1+1 relativistic object.

The Cluster Model

This model is based on the observation of colour pre-confinement. The color connected partons have a universal rapidly falling invariant mass spectrum which gives rise to a fragmentation model in which colour singlets, called `clusters`, are formed out of the partons emerging from the parton shower. In the cluster approach, final state gluons are forced to split into quark-antiquark pairs. Then each colour connected quark-antiquark pair is combined into a color-neutral cluster which is treated as a resonance that decays independently since one matches mass and flavour with a corresponding hadronic two-body system (or with a resonance) with the same flavour if the invariant mass of the colour connected pair is low enough [80]. A slightly modified version of this approach is implemented in Sherpa [92].

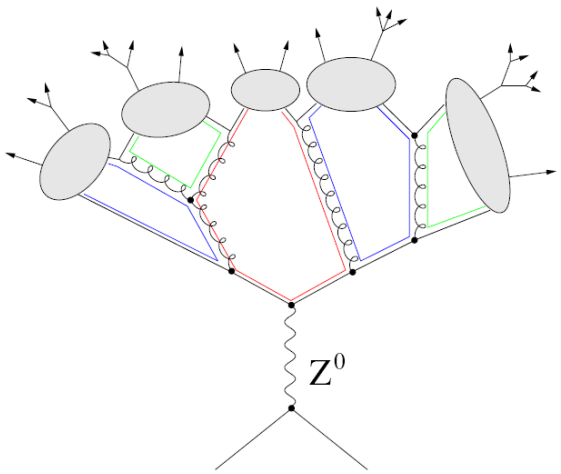


Figure 4.9: Schematic representation of the color field as the cluster form [80].

As discussed in the cluster model, the colour structure of the initial partons from the hard matrix element is fixed and the colour structure of the partons at the end of the parton shower is unambiguously determined.

4.2.5 Beam Remnants and Underlying Events

In hadron-hadron collisions each interaction parton leaves behind a beam remnant. These remnants do not take part in the initial state radiation and hard process, but they need to be colour connected to the rest of the event. In addition to that the composite structure of the two incoming hadrons implies the possibility that several pairs of partons can interact with simultaneous scatterings, called `Multiple-interactions`. These additional scatterings can be classified as hard or semi-hard which give non-negligible contributions to the total multiplicity, called the `Underlying Event`. The physics of the multiple-interactions is the least understood and its simulation relies on the used models, so extrapolation to different energies is difficult and uncertain. One possibility is simply to use a longitudinal phase space parametrization, as implemented in Herwig, with parameters fitted to describe data. Another is to allow a variable number of parton-parton interactions to take place within one and the same hadron-hadron collision. This is implemented in Pythia. Sherpa uses a simple model of the Pythia approach for the underlying event and multiple interactions. Further details related to the corresponding generators can be found in [77, 80].

4.3 Confronting Monte-Carlo predictions with recent Tevatron Data

This section reviews a comparison of DØ II data with the event generators MCFM, Alpgen, Sherpa and Pythia [28]. In Alpgen and Sherpa exact tree-level matrix elements are taken into account according to the MLM and CKKW prescriptions, respectively. The main objective is to show how accurately jet production in $Z^0(\rightarrow \mu\bar{\mu})/\gamma^*$ +jets +X events has been modeled by those two approaches.

The data used for the comparison plots was collected during October 2002 and February 2006 by the DØ at the Fermilab Tevatron collider at $\sqrt{s}$=1.96 TeV. The integrated luminosity corresponds to about 1 fb^{-1}. The measurement is carried out in a region of the di-muon mass of $65 < M_{\mu\bar{\mu}} < 115$ GeV in which the inclusive cross section for Z^0/γ^* production is approximately equal to that of pure Z boson production, and the measured distributions are corrected to the hadron level. The measured cross section in bins of the leading jet P_T^{Jet}, P_T^Z and rapidities for $Z^0(\rightarrow \mu\bar{\mu})/\gamma^*$ +jets +X events have been shown in Figures 4.10 and 4.11.

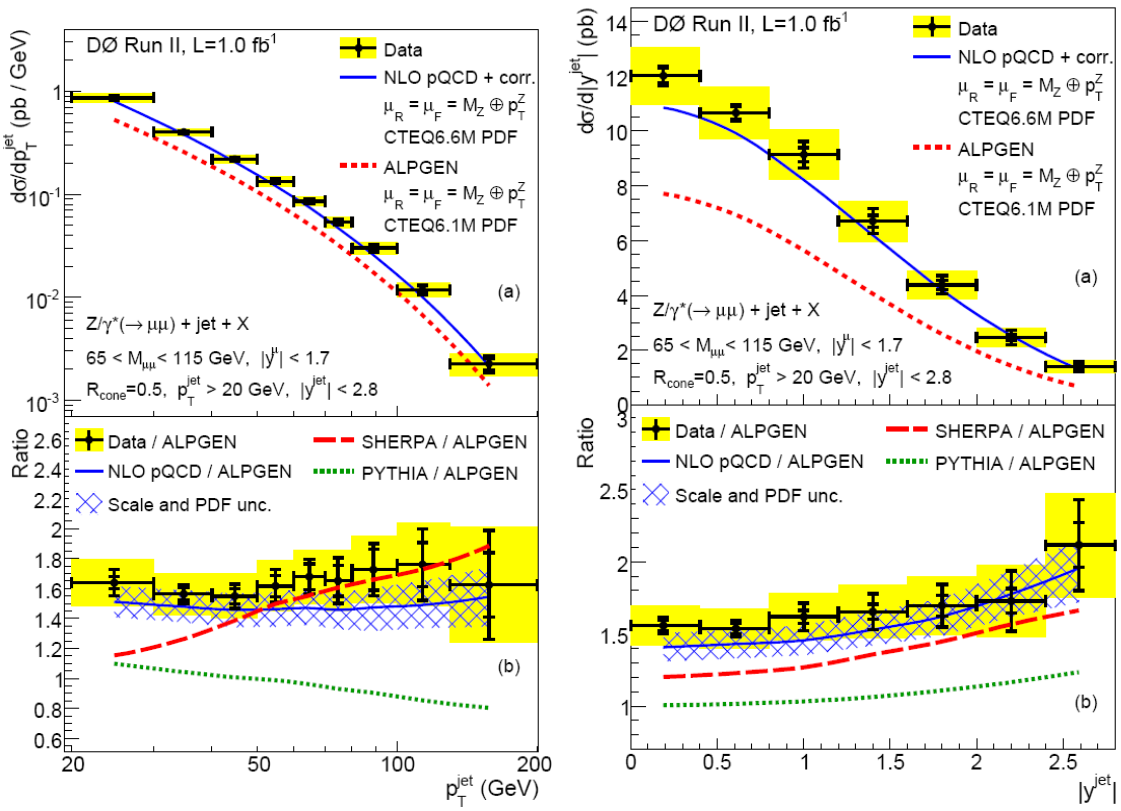


Figure 4.10: **Upper:** The measured cross section in bins of the leading jet P_T^{Jet} and rapidity for $Z^0(\rightarrow \mu\bar{\mu})/\gamma^*$ +jets +X events. Predictions from NLO and Alpgen are compared to the data. **Lower:** The ratio of data and predictions from NLO corrections, Sherpa and Pythia to the prediction from Alpgen [28].

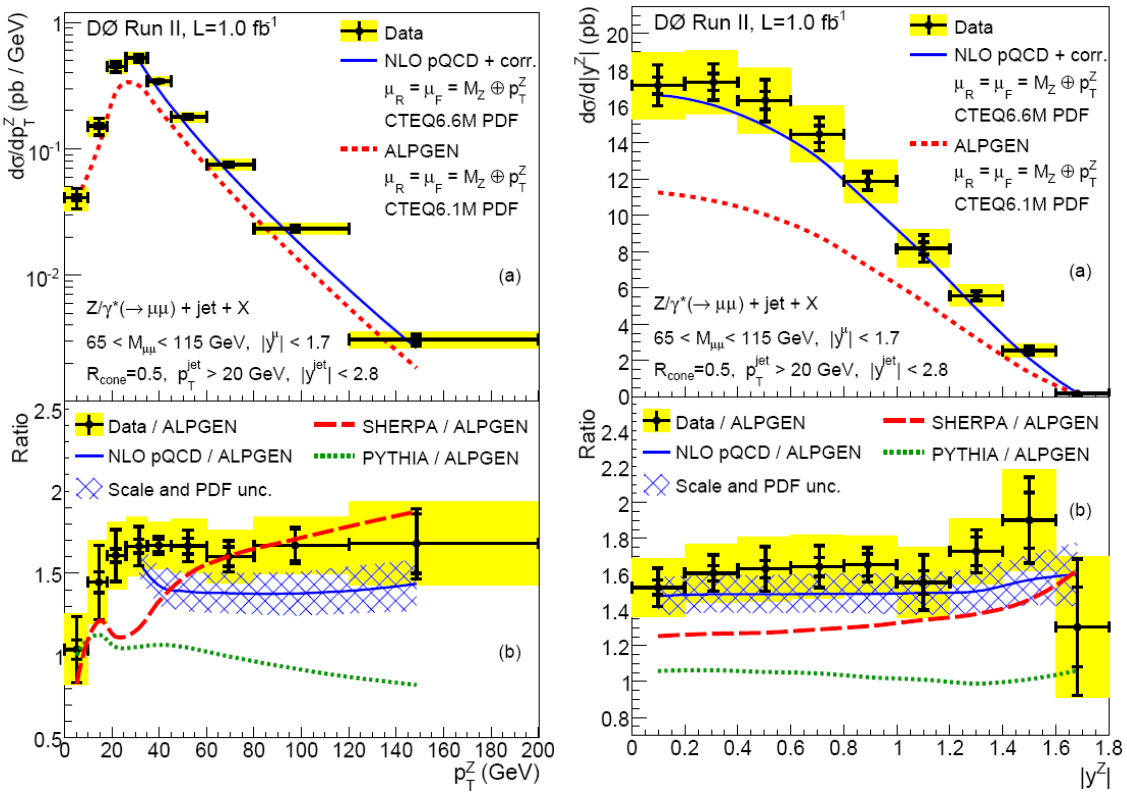


Figure 4.11: **Upper:** The measured cross section in bins of the leading jet P_T^Z and rapidity for $Z^0(\rightarrow \mu\bar{\mu})/\gamma^*$ +jets +X events. Predictions from NLO and Alpgen are compared to the data. **Lower:** The ratio of data and predictions from NLO corrections, Sherpa and Pythia to the prediction from Alpgen [28].

The results can be summarized as [28–31]:

- All uncertainties have been considered according to the tree-level exact matrix element with NLO pQCD.

- Although the distribution at low P_T^Z is dominated by non-perturbative processes, the shapes of the differential distributions are generally well described by NLO predictions.

- The shapes of the data distributions are generally well described by Alpgen except at low P_T.

- Jet rapidity distribution is narrower in Alpgen than in data, NLO pQCD, Sherpa and Pythia.

- Sherpa has a slope in P_T^{Jet} and P_T^Z relative to the data with more events at high P_T compared to the low P_T, Pythia shows an opposite behavior.

Chapter 5

Generator Comparison for LHC at $\sqrt{s} = 10$ TeV

The aim of this chapter is to compare Alpgen, Sherpa, Madgraph and MC@NLO for Z/γ^*+jets, W+jets, $t\bar{t}$+jets and VV+jets production in CMS operating at $\sqrt{s}$=10 TeV. The major characteristics of jets and leptons are studied at MC level and with CMS detector simulations. Differences between predictions are presented that can be interpreted as systematic uncertainties. Exclusive single vector boson samples, in which Z and W boson are forced to decay into muons without underlying event, are compared for Sherpa, Alpgen and MC@NLO simulations. Furthermore the inclusive samples that are decayed all leptons (for single and di-boson samples) and all possible channels for QCD and $t\bar{t}$ have been presented for Alpgen, Sherpa, Madgraph and MC@NLO generators.

All physics analysis at CMS use MC generators to design the event selection or extrapolate data to the region of interest. Most of these generators are tuned to existing experimental data available at $\sqrt{s} < 2$ TeV (see Section 4.3) but the extrapolation to the LHC energies introduces some systematic uncertainties which can be estimated by comparing MC generators with matching schemes, different higher order calculations and different factorization/renormalization scales. The model parameters will be tuned using LHC data, thus reducing different MC uncertainties. However, the knowledge which systematics will be the largest can help to select the observables and range of selection in different physics analysis starting from electroweak precision measurements to beyond the SM (BSM) searches to the calibration of data driven methods. In this chapter, the generator comparisons of SM background processes are discussed for the LHC. The effects of SM background uncertainties on SUSY observables will be extensively discussed in the next chapter.

From the description of the generators, the simulations of hard interaction at LHC can be divided in two steps:

1. Calculation of the ME at tree level or using higher order corrections (NLO, NNLO) The ME provides the kinematics of participating partons and the cross section. For most of $2 \rightarrow$ n, (n$\leq$ 6) processes the calculation at LO

can be done precisely but for final states with higher multiplicity the computation becomes difficult and different generators implement different methods producing somewhat different results.

2. After the ME calculation the PS and fragmentation are added. The radiation of soft gluons is usually described by the leading-log approximation (LLA) (see Section 4.2.2). The use of perturbative QCD for the exact calculation of the ME and approximate solutions for the PS cause double counting of jets. One has to ensure that no gluons harder than the ME gluons are generated at LLA, i.e. the LLA and ME phase space are matched. The matching is done differently: Alpgen uses the MLM scheme, Sherpa the CKKW method, and Madgraph uses both methods (see Section 4.2.3). The non-perturbative hadronization process is factorized using a parametrization for transverse and longitudinal momentum distributions of produced hadrons via fragmentation functions. Such schemes are implemented in general purpose generators like Pythia, which uses Lund string fragmentation, Herwig, which uses cluster fragmentation, and Sherpa, which uses modified cluster fragmentation (see Section 4.2.4).

The various differences between Alpgen, Sherpa and Madgraph make them very useful for the estimation of model uncertainties. The MC@NLO generator with the NLO ME calculations provides another independent check. Moreover the considered generators are relatively well tested with Tevatron data and are extensively used in the MC production for LHC [28–31, 123–125]. Clearly not all model systematics are covered by this limited choice. On the other hand after tuning the generators to data the differences will become smaller.

The main features of the generators used in the analysis are presented in Table 5.1. The simulation chain for different generators is different. Alpgen calculates all tree level diagrams and results are passed to Pythia to produce events. The MLM matching vetoes events with unwanted kinematics; i.e. when PS jets are harder than ME partons. The processes with different jet multiplicity are calculated separately and the exclusive samples are combined by weighting with the corresponding cross sections. Madgraph is also a hard process generator which uses Pythia the same way for covering shower, multiple-interactions and fragmentation processes. Although Madgraph uses a different ME algorithm, the corresponding Feynman diagrams in the processes are almost identical to Alpgen. The matching method used by this generator also shows a similar behavior as Alpgen because they use the same parton shower generator. The produced parton events via the Madgraph and Alpgen ME are passed to Pythia and the MLM method vetoes events with unwanted kinematics. The process with different jet multiplicity are calculated at parton level and the final events are to be combined externally by weighting with corresponding cross sections (see Section 4.2.3).

Sherpa is a standalone generator which includes PS and fragmentation. The CKKW matching is based on Sudakov rescaling which reweighs the mapped ME and PS history to calculate the survival probability for each jet. Sherpa can produce inclusive

samples internally with the appropriate scale for both ME and PS. MC@NLO can calculate the processes at NLO for 2→2 processes. Multi-jet events are produced via the PS as implemented in Herwig, i.e. the first jet is calculated at NLO while the softer jets are produced at LO in PS. The specific NLO matching scheme provides a smooth transition between hard and soft regions.

Table 5.1: Description of MC generators used in this analysis (see details in Chapter 4). The angular, virtuality and P_T ordered parton shower methods have been used by Herwig, Sherpa and Pythia, respectively. The Lund, Cluster and modified Cluster hadronization methods have been implemented in Pythia, Herwig and Sherpa.

	MC@NLO-3.3	Sherpa-1.1.3	Alpgen-2.13	MadGraph-v4
Matrix Element	NLO(2→2)	LO (2→4)	LO(2→6)	LO(2→9)
Parton Shower	Angular(θ^2)	Virtuality(Q^2)	$P_T(P_T^2)$	$P_T(P_T^2)$
Matching Scheme	NLO	CKKW	MLM	MLM
Matching parameters	Fixed Scale	P_T^{Jet}	P_T^{Jet}, η_{Jet}, ΔR	P_T^{Jet}, η_{Jet}, ΔR
Fragmentation	Clustered	Modified Clustered	Lund	Lund

5.1 Generator Comparison for SM Background in the CMS

This section is devoted to the exclusive and inclusive SM samples for the CMS simulation. Exclusive samples are considered for single vector bosons, Z($\mu\bar{\mu}$)+jets and W($\mu\nu$) +jets, productions where bosons are forced to decay into muons without underlying event. Inclusive samples are considered for single vector bosons (Z/γ^*+jets, W+jets) and double vector bosons (VV+jets) that are simulated with leptonic decays (e, μ, τ) and for $t\bar{t}$+jets and QCD+jets that are inclusively decayed with the default underlying event tunes. In addition Zinvisible+jets production is considered via exclusive ($\mu \leftrightarrow \nu$ replacement at generator level) and inclusive (Z decay to all ν's at MC level) in this analysis. The motivation of this section is to understand the behavior of hard processes (LO+NLL and NLO normalization) and fragmentation methods on important observables at generator and detector level simulations in CMS.

The jet multiplicity calculated at the ME element level for Sherpa has been limited to three jets for single vector bosons and one jet for the $t\bar{t}$ production due to limitations in the Sherpa ME. Higher jet multiplicity has been produced by state radiation in PS. MC@NLO calculates only 2→2/3 process at NLO. The same ME calculation has been applied in Sherpa when it is compared with MC@NLO. For Madgraph samples the official inclusive samples simulated by the CMS collaboration in Summer 2008 and Winter 2009 has been used [126]. For the Z+jets production the γ^* interference has been included with a cutoff of the di-muon invariant mass at the Z peak $M_{\mu\bar{\mu}} >$ 66 GeV and $M_{\mu\mu} > 50$ GeV for exclusive and inclusive samples, respectively.

For the simulations the default SM parameters have been used, see Table 5.2. The complete list of all used configurations is available in Appendix C.

The generators have their own free parameters, related to the non-perturbative QCD effects. First of all this is the renormalization scale μ_R, i.e. the scale where the initial value of the strong coupling constant is fixed, while the renormalization group equations provide the coupling constant at any other scale. The factorization scale, defined as μ_F, splits the cross section into the calculable short-distance and the universal long-distance functions. Those universal long-distance functions can be parametrized by a global fit to experimental data (see Section 4.1). In practice μ_R and μ_F are taken to be equal $\mu = \mu_R = \mu_F$ and the μ scale can be defined by the mass of the heaviest particle produced in hard interactions, like M_t or $M_{Z,W}$, or using the sum of the transverse masses (Σm_T) or momentum (Σp_T) of all particles. The scales can be set in the generators by the user. The default scales of generators are used in this analysis. In the study the default μ_F and μ_R scales have been varied for each generator by a factor two up and down; a typical test used in generator validations (see e.g. [123]). The different choices of the factorization scale can be varied independently of the renormalization scale and this has been done as well. Another parameter which affects the ME and PS matching is the jet threshold $P^{Jet}_{T,matching}$ where soft and hard jets are split. This threshold has been varied too. The summary of different scales and thresholds is presented in Table 5.3. The sensitivity to the PDFs has been investigated by considering LO(CTEQ5L-CTEQ6L1) and NLO (CTEQ6L) parametrization without any weighting method. The PDF sets via the LHAPDF interface are directly used in the generation.

Table 5.2: Default SM parameters used in this analysis.

Parameters	Value
M_W	80.419 GeV
M_Z	91.188 GeV
$M_b(M_b)$	4.8 GeV
$M_t(M_t)$	172.4 GeV
Γ_Z	2.49
Γ_W	2.06
$\alpha_s[M_Z]_{NLO}$	0.118[CTEQ6L]
$\alpha_s[M_Z]_{LO}$	0.127[CTEQ5L]
$\alpha_s[M_Z]_{LO}$	0.13[CTEQ6L1]

Sherpa also offers the possibility to enable or disable (default) spin correlation and to consider lepton production explicitly from bosons decays (default) or from all electroweak processes. In addition Sherpa includes Lund fragmentation/hadronization via the Pythia interface. These options have been tested internally as well.

Table 5.3: Summary of default matching thresholds and $\mu_{F,R}$ scales used in this analysis.

	P_T^{Jet}, GeV	μ_F and μ_R scales (default)	matching scheme
Alpgen	> 20	$M^2_{W,Z,t} + \Sigma m^2_T$	MLM
MC@NLO	> 20	$M^2_{Z,t}$	NLO
MadGraph	> 20	$M^2_{W,Z,t} + \Sigma m^2_T$	MLM
Sherpa	> 20	$M^2_{W,Z,t}$	CKKW

5.1.1 Analysis Framework and Monte-Carlo Data Samples

The simulation is based on CMSSW using standard interfaces to Alpgen, Madgraph, Sherpa and MC@NLO [127]. The simulation has been done with Full and Fast simulations of the CMS detector. In contrary to the full-simulations, the Fast simulation skips the time consuming simulation of the interaction of particles with the detecting material (GEANT4 - see Section 3.7) and digitization step. The simulation and reconstruction have been done in a standard way using the cmsDriver script and Physics Analysis Toolkit (PAT) methods for reconstructed objects [128, 129].

Table 5.4: Data samples used in this analysis. Vector bosons are forced to decay into muons.

Simulated samples	Cross section (pb)	Number of events	Generator settings
Alpgen-v2.13:			
W+jets	13384	500K	Default
	1236	1474K	Default
Z+jets	1230K	300K	$\mu = M_Z^2 + \Sigma m_T^2$
	1240	300K	$\mu = M_Z^2 + \Sigma p_T^2$
	1270	300K	$\mu = M_Z^2$
Sherpa-v1.1.3:			
W+jets	14180	2005	Default
	14010	610K	all electroweak
Z+jets	1260	2485K	CTE6L1
	1258	1000K	CTEQ5L
	1253	900K	Default
	1260	300K	Spin Correlation
	1260	500K	Lund Fragmentation
	1202	500K	Scales $\times 0.5$
	1307	500K	Scales $\times 2$
MC@NLO-v3.3:			
Z+jets	1266	600K	Default
W+jets	13756	1430K	Default

For the exclusive samples, jets have been reconstructed using the IterativeCone(IC) algorithm with the cone size $R = 0.5$. For the inclusive samples, jets have been reconstructed using the SisCone algorithm with the cone size $R = 0.5$. The reconstructed jet energy has been corrected using offset(L1), relative(L2) and absolute(L3) corrections (see Section 3.7.2).

Table 5.5: Data samples used in this analysis

Simulated samples	Cross section (pb)	Number of events	Generator settings
MadGraph-v4:			
$t\bar{t}$+jets	317	1000K	$P_{T,mlm}^{jets} \geq 30$ GeV
		660K	$P_{T,mlm}^{jets} \geq 10$ GeV
		730K	$P_{T,mlm}^{jets} \geq 30$ GeV
		552K	Scales Up (x2)
		630K	Scales Down (x1/2)
		935K	Larger ISR
		720K	Smaller ISR
W+jets	40000	5430K	$P_{T,mlm}^{jets} \geq 30$ GeV
		2900K	Scales Up (x2)
		6000K	Scales Down (x1/2)
Z+jets	3700	1010K	$P_{T,mlm}^{jets} \geq 30$ GeV
		5420K	$P_{T,mlm}^{jets} \geq 30$ GeV
		995K	Scales Down (x1/2)
		200K	Scales Up (x2)
ZInvisible+Jets	2000	1010K	Default
QCD+jets	15+E06	1000K	ht100to250
	400000	500K	ht250to500
	14000	400K	pt500to1000
	370	250K	pt1000toInfinity
A^*+Jets	6000	1200K	$P_{T,mlm}^{jets} \geq 30$
VV+Jets	20	900K	$P_{T,mlm}^{jets} \geq 30$

In addition to the calorimeter missing transverse energy E_{Calo}^{Miss}, the E_{Recoil}^{Miss} calculated as a recoil of selected reconstructed jets and muons has been used. The jets have been selected with $P_T^{jets} > 30$ GeV, $|\eta_j| < 3.0$ and the electrons and muons with $P_T^e > 15$ GeV and $P_T^\mu > 10$ GeV, $|\eta_{e,\mu}| < 3.0$. All reconstructed muons are required to be *global muons* (see Section 3.7.3 and Appendix A.1), i.e. they have to be made out of a tracker track and a muon chamber track. The E_T^{Miss} at MC level (E_{MC}^{Miss}) is calculated as a recoil of all stable tracks. In addition a E_T^{Miss} selection ($E_T^{Miss} > 50$ GeV) has been applied in order to avoid threshold and low bin effects.

The full list of used data samples is presented in Table 5.4[1], 5.6[2] and 5.5[2]. The details of used official data samples and configuration files can be found in Appendix C.

Table 5.6: Data samples used in this analysis

Simulated samples	Cross section (pb)	Number of events	Generator settings
Alpgen-v2.13:			
W+jets	43000	2000K	$P_{T,mlm}^{jets} \geq 30$ GeV
Z+jets	4000	1000K	$P_{T,mlm}^{jets} \geq 30$ GeV
$t\bar{t}$+jets	210	500K	$P_{T,mlm}^{jets} \geq 30$ GeV
Sherpa-v1.1.3:			
$t\bar{t}$+jets	204	500K	$P_{T,ckkw}^{jets} \geq 10$ GeV
	206	800K	$P_{T,ckkw}^{jets} \geq 30$ GeV
	206	800K	$P_{T,ckkw}^{jets} \geq 40$ GeV
	186	770K	Scales Up (x2)
	227	770K	Scales Down (x1/2)
	206	770K	Lund Fragmentation
W+jets	42000	5200K	Default
		4900K	$P_{T,ckkw}^{jets} \geq 30$ GeV
		4900K	1. Order EWeak Corr.
Z+jets	4000	900K	Default
		1000K	$P_{T,ckkw}^{jets} \geq 30$ GeV
		1420K	$P_{T,ckkw}^{jets} \geq 15$ GeV
		1000K	Scales Up (x2)
		1100K	Scales Down (x1/2)
ZInvisible+Jets	7000	3600K	Default
		1500K	Lund Fragmentation
		860K	Scales Up (x2)
		920K	Scales Down (x1/2)
QCD+jets	11.1+E06	1000K	pt40to120
	31145	500K	pt120to280
	181	400K	pt280to500
	5.2	250K	pt500to5000
A^*+Jets	6000	1200K	$P_{T,ckkw}^{jets} \geq 30$
VV+Jets	14.5	900K	$P_{T,ckkw}^{jets} \geq 30$
MC@NLO-v3.3:			
$t\bar{t}$+jets	366	460K	$P_{T,NLO}^{jets} \geq 30$

[1]Bosons in the Z+jets and W+jets production are forced to decay into muons. The underlying event and spin correlation are not taken into account.

[2]Bosons samples are forced to decay leptonically (e, μ, τ). The underlying events and spin correlation are taken into account.

5.2 Results for Generator Comparison

In the following section the cross sections, E_T^{Miss}, jet and lepton (muon) kinematics are compared for the produced data samples. For the comparison, the distributions from different samples have been normalized to unit area. The ratio plots shown in the figures present the bin by bin relative differences together with the range of statistical errors[2] which are demonstrated by the blue shaded areas. Since most of the P_T distributions are steeply decreasing, the normalization to unit area is equivalent to normalization to low P_T bins with highest statistics - this fact has to be kept in mind. Most distributions have been compared at MC and detector level in order to see the effect of CMS detector and physics object reconstruction. However, the instrumental systematic uncertainties, which have not been considered, can have an even larger effect [130, 131].

5.2.1 Cross Sections

The summary of cross section calculations for all studied channels is presented in Tables 5.7 and 5.8. The exclusive and inclusive cross sections are calculated in Alpgen and Sherpa at LO, the MC@NLO provides NLO cross sections.

Table 5.7: Summary of single vector boson Z($\mu\bar{\mu}$)+jets (< 4) and W($\mu\nu$)+jets (< 4) production cross sections (pb) for Alpgen and Sherpa. The sensitivity to scale variations are presented for the Sherpa event generator. The deviations for the inclusive cross section with respect to Sherpa default are shown.

processes	Alpgen-v2.13	Sherpa-v1.1.3	Sherpa ($\times$0.5)	Sherpa ($\times$2.0)
Z+0j	930	900	830	968
Z+1j	208	234	235	229
Z+2j	78	88.4	95.1	81
Z+3j	20	36	42	29
Σ_{Total}	1236 (1.78%)	1258.4	1202.1(4.2%)	1307(3.7%)
W+0j	10730	10339	9480	11040
W+1j	2264	2400	2377	2323
W+2j	690	1028	963	785
W+3j	200	413	390	255
Σ_{Total}	13384 (5.6%)	14180	13212(3.41%)	14400(1.52%)

[2]The statistical error is calculated:

$$(\frac{\sigma_R}{R})^2 = (\frac{\sigma_1}{N_1})^2 + (\frac{\sigma_2}{N_2})^2 \quad with \quad \sigma_{1,2} = \sqrt{N_{1,2}}$$

where the R is defined as the ratio between the generators and N is the number of events in the simulated samples.

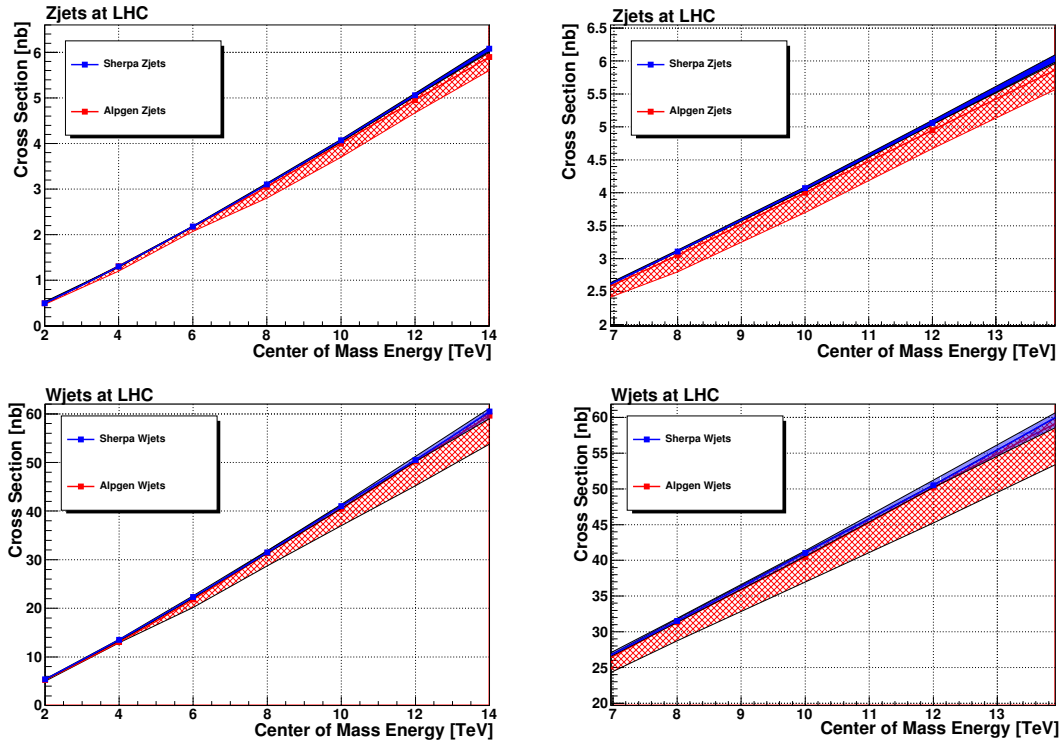


Figure 5.1: Total cross section for the $Z(l\bar{l}) + jets(< 4)$ (upper) and $W(l\nu) + jets(< 4)$ (lower) production versus center of mass energy. The light-shaded (red) area for Alpgen and dark-shaded (blue) for Sherpa show the sensitivity to different matching scales with $P^{Jet}_{T,matching}$ varying between 15, 20 and 30 GeV. Left: the cross section versus the center of mass energy between 2 and 14 TeV. Right: the cross section versus the center of mass energy between 7 to 14 TeV.

The sensitivity to scale variations is also shown for Sherpa as an example. In brackets the relative deviations in respect to the Sherpa inclusive cross section are presented. Differences in cross sections do not exceed 6% in case of Z($\mu\bar{\mu}$)+jets and W($\mu\nu$)+jets. For $t\bar{t}$+jets the difference reaches 11%. The cross section discrepancies between different generators tend to increase with the center of mass energy, as shown in Figure 5.1 for inclusive Z/W+jets. Sherpa usually gives larger cross section for higher jet multiplicities than Alpgen.

Table 5.8: Summary of $t\bar{t}$+jets ($N_j < 2$) cross section (pb) for Alpgen and Sherpa. $P^{T}_{jet,matching} >$ 30 GeV has been used for the simulation at parton level. The sensitivity to scale variations are presented for the Sherpa event generator. The deviations for inclusive cross section with respect to Sherpa default are shown.

process	Alpgen-v2.13	Sherpa-v1.1.3	Sherpa (×0.5)	Sherpa (×2.0)
$t\bar{t}$+0j	118	124	114.5	95.1
$t\bar{t}$+1j	69	80	115	82.6
Σ_{Total}	187(8.6%)	204	230.4(10.8%)	181(11%)

The cross section variation due to scale changes is within the range of differences between generators. However, the cross section of the $t\bar{t}$+jets production is very sensitive to the choice of factorization scale. It is crucial to apply the correct factorization scale for the comparison between the generators. The small sensitivity to the matching threshold in Sherpa compared to Alpgen is presumably due to the different treatment of the survival probability of soft gluons. The cross section presented in Table 5.8 are calculated at LO. The cross section of $t\bar{t}$ production for the Tevatron and LHC has been studied in detail [132].

5.2.2 Missing Transverse Energy

The generator level and detector level missing transverse energy E_{MC}^{Miss}, E_{Calo}^{Miss} and E_{Recoil}^{Miss} are studied in this section for different channels and samples.

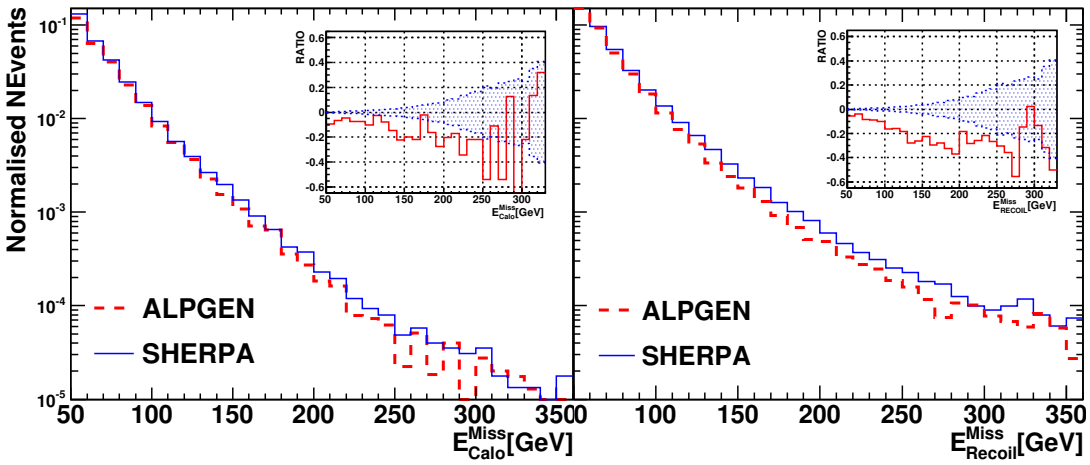


Figure 5.2: Detector level $E_{Calo}^{Miss} > 50$ GeV (left) and $E_{Recoil}^{Miss} > 50$ GeV (right) distributions for the exclusive $Z(\mu\bar{\mu})$+jets samples simulated with Alpgen and Sherpa. Ratio: relative difference of Alpgen with respect to Sherpa, (Alpgen/Sherpa) - 1. The (blue) shaded area represents the statistical uncertainty.

Figure 5.2 shows the difference between reconstructed E_{Calo}^{Miss} and E_{Recoil}^{Miss} for the exclusive Z($\mu\bar{\mu}$)+jets production. The missing transverse energy can be produced by calorimeter mis-calibration and noise for E_{Calo}^{Miss}, or jet and lepton mis-reconstruction for E_{Recoil}^{Miss}. A large difference between Alpgen and Sherpa, up to 20% to 30%, is observed for both distributions above 100 GeV.

In Figure 5.3, the ratio of intrinsic generator systematics (up and lower-left) and the comparison between them (lower-right) are presented for Alpgen and Sherpa. The comparison of various scale choices used to assess the systematic uncertainty for E_{Calo}^{Miss} and E_{Recoil}^{Miss} are shown after the detector simulation. The scale choices presented in Table 5.3 are used to predict the systematic uncertainties for Alpgen and Sherpa.

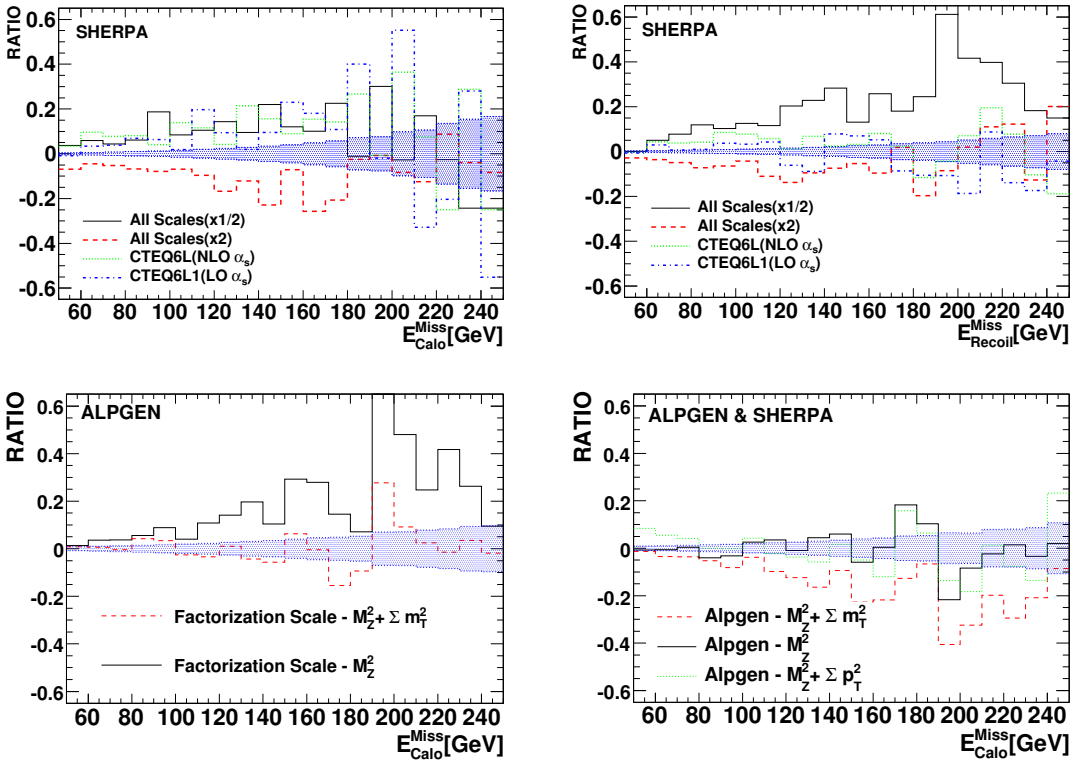


Figure 5.3: Upper: the relative differences of various Sherpa settings with respect to Sherpa default of $E^{Miss}_{Calo} > 50$ GeV (left) and $E^{Miss}_{Recoil} > 50$ GeV (right) distributions for the exclusive $Z(\mu\bar{\mu})$+jets production. Lower-left: the relative differences of various Alpgen factorization scales with respect to Alpgen default of $E^{Miss}_{Calo} > 50$ GeV. Lower-right: the relative differences of various Alpgen factorization scales with respect to Sherpa default of $E^{Miss}_{Calo} > 50$ GeV.

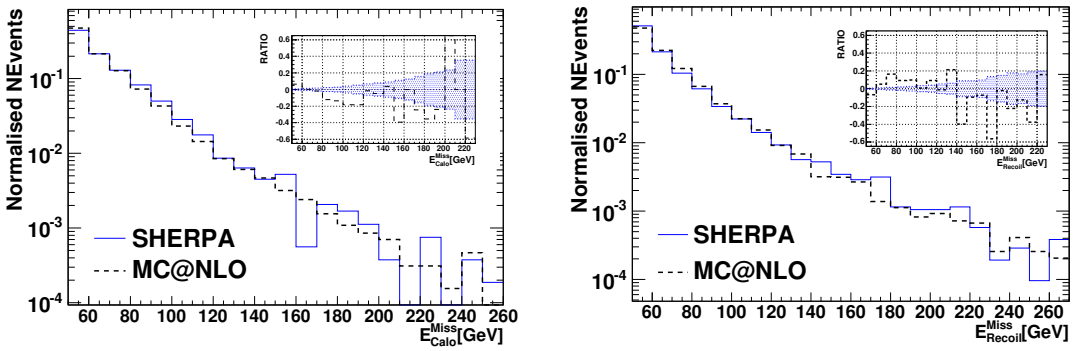


Figure 5.4: Detector level $E^{Miss}_{Calo} > 50$ GeV (left) and $E^{Miss}_{Recoil} > 50$ GeV (right) distributions for the exclusive $Z(\mu\bar{\mu})$+1jet samples simulated with MC@NLO and Sherpa. Ratio: relative difference of MC@NLO with respect to Sherpa. The (blue) shaded area represents the statistical uncertainty.

The largest deviation between Sherpa predictions, up to 20-25%, is observed for E^{Miss}_{Recoil} in Figure 5.3. For Sherpa, higher scales show a softer spectrum while lower scales show a harder spectrum. The largest differences between Alpgen factorization scales , up to 15-20%, appears as the differences between M^2_Z or p^2_T and $M^2_Z + \Sigma p^2_T$ or $M^2_Z + \Sigma m^2_T$. The factorization scales using M^2_Z or p^2_T tend to show a harder spectrum than $M^2_Z + \Sigma p^2_T$ and $M^2_Z + \Sigma m^2_T$ scale choices since Sherpa shows harder spectrum than Alpgen as well. The larger scale implies a smaller strong coupling constant (α_s) which leads to a softer spectrum in the tail of the distribution. Therefore, it is crucial to understand the scale variations of the generators at LHC energies since the tail of the distributions is mainly used for beyond the SM searches.

In Figure 5.4, the comparison of Sherpa and MC@NLO for Z($\mu\bar{\mu}$)+1jet is presented for E^{Miss}_{Calo} and E^{Miss}_{Recoil}. The comparison between LO-ME and NLO-ME shows a good agreement with deviations around 10%. Due to the ME limitation in MC@NLO, the distributions are compared with an extra parton in the ME.

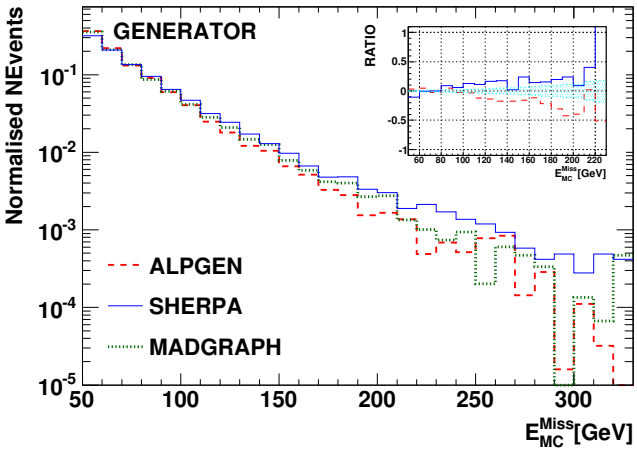


Figure 5.5: Generator level $E^{Miss}_{MC} > 50$ GeV distribution for the inclusive Z($l\bar{l}$)+jets ($l = e, \mu, \tau$) production simulated with Alpgen, Sherpa and Madgraph. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

In Figure 5.5 and 5.6 E^{Miss}_{MC} and E^{Miss}_{Calo} are compared for the inclusive Z($l\bar{l}$)+jets production using Alpgen, Sherpa and Madgraph. The comparison between the generators shows reasonable agreement with 10-15% differences at MC level. Alpgen shows a slightly softer spectrum at lower energies after reconstruction. Sherpa shows a harder spectrum starting at $E^{Miss}_{Calo} > 80$ GeV at detector level including different hard scales and a fragmentation function from Alpgen and Madgraph. In Figure 5.7 and 5.8, the internal systematics and the reconstruction validation for Madgraph are presented for the Z($l\bar{l}$)+jets production at MC and detector level. The validation of fast and full simulated samples shows a good agreement for the detector simulation. The scale variations of Madgraph simulations show 20-30% deviation from the default scale. They behave like the Sherpa deviation in Figure 5.8.

The Z($\nu\bar{\nu}$)+jets channel is an important background for the BSM searches because of

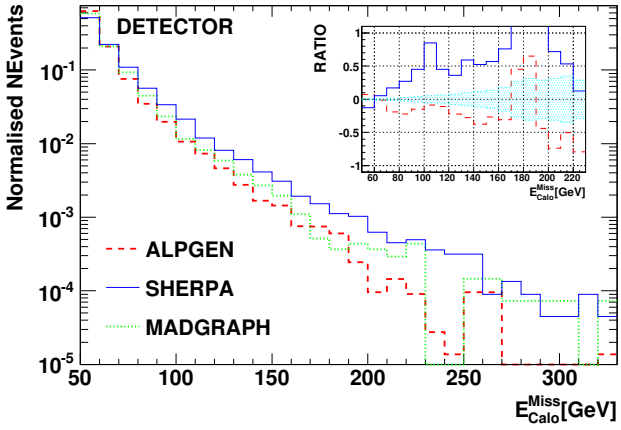


Figure 5.6: Detector level $E_{Calo}^{Miss} > 50$ GeV distribution for the inclusive Z($l\bar{l}$)+jets ($l = e, \mu, \tau$) production simulated with Alpgen, Sherpa and Madgraph. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

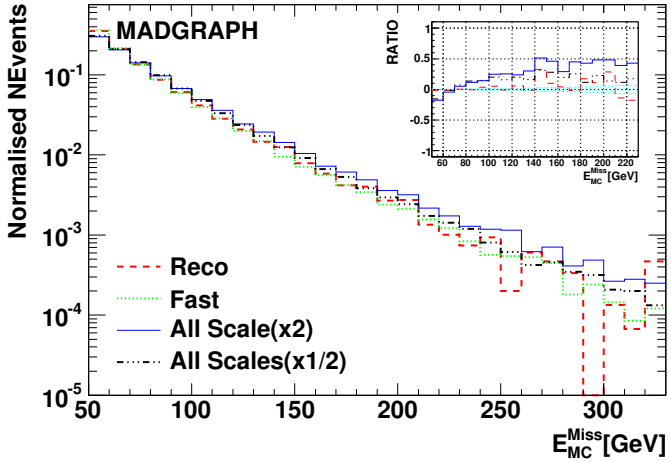


Figure 5.7: Generator level $E_{Calo}^{Miss} > 50$ GeV distribution for the inclusive Z($l\bar{l}$)+jets ($l = e, \mu, \tau$) production simulated with various Madgraph settings. Ratio: relative differences of various Madgraph settings with respect to Madgraph default (fast).

the large E_T^{Miss} signature. The E_{MC}^{Miss} for Z($\nu\bar{\nu}$)+jets at generator level is presented in Figure 5.9 for Alpgen-Sherpa (left) and Sherpa-Madgraph (right). The left plot has been produced with Z($\mu\bar{\mu}$)+jets where muons have been replaced by neutrinos and E_T^{Miss} is recalculated. This is one of the most common ways to determine the irreducible background via detector simulation since the corresponding distributions for Z($\mu\bar{\mu}$)+jets can be rescaled according to the irreducible background cross-section calculated at MC level. Both generators are in fair agreement above 100 GeV. Sherpa shows a harder spectrum than Alpgen with 15-20% deviation because of the default scale choices. E_{MC}^{Miss} for inclusive Z($\nu\bar{\nu}$)+jets production is presented in Figure 5.9

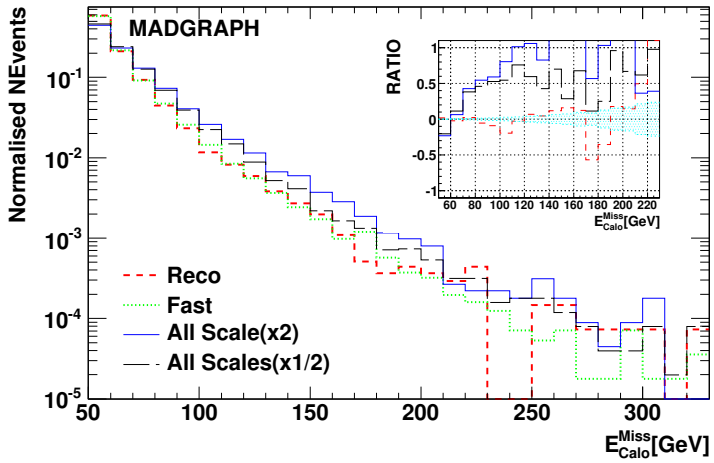


Figure 5.8: Detector level $E^{Miss}_{Calo} > 50$ GeV distribution for the inclusive Z($l\bar{l}$)+jets ($l = e, \mu, \tau$) production simulated with various Madgraph settings. Ratio: relative differences of various Madgraph settings with respect to Madgraph default (fast).

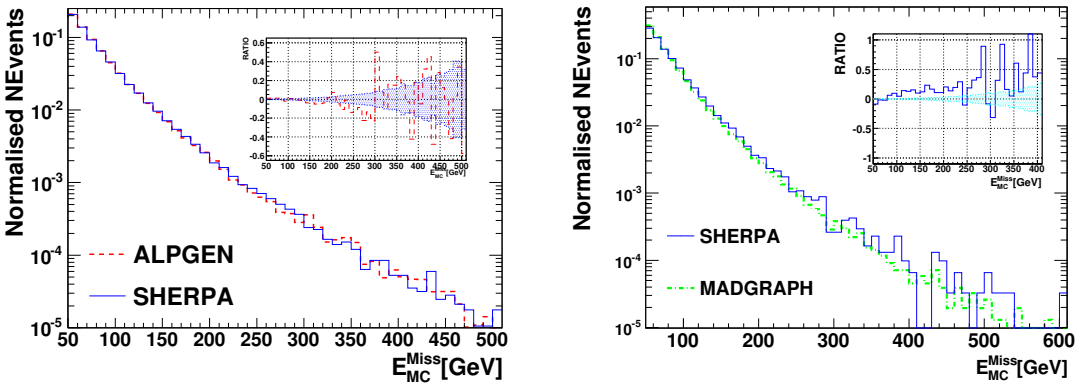


Figure 5.9: Generator level $E^{Miss}_{MC} > 50$ GeV for $Z^0(\nu\bar{\nu})$+jets (left) predicted via the exclusive $Z^0(\mu\bar{\mu})$+jets production for Alpgen and Sherpa samples. Ratio (left): relative difference of Alpgen with respect to Sherpa. Detector level $E^{Miss}_{Calo} > 50$ GeV distribution (right) for direct Z($\nu\bar{\nu}$)+jets production simulated with Sherpa and Madgraph. Ratio(right): relative difference of Sherpa with respect to Madgraph, (Sherpa/Madgraph) - 1.

for Madgraph and Sherpa (right). In this case the Z($\nu\bar{\nu}$)+jets (ν=ν_e, ν_μ, ν_τ) channel has been simulated directly. Both cases show the same differences on both level since there is no big effect on calorimeter missing transverse energy from hadronization effects. In Figure 5.10 the detector level comparison for the irreducible background is presented for Madgraph and Sherpa. Sherpa shows a harder spectra with 20% systematics observed at detector level.

The sensitivity of E^{Miss}_{MC} to different generator settings for the Z($\nu\bar{\nu}$)+jets production predicted via exclusive $Z^0(\mu\bar{\mu})$+jets production is shown in Figure 5.11. The scale variation show a small effect in Sherpa.

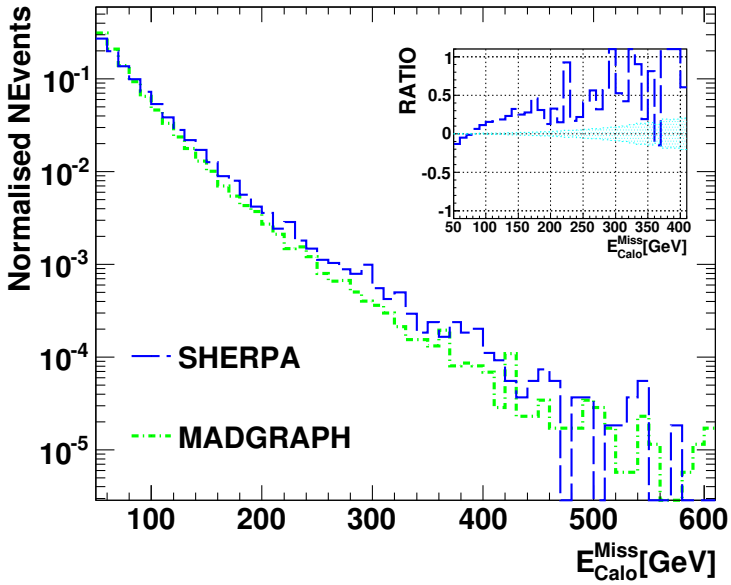


Figure 5.10: Detector level $E^{Miss}_{Calo} > 50$ GeV distribution for direct Z($\nu\bar{\nu}$)+jets production simulated with Sherpa and Madgraph. Ratio: relative difference of Sherpa with respect to Madgraph.

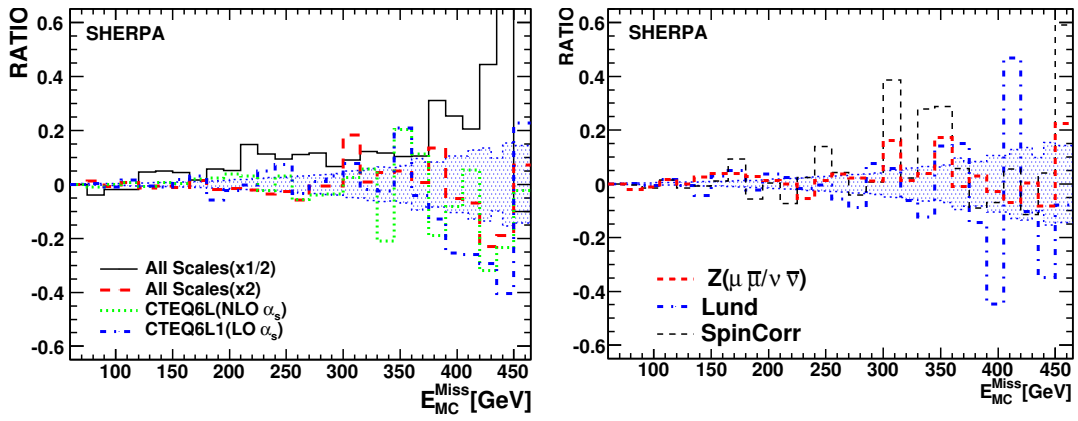


Figure 5.11: The relative differences of various Sherpa settings with respect to Sherpa default of $E^{Miss}_{MC} > 50$ GeV distribution for $Z^0(\nu\bar{\nu})$+jets production predicted via the exclusive $Z^0(\mu\bar{\mu})$+jets production. Systematics are taken into account for the exclusive Sherpa samples using different settings: scale variations, Lund fragmentation model, PDFs, include spin correlations. Ratio: relative differences of Sherpa systematics with respect to Sherpa default.

The Lund (Pythia), full spin correlation and clustered fragmentation models implemented in Sherpa show almost no difference at generator level. Also different PDFs have little influence on E^{Miss}_{MC}. The E^{Miss}_{MC} distribution obtained from the Z($\mu\bar{\mu}$)+jets production replacing muons with neutrinos is validated with the one from direct Z($\nu\bar{\nu}$)+jets production. They show a very good agreement.

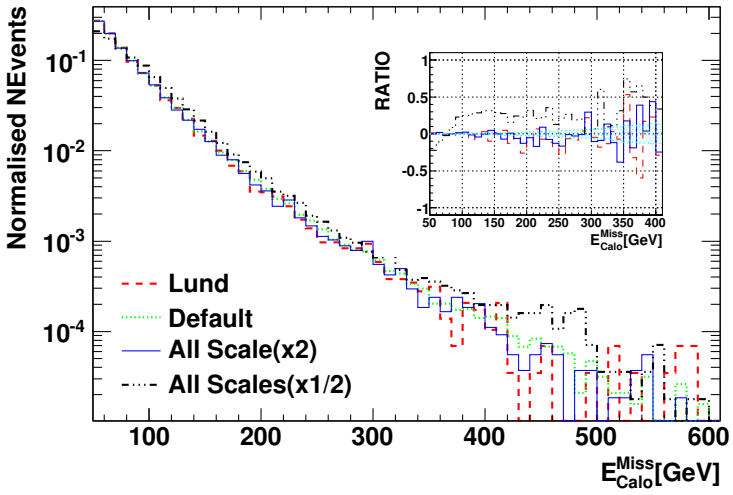


Figure 5.12: The relative differences of various Sherpa settings with respect to Sherpa default of $E_{Calo}^{Miss} > 50$ GeV distribution for direct $Z^0(\nu\bar{\nu})$+jets production. Systematics are taken into account for the exclusive Sherpa samples using different settings: scale variations, Lund fragmentation model, PDFs, include spin correlations. Ratio: relative differences of Sherpa systematics with respect to Sherpa default.

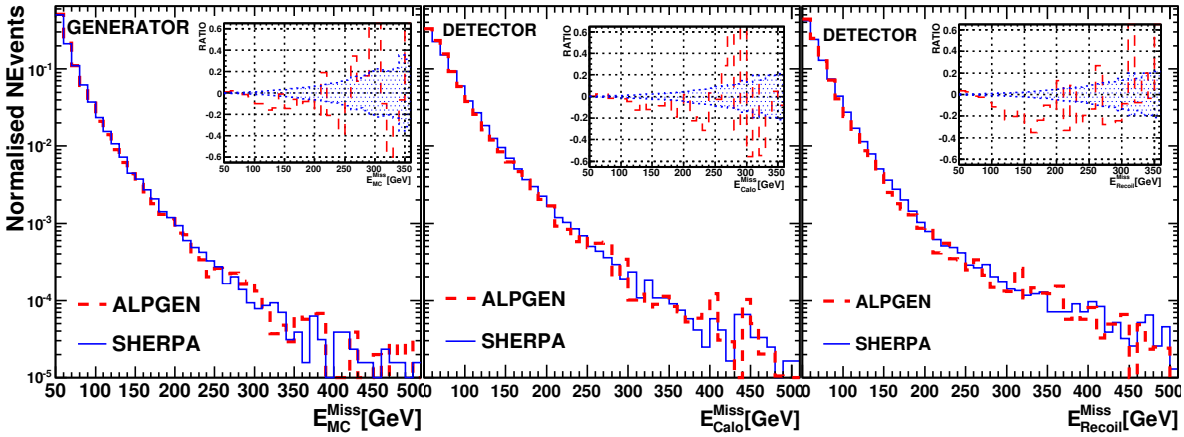


Figure 5.13: The generator level $E_{MC}^{Miss} > 50$ GeV (left) and detector level $E_{Calo}^{Miss} > 50$ GeV and $E_{Recoil}^{Miss} > 50$ GeV (middle and right) distributions for the exclusive $W(\mu\nu)$+jets production simulated with Alpgen and Sherpa. Ratio: relative difference of Alpgen with respect to Sherpa, (Alpgen/Sherpa) - 1. The (blue) shaded area represents the statistical uncertainty.

The generator systematics of the reconstructed E_{Calo}^{Miss} for inclusive Z($\nu\bar{\nu}$)+jets are shown in Figure 5.12 for Sherpa. The scale variation (x1/2) causes 20% deviation from the default setting. The other systematics show almost negligible effects.

The comparison of E_{MC}^{Miss} and E_{Calo}^{Miss}, E_{Recoil}^{Miss} for the exclusive W($\mu\nu$)+jets production is given in Figure 5.13. The difference between Alpgen and Sherpa shows up already at MC level and has a similar range of 10% to 20% above 100 GeV after detector simulation. E_{Calo}^{Miss} and E_{Recoil}^{Miss} behave similar although the difference is slightly larger for the E_{Recoil}^{Miss} due to different jet P_T spectra.

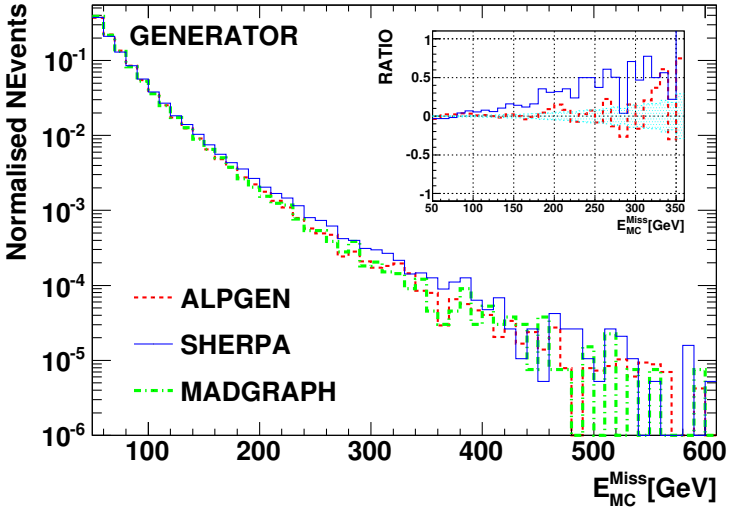


Figure 5.14: Generator level $E_{MC}^{Miss} > 50$ GeV distribution for the inclusive W($l\nu$)+jets ($l = e, \mu, \tau$) production simulated with Alpgen, Sherpa and Madgraph. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

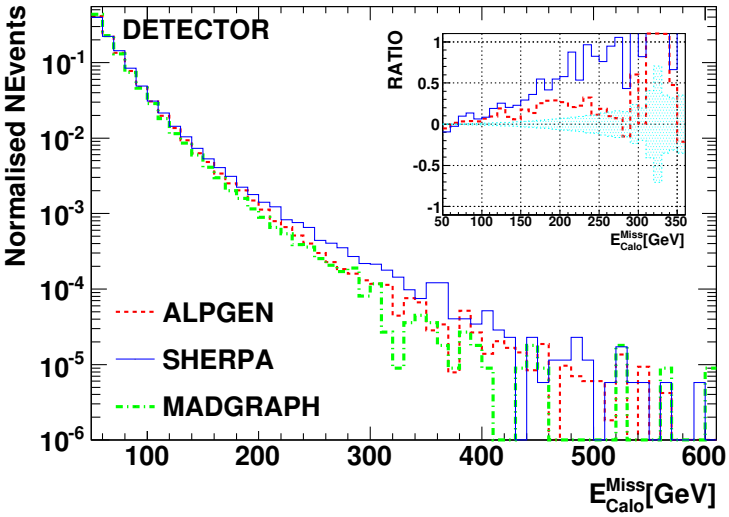


Figure 5.15: Detector level $E_{Calo}^{Miss} > 50$ GeV distribution for the inclusive W($l\nu$)+jets ($l = e, \mu, \tau$) production simulated with Alpgen, Sherpa and Madgraph. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

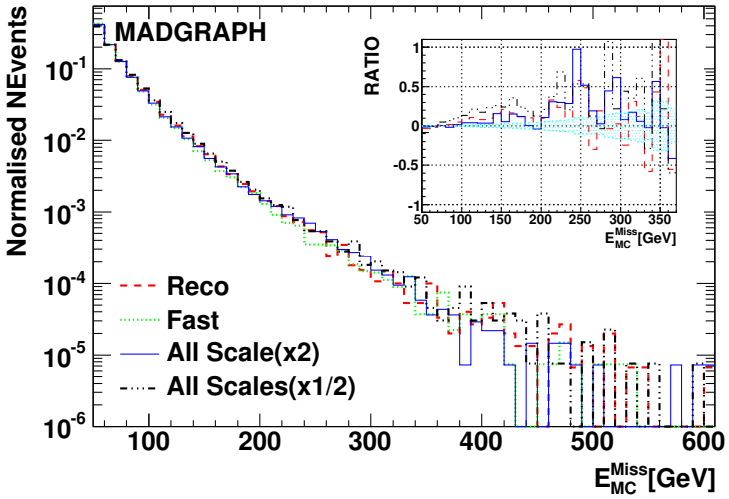


Figure 5.16: Generator level $E^{Miss}_{MC} > 50$ GeV distribution for the inclusive W($l\nu$)+jets ($l = e, \mu, \tau$) production simulated with various Madgraph settings. Ratio: relative differences of various Madgraph settings with respect to Madgraph default (fast).

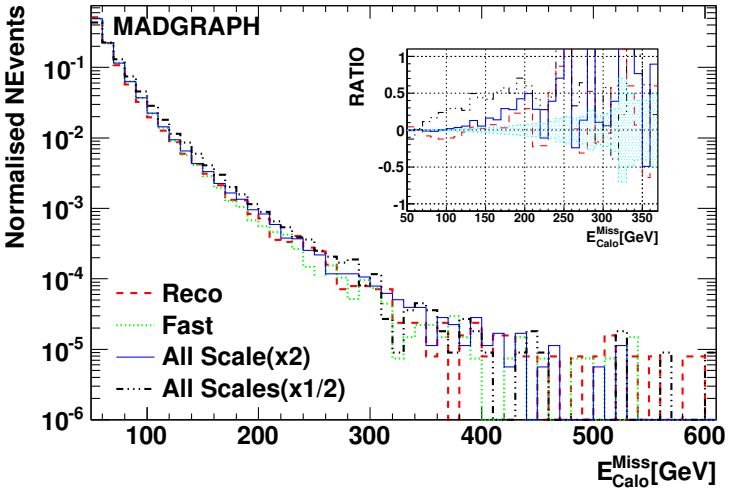


Figure 5.17: Detector level $E^{Miss}_{Calo} > 50$ GeV distribution for the inclusive W($l\nu$)+jets ($l = e, \mu, \tau$) production simulated with various Madgraph settings. Ratio: relative differences of various Madgraph settings with respect to Madgraph default (fast).

In Figure 5.14 and 5.15, the comparison of generated and reconstructed inclusive W($l\nu$)+jets production is presented for Alpgen, Sherpa and Madgraph. Alpgen has a slightly harder spectra in the initial case but it shows generally good agreement. However, Sherpa has a significantly harder spectrum than Alpgen and Madgraph. This deviation, which is clearly visible at MC level, is caused by the renormalization scale choices of the Sherpa generator since the scale dependence of the Madgraph

generator in Figure 5.16 and 5.17 shows comparable systematics such as Sherpa prediction at MC and detector level. In addition to that, Sherpa has a different fragmentation method and hadron decay structure which slightly affect the final spectra of the W($l\nu$)+jets production after detector reconstruction.

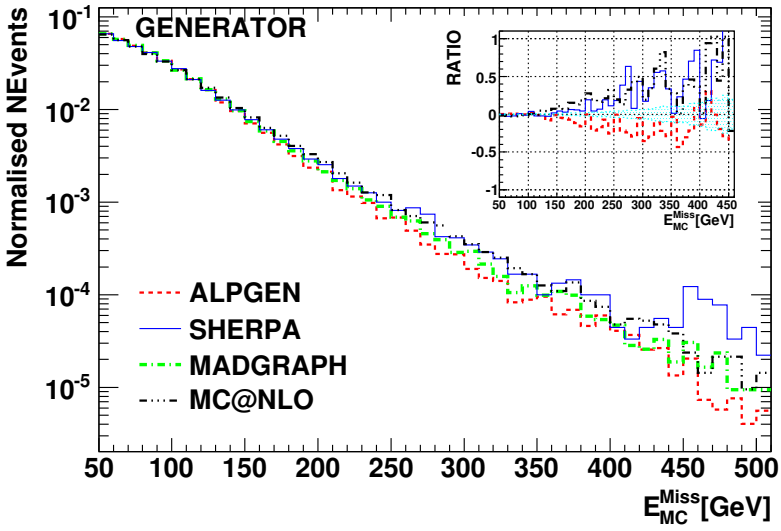


Figure 5.18: Generator level $E_{MC}^{Miss} > 50$ GeV distribution for the inclusive $t\bar{t}$+jets production simulated with Alpgen, Sherpa, MC@NLO and Madgraph. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph. The (blue) shaded area represents the statistical uncertainty.

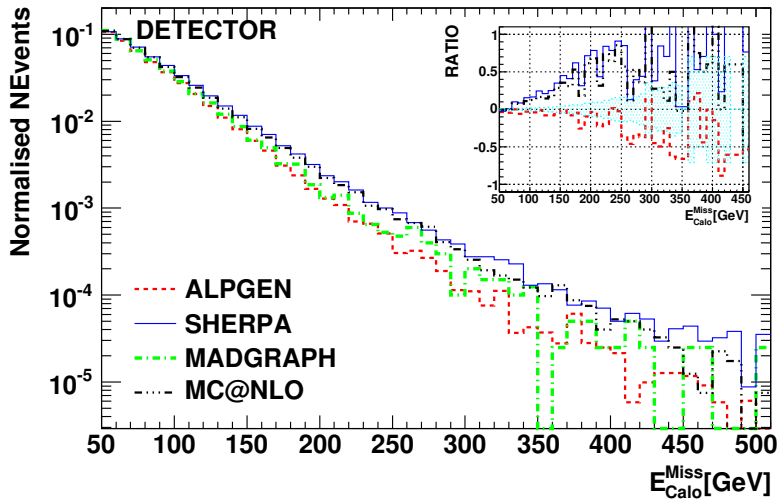


Figure 5.19: Detector level $E_{Calo}^{Miss} > 50$ GeV distribution for the inclusive $t\bar{t}$+jets production simulated with Alpgen, Sherpa, MC@NLO and Madgraph. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

E_T^{Miss} for the $t\bar{t}$+jets production is shown in Figure 5.18 and 5.19 at generator and detector level. Four generators are compared: Alpgen, Sherpa, Madgraph and MC@NLO. Sherpa and MC@NLO have been simulated with one additional extra jet due to limitation at ME.

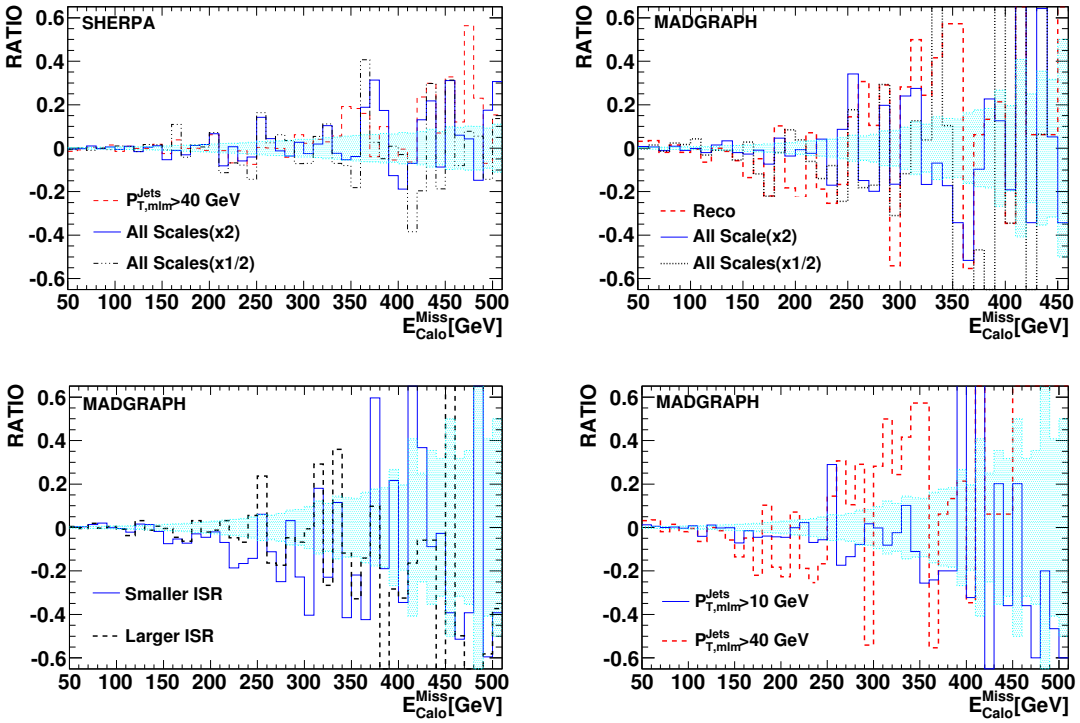


Figure 5.20: The relative differences of detector level $E_{Calo}^{Miss} > 50$ GeV distributions for the $t\bar{t}$+jets production simulated with various Sherpa and Madgraph settings. Upper-left: the ratio of relative differences of various Sherpa settings with respect to Sherpa default. Upper-left and lower: the relative differences of various Madgraph settings with respect to Madgraph default (fast). The (blue) shaded area represents the statistical uncertainty.

At generator and detector level Alpgen and Madgraph are in good agreement in the whole range, while Sherpa and MC@NLO deviate substantially in the slope reaching 20% above 100 GeV at detector level. The scale choice of higher order corrections and different implementation of the Sudakov rescaling from ME to PS implementation are possibly the reason of this shape difference between Sherpa-MC@NLO and Alpgen-Madgraph simulations. After detector simulation the differences are smeared, but the behavior is still clearly visible. Note that Sherpa and MC@NLO use the cluster hadronization approach while Alpgen and Pythia use the Lund scheme. The sensitivity to different Sherpa and Madgraph settings for $t\bar{t}$+jets production is presented in Figure 5.20. The differences on the matching, factorization/renormalization scales and the ISR and FSR have been studied for the $t\bar{t}$+jets production. For Sherpa and Madgraph, the E_{Calo}^{Miss} distributions is stable in the whole range. However, for Madgraph differences about 20% have been observed for smaller ISR and different matching scales.

E_T^{Miss} for the inclusive VV+jets and A^*+jets production are shown in Figure 5.21 and 5.22, respectively. Two generators are compared: Sherpa and Madgraph. At generator and detector level for VV+jets production, Madgraph shows a harder spectrum than Sherpa due to the higher jet-multiplicity at ME, while the Sherpa deviates substantially softer in the slope reaching 20% above 200 GeV. The differences in the lower region arises from the scale differences between the generators. The comparison of A^*+jets between Sherpa and Madgraph shows a difference starting from 70 GeV. Sherpa shows a harder spectrum at higher energy. The implementation of the γ^* interference to the ME and scale differences can be the reason for this differences. The sensitivity to scale variations for VV+jets and A^*+jets are expected to be similar to the Z/W+jets production.

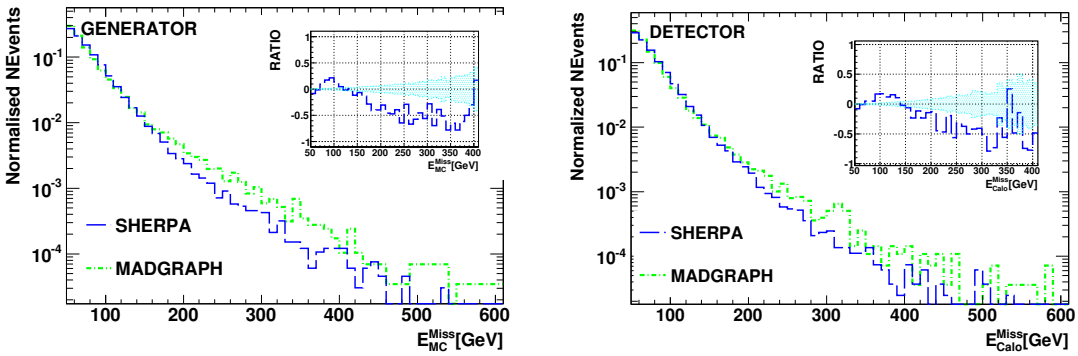


Figure 5.21: Generator level $E_{MC}^{Miss} > 50$ (left) and detector level $E_{Calo}^{Miss} > 50$ GeV (right) distributions for the inclusive VV+jets production simulated with Sherpa and Madgraph. Ratio: relative difference of Sherpa with respect to Madgraph, (Sherpa/Madgraph) - 1. The (blue) shaded area represents the statistical uncertainty.

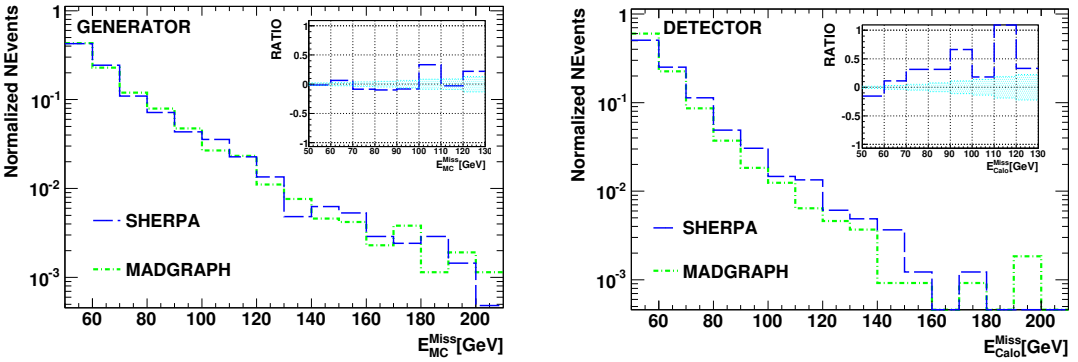


Figure 5.22: Generator level $E_{MC}^{Miss} > 50$ (left) and detector level $E_{Calo}^{Miss} > 50$ GeV (right) distributions for the inclusive A^*+jets production simulated with Sherpa and Madgraph. Ratio: relative difference of Sherpa with respect to Madgraph.

In summary, the largest deviation for the exclusive samples is observed between Sherpa and Alpgen for channels with intrinsic E_T^{Miss} like the W($\mu\nu$)+jets production. In addition the E_{Recoil}^{Miss} distribution shows a difference for the Z($\mu\bar{\mu}$)+jets production. The variation increases with E_T and can reach 20% above 200 GeV. One of the reason is that the default tune of factorization scales is different between Alpgen and Sherpa. Additionally Sherpa and Alpgen use different parton shower and hadronization modules for the final events which affect the distributions of E_{Calo}^{Miss} and E_{Recoil}^{Miss}. However, there is good agreement between Sherpa and MC@NLO for Z($\mu\bar{\mu}$)+1jet samples. Finally the influence of parton shower and hadron decay differences are visible in the range of 15-20% for E_{Calo}^{Miss} and E_{Recoil}^{Miss}. The influence of the renormalization/fragmentation scales, matching thresholds and PDF has a reasonable range within 10% to 20%.

The deviation observed in the exclusive samples for the single vector boson (Z($\mu\bar{\mu}$)+jets and W($\mu\nu$)+jets) increases for inclusive samples (Z($l\bar{l}$)+jets and W($l\nu$)+jets where $l = e, \mu, \tau$). The variation increases with E_T and reaches significant deviations above 100 GeV for the single and double-vector bosons production. Zinvisible+jets samples are generally in a good agreement between generators. However, a deviation of 20% in the tails of the E_T^{Miss} distribution arises from the choice of the factorization/renormalization scales. Also internal systematics of Sherpa with respect to Sherpa default shows a similar behavior.

The comparison of inclusive t$\bar{\text{t}}$+jets samples has shown differences between Sherpa-MC@NLO and Alpgen-Madgraph approaches. Alpgen and Madgraph agree with less than 10% for the whole spectra, also Sherpa and MC@NLO agree well. Sherpa and MC@NLO methods show a good agreement since the implementation of the Sudakov mapping from ME to PS for Sherpa gives a good agreement with NLO normalization for t$\bar{\text{t}}$+jets production. Madgraph and Alpgen show a very good agreement for the prediction with multi-jet at ME. They both have similar showering and matching methods. The effects of systematics from all generators are in the range of 10%-20% for the inclusive samples.

Differences on missing transverse energy distribution arise from various the MC estimation methods, which are implemented in the event generators (see Table 5.1), from ME+PS matching to hadronization schemes. The shape of the distributions have been determined by the ME and PS matching. The CKKW and MLM methods show differences between Sherpa and Alpgen-Madgraph. In addition the fact that hadronization schemes, Lund and Cluster, show the important differences between generator and detector level distributions on missing transverse energy. Those effects are clearly visible for the Z+jets, W+jets and $t\bar{t}$+jets production which are important for SUSY searches.

5.2.3 Jet Observables

In this analysis the jet observables have been studied after detector simulation and reconstruction. Remember that the exclusive samples for Alpgen, Sherpa and MC@NLO are calculated for Z($\mu\bar{\mu}$)/W($\mu\nu$)+jets. The transverse momentum (P_T) and pseudo-rapidity (η) distributions of the first three leading jets are shown in Figure 5.23 and 5.24. While the kinematics of the first leading jet in the Z+jets samples agrees well for Sherpa and Alpgen, the second and the third jet show a significant difference. Sherpa produces harder and less central jets compared to Alpgen. The difference in the P_T distribution of the sub-leading jets reaches up to 30% after $P_T > 100$ GeV. Similar behavior is observed for W+jets production where the disagreement is also in the P_T distribution of the leading jet.

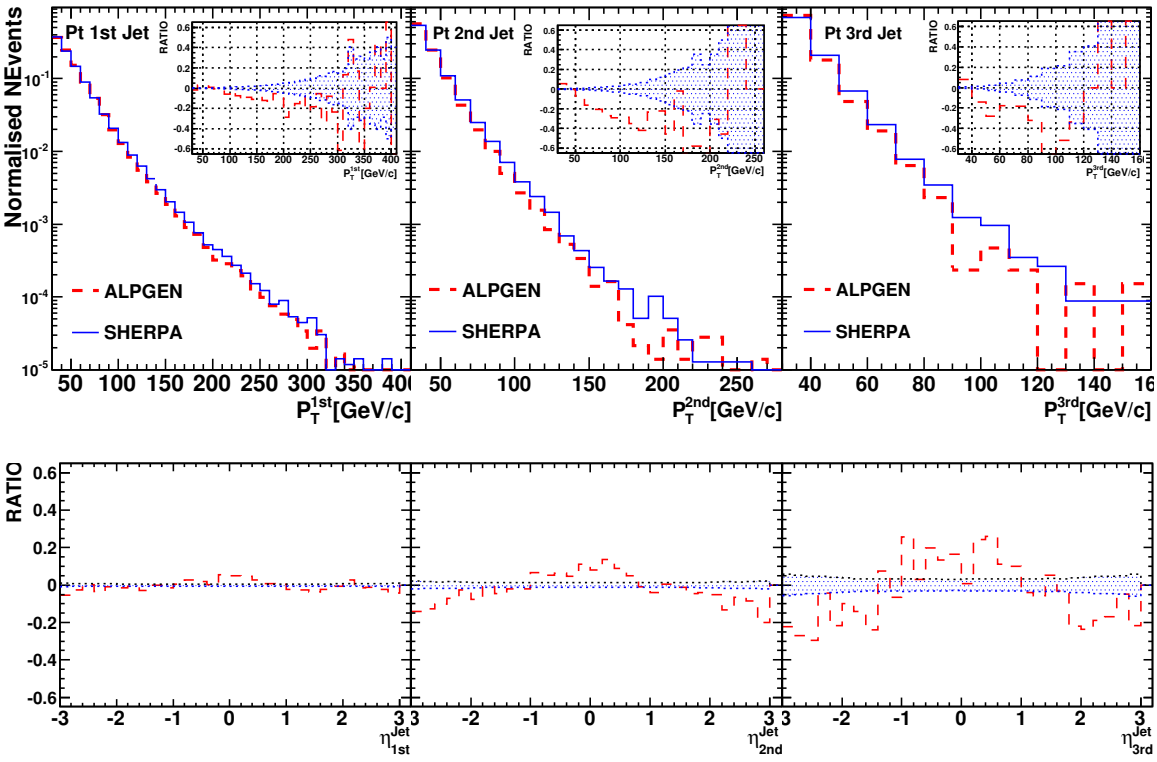


Figure 5.23: Upper: P_T distribution of the three leading jets for exclusive Sherpa and Alpgen $Z(\mu\bar{\mu})$+jets samples at detector level. Ratio: relative difference of Alpgen with respect to Sherpa. Lower: the same ratio for the η spectra for the leading jets. The (blue) shaded area represents the statistical uncertainty.

The prediction of MC@NLO and Sherpa for Z+jets(0,1) agree very well for the shape of the leading jet P_T and η, as shown in Figure 5.25. The fact that Sherpa and MC@NLO agree well, shows that the matching in Sherpa works correctly for NLO normalization at low jet multiplicities. The application of the Sudakov rescaling between ME and PS information are similar for both methods.

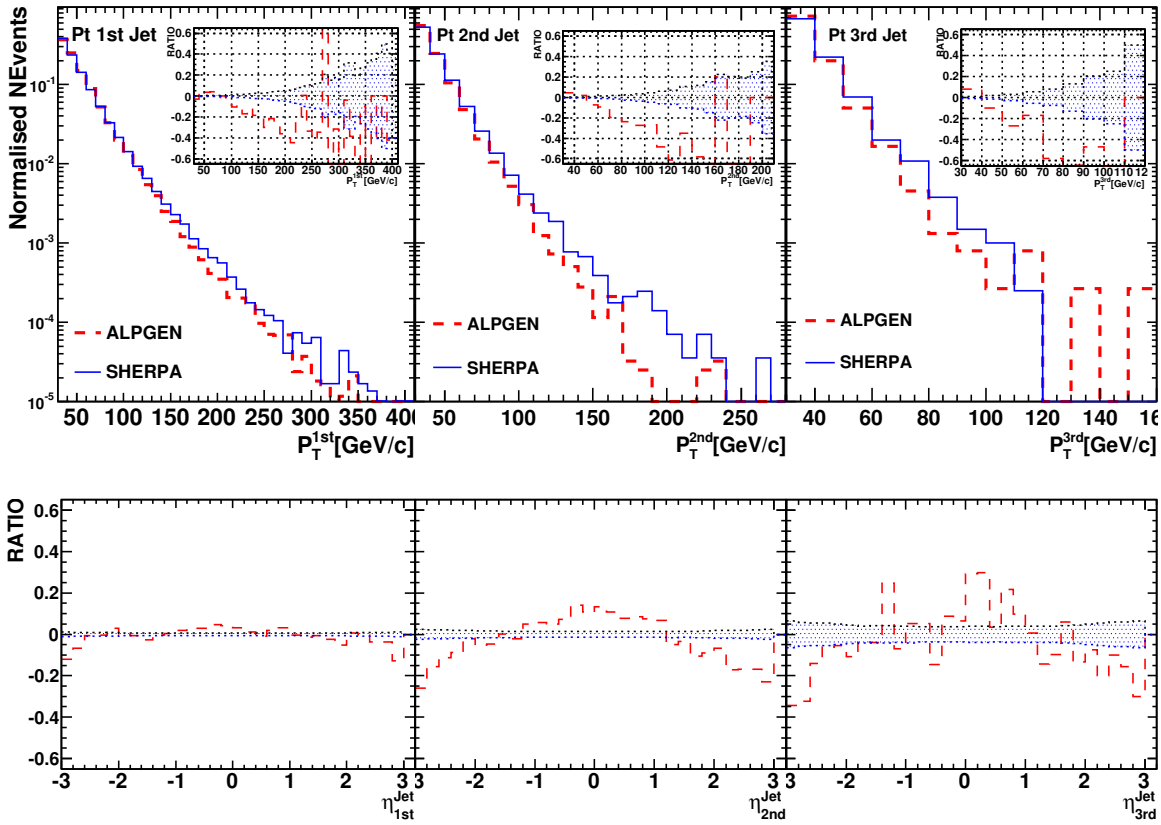


Figure 5.24: Upper: P_T distribution of the three leading jets for exclusive Sherpa and Alpgen $W(\mu\bar{\nu})$+jets samples at detector level. Ratio: relative difference of Alpgen with respect to Sherpa. Lower: the same ratio for *eta* spectra for the leading jets.

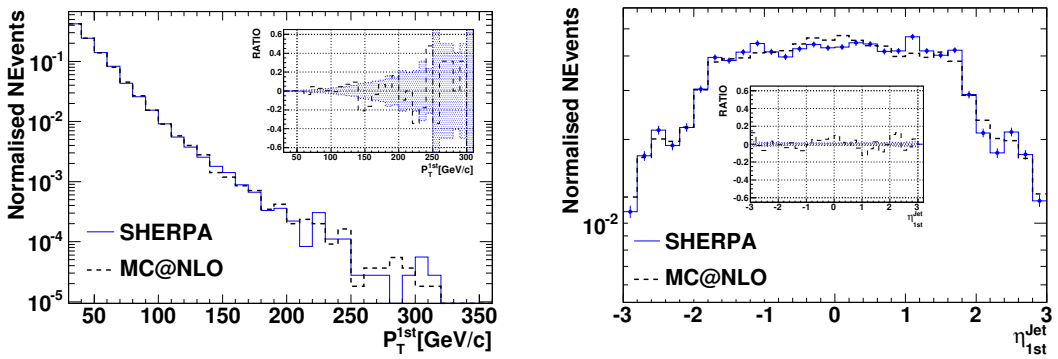


Figure 5.25: P_T (right) and η (left) distributions of the leading jet for exclusive Sherpa and MC@NLO $Z(\mu\bar{\mu})$+1jet samples at detector level. Ratio: relative difference of MC@NLO with respect to Sherpa.

For the inclusive samples of the Z($l\bar{l}$)+jets and W($l\bar{\nu}$)+jets production ($l = e, \mu, \tau$), the transverse momentum (P_T) and pseudo-rapidity (η) distribution of the three leading jets are presented in Figure 5.26. The differences between Alpgen and Sherpa for exclusive and inclusive samples are of the same order. Alpgen usually shows softer jet P_T spectra and central pseudo-rapidity distribution than Sherpa.

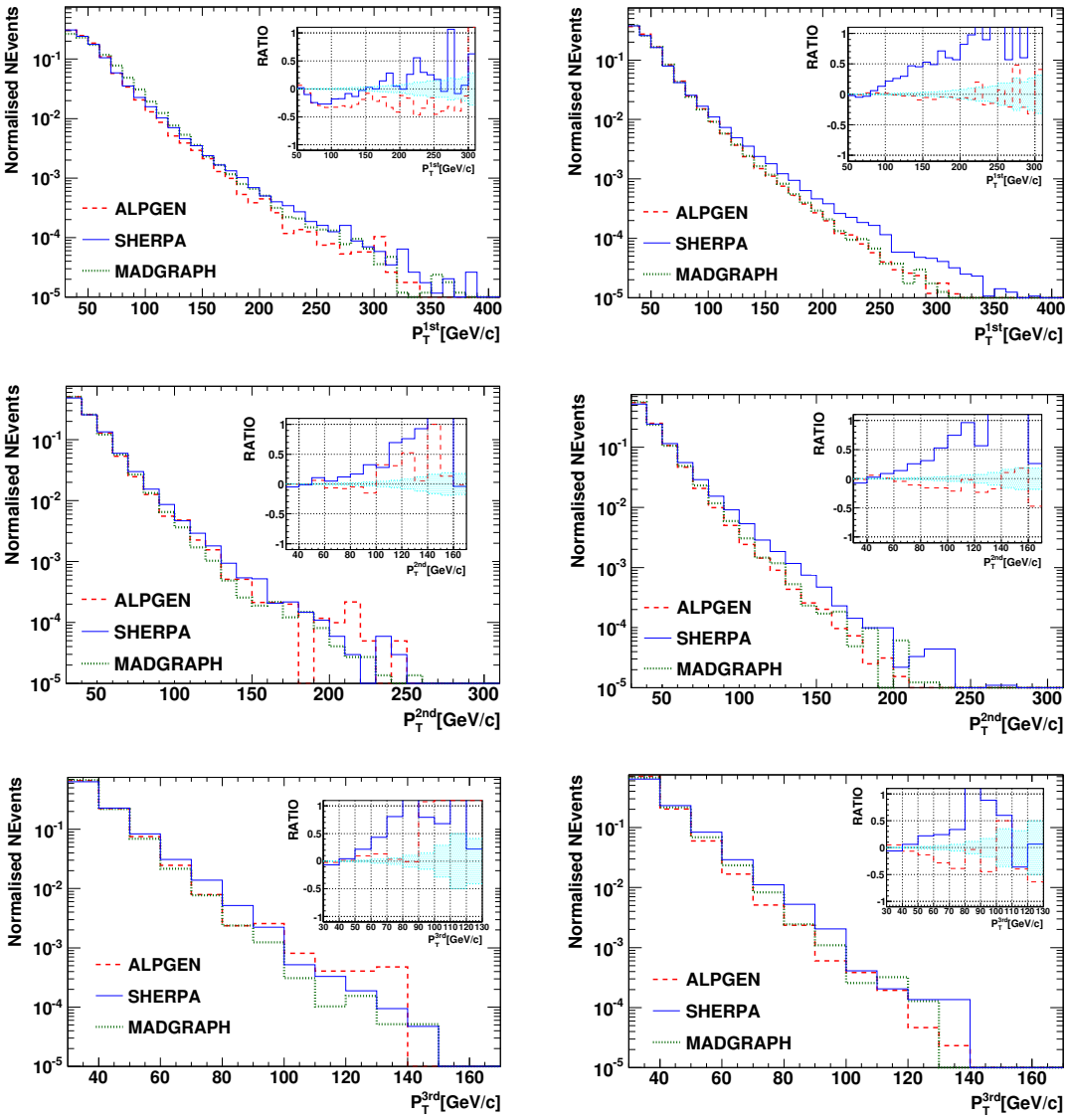


Figure 5.26: P_T distribution of the three leading jets for the inclusive Alpgen, Sherpa and Madgraph Z($l\bar{l}$) +jets (left) and W($l\nu$)+jets (right) ($l = e, \mu, \tau$) production at detector level. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph. The (blue) shaded area represents the statistical uncertainty.

However, Sherpa shows a larger difference for W($l\nu$)+jets. The observed difference is slightly larger than the difference for the exclusive samples. The Alpgen and Madgraph spectrum are in good agreement. Sherpa generally shows a harder spectrum.

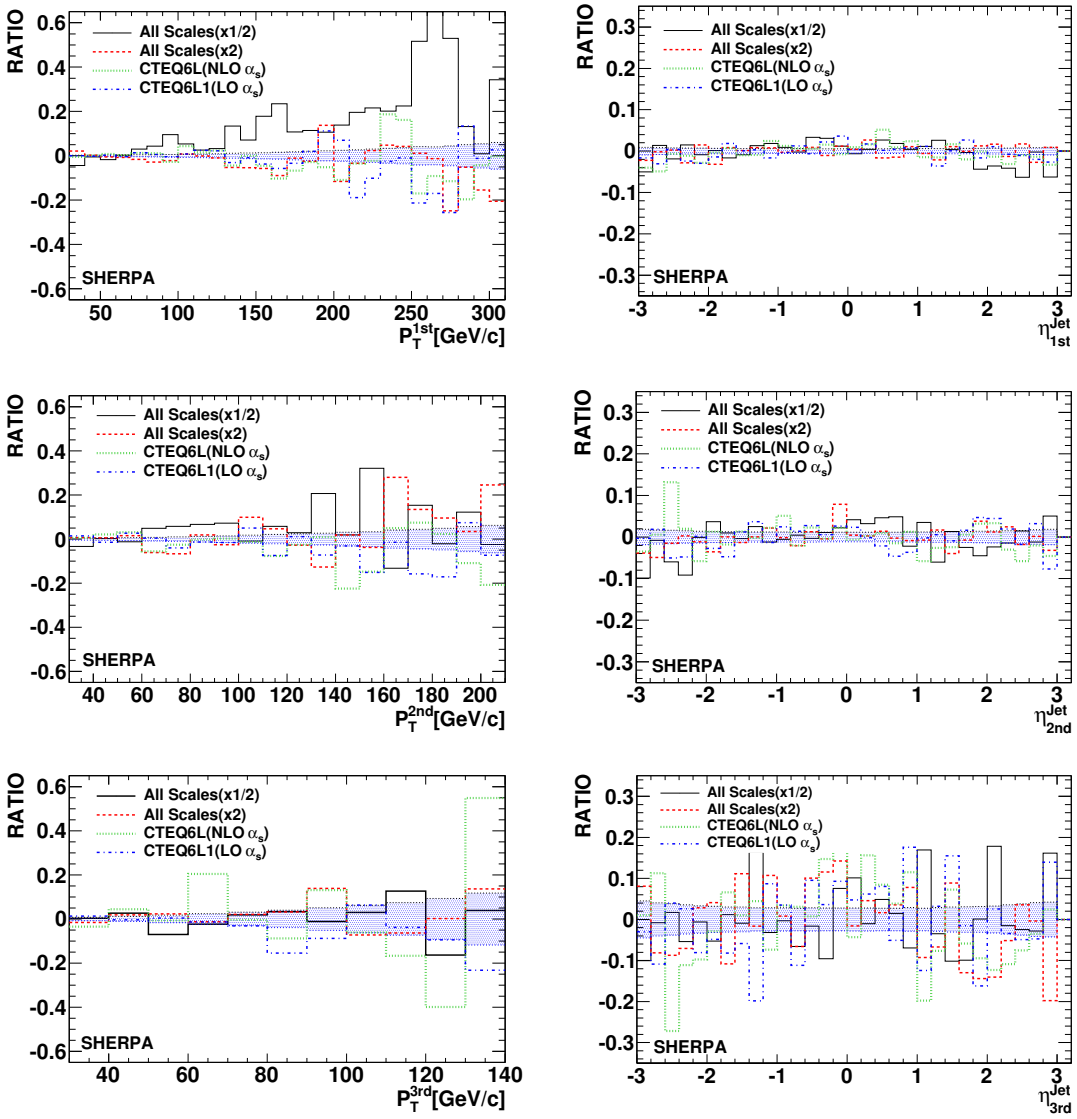


Figure 5.27: The ratio P_T (left) and η (right) distribution of the three leading jets systematics for the Sherpa Z($\mu\bar{\mu}$)+jets production at detector level. Ratio: relative differences of various Sherpa settings with respect to Sherpa default.

For Z($l\bar{l}$)+jets at lower energy, there is deviation of 20% visible on the leading jet spectrum which is most likely an effect of the detector simulation. Sherpa produces harder and less central jets in comparison with Alpgen and Madgraph predictions.

The systematics of Sherpa are presented in Figure 5.27 for exclusive Z($\mu\bar{\mu}$)+jets. The effect of varying the scales in the PDF and strong coupling evaluations by a factor of 0.5 (2.0) is that for the lower (higher) scale choices the leading jets P_T becomes slightly harder (softer). Almost no big sensitivity to these settings is observed for Z($\mu\bar{\mu}$)+jets in the whole range of P_T and η. For this kind of observables the uncertainties given by scale variations dominate the ones emerging through variations

of the internal separation cut. This is mainly due to a reduced (enhanced) suppression of hard-jet radiation through the α_s rejection weights. Moreover, considering the pseudo-rapidity of the leading jet, this distribution shows a very stable behavior under the studied variations, since they are indirectly influenced by the cut scale only and these effects are not visible at detector level reconstruction.

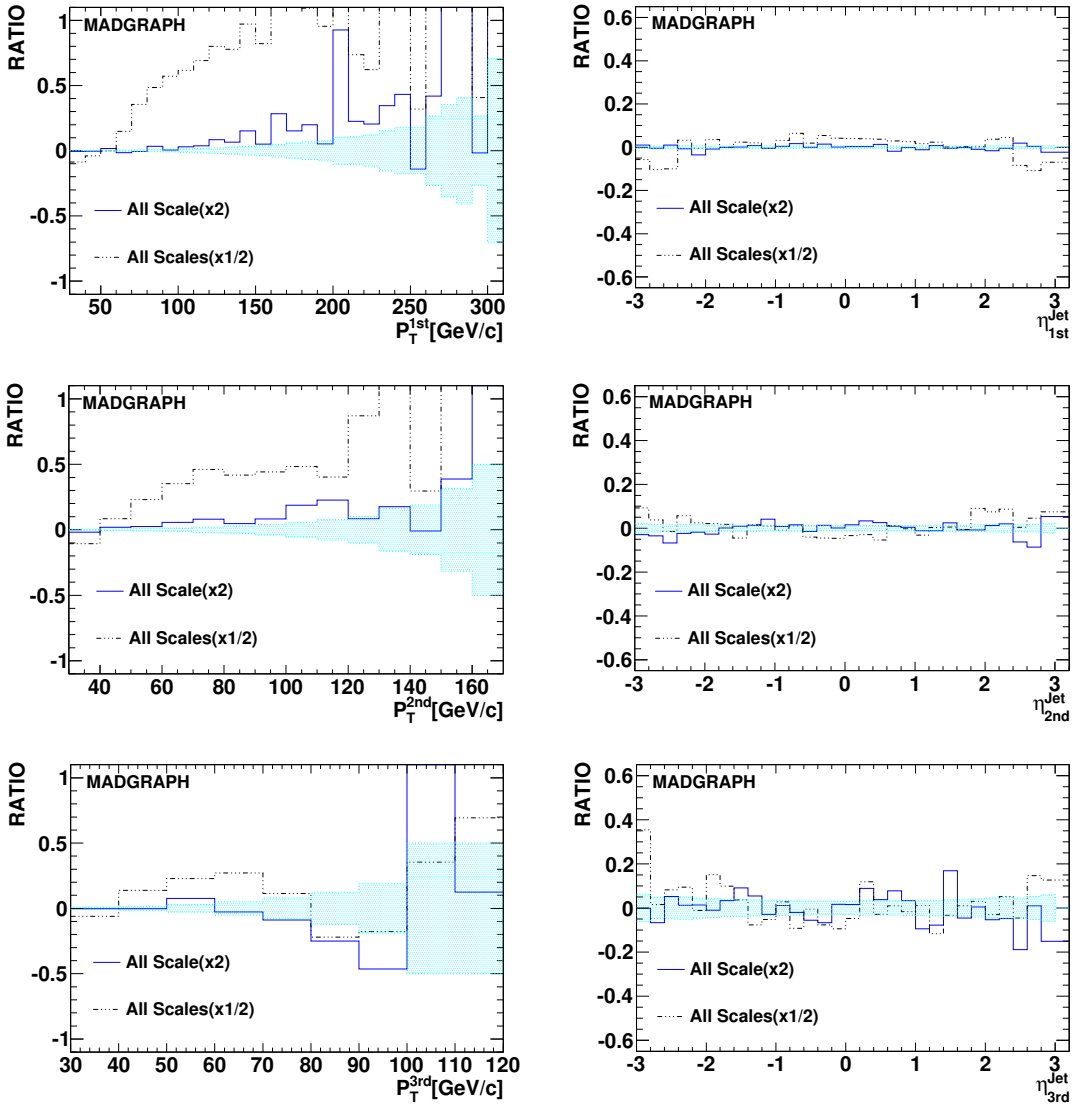


Figure 5.28: The ratio P_T (left) and η (right) distribution of the three leading jets systematics for the inclusive Madgraph W($l\nu$)+jets production ($l = e, \mu, \tau$). Ratio: relative differences of various Madgraph settings with respect to Madgraph default.

The systematics of Madgraph is presented in Figure 5.28 for W($l\nu$)+jets. The effect of varying the scales in the PDF and strong coupling evaluations by a factor of 0.5 is that for the lower scale choice the P_T of the leading jets shows a harder spectrum compared to the default settings. The differences are getting smaller for the sub-leading jets. However, there is a negligible difference for the higher scale

variation (x2). Almost no sensitivity to these settings is observed for the W($l\nu$)+jets production in the whole range of η.

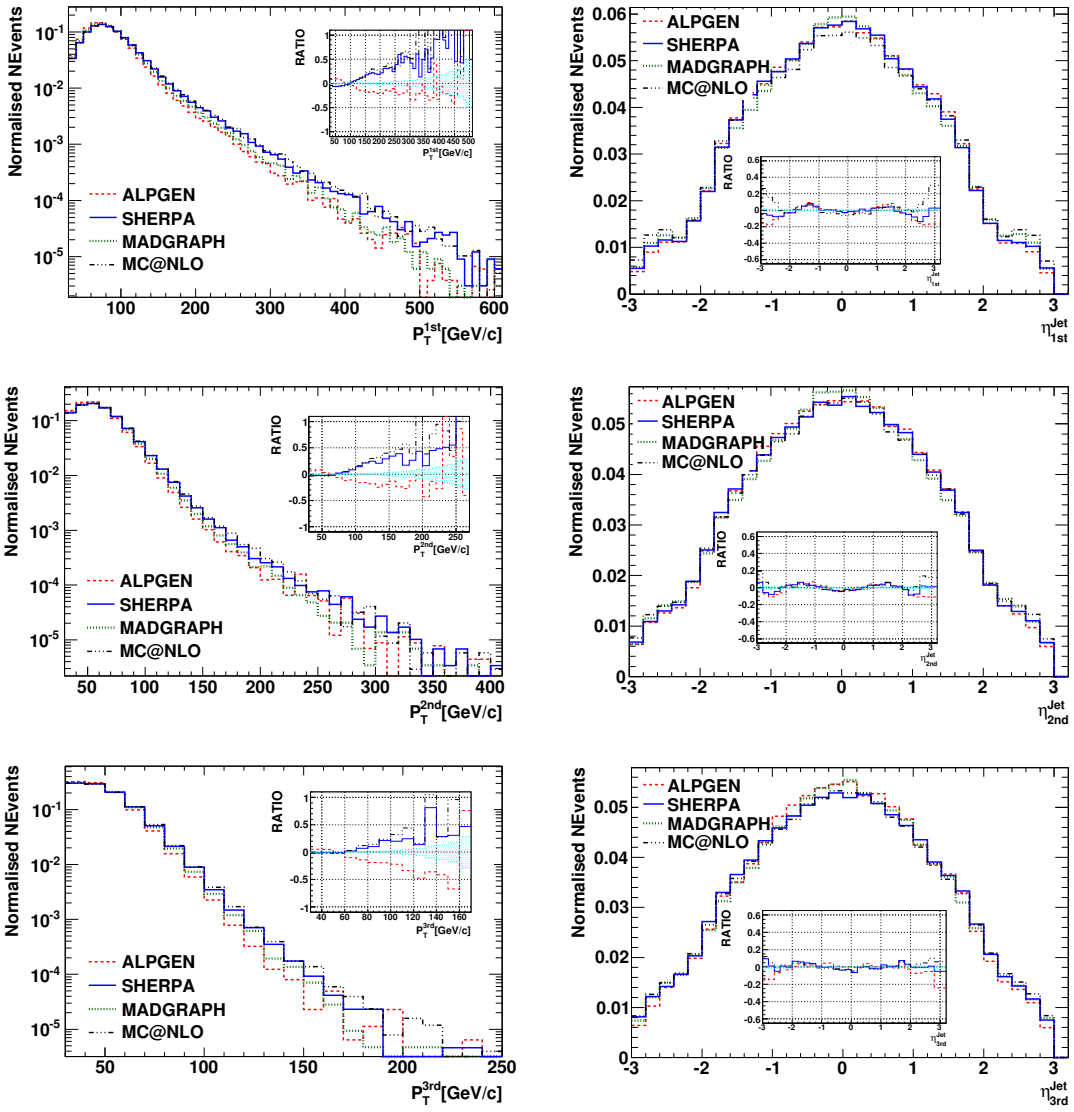


Figure 5.29: P_T (left) and η (right) distribution of the three leading jets for the inclusive Alpgen, Sherpa and Madgraph $t\bar{t}$+jets production at detector level. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

The transverse momentum (P_T) and η distribution of the leading jets are shown in Figure 5.29 for the $t\bar{t}$+jets production. The Madgraph and Alpgen predictions generally show a good agreement. There are slight differences in the range of 10% to 30% for the distributions of the sub-leading jets. Sherpa produces harder and less central jets compared to Alpgen and Madgraph and the difference for the leading and sub-leading jets reaches 30%-40% after $P_T > 150$ GeV. This observation agrees with the Sherpa and Alpgen comparison for Z/W+jets. Due to the Sudakov rescaling between ME and PS via CKKW, Sherpa shows a different shape. However, Sherpa

and MC@NLO are in a good agreement.

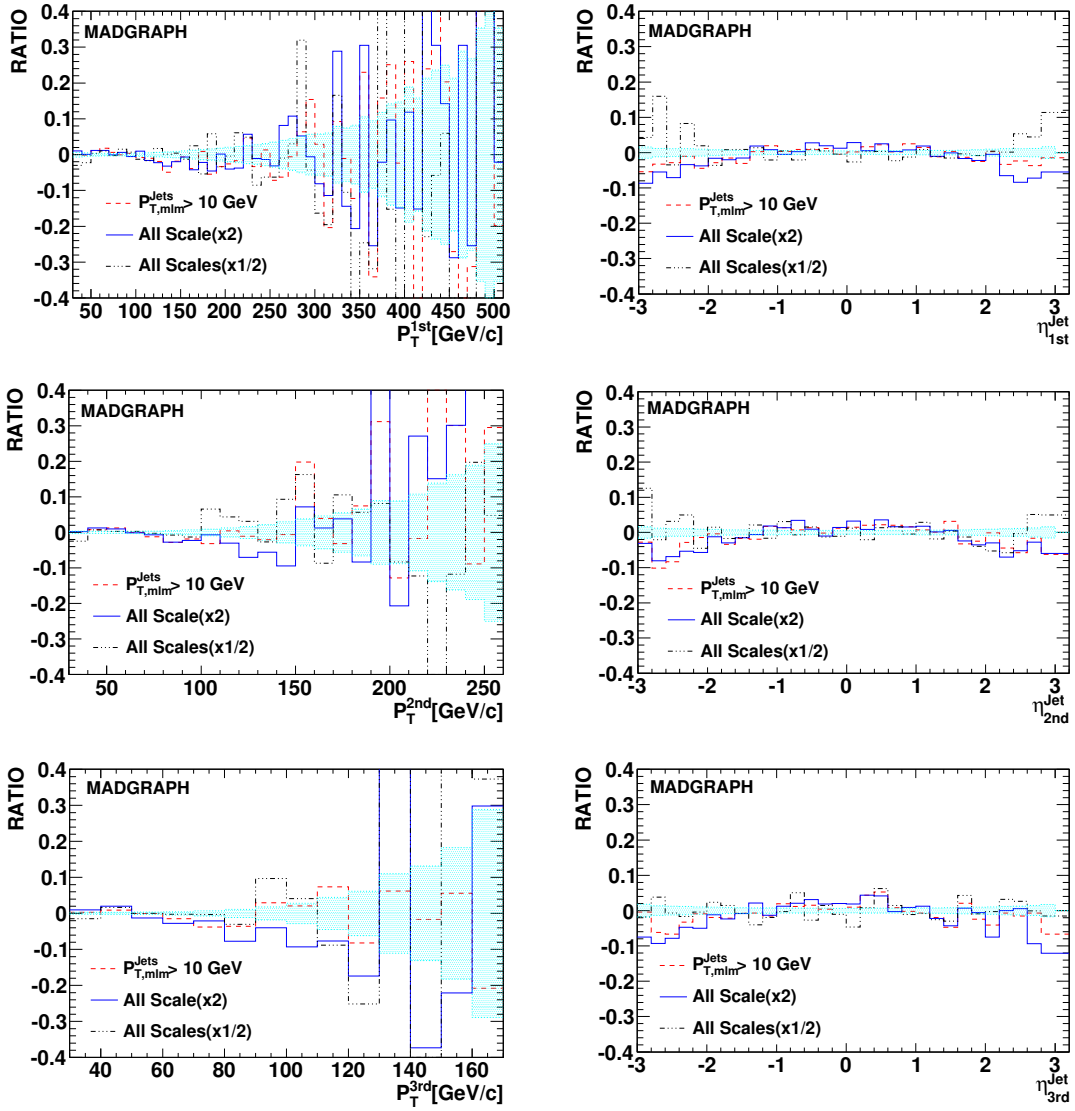


Figure 5.30: The ratio P_T (left) and η (right) distribution of the three leading jets systematics for the inclusive $t\bar{t}$+jets production. Ratio: relative differences of various Madgraph settings with respect to Madgraph. The smaller and larger ISR effects are not presented due to the negligible effects (below 3%).

The important of Madgraph for the inclusive $t\bar{t}$+jets production are presented in Figure 5.30. The effect of varying the scales in the PDF by a factor of 0.5 (2) is that for the lower (higher) scale choices the leading jets P_T and η show a negligible deviation below 10%. The matching scale variation is also of negligible order for leading jets P_T and η spectra.

The sum of E_T of all jets (sumEt) is shown in Figure 5.31. The Madgraph and Alpgen prediction generally show a good agreement. There is a slight difference for the Z+jets production at higher energy in the range of 20% to 30% because

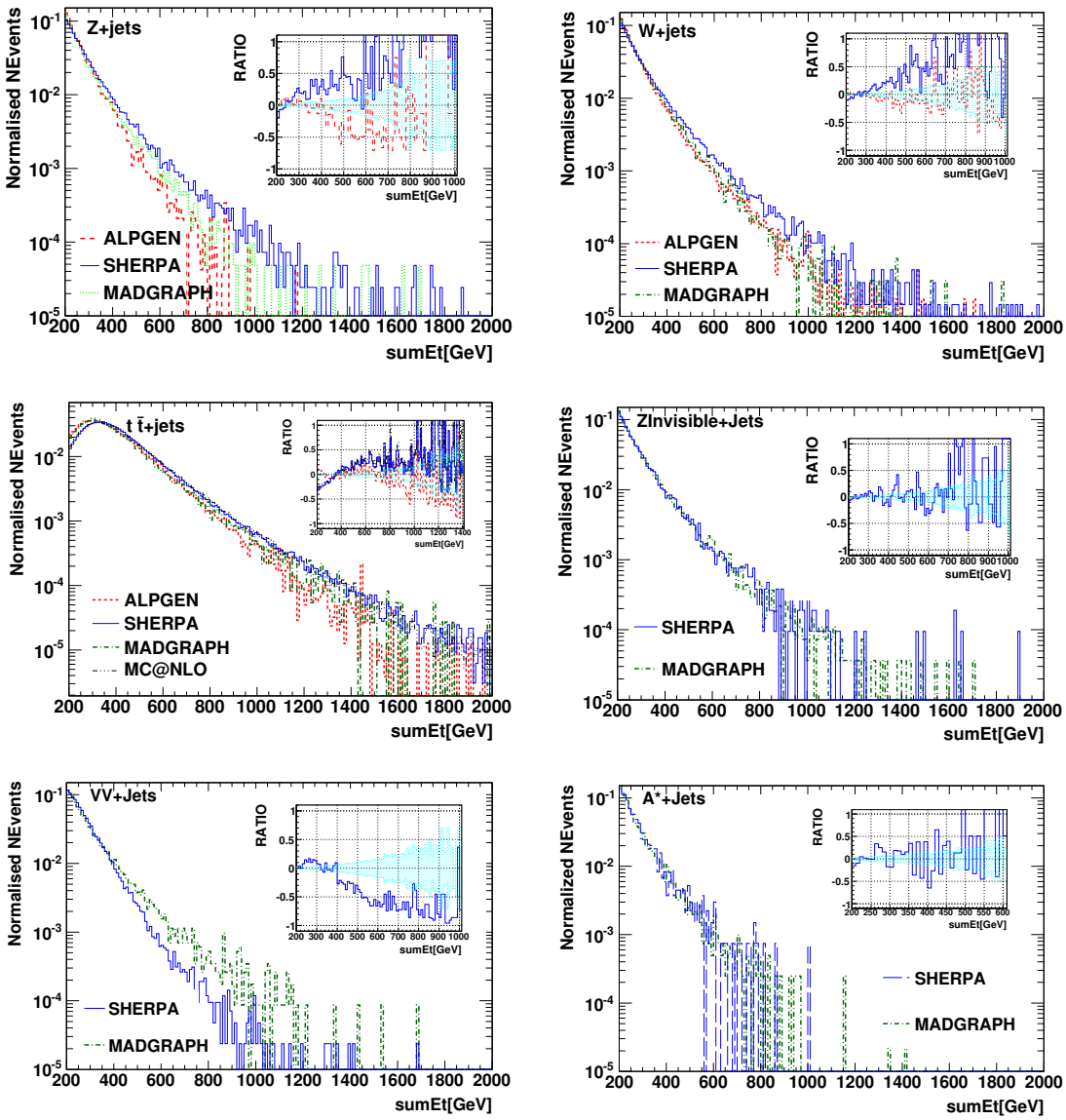


Figure 5.31: The sum of E_T of all jets (> 200 GeV) for Z($l\bar{l}$)+jets, W($l\nu$)+jets, $t\bar{t}$+jets, Zinvisible+jets, VV+jets and A*+jets production. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph. The shaded area represents the statistical uncertainty.

of higher jet multiplicities in Madgraph samples. Sherpa produces harder spectra compared to Alpgen and Madgraph and the difference reaches 50% for higher values of sumEt. Sherpa has differences in the jet production because of the different matching method. Sherpa and MC@NLO are in a good agreement for $t\bar{t}$+jets since both generators uses the similar implementation of Sudakov form factor between ME and PS. However, Sherpa gives a softer spectrum for the VV+jets production due to lower jet multiplicity at ME compared to Madgraph.

The jet multiplicity for the reconstructed jets with $P_{Tj} > 30$ GeV and $|\eta| < 4$ is presented in Figure 5.32 for the exclusive Z($\mu\bar{\mu}$)+jets and W($\mu\nu$)+jets production.

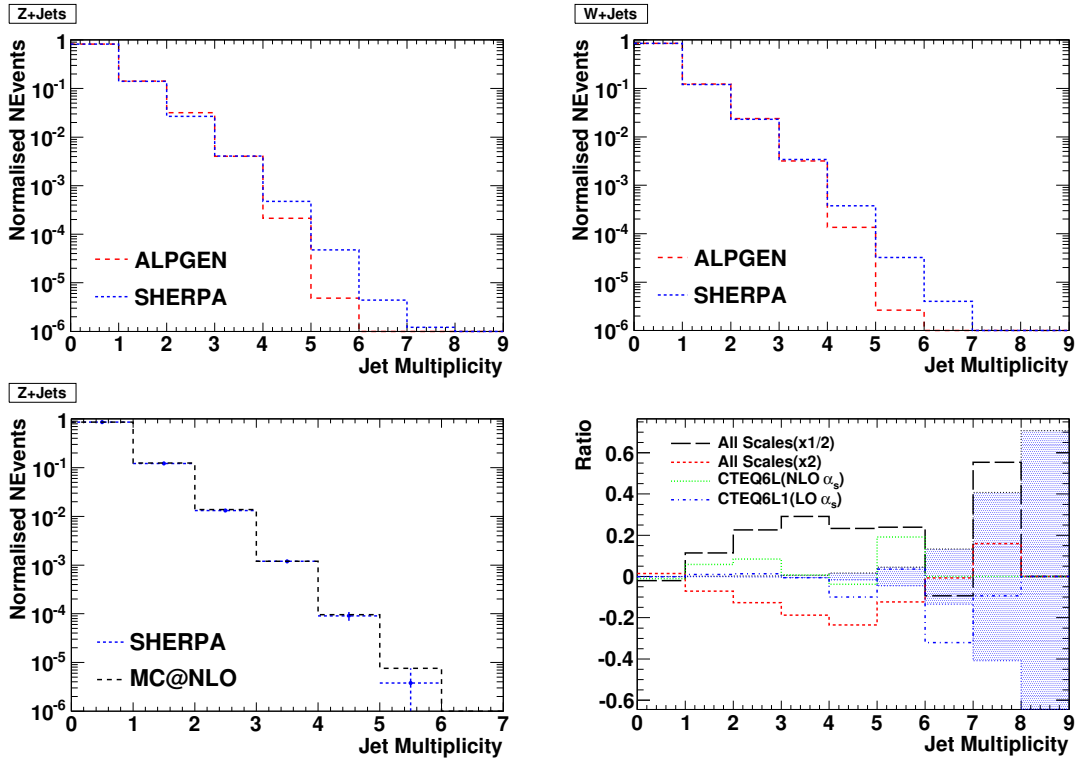


Figure 5.32: Jet multiplicities at detector level ($P_{Tj} > 30$) for exclusive Z($\mu\bar{\mu}$)+jets, W($\mu\nu$)+jets production simulated with Alpgen, Sherpa and MC@NLO. Sherpa and MC@NLO generators are only compared with an additional jet at parton level. Lower-left: relative differences of various Sherpa settings with respect to Sherpa default.

In spite of the difference in the P_T distributions for the sub-leading jets and differences in the matching schemes, the jet multiplicity is in good agreement up to four jets. However, higher jet multiplicities, mostly defined by initial and final state radiation, differ significantly. Sherpa shows a higher jet-multiplicity since the parton shower scale for the initial and final state via Sudakov implementation is different in Sherpa. However, Sherpa shows a very good agreement with MC@NLO for Z+1jet. Although they have different ME and PS approaches both generators use similar Sudakov implementation for the ME and PS merging and normalization. The jet multiplicity shows similar behavior for Sherpa and MC@NLO. The effect of varying the scales in the PDF and strong coupling evaluations by a factor of 0.5 (2.0) is that for the lower (higher) scale choices the jet multiplicity becomes slightly harder (softer). A larger jet multiplicity is observed for the variation of scales 0.5, whereas a variation of 2.0 causes a lower jet multiplicity.

The jet multiplicity for the reconstructed jets with $P_{Tj} > 30$ GeV and $|\eta| < 4$ is presented in Figure 5.33 for all inclusive samples. In spite of the differences in the P_T distributions for the sub-leading jets, differences in matching schemes and jet multiplicity at matrix element the jet multiplicity is in good agreement up to four or five jets for the Z+jets, W+jets, Zinvisible+jets and VV+jets production. The $t\bar{t}$+jets production shows good agreement up to eight or nine jets for all generators.

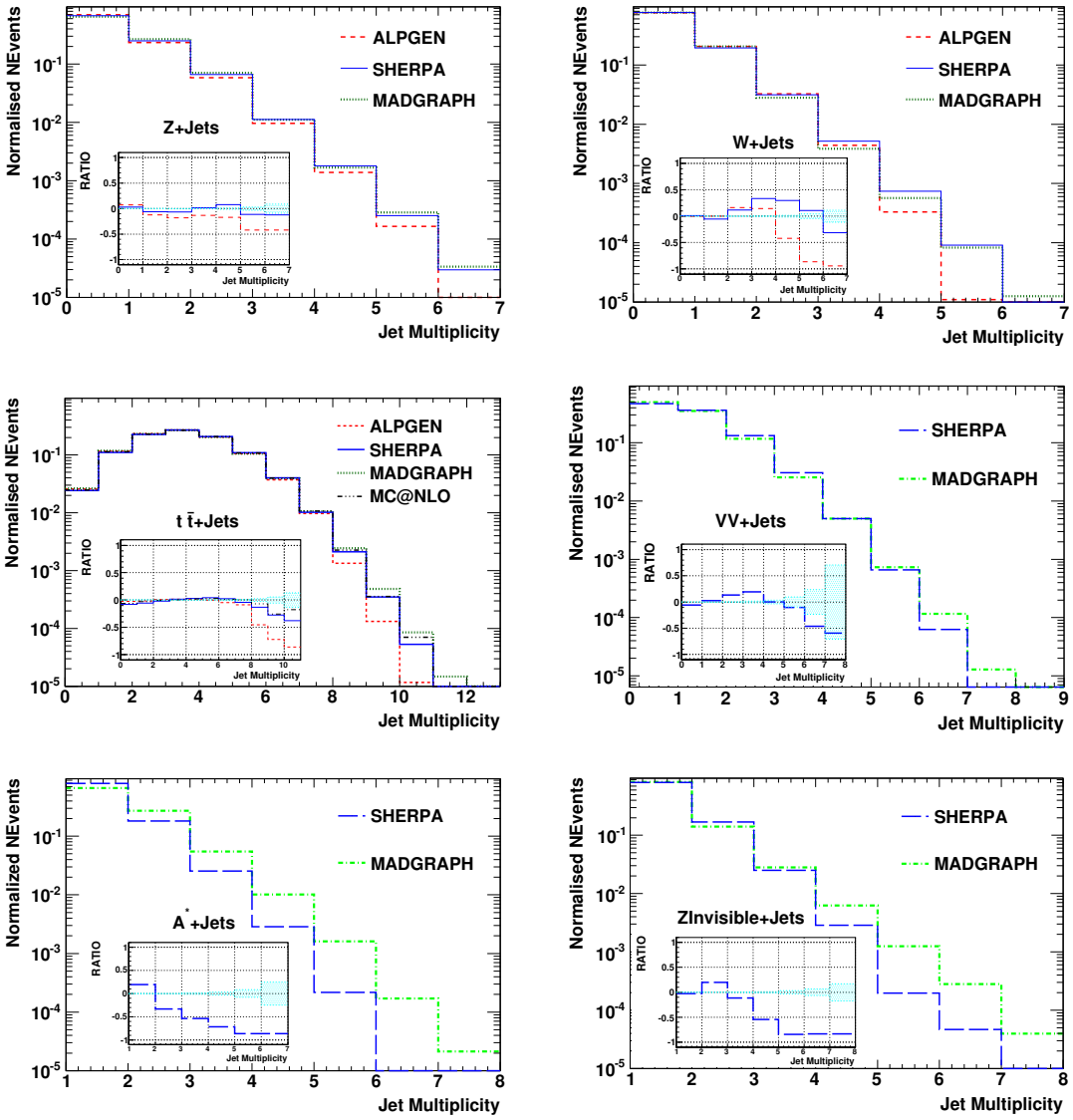


Figure 5.33: Jet multiplicities at detector level ($P_{Tj} > 30$ GeV) for the inclusive Z($l\bar{l}$)+jets, W($l\nu$)+jets ($l = e, \mu, \tau$), $t\bar{t}$+jets, VV+jets A*+jets and Zinvisible+jets production. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

There is a difference between Madgraph and Sherpa for the jet multiplicity of the A^*+jet production. The difference arises from higher jet multiplicity production at ME level in Madgraph. Additionally there can be different implementations of the γ^* interference for the initial state radiation between generators. Sherpa and Madgraph generally show a harder jet multiplicity than Alpgen. However, higher jet multiplicities, mostly defined by initial and final state radiation, differ significantly as Sherpa shows a higher jet-multiplicity. Although Madgraph uses the same shower algorithm and matching methods as Alpgen, it simulates more jets at parton level which affect the tail of the distribution.

The angular correlations between the missing transverse energy and the leading jets have been presented for the inclusive samples. The azimuthal angle between reconstructed leading jets and missing transverse energy is presented in Figures 5.34, 5.34, 5.36, 5.37 and 5.38 for the Z($l\bar{l}$)+jets, W($l\nu$)+jets, $t\bar{t}$+jets, Zinvisible+jets, VV+jets and A*+jets production, respectively. The angular separation is extensively used for hadronic SUSY searches for separating signal from QCD background. All generators shows a very good agreement for Z+jets, W+jets, $t\bar{t}$+jets, VV+jets and A*+jets. There are small deviations where the ISR occurs ($\Delta\phi > 3$) due to the different implementation of the matching methods with initial state radiation. The systematics for the reconstructed azimuthal angle between the leading jets and calorimeter transverse energy are almost negligible for all generators. However, there is a significant difference between Sherpa and Madgraph for the leading jet in Zinvisible+jets. Madgraph gives harder spectrum for the azimuthal angle between the leading jet and the missing transverse energy.

The cone separation ΔR=$\sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the four leading jets is presented in Figure 5.35 and 5.36 for Z($l\bar{l}$)+jets, W($l\nu$)+jets and $t\bar{t}$+jets, respectively. Alpgen and Madgraph generally show a very good agreement for the different jet combinations. However, Alpgen-Madgraph and Sherpa show differences in the region of ΔR$>$ 3.3. Sherpa and Alpgen-Madgraph generators have different implementation of ISR and FSR and fragmentation methods which has influence on the detector simulations. However, Sherpa and MC@NLO agree well for the azimuthal angle between leading jets and calorimeter E_T^{Miss} and the cone separation for the leading two jets in Figure 5.39. The second jet for Z($\mu\bar{\mu}$)+1jet naturally comes from parton shower that means Sherpa CKKW and MC@NLO matching methods work well. Because Sherpa and MC@NLO have almost the same fragmentation functions, both show a good agreement at detector level.

In summary, the largest deviations for the exclusive and inclusive samples are observed between Sherpa and Alpgen-Madgraph for the P_T of the three leading jets. Sherpa produces harder and less central predictions compared to Alpgen and Madgraph. The variation increases with sub-leading jets P_T and can reach $30 - 40\%$ above P_T >100 GeV. The application of the Sudakov rescaling between ME and PS via CKKW method gives different shapes for the jet P_T predictions with respect to the MLM method (see Section 4.2.3). Additionally Sherpa and Alpgen-Madgraph use different hadronization tools (see Table 5.1) for the final events which also affect the jet reconstruction. Alpgen also tends to produce lower jet multiplicities with respect to Sherpa and Madgraph for N$_j \geq 4$ (for $t\bar{t}$+jets N$_j \geq 8$), i.e. in the region where state radiation jets contribute. However, there is good agreement between Alpgen and Madgraph. The same behavior is also observed between Sherpa and MC@NLO. Although the influence of the renormalization/fragmentation scales, matching thresholds and PDFs for Sherpa predictions has a reasonable range within 10% to 20%, Madgraph scale variations for the single vector boson production show larger deviations for the leading jet P_T distribution. These deviations are getting softer for the P_T of the sub-leading jets. Azimuthal angle and ΔR separation for all generators are in good agreement.

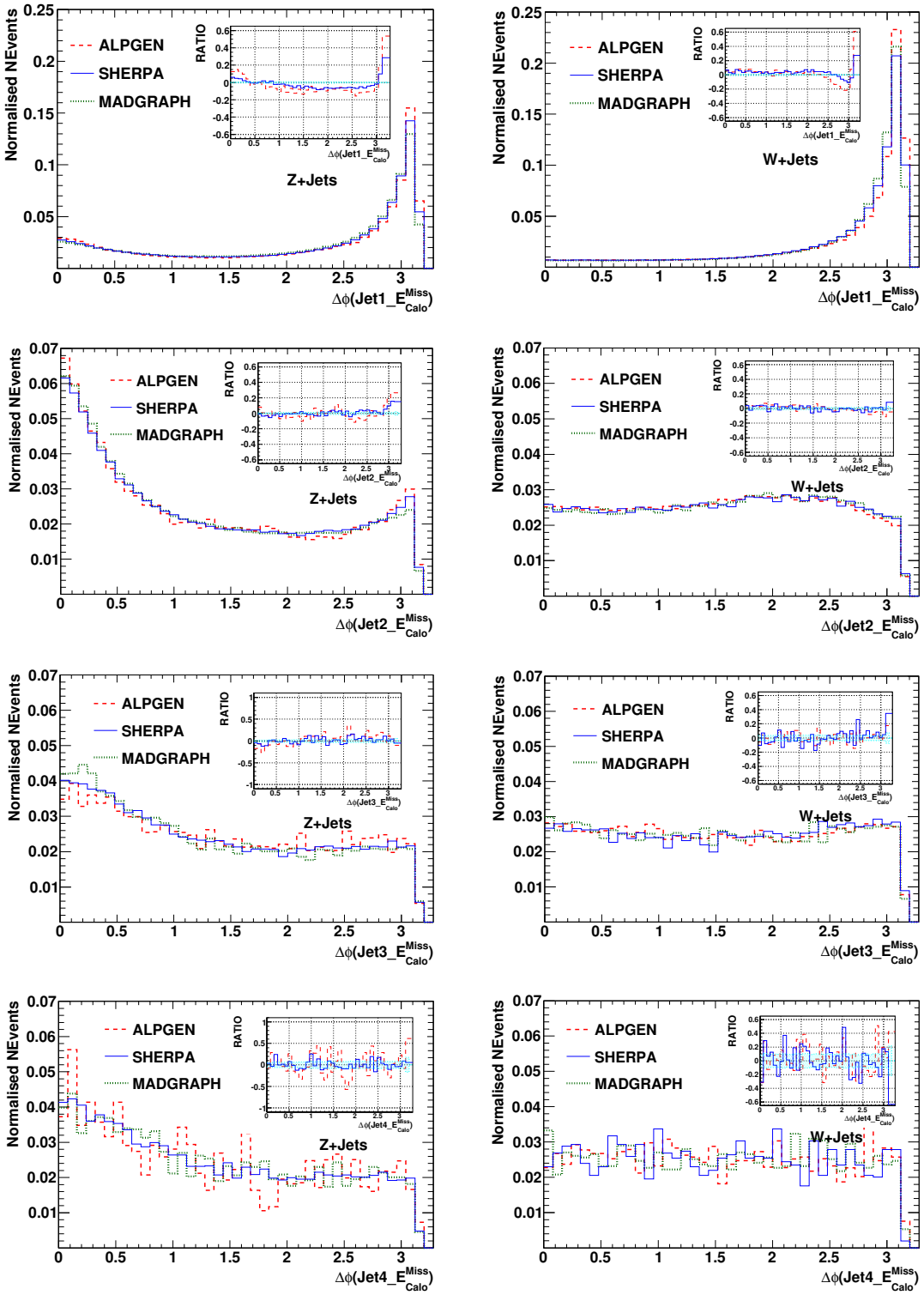


Figure 5.34: Azimuthal angle ($\Delta\phi$) between leading jets and calorimeter missing transverse energy for the inclusive Z($l\bar{l}$)+jets (left) and W($l\nu$)+jets (right) production ($l = e, \mu, \tau$) at detector level. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

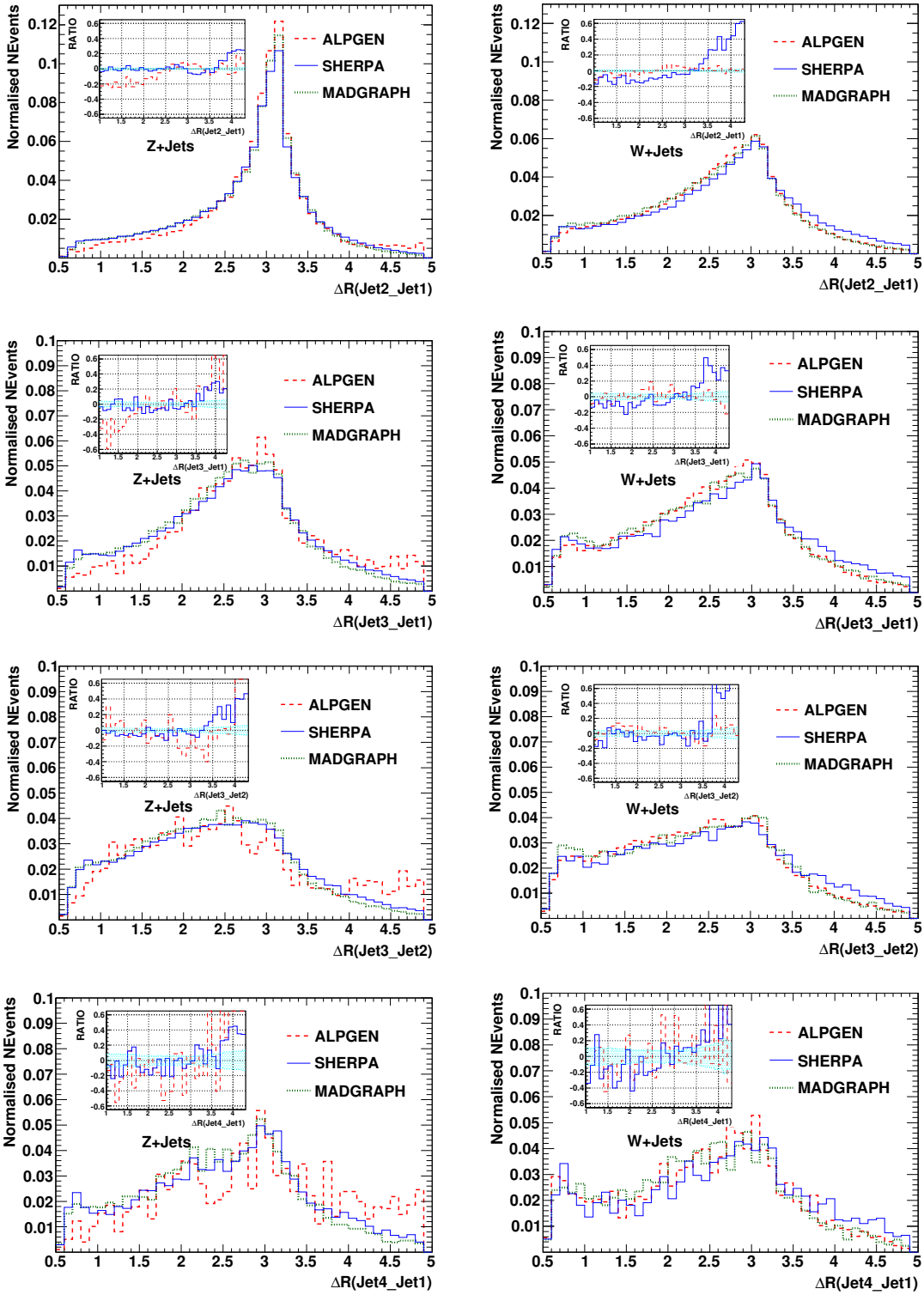


Figure 5.35: ΔR separation between leading jets for the inclusive Z($l\bar{l}$)+jets (right) and W($l\nu$)+jets (left) ($l = e, \mu, \tau$) production at detector level. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

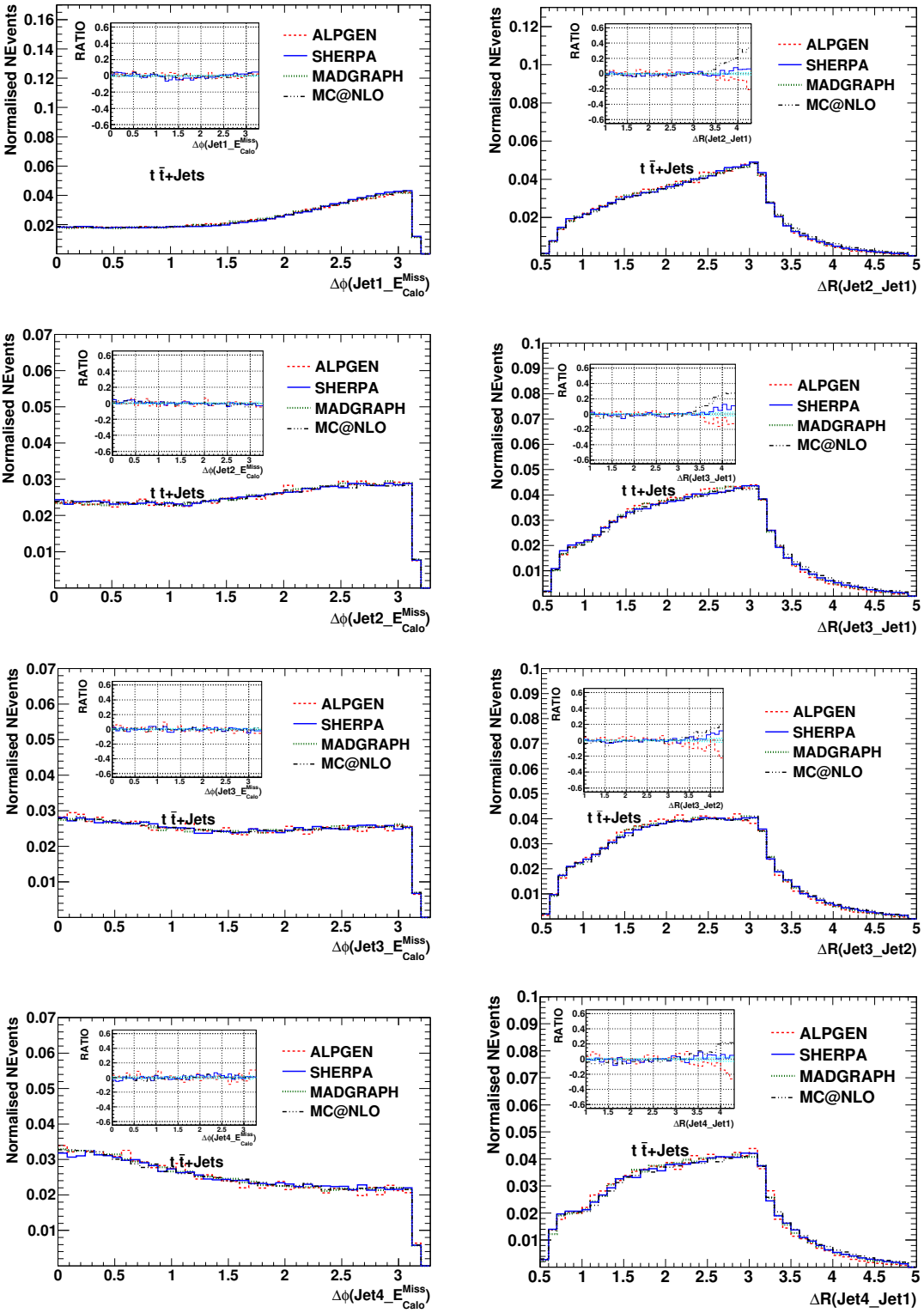


Figure 5.36: Left: Azimuthal angle ($\Delta\phi$) between leading jets and calorimeter missing transverse energy for inclusive $t\bar{t}$+jets. Right: ΔR separation between leading jets. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

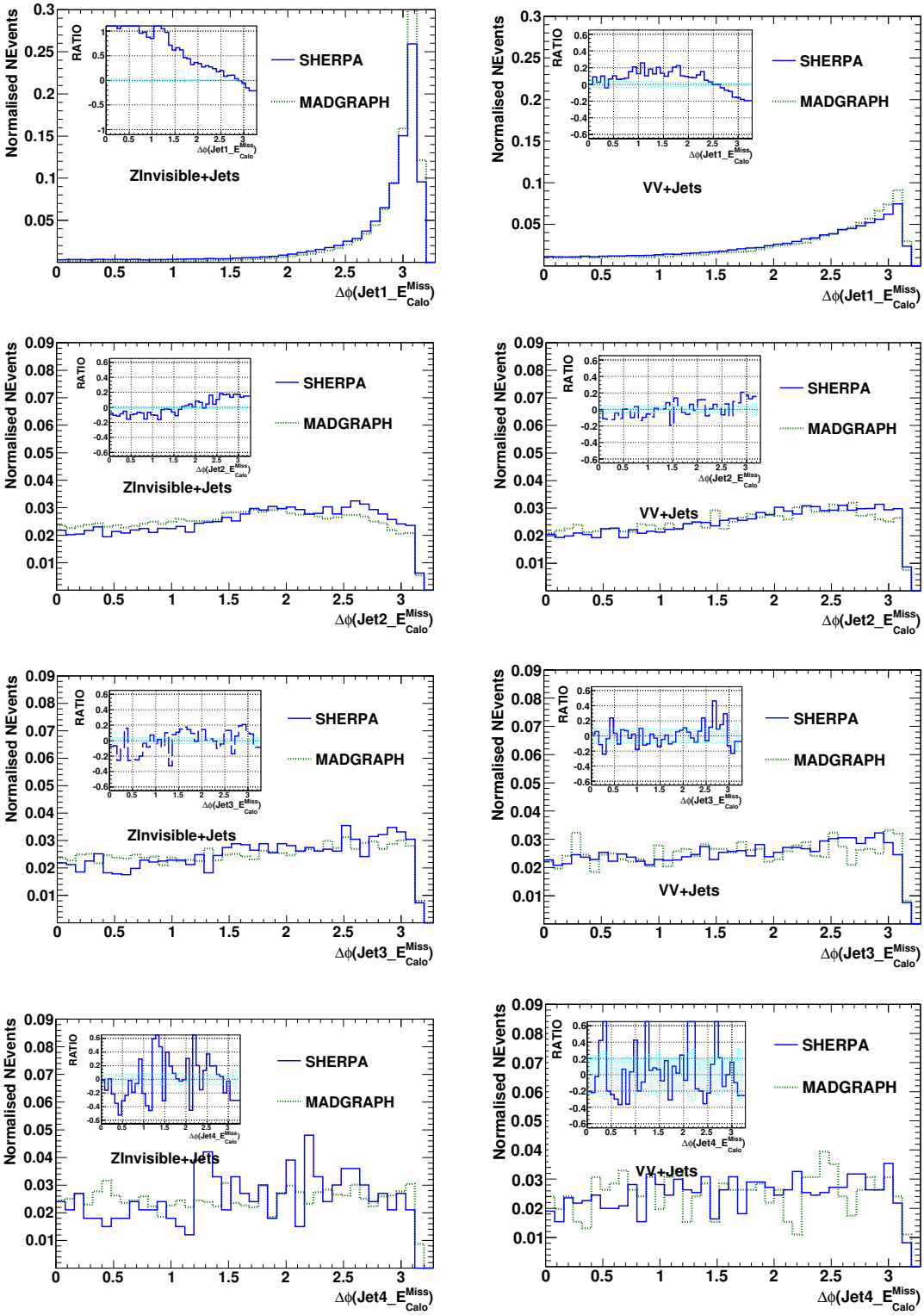


Figure 5.37: Azimuthal angle ($\Delta\phi$) between leading jets and calorimeter missing transverse energy for the inclusive VV+jets and Zinvisible+jets production at detector level. Ratio: relative difference of Sherpa with respect to Madgraph.

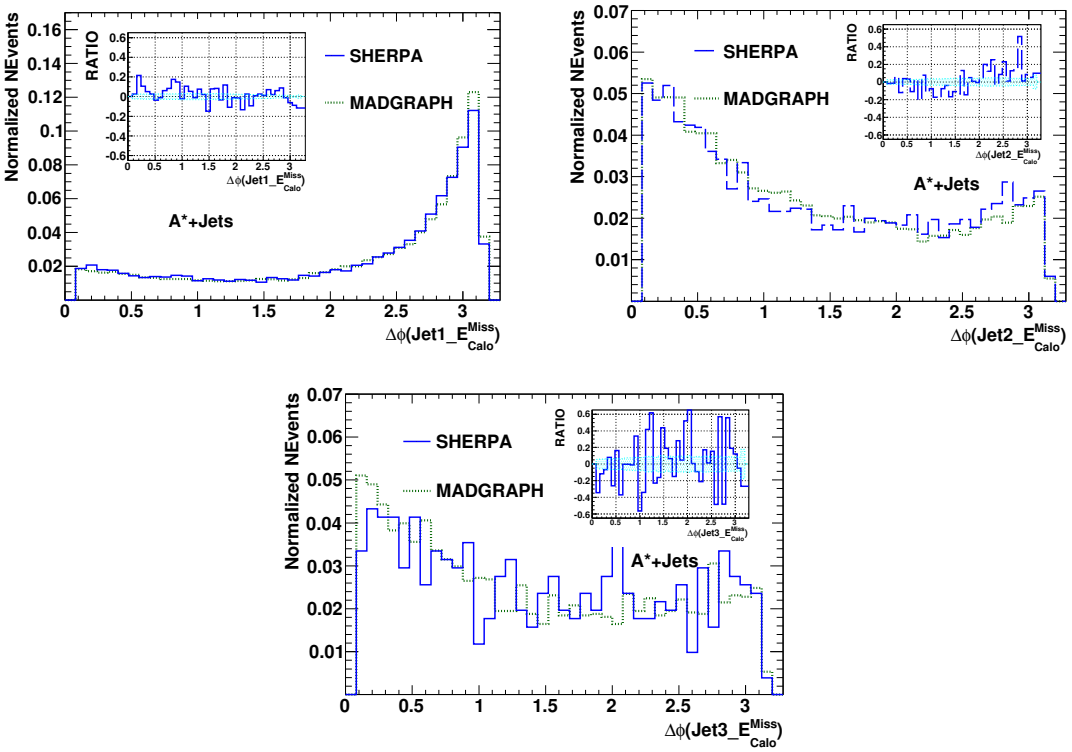


Figure 5.38: Azimuthal angle ($\Delta\phi$) between leading jets and calorimeter missing transverse energy for inclusive A*+jets at detector level. Ratio: relative difference of Sherpa with respect to Madgraph.

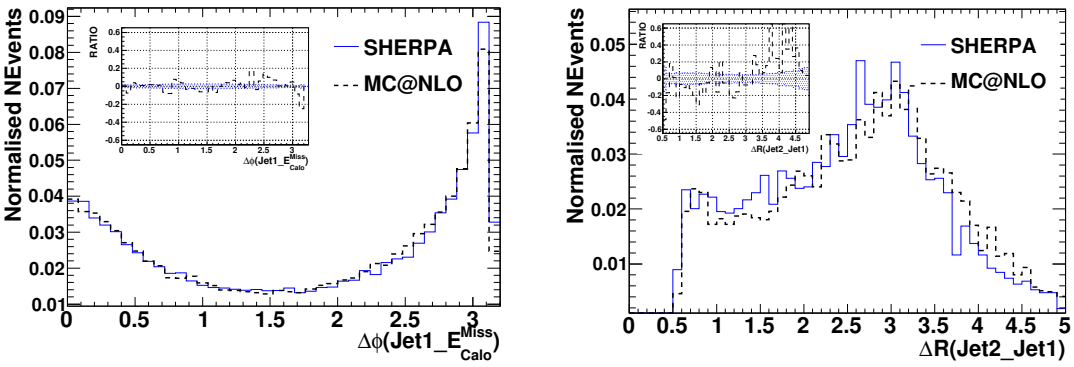


Figure 5.39: Left: Azimuthal angle ($\Delta\phi$) between leading jets and calorimeter missing transverse energy for the exclusive Z($\mu\bar{\mu}$)+ 1jet production at detector level. Right: ΔR separations between reconstructed leading jets. Ratio: relative difference of MC@NLO with respect to Sherpa.

5.2.4 Lepton Observables

For studying lepton observables, the muons from Z and W decays have been considered. Figures 5.40 and 5.41 show the P_T distribution of muons at generator and detector level for samples produced with Alpgen and Sherpa. The muon P_T spectra simulated with Alpgen and Sherpa are in good agreement.

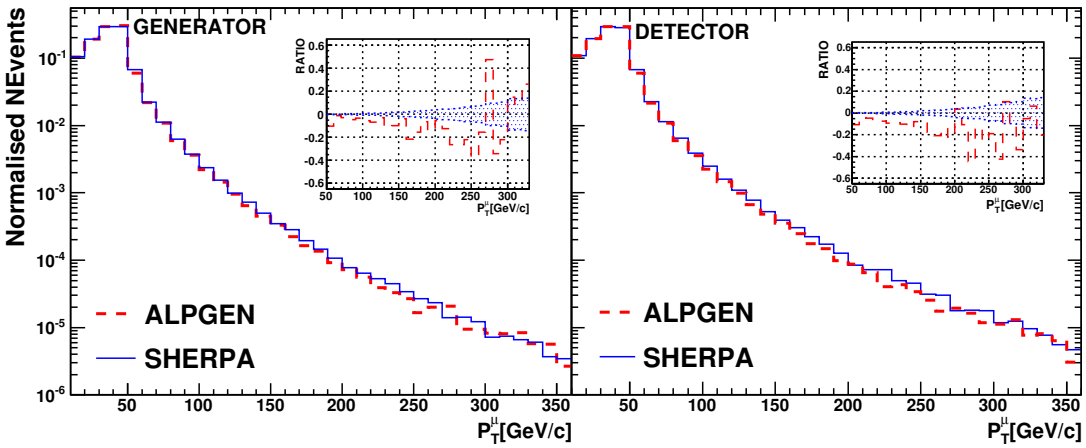


Figure 5.40: Generator (left) and detector (right) level P_T distribution of muons for the exclusive Z($\mu\bar{\mu}$)+jets production simulated with Alpgen and Sherpa. Ratio: relative difference of Alpgen with respect to Sherpa. The shaded area represents the statistical uncertainty.

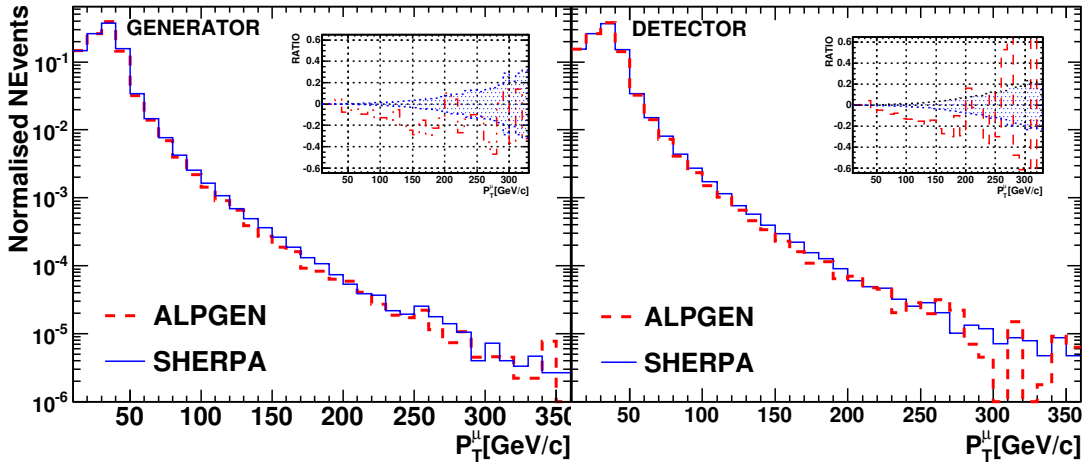


Figure 5.41: Generator (left) and detector (right) level P_T distribution of muons for the exclusive W($\mu\nu$)+jets production simulated with Alpgen and Sherpa. Ratio: relative difference of Alpgen with respect to Sherpa.

The P_T and η_μ distribution of muons from the inclusive Z($l\bar{l}$)+jets, W($l\nu$)+jets and $t\bar{t}$+jets production are presented in Figure 5.42. The P_T spectra is in good agreement

for Alpgen and Madgraph. The largest deviation in the P_T spectra is observed for Sherpa for the Z+jets and W+jets production.

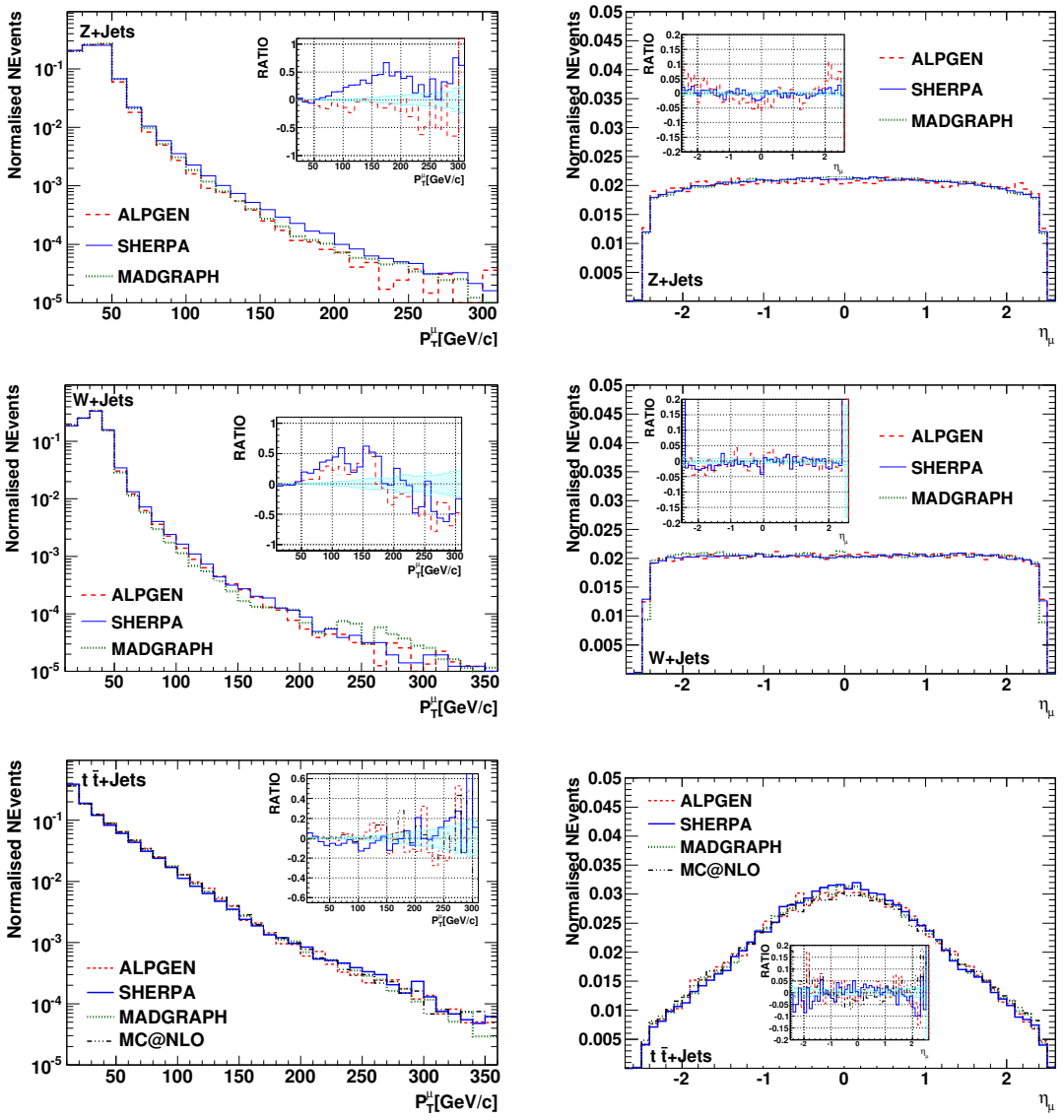


Figure 5.42: Left: detector level P_T distribution of muons for the Z($l\bar{l}$)+jets, W($l\nu$)+jets ($l = e, \mu, \tau$) and $t\bar{t}$+jets production simulated with Alpgen, Madgraph, MC@NLO and Sherpa. Right: detector level η distribution of muons for Z($l\bar{l}$)+jets, W($l\nu$)+jets and $t\bar{t}$+jets calculated with Alpgen, Madgraph and Sherpa. Ratio: relative difference of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

In Figures 5.43 and 5.44 the transverse momentum of the Z boson is presented for exclusive and inclusive samples. The generator and detector level distributions of Sherpa show a harder spectrum compared to Alpgen and Madgraph. Madgraph generally gives a softer spectrum on both generator and detector level. The differences decrease at detector level and the distribution are comparable with the Alpgen prediction. Figures 5.45 and 5.46 show the invariant distribution of opposite sign

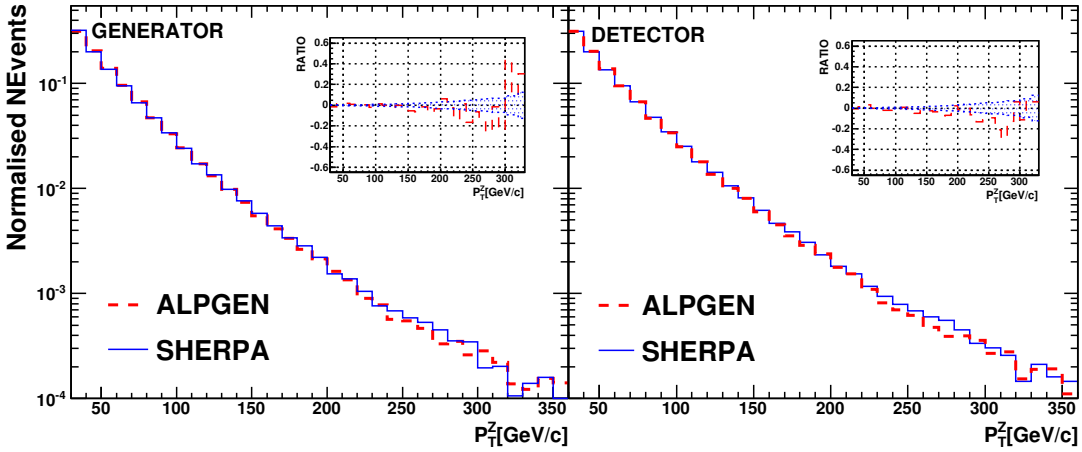


Figure 5.43: Generator (left) and detector (right) level P_T^Z distribution for the Z($\mu\bar{\mu}$)+jets production for Alpgen and Sherpa. Ratio: relative difference of Alpgen with respect to Sherpa.

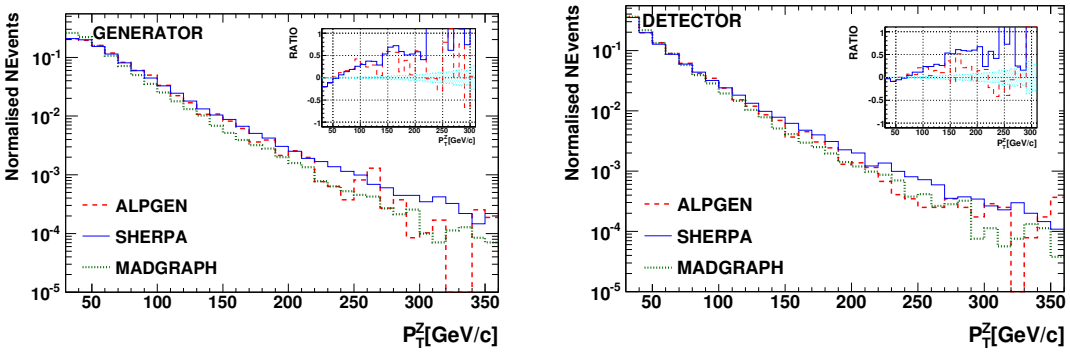


Figure 5.44: Generator (left) and detector (right) level P_T^Z distribution for the Z($l\bar{l}$)+jets ($l = e, \mu, \tau$) production for Alpgen, Madgraph and Sherpa. Ratio: relative difference of Alpgen and Sherpa with respect to Madgraph.

(OS) muons from Z decay for Alpgen, Sherpa and Madgraph at generator and detector level. The invariant mass of OS muons is shown with $66 < M_{\mu\mu} < 116$ GeV for the exclusive Alpgen and Sherpa samples. The distribution for the inclusive Z($l\bar{l}$)+jets samples are presented with $50 < M_{\mu\mu} < 200$ GeV for Alpgen, Madgraph and Sherpa. Additionally the η_Z distribution is shown. Alpgen shows a less central prediction compared to Sherpa and Madgraph.

As mentioned for Sherpa in the previous sections, the observed differences depend on the scale choice of the matching schemes and the different shower implementations with Sudakov rescaling approaches between the generators. The differences spotted in Figure 5.47 has been considered as the sensitivity between angular and virtuality ordered showers for Herwig and Sherpa. Both generators use mainly the same approach for the fragmentation.

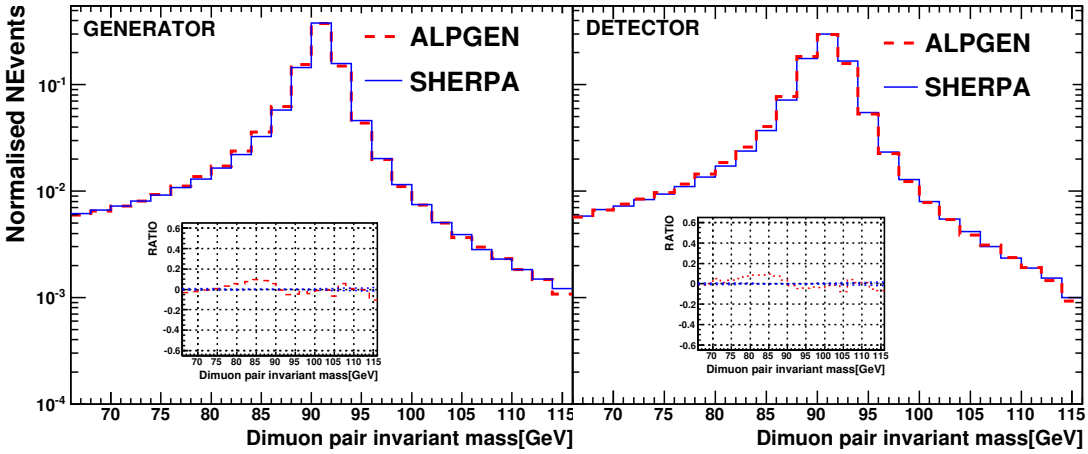


Figure 5.45: Generator (left) and detector (right) level invariant mass of OS muons ($M_{\mu\bar{\mu}} > 66$ GeV) distributions calculated with Alpgen and Sherpa for the exclusive Z($\mu\bar{\mu}$)+jets production. Ratio: relative difference of Alpgen with respect to Sherpa.

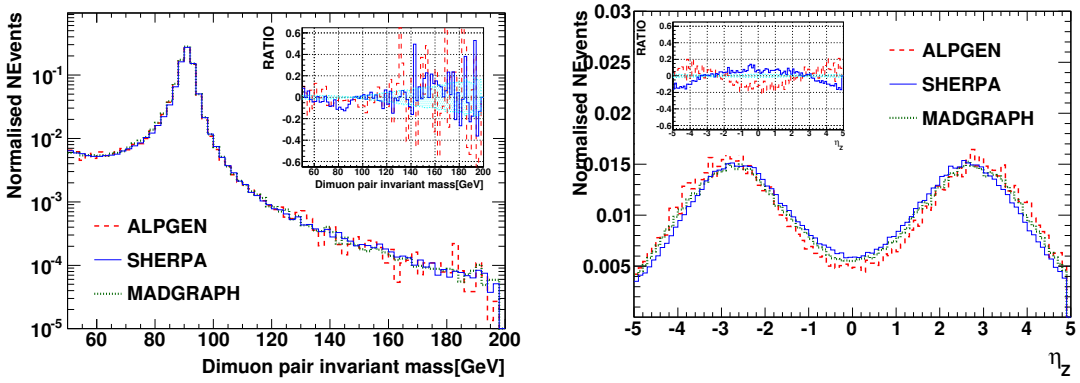


Figure 5.46: Invariant mass of OS muons (left) ($M_{\mu\bar{\mu}} > 50$ GeV) and η_Z (right) distributions for the inclusive Z($l\bar{l}$)+jets production ($l = e, \mu, \tau$) simulated with Alpgen, Sherpa and Madgraph at detector level. Ratio: relative difference of Alpgen and Sherpa with respect to Madgraph.

Therefore, the difference of the invariant mass of OS muons and P_T^{μ} depends on the P_T spectrum of the Z. Both generators have different ISR radiation and the γ^* implementation at ME. This has an impact on the slope of the invariant mass distribution between 66-86 GeV. Differences can also arise from the default settings of MC@NLO (γ^* interference).

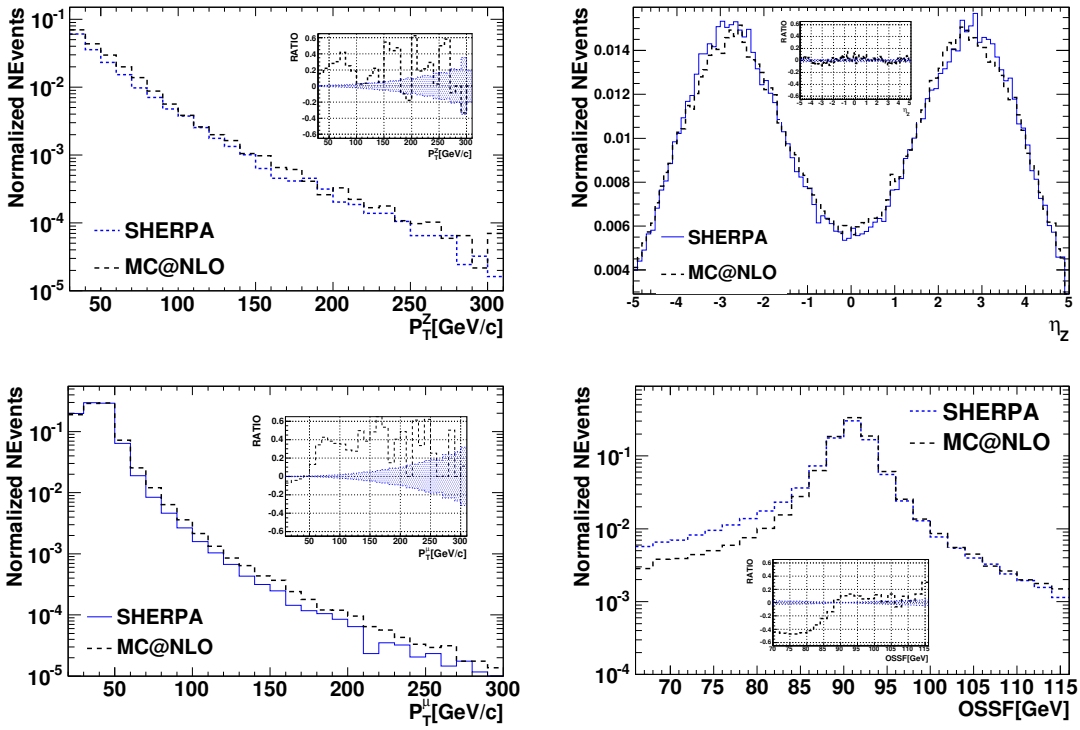


Figure 5.47: P_T^Z (up-left), η_Z (up-right), P_Z^μ (down-left) and OSSF (down-right) simulated with MC@NLO and Sherpa. Invariant mass of OS muons ($M_{\mu\bar{\mu}} > 66$ GeV) calculated with MC@NLO and Sherpa for Z+1jet. Ratio: relative difference of MC@NLO with respect to Sherpa.

5.3 Summary

The goal of this chapter was to spot the differences between various MC generators in the typical observables used in physics analysis, like missing transverse energy, jet P_T and pseudo-rapidity (η), jet multiplicity, the summation of all jet E_T(sumET or HT), leptons P_T and reconstruction parameters in the CMS detector. The MC generators are used as a 'black box' without deep investigation of the sources of discrepancies. Four generators have been tested: Sherpa, Alpgen(Pythia), Madgraph(Pythia) and MC@NLO(Herwig) for five SM processes: Z+jets, W+jets, $t\bar{t}$+jets, VV+jets and A^*+jets production at $\sqrt{s}$=10 TeV. The comparison has been done at generator level and using CMS detector reconstruction simulation.

The cross sections of the considered channels are in a good agreement; $\sim 6\%$ for Z/γ^*+jets, W+jets and $\sim 11\%$ for $t\bar{t}$+jets. In general the P_T distribution of jets, leptons and E_T^{Miss} agree within $10 - 20\%$ for Alpgen and Madgraph predictions. However, some important deviations have been observed for Sherpa predictions. The major difference, up to $40 - 50\%$, between Alpgen-Madgraph and Sherpa occurs in the jets P_T tail above 100 GeV and as a result, in sumEt. The jets are softer and more central for Alpgen and Madgraph as compared to Sherpa, which is in turn in good agreement with MC@NLO. Similarly, the tail of the E_T^{Miss} distribution for Sherpa is harder than for Alpgen and Madgraph. Alpgen also tends to produce lower jet multiplicities for $N_j \geq 4$, i.e. in the region where state radiation jets contribute. The angular characteristics of jets and muons are in a good agreement.

The effect of renormalization and factorization scales, PDF, hadronization models are small compared with the major difference between Alpgen and Sherpa in the exclusive samples ($Z(\mu\bar{\mu})$+jets and $W(\mu\nu)$+jets without underlying event). The differences are larger for the inclusive samples ($Z(l\bar{l})$+jets and $W(l\nu)$+jets where $l = e, \mu, \tau$ with underlying event). The observed deviations can be traced to the differences in the matrix element and parton shower matching. The simulated number of partons and their rescaling factor with matching show differences at MC level simulation. The deviations between Alpgen, Madgraph and Sherpa can be interpreted as systematic uncertainties for a particular observable which can be propagated to the final number of selected events and used in the statistical model for hypothesis testing. This is done in the next chapter.

Chapter 6

Systematic Uncertainties for SUSY Searches in CMS

In the previous chapter the difference between various Monte-Carlo generators for variables used in SUSY searches have been studied. The differences were found to be appreciable in MC and CMS detector simulations, but it did not answer the question of how much background is expected in the different SUSY searches *after* applying all cuts. In this chapter typical cuts of the hadronic and multi-leptonic SUSY searches have been applied and evaluated how much SM background one gets for the different search channels. The variation of background estimates from the various generators can be considered as *theoretical systematic uncertainties*. These in turn are important, if one wants to calculate and/or combine the statistical significance of the various channels. This thesis uses the mSUGRA benchmark points (see Section 6.3), to illustrate the discovery reach in the CMS detector.

6.1 SUSY Searches and MSSM signals

At LHC energies squark and gluino production are expected to dominate the SUSY particle production. Since squarks and gluinos are colored particles they are predicted to be produced via strong interaction with a cross-section similar to the QCD processes at the same scale. To a good approximation the production cross section depends only on the masses of the particles and are therefore independent of the SUSY model. In contrast the decays of SUSY particles depend critically on the details of the SUSY model under considerations, such as mass spectra and branching ratios. However, some generic properties can be identified. Assuming R-parity is conserved (see Section 2.2.1), the decay chains end with the lightest SUSY particle which escapes undetected, thus leading to missing transverse energy. Assuming that squarks and gluinos are heavy, long decay chains are expected, with several jets and/or leptons involved. The typical LO squark and gluino production is shown in Figure 6.1. A generic search of the type described above can be constructed using sample event selections to categorize those with large E_T^{Miss} by the number of leptons

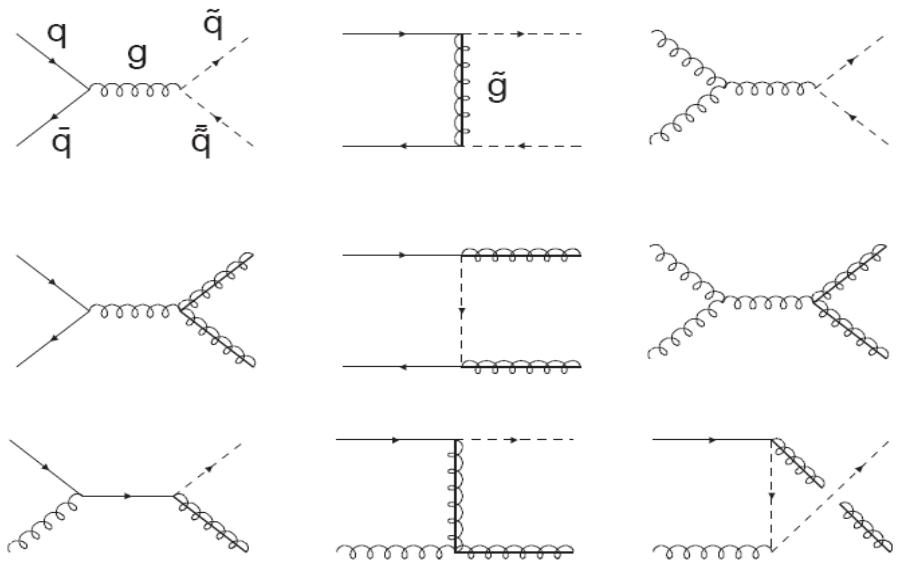


Figure 6.1: Examples of production diagrams for squarks and gluinos. The leading order diagrams involving $\tilde{q}\tilde{q}$ and $\tilde{g}\tilde{g}$ production and associated $\tilde{g}\tilde{q}$, respectively.

and jets (see Section 2.2.4). An excess of events from low mass SUSY benchmark signals can be observed in many event topologies with as little as $\mathcal{L}$=100 pb^{-1} of data, taken at $\sqrt{s}$=10 TeV. It is important to understand the theoretical uncertainties in the SM predictions for SUSY discovery. The search strategy of the CMS collaboration is to focus on simple and robust event signatures, which are common to a wide variety of models discussed in the next sub-sections.

Search for SUSY at the LHC

SUSY searches at the LHC are based on the assumption that the masses of super-partners indeed are in the region of $\sim$ 1 TeV (see Section 2.2.2) so that they might be created on mass shell with cross sections big enough to distinguish them from the background of ordinary SM particles. Calculation of the background in the framework of the SM thus becomes essential since the secondary particles in all cases will be the same. There are many possibilities to create super-partners at hadron colliders, besides the usual annihilation channel, and numerous processes of gluon fusion, quark-anti-quark and quark-gluon scattering. The maximal cross-section of the order of few picobarn is obtained in the process of gluon fusion.

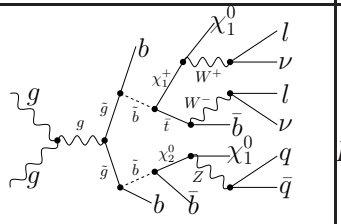
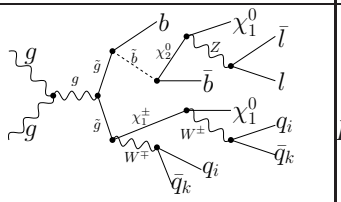
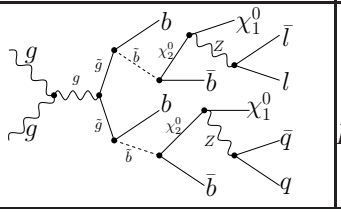
As a rule all super-partners are short lived and decay into ordinary particles and the lightest super-particle [133]. The main decay modes of super-partners, i.e the experimental manifestation of SUSY at LHC, are presented in Table 6.1. The main difference from the SM background process can be obtained by missing transverse energy from escaping neutralinos.

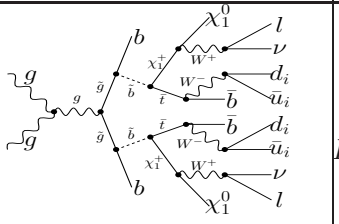
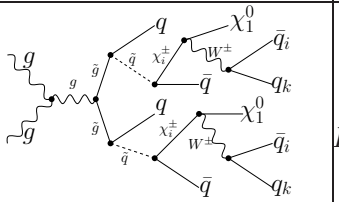
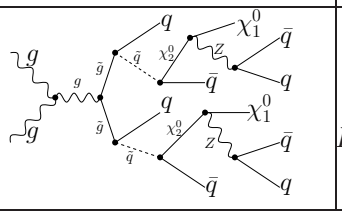
As discussed for the SM process before, the missing transverse energy carried away

Table 6.1: Creation of super-partners and the main decay modes [133].

Creation	The main decay modes	Signature
• $\tilde{g}\tilde{g}, \tilde{q}\tilde{q}, \tilde{g}\tilde{q}$	$\left.\begin{array}{l}\tilde{g} \to q\bar{q}\tilde{\chi}_1^0 \\ \quad q\bar{q}'\tilde{\chi}_1^\pm \\ \quad g\tilde{\chi}_1^0\end{array}\right\} m_{\tilde{q}} > m_{\tilde{g}}$ $\left.\begin{array}{l}\tilde{q} \to q\tilde{\chi}_i^0 \\ \tilde{q} \to q'\tilde{\chi}_i^\pm\end{array}\right\} m_{\tilde{g}} > m_{\tilde{q}}$	E_T^{Miss} + multijets (+leptons)
• $\tilde{\chi}_1^\pm\tilde{\chi}_2^0$	$\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0\ell^\pm\nu,\ \tilde{\chi}_2^0 \to \tilde{\chi}_1^0\ell\ell$ $\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0 q\bar{q}', \tilde{\chi}_2^0 \to \tilde{\chi}_1^0\ell\ell,$	trilepton + E_T^{Miss} dileptons + jet + E_T^{Miss}
• $\tilde{\chi}_1^+\tilde{\chi}_1^-$	$\tilde{\chi}_1^+ \to \ell\tilde{\chi}_1^0\ell^\pm\nu$	dilepton + E_T^{Miss}
• $\tilde{\chi}_i^0\tilde{\chi}_i^0$	$\tilde{\chi}_i^0 \to \tilde{\chi}_1^0 X, \tilde{\chi}_i^0 \to \tilde{\chi}_1^0 X'$	dilepton + jet + E_T^{Miss}
• $\tilde{t}_1\tilde{t}_1$	$\tilde{t}_1 \to c\tilde{\chi}_1^0$ $\tilde{t}_1 \to b\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \to \tilde{\chi}_1^0 q\bar{q}'$ $\tilde{t}_1 \to b\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \to \tilde{\chi}_1^0\ell^\pm\nu,$	2 non-collinear jets + E_T^{Miss} single lepton + $E_T^{Miss} + b's$ dilepton + $E_T^{Miss} + b's$
• $\tilde{l}\tilde{l}, \tilde{l}\tilde{\nu}, \tilde{\nu}\tilde{\nu}$	$\tilde{\ell}^\pm \to \ell^\pm\tilde{\chi}_i^0, \tilde{\ell}^\pm \to \nu_\ell\tilde{\chi}_i^\pm$ $\tilde{\nu} \to \nu\tilde{\chi}_1^0$	dilepton + E_T^{Miss} single lepton + E_T^{Miss}

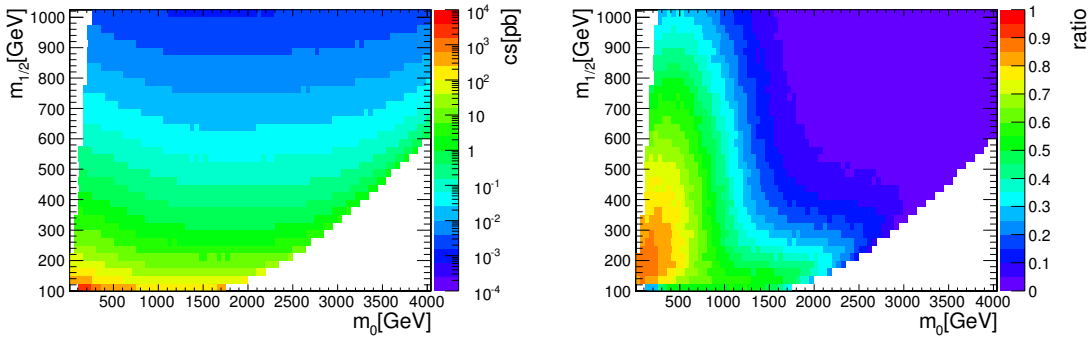
Table 6.2: Creation of a pair of gluinos with further cascade decay [133].

Process	final states
	2ℓ 2ν $6j$ E_T^{Miss}
	2ℓ $6j$ E_T^{Miss}
	2ℓ $6j$ E_T^{Miss}

Process	final states
	2ℓ 2ν $8j$ E_T^{Miss}
	$8j$ E_T^{Miss}
	$8j$ E_T^{Miss}

by a heavy particle with the mass of order 100 GeV is essentially different from the processes with a neutrino in the final state. In hadron collisions the super-partners are always created in pairs and then further quickly decay creating a cascade of ordinary quarks and/or leptons. The standard SUSY searches can be categorized

by lepton multiplicity of the corresponding SUSY signature: Zero-lepton plus missing transverse energy (hadronic SUSY searches), single and di-lepton plus missing transverse energy and tri-lepton search. Figure 6.2(a) presents the inclusive SUSY cross section at $\sqrt{s} = 10$ TeV for $tan\beta = 10$, $A_0 = 0$, $\mu > 0$ in m_0-$m_{1/2}$ plane. The cross section decreases with increasing mass of SUSY particles and shows a strong dependence on the value of $m_{1/2}$. Large values of the cross section are observed at m_0, $m_{1/2} < 200$ GeV, where the SUSY masses are low.



(a) Total SUSY cross section (b) $pp \rightarrow \tilde{g}\tilde{g}, \tilde{g}\tilde{q}, \tilde{q}\tilde{q}$

Figure 6.2: Total SUSY cross section (a) at the LHC for $tan\beta = 10$, $A_0 = 0$, $\mu > 0$ and the corresponding cross section fraction of gluino and squark production (b) [44].

The cross section for $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$ and $\tilde{q}\tilde{q}$ is shown as function of m_0 and 1/2 in Figure 6.2(b). The SUSY discovery reaches with the corresponding experimental constraints (see Section 2.2.4) in the m_0-m1/2 plane are discussed in section 6.6.

6.2 Standard Model Background

The search for SUSY can be simply done as a standard counting experiment, where numerical excess is searched for over the SM background. In order to comprehend the deviation between generators, the SUSY signal topologies over the SM background need to be understood. In the following, the SM background for hadronic and multi-leptonic SUSY searches are discussed with the corresponding SUSY signal topologies.

6.2.1 Hadronic Searches - Jets+E_T^{Miss}

If SUSY exists CMS has the potential to discover it in early LHC running with E_T^{Miss}+Jets search strategies. Such strategies, as outlined in the CMS Physics TDR2 [51], search for events with high E_T^{Miss} and several high-energy jets, along with topological cuts to reject the background from QCD events and a lepton veto to reject W, Z, and $t\bar{t}$ events.

Most top quarks are produced as $t\bar{t}$ pairs and the dominant production mechanism is gluon-gluon fusion($\sim$ 90%). Quark pair annihilation is $\sim$ 10%. Within the SM the top quark decays almost exclusively to a W boson and a b quark (see Figure 6.5). The decays of the $t\bar{t}$ system are then classified according to the decays of the W systems as dileptonic, semi-leptonic or fully hadronic with branching fraction of $\sim$ 11% for the di-lepton and $\sim$ 44.4% for the semi-leptonic and hadronic mode. Neutrinos arising from leptonic W decays production, give a significant missing transverse energy with high P_T hadronic activity for $t\bar{t}$ events. Monte-Carlo predictions show that the $t\bar{t}$ contribution to the SM background in the non-lepton mode is about $\sim$ 65% [51, 134]

An irreducible background from these searches is the Z boson production in conjunction with energetic jets, where the Z decays to neutrinos giving real missing transverse energy. The Z$\rightarrow \mu\bar{\mu}$ events can be identified. The calorimetric deposits of leptons are removed from the event, the remaining events a good approximation for Z$\rightarrow \nu\bar{\nu}$, as discussed in section 5.2.2. However, this method suffers from the small number of Z$\rightarrow \mu\bar{\mu}$ events expected during the early LHC running, because of the lower branching ratio of Z to charged leptons with respect to Z to neutrinos. Alternatively, one can study the transverse missing energy resolution from expected energy balanced in γ^*+jets events [51, 135].

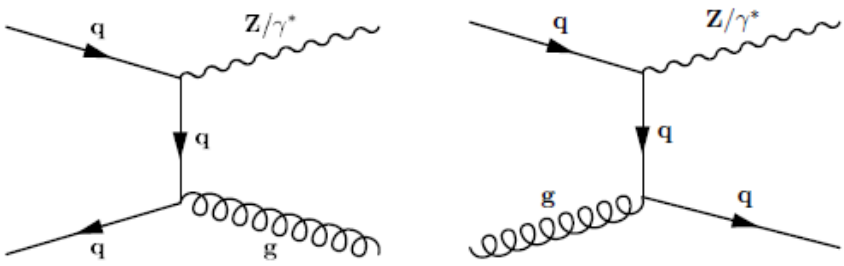


Figure 6.3: Typical diagrams for the Z+jets production [136].

6.2.2 Multi-Leptonic Searches - Jets+E_T^{Miss}+NLeptons

Multi-leptonic SUSY searches aim to reject QCD events and events with leptons from W, Z/γ^*, WW, ZZ, ZW and $t\bar{t}$ events. Due to the lepton requirement, it is relatively to reject QCD background. However, any reaction which produce lepton signatures is a potential background for the multi-leptonic SUSY searches, as discussed in the following sub-sections.

SM Background with single Prompt Muons

W+jets production is one of the most important backgrounds for the single lepton SUSY search, since the neutrinos leads to missing energy in the detector. Typical

diagrams are shown in Figure 6.4. Although there is only one prompt muon produced at the hard scale, there can also be fake muons produced after hadronization (see Appendix A.1). Events with fake muons can also contribute to the di- and tri-muon backgrounds.

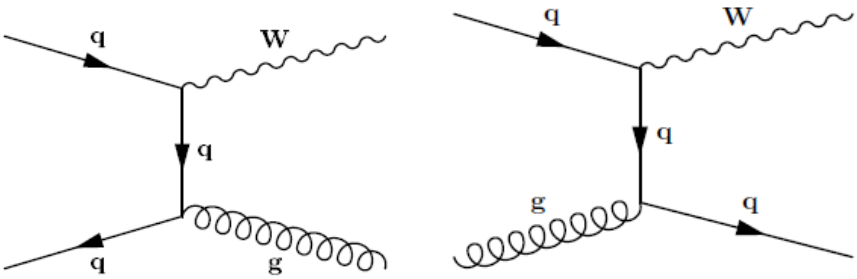


Figure 6.4: Typical diagrams for the W+jets production [136].

SM Background with two Prompt Muons

The most important background of this kind are Z boson and $t\bar{t}$ production with two prompt muons in the final state. The Z/γ^*+jets production in Figure 6.3 includes Drell-Yan processes and the associated production of a Z boson with light quarks (u,d,c,s) and heavy quarks (c,b). The $t\bar{t}$ events, illustrated in Figure 6.5, are produced with heavy quarks, since the t-quark decays exclusively to a W boson and b-quark with escaping energy from the neutrinos produced in leptonic decays of the W boson. In addition to the produced prompt muons additional muons can be produced via parton showers and hadronization, so that these channels contribute to the SM background of the tri-muon SUSY search as well. The background from WW+jets di-boson production has a smaller cross-section and is less important.

SM Background with three Prompt Muons

For the tri-muon search it is necessary to understand the di-boson productions (VV+jets). One of the important contribution comes from the WZ/γ^* production. The SM partner process of the direct $\chi_2^0\chi_1^\pm$ pair production gives three prompt isolated muons in the final state and forms the major irreducible background of the tri-muon search [44]. The ZZ production with the subsequent decay of Z boson to muons gives four prompt muons in the final state and appears as a trimuon final state, if one of the muons is produced outside of the detector acceptance or is rejected by the muon selection. Most of these events can be suppressed by a veto on the Z boson di-muon invariant mass. Figure 6.6 shows typical diagrams for the WZ and ZZ production. Due to the small cross section, these channel are not important for the single and di-muon SUSY searches.

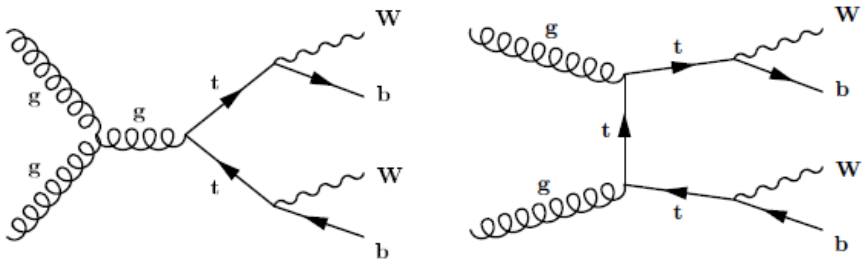


Figure 6.5: Typical diagrams for the $t\bar{t}$ production [136].

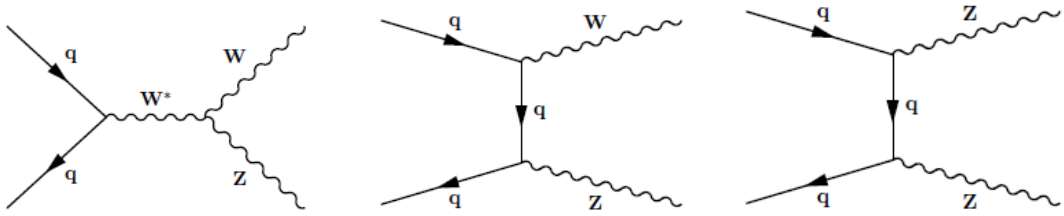


Figure 6.6: Typical diagrams for the ZW and ZZ diboson production [136].

To summarize: the gauge boson production is relatively well understood for Tevatron studies and the extrapolation to LHC energies looks promising for the analyses. However, the differences and extrapolation to LHC energies of the various MC simulations and the rate of fake muons from jets depend on numerous factors including uncertainties in soft gluon radiation and gluon splitting, jet matching, fragmentation and reconstruction of muon isolation efficiency. These factors introduce significant systematic uncertainties in the estimation of these SM backgrounds.

6.3 Monte Carlo Data Samples

The SUSY signal and the SM background have been analyzed using the official Madgraph data samples available in the CMS database. The Alpgen, Sherpa and MC@NLO samples have been produced with the official CMS software release CMSSW2_1_X and CMSSW2_2_X in the Summer08, Fall08 and Winter09 CMS production for physics at $\sqrt{s} = 10$ TeV (see Tables 5.4, 5.6 and 5.5). The summary of the SM background channels has been presented in the previous chapter. The summary of the signals samples of the corresponding benchmark points is presented in Table below:

SUSY Data Samples(Summer08)				
SUSY Signals	Xsec(pb)	Events	Datasets	m_0,$m_{1/2}$,tan β, A_0, sgn μ
LM0	110	200K	SUSY-LM0-sftsht-v1	200,160, 10, −400, 1
LM1	16.06	202K	SUSY-LM1-sftsht-v1	60,250, 15, 0, 1
LM2	2.42	130K	SUSY-LM2-sftsht-v1	185, 350, 35, 0, 1
LM3	11.79	113K	SUSY-LM3-sftsht-v1	330, 240, 20, 0, 1
LM4	6.7	80K	SUSY-LM4-sftsht-v1	210, 285, 10, 0, 1
LM5	1.94	171K	SUSY-LM5-sftsht-v1	230, 360, 10, 0, 1
LM6	1.28	134K	SUSY-LM6-sftsht-v1	85, 400, 10, 0, 1
LM7	2.9	63K	SUSY-LM7-sftsht-v1	3000, 230, 10, 0, 1
LM8	2.86	211K	SUSY-LM8-sftsht-v1	500, 300, 10, −300, 1
LM9	11.58	202K	SUSY-LM9-sftsht-v1	1450, 175, 50, 0, 1
LM9t175	11.09	214K	SUSY-LM9-sftsht-v1	1450, 175, 50, 0, 1
LM10	0.065	202K	SUSY-LM10-sftsht-v1	3000, 500, 10, 0, 1
LM11	3.24	208K	SUSY-LM11-sftsht-v1	250, 325, 35, 0, 1

The number of expected events for signal and backgrounds have been calculated with leading-order cross sections. All mSUGRA benchmark points and SM background processes have been simulated with the full and fast detector simulation. The full simulation has used the complete CMS detector in the ideal-tag configuration.

The SUSY signal has been calculated in a few steps. With the given mSUGRA parameters (m_0, $m_{1/2}$, $tan\beta$, A_0, $\mu > 0$) the sparticle mass spectrum has been calculated at the electroweak scale using the renormalization group equations implemented in the SOFTSUSY (v. 2.18) package. The radiative corrections to decays of sparticles have been calculated with the SUSYHIT (v1.3) code. The Monte Carlo events have been simulated with PYTHIA (v6.4) using the CTEQ6l PDF's.

6.4 Hadronic SUSY Searches in CMS

This section presents search strategies for a possible discovery of SUSY signatures at the LHC using hadronic events. The exclusive and inclusive hadronic SUSY searches have been analyzed. The difference between those two searches arises from the used selection cuts and observables. In the following sub-sections, these selection cuts and observables are discussed for the corresponding hadronic SUSY searches.

6.4.1 Exclusive Hadronic SUSY Search

A new approach to SUSY searches with di-jet and n-jet events was proposed in [137, 138]. The used observable α_T was first defined for di-jet events and extended to n-jet events later [139]. In hadronic SUSY searches it is based on the assumption that squarks are pair produced and subsequently decay directly to a quark and the χ_1^0. This approach is most promising for points in a SUSY parameter space where squarks have large branching ratios to decay directly to the LSP. This requires the gluino to

be heavier than the squarks, thus avoiding cascade decays of squarks via the gluino. The event topology consists of two or more high pt jets and two invisible neutralinos which lead to a missing transverse energy signature. The main background processes for this topology are QCD di-jet/n-jet events and Z+jets where Z decays into two neutrinos.

Kinematics in the Transverse Plane and α_T

The α_T observable is calculated using the variables discussed in the following. H_T has been defined for the search of n jets(j) i as:

$$H_T = \sum_{i=1}^{n} E_T^{j_i}. \tag{6.1}$$

It sets the scale of the interaction in the transverse plane. It is assumed that $E_T^{j_i}$=$P_T^{j_i}$ holds for final state objects. The missing HT is defined as

$$\not{H}_T = -\sum_{i=1}^{n} P_T^{j_i} = -h_T, \tag{6.2}$$

where $P_T^{j_i}$ is the projection of the momentum in the transverse plane. It describes the scale and direction of the missing transverse energy as defined by the n object system. The transverse mass of the system is the sum of all calorimeter measurements:

$$M_T = \sqrt{(\sum_{i=1}^{n} E_T^{j_i})^2 - (\sum_{i=1}^{n} P_x^{j_i})^2 - (\sum_{i=1}^{n} P_y^{j_i})^2} = \sqrt{H_T^2 - |\not{H}_T|^2}. \tag{6.3}$$

It shows a system where energy and momentum are conserved and all constituent objects are included and have been measured perfectly. For such a perfect system $|\not{H}_T|$=E_T^{Miss}=0, and M_T is at its maximal value, H_T.

α_T is defined as the ratio of the P_T of the second hardest jet and the invariant mass formed from the two hardest jets

$$\alpha_T = \frac{min(E_T^{j_1}, E_T^{j_2})}{M_T^{j_1,j_2}}. \tag{6.4}$$

Equation 6.4 can be re-written in terms of the variables shown in Eq. 6.1 and 6.2 as:

$$\alpha_T = \frac{\frac{1}{2}(E_T^{j_1} + E_T^{j_2} - |E_T^{j_1} - E_T^{j_2}|)}{M_T^{j_1,j_2}} = \frac{\frac{1}{2}(H_T - |\Delta H_T|)}{\sqrt{H_2^2 - |\not{H}_T|^2}}. \tag{6.5}$$

In a perfect di-jet system, ΔH_T=0, $|\not{H}_T|$=0 and so α_T=$\frac{1}{2}$.

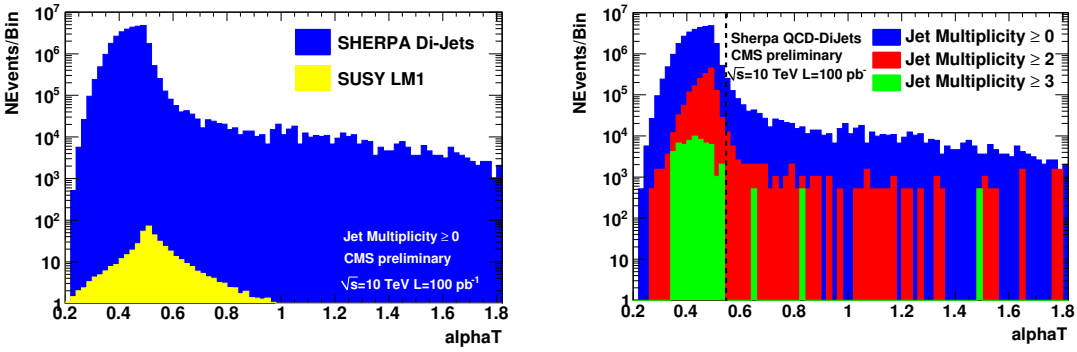


Figure 6.7: The α_T distribution for different jet multiplicity selections using Sherpa. Left: α_T distribution with Sherpa-Dijet and SUSY LM1 point. Right: α_T distribution with different jet multiplicity. Dashed line shows a typical selection cuts ($\alpha_T > 0.55$) in the exclusive hadronic search.

However all possible combinations of 2→ n jets need to be considered and the one that minimizes ΔH_T has to be selected. Hence the definition of α_T is projected for n-jet:

$$\alpha_T = \frac{\frac{1}{2}(H_T - |\Delta H_{T_{(n)}}|)}{\sqrt{H_2^2 - |\not{H}_T|^2}} \tag{6.6}$$

where $\Delta H_{T_{(n)}}$ is the minimum ΔH_T obtained by considering all possible 2 → n jet combinations [139].

The α_T variable is of phenomenological interest since it constitutes a strong discriminator in the exclusive hadronic SUSY searches for separating the signal from QCD background. In the following the variables, such as α_T, for the important SM background in the exclusive hadronic SUSY search are discussed.

Data Samples

The datasets (see Sections 5.1.1 and 6.3) used in this analysis are summarized below:

- **QCDjets** - Madgraph and Sherpa
- $t\bar{t}$**+jets** - Madgraph, Alpgen, Sherpa and MC@NLO
- **Z+jets (Z** → $\nu\bar{\nu}$ **or Zinvisible), W+jets** - Madgraph, Alpgen and Sherpa
- **SUSY LM** samples - SoftSusy+SusyHit+Pythia

Physics Objects Selection

- **Jet definitions:** As jet collection, SiSCone ΔR=0.5 with L1 Offset, L2 and L3 relative and absolute jet corrections (see Section 3.7.2) have been used. Also a selection on the electromagnetic fraction (F_{em}) has been applied.
- **Missing transverse energy and jets:** The corrected calorimeter missing transverse energy is considered for this analysis (see Section 3.7.5) together with the MHT ($\not{H}_T$) observable.
- **Leptons:** Global muons (see details in A.1) and loose electrons (see details A.2) have been used for the event selections.

The event selection for the exclusive hadronic search in CMS is splitted in four steps [140]:

Event Selection

- **Pre-Selection:** $N_{jets} \geq 2$, $P_{T_j} > 50$ GeV, $|\eta_j| < 5$ with $F_{em} < 1$, $P_{T_e} > 10$ GeV, $|\eta_e| < 2.5$ and $P_{T_\mu} > 10$ GeV, $|\eta_\mu| < 2.4$, $E^{Miss}_{Calo} > 50$ and MHT > 50 GeV
- **Lepton Veto:** $N_{leptons}$=0, $|d^e_0| \leq 0.2$, electron relative isolation< 0.5, $|d^\mu_0| \leq 0.2$, muon relative isolation< 0.1, $N_{Hits} \geq 11$, χ^2/dof< 10.
- E^{Miss}_{Calo} **and MHT:** MHT > 200 GeV and $E^{Miss}_{Calo} > 200$ GeV, $P^{1st}_{T_j} > 100$ GeV, $P^{2nd}_{T_j} > 50$ GeV $|\eta_j| < 3$ with $0.1 < F_{em} < 0.9$
- H_T **and** α_T**:** $H_T > 350$ GeV and $\alpha_T > 0.55$

6.4.2 Results for Exclusive Hadronic SUSY Search

The event selection introduced in the last sub-section has been applied on the corresponding SM background processes for this search. The effect of the generators and their systematics have been studied with the following parameters: the hard scale choices (factorization/renormalization), initial and final state radiations (ISR-FSR), the matching scales with the different matching methods.

In Table 6.3, final number of predicted events after the different event selection steps of the exclusive hadronic SUSY search for the $t\bar{t}$+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$ are presented. To understand these differences the following observables have been shown for the $t\bar{t}$+jets production in Figs. 6.8, 6.9 and 6.10: missing transverse energy from the calorimeters (E^{Miss}_{Calo}), missing HT (MHT), α_T and effective mass (missing transverse energy plus jets E_T).

Table 6.3: Final number of predicted events after the different event selection steps of the exclusive hadronic SUSY search for $t\bar{t}$+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Alpgen, Sherpa, Madgraph and MC@NLO are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 317$pb.

Generators	Datasets	Pre-Selection	Lepton Veto	$E_{Calo}^{Miss} > 200$	$\alpha_T > 0.55$
ALPGEN	PtJet> 30 GeV	8862	4215	74	31
SHERPA	PtJet> 30 GeV	11175	5687	136	53.29
	PtJet> 40 GeV	11167	5641	129	52.3
	Scale Up	11209	5648	131	53.14
	Scale Down	11072	5665	132	49.53
MADGRAPH	PtJet> 10 GeV	9548	4830	80	35.45
	PtJet> 30 GeV	9565	5403	87.8	39.53
	PtJet> 40 GeV	9636	4574	82	37
	Scale Up	9515	4522	80.8	35.26
	Scale Down	9603	5159	82.6	39.28
	Smaller ISR	9711	5350	90.2	41.81
	Larger ISR	9638	5083	88.7	39.15
MC@NLO	NLO	10905	5388	120	47.55

One observes a significant difference between the generators especially Sherpa and MC@NLO predicts significantly more background than Alpgen. As expected from the previous chapter the Zinvisible+jets production shows very similar results for Sherpa and Madgraph. However, $t\bar{t}$+jets and W+jets predictions from Sherpa/MC@NLO are higher than predicted from Alpgen/Madgraph because of larger P_T tails.

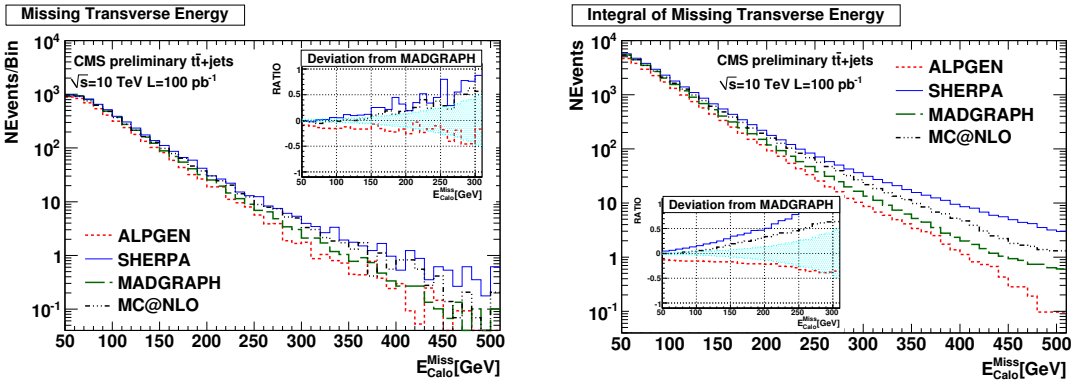


Figure 6.8: Differential (left) and integral (right) E_{Calo}^{Miss} distribution for $t\bar{t}$+jets production simulated with Alpgen, Sherpa, Madgraph and MC@NLO after applying the lepton-veto selection step. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph. The (blue) shaded area represents the statistical uncertainty.

For $t\bar{t}$+jets, Sherpa and MC@NLO show a harder missing transverse energy spectrum at increasing energies (see Section 5.2.2). The integrals of E_{Calo}^{Miss} and MHT distribution show 50% deviation after applying the final selection step (> 200 GeV). In addition to that another difference occurs up to $30-40\%$ in the jets P_T (see Section 5.2.3) between Alpgen/Madgraph and Sherpa/ MC@NLO and as a result, in effective mass. Thus, these differences have been clearly observed at the predictions of events after final selection step. The various settings of Sherpa and Madgraph predictions with respect to their default values are in good agreement ($\sim$ below 10%). The jet multiplicity selection does not change the selected number of predicted events since all generators agree up to 8 jets (see Fig. 5.33). Madgraph shows a harder spectrum for increasing α_T values (> 0.55) because of higher jet multiplicity at the ME calculation.

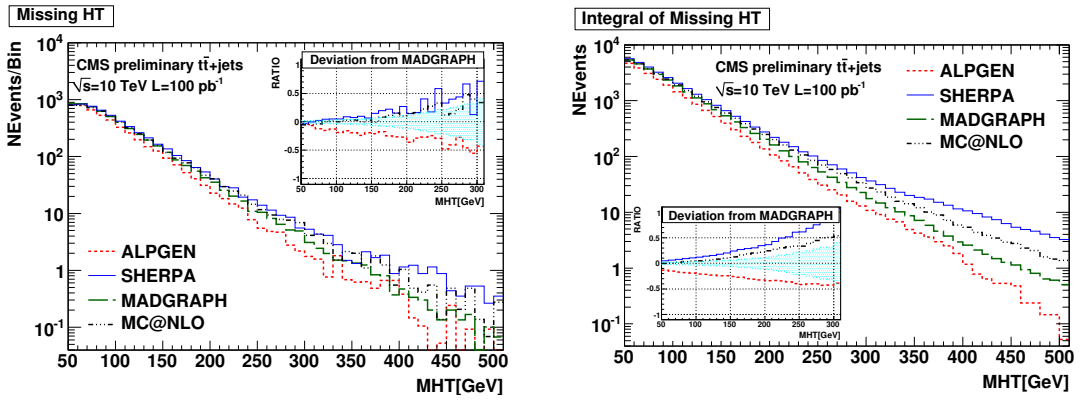


Figure 6.9: Differential (left) and integral (right) MHT distribution for $t\bar{t}$+jets production simulated with Alpgen, Sherpa, Madgraph and MC@NLO after applying the lepton-veto selection step. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

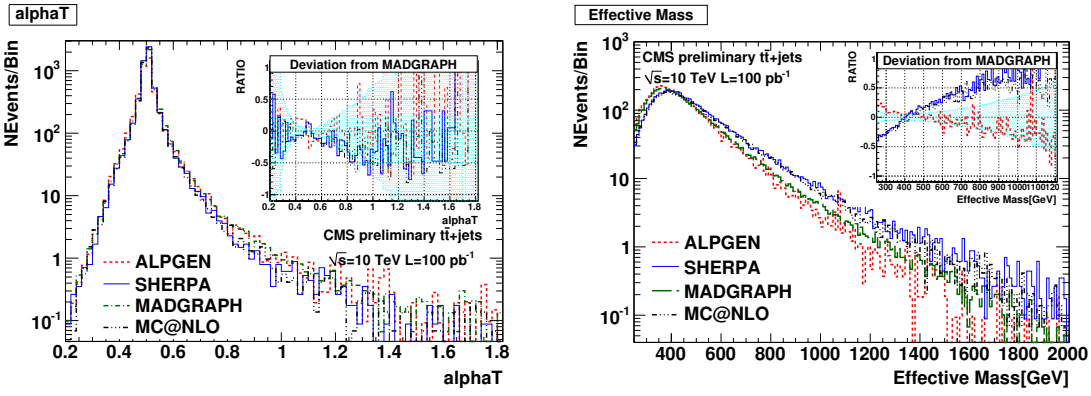


Figure 6.10: α_T (left) and effective mass $M_{\text{eff}} = E_{Calo}^{Miss} + \sum_{i=5}^{n} P_T^{ji}$ (right) distribution for $t\bar{t}$+jets production simulated with Alpgen, Sherpa, Madgraph and MC@NLO after applying the lepton-veto selection step. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

Table 6.4: Final number of predicted events after the different event selection steps of the exclusive hadronic SUSY search for Zinvisible+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Sherpa and Madgraph are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 2000$pb.

Generators	Datasets	Pre-Selection	Lepton Veto	$E_{Calo}^{Miss} > 200$	$\alpha_T > 0.55$
SHERPA	PtJet> 20 GeV	3143	3085	199	30
	Lund	3104	3058	180	27.56
	Scale Up	2727	2675	178	25.8
	Scale Down	3230	3173	211	34
MADGRAPH	PtJet> 30	3160	3076	165	27

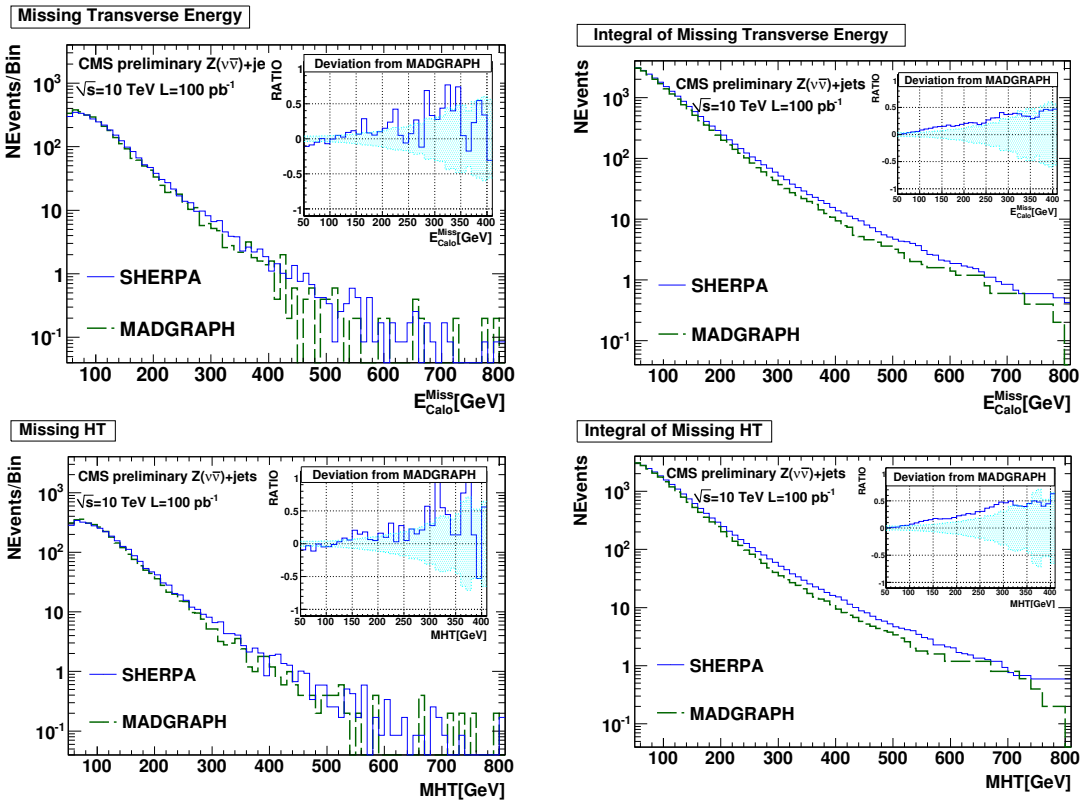


Figure 6.11: Differential (left) and integral (right) E_{Calo}^{Miss} and MHT distribution for Zinvisible+jets production simulated with Sherpa and Madgraph after applying the lepton-veto selection step. Ratio: relative differences of Sherpa with respect to Madgraph.

The Zinvisible+jets production is discussed in Table 6.4 for Madgraph and Sherpa. A good agreement has been generally observed for all observables between Sherpa and Madgraph predictions, as can be seen in Figures 6.11 and 6.12. Although Sherpa shows a harder spectrum than Madgraph, the differences are in the range of internal and statistical uncertainties.

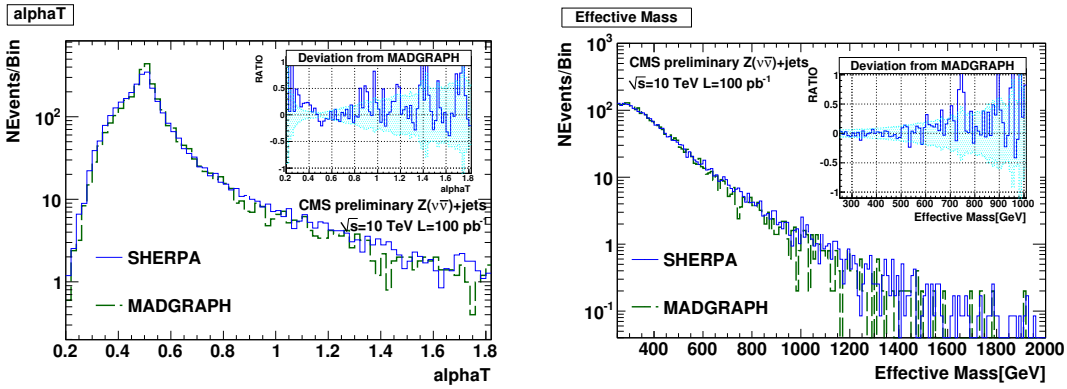


Figure 6.12: α_T (left) and effective mass $M_{\text{eff}} = E_{Calo}^{Miss} + \sum_{i=5}^{n} P_T^{ji}$ distribution for the Zinvisible+jets production simulated with Sherpa and Madgraph after applying the lepton-veto selection step. Ratio: relative differences of Sherpa with respect to Madgraph.

Table 6.5: Final number of predicted events after the different event selection steps of the exclusive hadronic SUSY search for W+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Alpgen, Sherpa and Madgraph are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 40000$pb.

GENERATORS	DATASETS	Pre-Selection	Lepton Veto	$E_{Calo}^{Miss} > 200$	$\alpha_T > 0.55$
ALPGEN	PtJet> 30	30679	18916	341	52.14
SHERPA	ewk1	33326	19080	630	75
	ewk2	30276	17741	480	62
MADGRAPH	PtJet> 30	30568	18400	312	38.78
	PtJet> 30(R)	20679	12319	320	43.52
	Scale Up	32990	21900	317	39.54
	Scale Down	31467	19800	332	68.14

The W+jets production is discussed in Table 6.5. Sherpa usually has a harder spectrum than Madgraph and Alpgen predictions because of the hard scale choices in CKKW method (discussed in Section 5.2) and the jet P_T simulated at parton level. The differential and integral distribution of E_{Calo}^{Miss}, MHT and the differential α_T and effective mass distribution are presented in Figures 6.13 and 6.14, respectively. There are 20%-30% deviation on E_{Calo}^{Miss} and MHT in the region between $150 - 200$ GeV for Sherpa and Alpgen-Madgraph. Alpgen and Madgraph show a very good agreement for the spectra and the final number of selected events. The scale down choice of Madgraph sample shows a good agreement with Sherpa. The Sherpa prediction of the final number of selected events is clearly different than other ones.

The QCDjets production has been analyzed for Sherpa and Madgraph. For Sherpa, the phase space selections on these samples are applied with pt of the jet between 40 and 5000 GeV. QCDjets predictions for Sherpa and Madgraph are presented in Table 6.6.

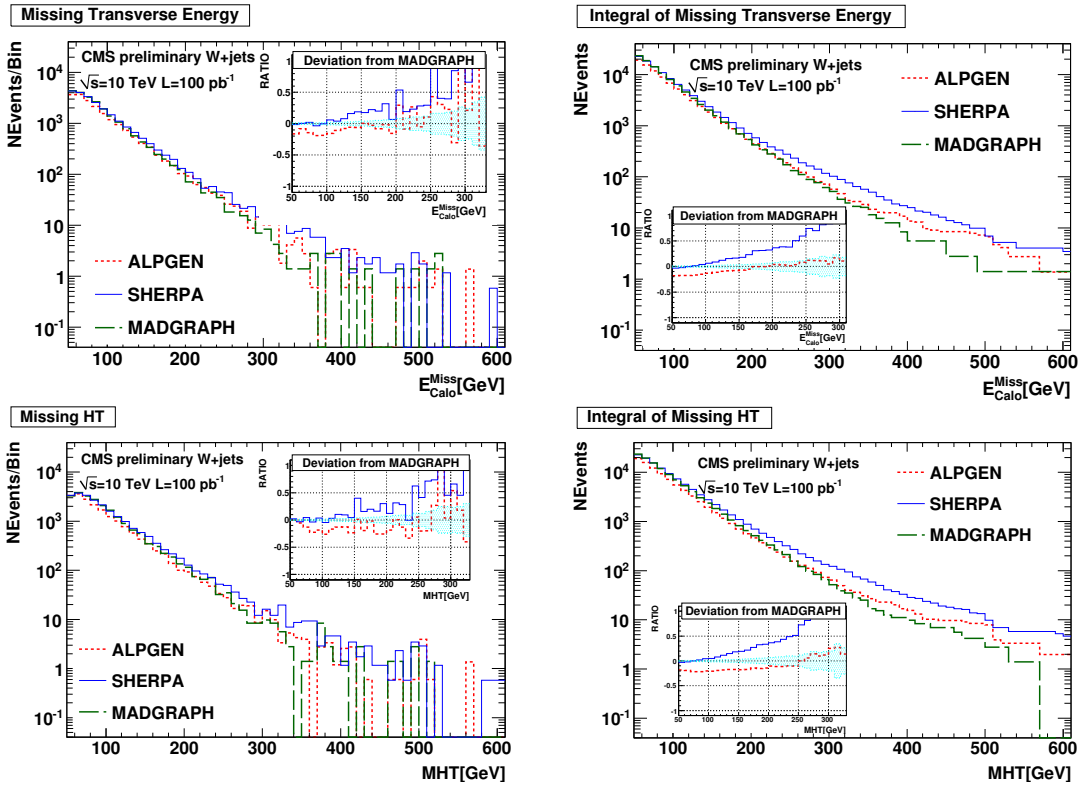


Figure 6.13: Differential (left) and integral (right) E_{Calo}^{Miss} and MHT distribution for W+jets production simulated with Alpgen, Sherpa and Madgraph after applying the lepton-veto selection step. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

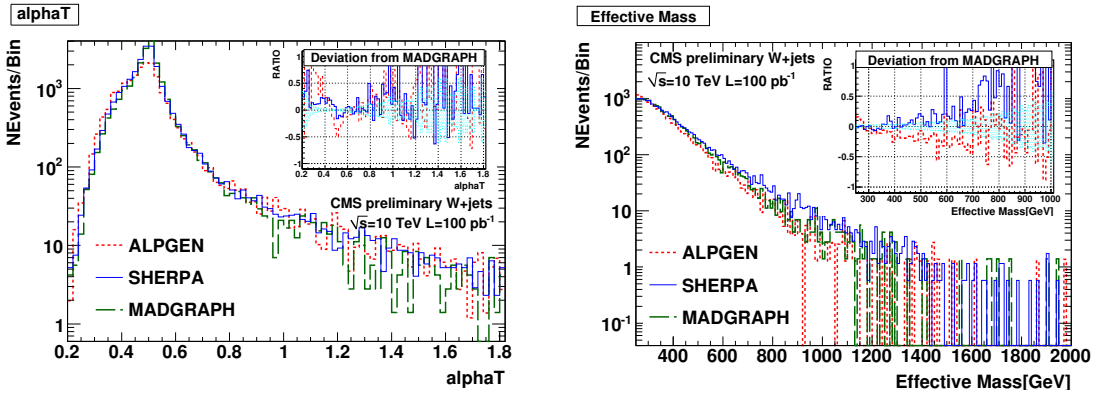


Figure 6.14: α_T (left) and effective mass $M_{\text{eff}} = E_{Calo}^{Miss} + \sum_{i=5}^{n} P_T^{j_i}$ (right) distribution for W+jets production simulated with Alpgen, Sherpa and Madgraph after applying the lepton-veto selection step. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

Table 6.6: Final number of predicted events after the different event selection steps of the exclusive hadronic SUSY search for QCDjets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Sherpa and Madgraph are compared. They are produced with different phase space cuts (P_T and HT of the jet). Sherpa has limited jet multiplicity (Njets < 5) at ME calculation.

Generators	Datasets	Pre-Selection	Lepton Veto	$E_{Calo}^{Miss} > 200$	$\alpha_T > 0.55$
SHERPA	Pt40 − 120 GeV	$3.25 + E06$	$3.19 + E06$	0	0
	Pt120 − 280	307946	303757	794.1	0
	Pt280 − 500	5328	5226	287	0
	Pt500 − 5000	295	287	51	0.1
MADGRAPH	HT100 − 250	$2.5 + E06$	$2.46 + E06$	0	0
	HT250 − 500	$1.7 + E06$	$1.63 + E06$	0	0
	HT500 − 1000	193692	191462	65.1	8.36
	HT1000-Inf	12223	12201	72.4	2.39

The numbers are calculated with the corresponding cross-section predictions of each generators. The differences between pre- and lepton-veto selections are of a comparable range for Sherpa and Madgraph. For the prediction of the final number of events, it would be required to simulate Sherpa QCD samples with higher statistics. However, Madgraph has more jets at matrix element calculation. As shown in Figure 6.15, this effect is clearly visible in the α_T and jet multiplicity distribution.

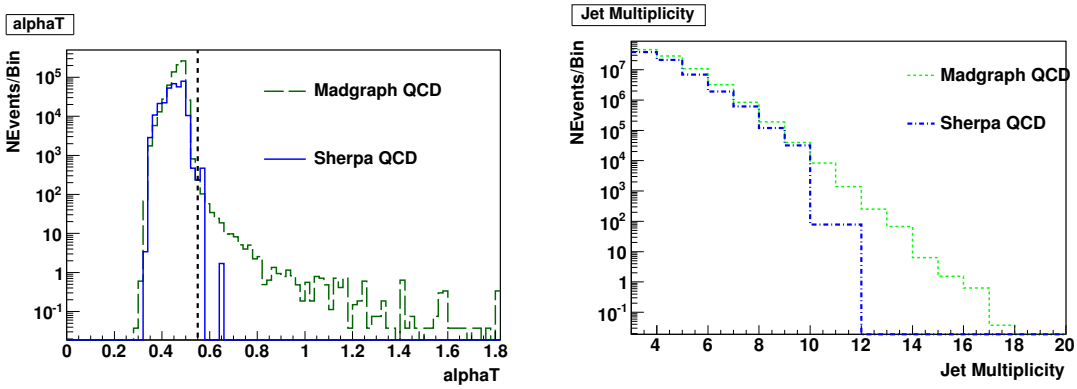


Figure 6.15: α_T and jet multiplicity distribution for $t\bar{t}$+jets production simulated with Sherpa and Madgraph after applying the lepton-veto selection step.

6.4.3 Inclusive Hadronic SUSY Search

This analysis searches for new physics in the inclusive final state with at least three jets, missing transverse energy, and a veto for muons and electrons. In the inclusive search the selection of the SUSY signal from a large amount of QCD background requires another observable depending on the angular correlation between the leading jets and E_T^{Miss}. In addition to that hard selection cuts are applied to the three leading jets. The selection cuts are discussed in detail below and are currently based upon the CMS PTDR2 selection [51, 135] .

$\Delta\phi$(Jets,E_T^{Miss})-Cuts for QCD Background Separation

This part of the analysis aims to reduce the QCD background with large fake E_T^{Miss}. $\Delta\phi$ is defined as the difference in azimuthal angle ϕ between E_T^{Miss} and the four leading jets.

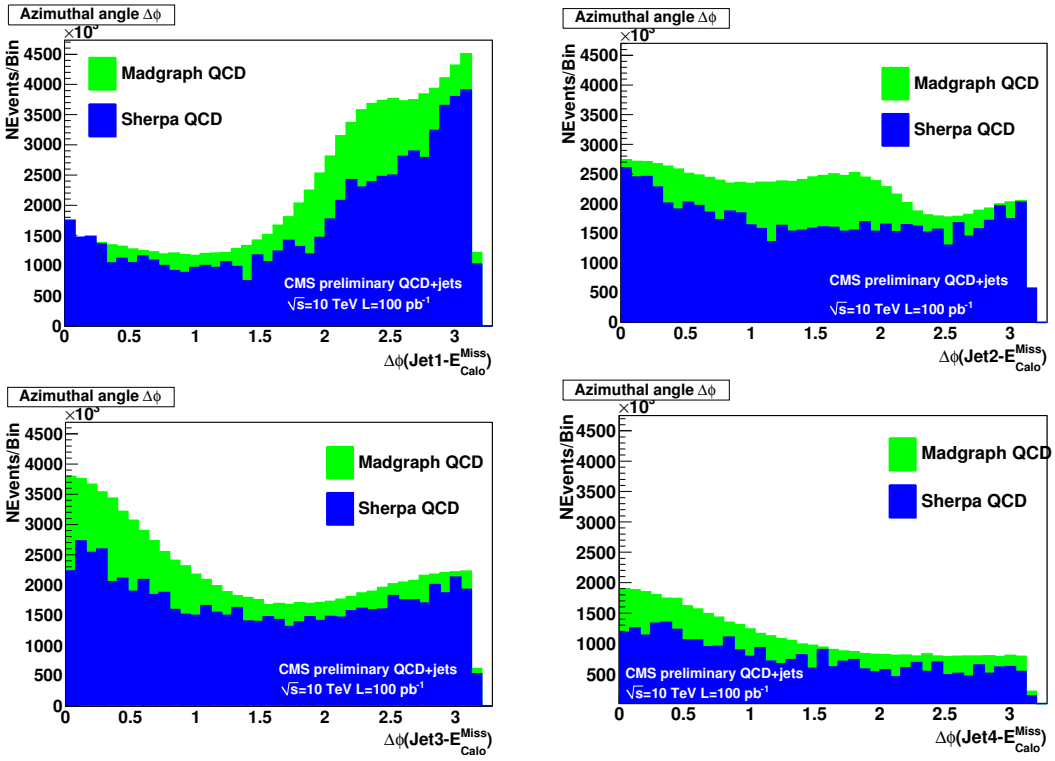


Figure 6.16: $\Delta\phi$ distribution for QCDjets production simulated with Sherpa and Madgraph after applying the lepton-veto selection step. The $\Delta\phi$ variable between missing transverse energy and the leading jet P_T distribution are applied after $E_{Calo}^{Miss} > 200$ GeV requirement. The typical value is $\Delta\phi$(Jets,E_{Calo}^{Miss})> 0.3 for all the leading jets.

Events with large missing transverse energy usually come from a mis-measured high-P_T jet. E_T^{Miss} will be pulled into the direction of this jet and ϕ turns to be small. Figure 6.16 show the $\Delta\phi$ for the first four leading jets.

The dependence of $\Delta\phi$ between leading jets and missing transverse energy is important for separating QCDjets after all selection cuts are applied.

Data Samples

The datasets (see Sections 5.1.1 and 6.3) used in this analysis are summarized below:

- **QCDjets** - Madgraph and Sherpa
- $t\bar{t}$**+jets** - Madgraph, Alpgen, Sherpa and MC@NLO
- **Z+jets (Z** $\rightarrow \nu\bar{\nu}$**), W+jets** - Madgraph, Alpgen and Sherpa
- **SUSY LM** samples - SoftSusy+SusyHit+Pythia

The physics objects selection is identical to the one used in the exclusive SUSY hadronic search. The event selection for inclusive hadronic search in CMS is splitted in four selection steps [135]:

- **Pre-Selection:** $N_{jets} \geq 3$, $P_{T_j} > 50$ GeV, $|\eta_j| < 5$ with $F_{em} < 1$, $P_{T_e} > 15$ GeV, $|\eta| < 2.5$ and $P_{T_\mu} > 10$ GeV, $|\eta| < 2.4$
- **Lepton Veto:** $N_{leptons}$=0, $|d_0^e| \leq 0.2$, electron relative isolation< 0.5, $|d_0^\mu| \leq 0.2$, muon relative isolation< 0.1, $N_{Hits} \geq 11$, χ^2/dof< 10.
- **Jet Final:** $P_{T_j}^{1st} > 180$ GeV, $P_{T_j}^{2nd} > 150$ GeV and $P_{T_j}^{2nd} > 50$ GeV , $|\eta_j| < 2.5$ with $0.05 < F_{em} < 0.95$.
- $\Delta\phi$**(LeadingJets,**E_{Calo}^{Miss}**):** MHT > 200 GeV and $E_{Calo}^{Miss} > 200$ GeV have been considered. It is applied for the first three leading jets and E_{Calo}^{Miss}. The $\Delta\phi$ (Jets,E_{Calo}^{Miss})> 0.3 has been chosen.

6.4.4 Results for Inclusive Hadronic SUSY Search

The event selection introduced in the last sub-section have been applied on the corresponding SM background processes for this search. Events with less than three jets have been rejected in the pre-selection step. The effect of the generators and their systematics have been studied with the following parameters: the hard scale choices (factorization/renormalization), initial and final state radiations (ISR-FSR), the matching scales with the different matching methods.

Table 6.7: Final number of predicted events after the different event selection steps of the inclusive hadronic SUSY search for $t\bar{t}$+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Alpgen, Sherpa, Madgraph and MC@NLO are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 317$pb.

Generators	Datasets	Pre-Selection	Lepton Veto	Jet	Final $\Delta\phi$(Jets,E^{Miss}_{Calo})$>$ 0.3
ALPGEN	PtJet$>$ 10 GeV	21397	12157	886.2	6.9
SHERPA	PtJet$>$ 30 GeV	19920	9977	1117	7.04
	PtJet$>$ 40 GeV	20235	10146	1134	3.8
	Scale Up	18927	9378	1116	6.2
	Scale Down	20235	9685	1164	6.93
MADGRAPH	PtJet$>$ 10 GeV	23930	12011	924	11.35
	PtJet$>$ 30 GeV	18254	12035	949	11.2
	PtJet$>$ 40 GeV	18247	11743	918.6	13
	Scale Up	16948	11718	953.9	8.2
	Scale Down	17753	12394	802	11.5
	Smaller ISR	20182	12617	1042	8.96
	Larger ISR	19985	12283	960	7.8
MC@NLO	NLO	21029	10465	793	3.4

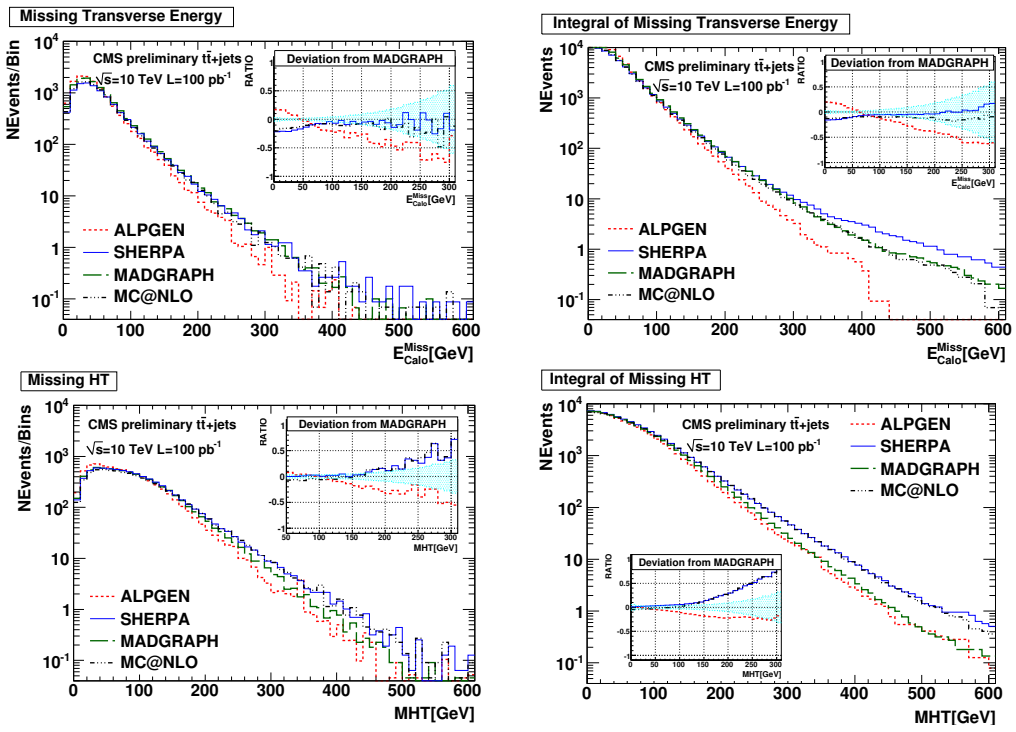


Figure 6.17: Differential (left) and integral (right) E^{Miss}_{Calo} and MHT distribution for the $t\bar{t}$+jets production simulated with Alpgen, Sherpa, Madgraph and MC@NLO after applying the lepton-veto selection step. Ratio: relative differences between Alpgen, Sherpa and MC@NLO with respect to Madgraph.

Generators have their own parameters related to the non-perturbative QCD effects (see Section 4.2 and 5.1). Variations of these parameters are considered for the inclusive hadronic SUSY search. In Table 6.7 the number of predicted events after the different event selection steps of the inclusive hadronic SUSY search for $t\bar{t}$+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$ are presented. The different generators and their settings have been studied using the four main selection steps. The following observables are shown for the $t\bar{t}$+jets production in Figures 6.17 and 6.18: missing transverse energy from the calorimeters (E_{Calo}^{Miss}), missing HT (MHT), $\Delta\phi$ (Jets,E_{Calo}^{Miss}) and effective mass.

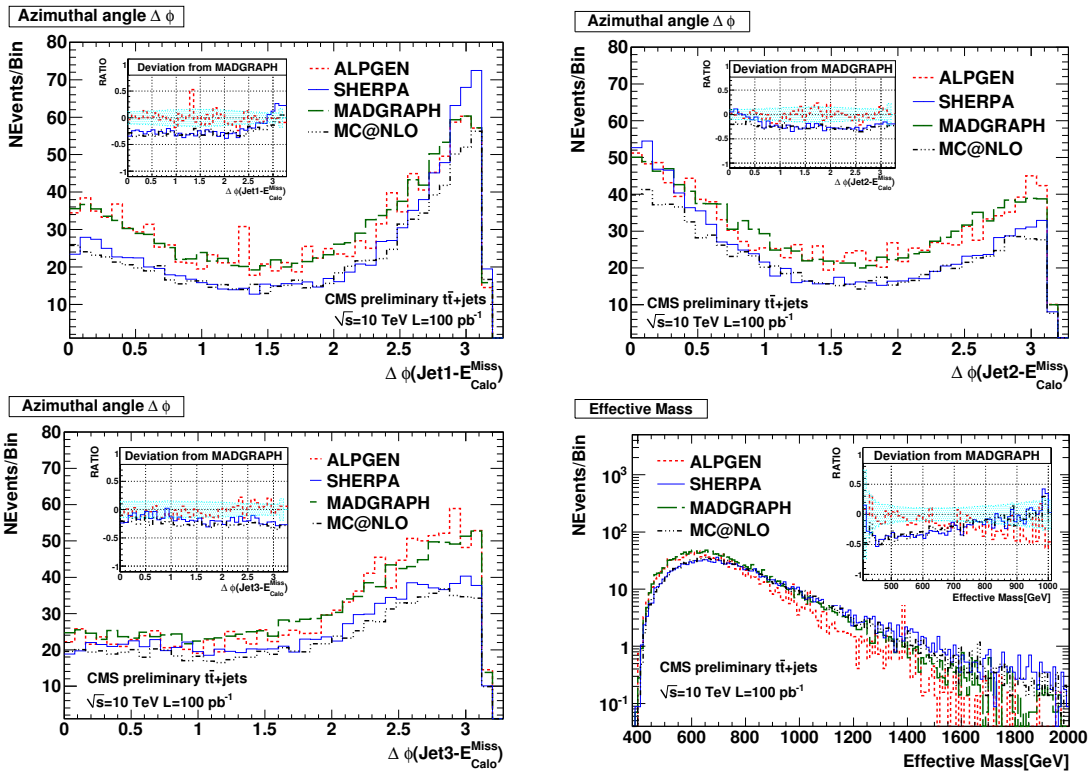


Figure 6.18: Azimuthal angle $\Delta\phi$ between three leading jets and missing transverse energy (top and lower-left) and effective mass $M_{\text{eff}} = E_{Calo}^{Miss} + \sum_{i=5}^{n} P_T^{ji}$ (lower-right) are shown for $t\bar{t}$+jets before the final selection step. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

The differential plots in Figure 6.17 show a good agreement between generators up to 200 GeV. Alpgen gives a softer spectrum. Sherpa and MC@NLO show good agreement with Madgraph. However, Sherpa and MC@NLO show significant deviations of the integral distribution due to the differences in tails of the E_{Calo}^{Miss} and MHT distribution. This difference has to be taken into account when a hard jet selection is applied. As discussed in the previous chapter, Sherpa generally has a harder leading jet P_T spectrum (see Section 5.2.3). Therefore Sherpa produces more events at this selection step. On the other hand these differences between generators are somewhat compensated by the final $\Delta\phi$ (Jets,E_{Calo}^{Miss}) selection.

Table 6.8: Final number of predicted events after the different event selection steps of the inclusive hadronic SUSY search for Zinvisible+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Sherpa and Madgraph are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 2000$pb.

Generators	Datasets	Pre-Selection	Lepton Veto	Jet Final	$\Delta\phi$(Jets,E^{Miss}_{Calo})> 0.3
	PtJet> 20 GeV	2270	2224	84	13.56
SHERPA	Lund	2291	2269	83	10.91
	Scale Up	2085	2064	64	10.93
	Scale Down	2720	2382	99	14.65
MADGRAPH	PtJet> 20 GeV	2476	2454	84	10.76

The Zinvisible+jets production is discussed in Table 6.8. A good agreement between Sherpa and Madgraph predictions has been observed for differential E^{Miss}_{Calo} and MHT distribution, as can be seen in Figure 6.19.

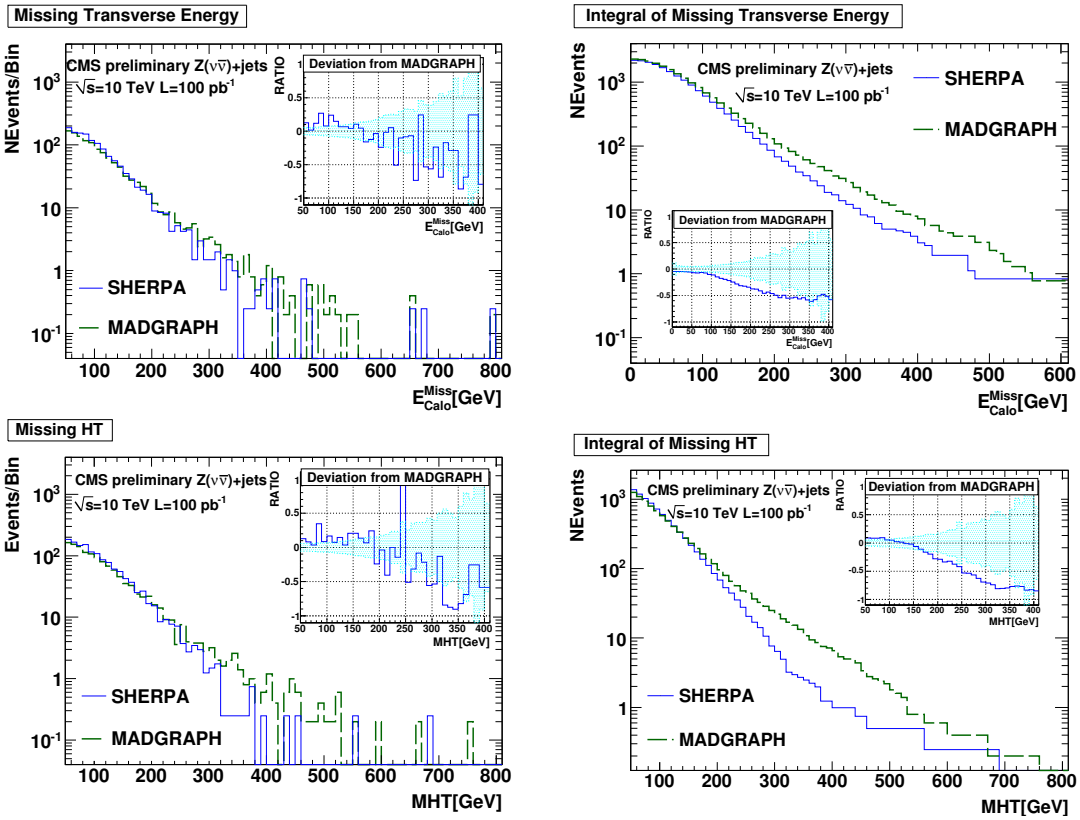


Figure 6.19: Differential (left) and integral (right) E^{Miss}_{Calo} and MHT distribution for Zinvisible+jets production simulated with Sherpa and Madgraph after applying the lepton-veto selection step. Ratio: relative difference of Sherpa with respect to Madgraph.

Although Sherpa shows a softer spectrum than Madgraph, the differences in the final number of events are in the range of internal uncertainties. For azimuthal angles ($\Delta\phi$) between three leading jets and missing transverse energy are generally in good agreement for all selection steps (see Figure 5.34). Sherpa gives a slightly softer spectrum because of jet multiplicity requirement in the pre-selection (see Figure 5.33). Remember that the Zinvisible+jets samples for Sherpa have been simulated with up to three partons in the ME. There are more partons for Madgraph in the ME. Therefore, it gives a harder spectrum after these selection requirements.

Table 6.9: Final number of predicted events after the different event selection steps of the inclusive hadronic SUSY search for QCDjets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Sherpa and Madgraph are compared. They are produced with different phase space cuts(P_T and HT of the jet). Sherpa has limited jet multiplicity (Njets < 5) at parton level production.

Generators	Datasets	Pre-Selection	Lepton Veto	Jet	Final $\Delta\phi$(Jets,$\not\!\!E > 0.3$)
SHERPA	Pt40-120 GeV	2.78+E07	2.76+E07	3143	0
	Pt120-280	310889	309986	18314	15.57
	Pt280-500	1725.3	1717	197	1.81
	Pt500-5000	56	54	6.7	0.168
MADGRAPH	HT100-250	1.05+E08	1.04+E08	0	0
	HT250-500	2.9+E07	9.06+E06	168811	0
	HT500-1000	1.17+E06	292953	542252	4.99
	HT1000-Inf	35283	33959	30333	0.29

Madgraph QCDjets are compared with Sherpa. Madgraph samples have been simulated by the CMS collaboration with high statistics. In order to compare the MLM matching methods with the CKKW matching method, Sherpa QCD samples have been simulated for this analysis. The phase space selections on these simulated samples are applied with P_T (Sherpa) and HT (Madgraph) at parton level. QCDjets samples are presented in Table 6.9 for Sherpa and Madgraph predictions. The numbers are calculated with the corresponding cross sections in Sherpa and Madgraph predictions at an integrated luminosity $\mathcal{L}$=$100pb^{-1}$. The differences between pre- and lepton veto selection steps are not comparable between generators. On the other hand the number of events after the $\Delta\phi$(Jets,E^{Miss}_{Calo})> 0.3 selection step with $E^{Miss}_{Calo} > 200$ GeV shows an agreement. Sherpa gives more events than the Madgraph prediction. This difference is an comparable order with respect to other generator comparison studies [135].

The W+jets production is discussed in Table 6.10. The differential and integral distribution of E^{Miss}_{Calo} and MHT are compared in Figure 6.21. Sherpa shows the same behavior observed in the Zinvisible+jets distribution. Sherpa and Alpgen give a softer spectrum compared to the Madgraph prediction. The reason, also discussed in Zinvisible+jets, is that the generated jet multiplicity at ME level for Sherpa and Alpgen is equal to the jet multiplicity requirement at pre-selection. The rest of the jet spectra are mainly driven by the corresponding parton showers.

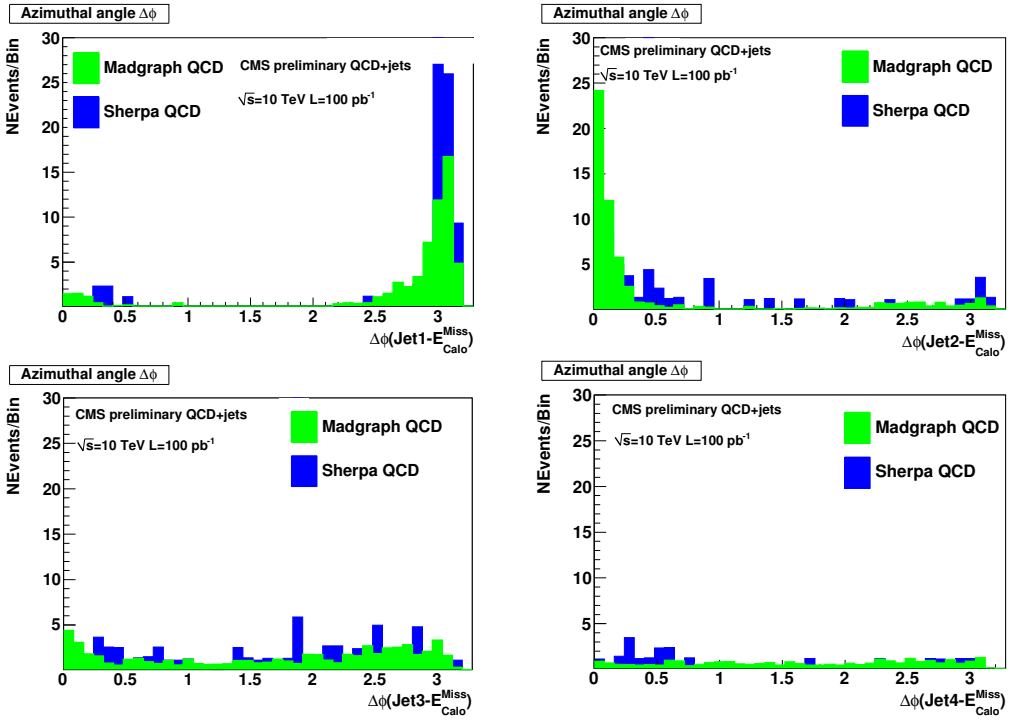


Figure 6.20: Azimuthal angle $\Delta\phi$ between leading jets and E_{Calo}^{Miss} distribution for QCDjets production simulated with Sherpa and Madgraph before applying the final selection step.

Table 6.10: Final number of predicted events after the different event selection steps of the inclusive hadronic SUSY search for W+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Alpgen, Sherpa and Madgraph are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 40000$pb.

GENERATORS	DATASETS	Pre-Selection	Lepton Veto	Jet	Final $\Delta\phi$(Jets,E_{Calo}^{Miss})> 0.3
ALPGEN	PtJet> 30	21100	11130	234	6.5
	PtJet> 20(R)	24159	11652	586	20
SHERPA	PtJet> 30	24199	10442	564	28
	PtJet> 30	17234	12068	192	10
MADGRAPH	PtJet> 30(R)	24047	12929	341	14
	Scale Up	22115	16787	254	8
	Scale Down	23142	17456	570	22

On the other side Madgraph has more partons in the ME production. Therefore, it gives a bit harder spectrum and includes somewhat harder jets for the inclusive hadronic SUSY analysis. There is a 20%-30% deviation starting from $E_{Calo}^{Miss} > 50$ GeV and MHT> 50 GeV. Sherpa and Alpgen show a softer spectra. Alpgen gives a much softer spectrum for increasing energies.

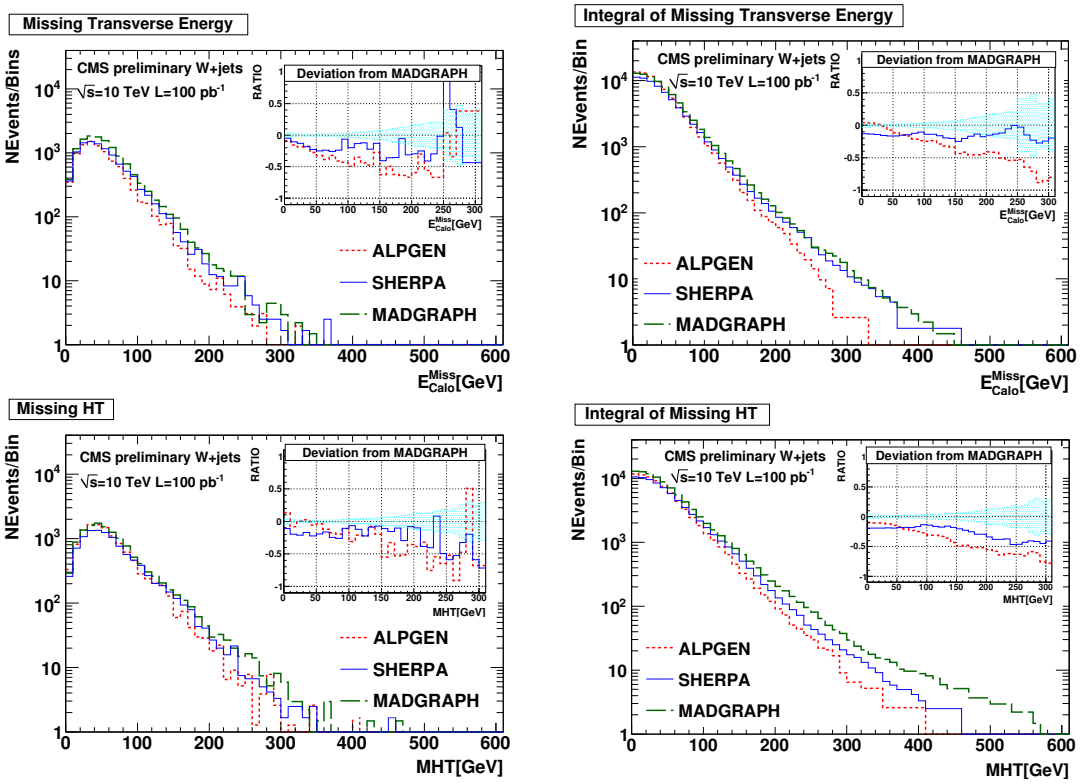


Figure 6.21: Differential (left) and integral (right) E_{Calo}^{Miss} and MHT distribution for W+jets production simulated with Alpgen, Sherpa and Madgraph after applying the lepton-veto selection step. Ratio: relative differences of Alpgen, Sherpa with respect to Madgraph

Madgraph generally shows a harder spectrum for all distributions. The effects of all jet P_T, as can be seen from the sumEt distribution, reach 50% (see Figure 5.31). W+jets samples simulated with various settings at generator level (including first order electro-weak corrections) in Sherpa show a larger number of events after the final selection step. The same behavior is also observed for the Madgraph scale-down sample. Simulation effects and missing higher order effects can be the reason for these discrepancies. The default scale variations of the generators from their Tevatron tunes are significantly different for the inclusive hadronic search because of hard jet P_T selection cuts. Therefore, a large variation for the W+jets production has been observed (discussed in Section 5.2).

6.4.5 SUSY Signals for Hadronic Searches

The SUSY signal samples in Table 6.3 have been analyzed for the exclusive and inclusive hadronic SUSY searches at $\sqrt{s}$=10 TeV. The corresponding final number of events are presented in Table 6.11. Bold numbers represent the convenient number of events for possible SUSY discovery above SM background at $\mathcal{L}$=100 pb^{-1}. The background estimations via various generators and settings for the hadronic searches with/without uncertainties, discussed in the appendix B.2.1, is discussed with SUSY discovery reaches at the end of this chapter. E^{Miss}_{Calo}, HT, Effective mass and the number of jet multiplicity have been shown for the illustration in Figure 6.22 with the lepton-veto selection step via inclusive hadronic SUSY searches.

Table 6.11: Final number of predicted events at the final event selection step of the exclusive and inclusive hadronic SUSY searches for SUSY LM samples production scaled to an integrated luminosity of $\mathcal{L}$=100pb^{-1}. All samples are normalized to the corresponding cross section at $\sqrt{s}$=10 TeV.

Generators	Benchmark points	Exc. search $\alpha_T > 0.55$	Inc. search $\Delta\phi$(Jets,$\not\!\!E_{T>200}$)
SOFTSUSY + SUSYHIT +PYTHIA	LM0	**623**	**476**
	LM1	**257**	**249**
	LM2	38	51
	LM3	103	**161**
	LM4	78	**126**
	LM5	14	28
	LM6	15	21
	LM7	3	8
	LM8	15	33
	LM9	24	40
	LM9p	4	12
	LM9t175	26	46
	LM10	0	0
	LM11	44	**68**

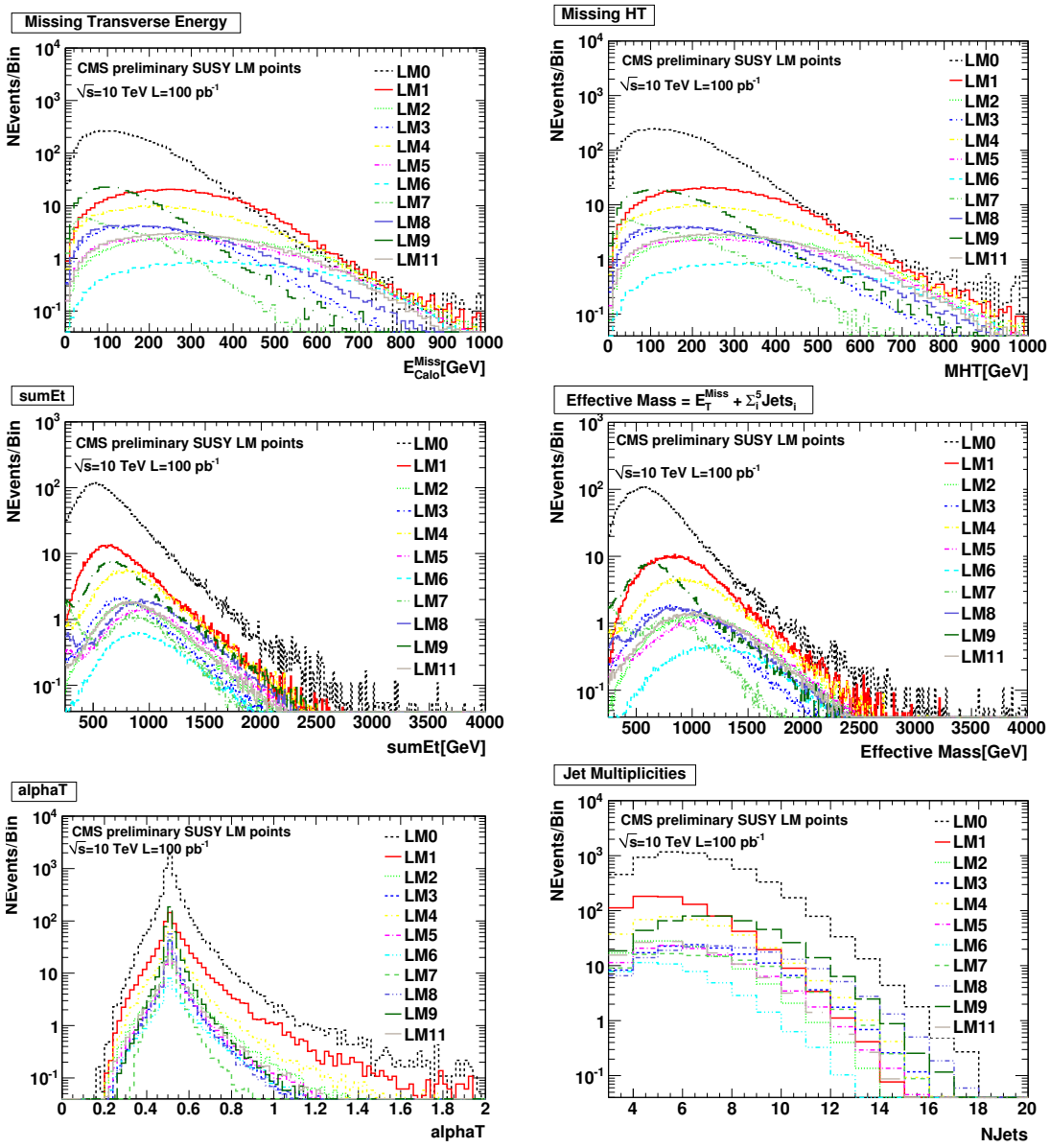


Figure 6.22: E_{Calo}^{Miss} (top-left), MHT(top-right), sumEt(middle-left), effective mass(middle-right), alphaT(lower-left) and jet multiplicity (lower-right) have been presented for all SUSY LM points at $\mathcal{L}$=100 pb^{-1}. The three jets requirement with direct lepton veto has been applied for the selection

6.5 Multi-Leptonic SUSY Searches in CMS

Important event topologies for SUSY searches are the signatures with a single, two and three isolated muons; three jets above a P_T threshold, and a quantity such as missing transverse energy. In this section different MC predictions are analyzed for multi-leptonic SUSY searches and the theoretical systematical uncertainties on SM background are studied.

6.5.1 Data Samples

The datasets (see Section 5.1.1 and 6.3) used in this analysis are summarized below:

- $t\bar{t}$**+jets** - Madgraph, Alpgen, Sherpa and MC@NLO
- **Z+jets, W+jets** - Madgraph, Alpgen and Sherpa
- **VV+jets, A*+jets (or γ^*+jets)** - Madgraph and Sherpa
- **SUSY LM** samples - SoftSusy+SusyHit+Pythia

6.5.2 Physics Objects Selection

- **Jet definitions:** A jet is reconstructed with the SISCone algorithm using a ΔR=0.5 and energy is corrected as defined in section 3.7.2 with L1 Offset, L2 and L3 relative and absolute jet corrections.
- **Missing transverse energy based on calorimetry:** The corrected missing transverse energy is considered for the analysis (see section 3.7.5)
- **Leptons** `GlobalPromptTight` muons and `RobustTight` electrons have been used for this analysis (see details in sections 3.7.3, 3.7.4 and appendix A.1, A.2).

The event selection for multi-muonic searches in CMS is splitted in two main selection steps:

- **Muon Selection**

  `Single and Di-Muons:` N_μ=1, 2, $N_{jets} \geq 3$, $P_{T_j} > 30$ GeV, $|\eta_j| < 2.4$ with $F_{em} \geq 0.1$. Three leading jets with $P_{T_j} > 50$ GeV. $P_{T_e} > 20$ GeV, $|\eta| < 2.5$ and $P_{T_\mu} > 10$ GeV, $|\eta| < 2.1$

  `Tri-Muons:` N_μ=3, $P_{T_j} > 30$ GeV and $P_{T_\mu} > 8$ GeV, $|\eta| < 2.1$

  `Lepton Isolation:`

  `Single and Di-Muons:` $|d_0^\mu| \leq 0.2$, relative muon isolation< 0.1, $N_{Hits} \geq 11$, χ^2/dof< 10, Track/HCal/ECal Isolation< 6 for muons. Electron veto with $|d_0^e| \leq 0.2$, relative electron isolation< 0.1.

- **Final Selection:**

  `Single and Di-Muons:` $E_{Calo}^{Miss} > 100$ GeV and $E_{Calo}^{Miss} > 200$ GeV for single and di-muons, respectively.

  `Tri-Muons:` N_{μ}=3, impact parameter significance sigDxy=4, $|d_0^{\mu}| \leq 0.2$, muon relative isolation< 0.15, $N_{Hits} \geq 11$, χ^2/dof< 10. Opposite sign(OS) muon mass between 20-86 GeV. No electron veto.

6.5.3 Results for Single-Muon and OSSF-Di+E_T^{Miss} and Tri-Muon Signatures

Differences of the generators and their theoretical systematic uncertainties on the leptonic SUSY selections have been analyzed for multi-muonic SUSY searches. Tables 6.12, 6.13, 6.14 and 6.15 present the results of the generator comparisons for $t\bar{t}$+jets, Z+jets, W+jets, A^*+jets and VV+jets, respectively.

Table 6.12: Final number of predicted events after the different event selection steps of the multi-muonic SUSY searches for $t\bar{t}$+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Alpgen, Sherpa, Madgraph and MC@NLO are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 317$pb.

Generators	Datasets	Single μ	Final	OSSF μ	Final	Tri μ	Final
ALPGEN	PtJet$>$ 30 GeV	1504	216	20.66	6.68	166	0.1
SHERPA	PtJet$>$ 30 GeV	1619	323.2	23.26	8.38	186	0.21
	PtJet$>$ 40 GeV	1630	320	26.85	10.1	182	0.7
	Scale Up	1630	338.8	26.27	10.5	192	0.26
	Scale Down	1624	3272	24.76	10.2	178	0.39
MADGRAPH	PtJet$>$ 10 GeV	1466.73	249.7	25.78	8.1	117	0.12
	PtJet$>$ 30 GeV	1424.6	239.4	24.1	7.8	157	0.21
	PtJet$>$ 40 GeV	1480.4	240.6	26.3	7.6	122	0.3
	Scale Up	1444	234	24.29	7.86	123.5	0.51
	Scale Down	1481	227	25.61	7.4	101	0.1
	Smaller ISR	1530	250	29.42	8.7	149	0
	Larger ISR	1515	249.5	28.08	9.1	151	0.04
MC@NLO	NLO	1624	355	34	11.3	140.1	0.48

In Table 6.12, the predicted number of events after the different event selection steps of the multi-muonic SUSY searches for the $t\bar{t}$+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$ are presented. Predictions from the different generators and their settings have been shown using the two main selection steps. The differences between Alpgen and Madgraph after the muon selection step for single and di-muons searches are in a good agreement ($\sim$ 10 - 15%). Alpgen and Madgraph also give comparable numbers for the tri-muon search.

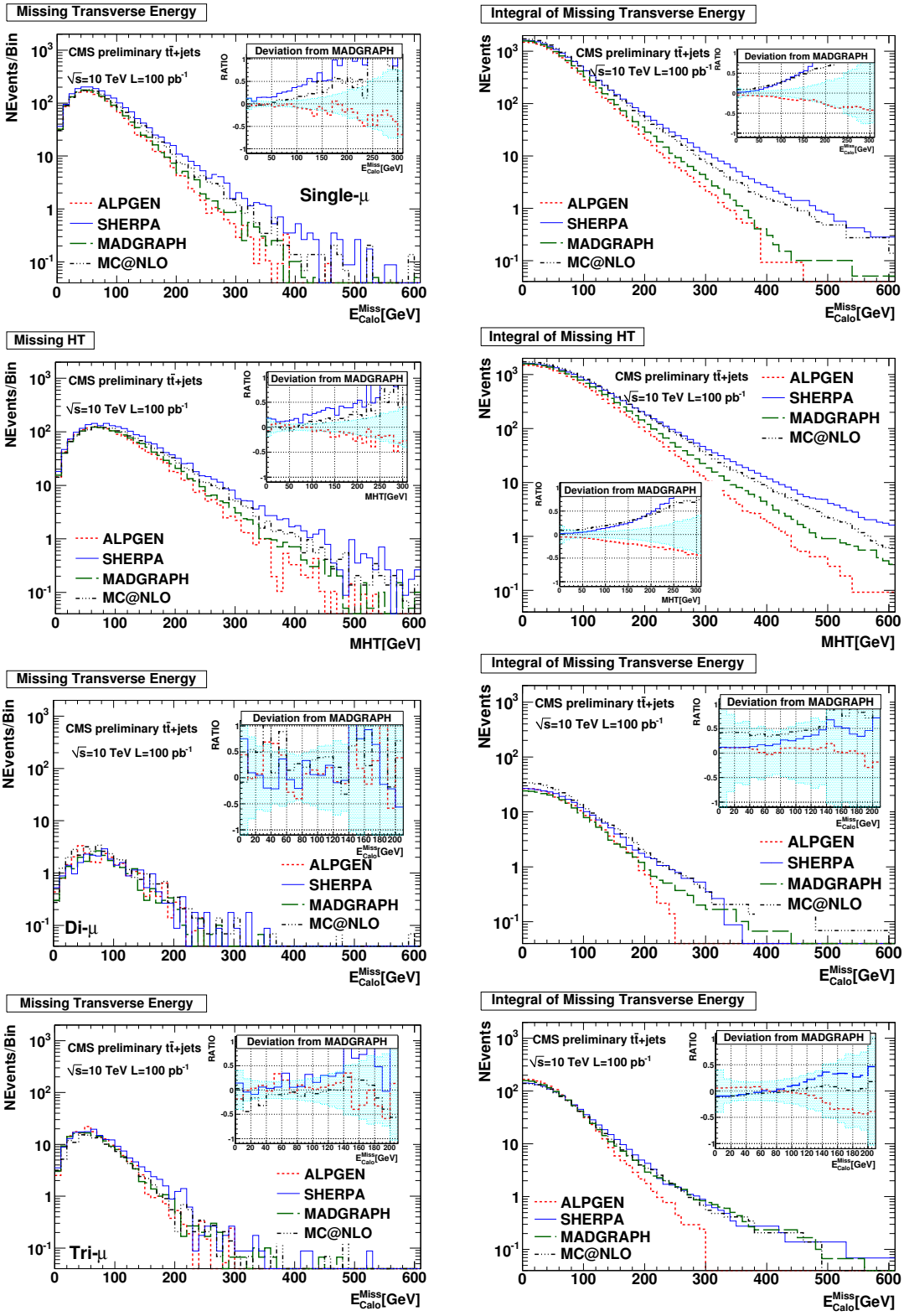


Figure 6.23: Differential (left) and integral (right) E_{Calo}^{Miss} and MHT (second-row) distribution for $t\bar{t}$+jets production simulated with Alpgen, Sherpa, Madgraph and MC@NLO after applying the muon selection step. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

However, Sherpa shows a larger numbers $(20-25\%)$ because of the hard jet and lepton P_T spectrum. Although Sherpa does not show a difference with respect to other generators for muon P_T (see Figure 5.42), Sherpa generally has a harder leading jet P_T spectrum for the predictions (see Section 5.2.3). The same behavior shows in MC@NLO for single, di and tri-muons searches. Although it gives a good agreement within the tri-muons search, MC@NLO generally gives a harder spectrum for all selection steps. Differential and integral distribution of E_{Calo}^{Miss} and MHT are shown in Figure 6.23 for $t\bar{t}$+jets. MHT is presented only for the single muon search due to the same jet structures and similar behavior for the di-muon selection. These differences observed in single muon search are not the same order for the di-muons search. The systematical uncertainties between generators are almost of negligible order for all selection steps. The corresponding distributions for differential and integral E_{Calo}^{Miss} show a good agreement.

The jet and E_{Calo}^{Miss} dependence of the generators can be seen in the effective mass distribution for the single muon search in Figure 6.24. In this case the effective mass has been defined as missing transverse energy plus five leading jets at final. This is important for single and di-muon SUSY searches. However, the tri-muon search does not have a jet and missing transverse energy requirement. A similar behavior is observed for Sherpa and MC@NLO with respect to Alpgen and Madgraph.

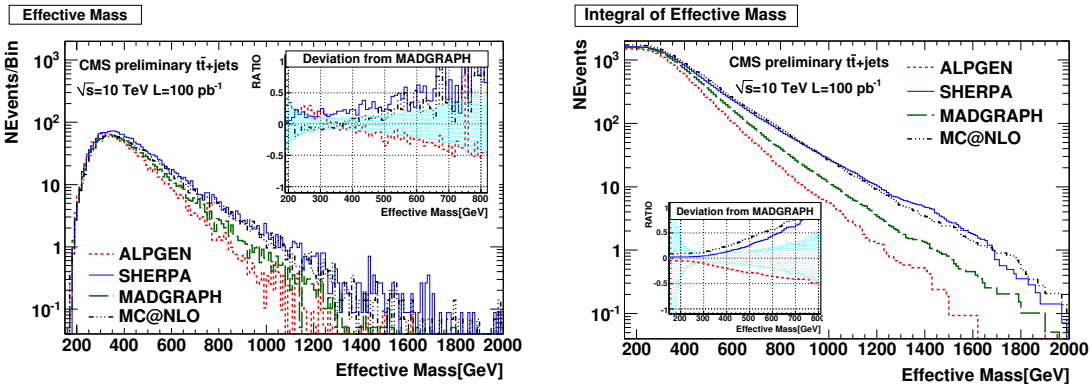


Figure 6.24: Differential (left) and integral (right) effective mass $M_{\text{eff}} = E_{Calo}^{Miss} + \sum_{i=5}^{n} P_T^{ji}$ distribution for $t\bar{t}$+jets production simulated with Alpgen, Sherpa, Madgraph and MC@NLO after applying the single muon selection step. Ratio: relative differences of Alpgen, Sherpa and MC@NLO with respect to Madgraph.

The number of events after the different muon selections are presented in Table 6.13 for the Z+jets production. The predictions from Sherpa are significantly different than those of other generators for all searches after the muon selection steps. The differences can be traced from the typical behaviors of the generators (see Section 5.1). Sherpa generally produces more events while Alpgen and Madgraph have an very good agreement for the final number of selected events. The number of events is small for the tri-muon selection steps.

Table 6.13: Final number of predicted events after the different event selection steps of the multi-muonic SUSY searches for Z+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Alpgen, Sherpa and Madgraph are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 3700$pb.

Generators	Datasets	Single μ	Final	OSSF	Final	Tri μ	Final
ALPGEN	PtJet> 30 GeV	106	3.8	40.24	0	–	–
SHERPA	PtJet> 15 GeV	366	15.5	72	1.4	147	0.52
	PtJet> 20(R) GeV	352	11	60	2.2	143	0.86
	PtJet> 30 GeV	183	14.77	71	1.6	121	0.25
MADGRAPH	PtJet> 30(R) GeV	133	6.77	48.52	0	97.4	1.02
	PtJet> 30 GeV	115	4.48	44.25	0.2	113	0.2

The SUSY search with a single muon is presented in Table 6.14 for the W+jets production. The W+jets production has a negligible contribution for the di- and tri-muons searches.

Table 6.14: Final number of predicted events after the different event selection steps of the single muon SUSY search for W+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Alpgen, Sherpa and Madgraph are compared for different generator settings. All samples are normalized to the Madgraph cross section $\sigma = 40000$pb.

Generators	Datasets	Single μ	Final
ALPGEN	PtJet> 30 GeV	523	119
SHERPA	PtJet> 20(R) GeV	1246.8	259
	PtJet> 30 GeV	1029	248.9
	Eweak1	1227	268
MADGRAPH	PtJet> 30(R) GeV	581	117
	PtJet> 30 GeV	793	146.8
	Scale Up	532	86.45
	Scale Down	1258	242.8

The corresponding Z+jets and W+jets comparisons between generators are shown in Figure 6.25. Similar results have been observed for the used generators. While Madgraph and Alpgen predictions show good agreement, Sherpa gives a harder spectrum after jet multiplicity and muon requirements.

The global muon P_T distribution are presented in Figure 6.26 at detector level. Except for VV+jets in the tri-muons search, $t\bar{t}$+jets, Z+jets and W+jets have been discussed for the single muon search. The selection on the muon selection does not affect the muon P_T distribution between different multi-muonic SUSY searches. Sherpa gives a harder spectrum for all predictions with 20% uncertainties.

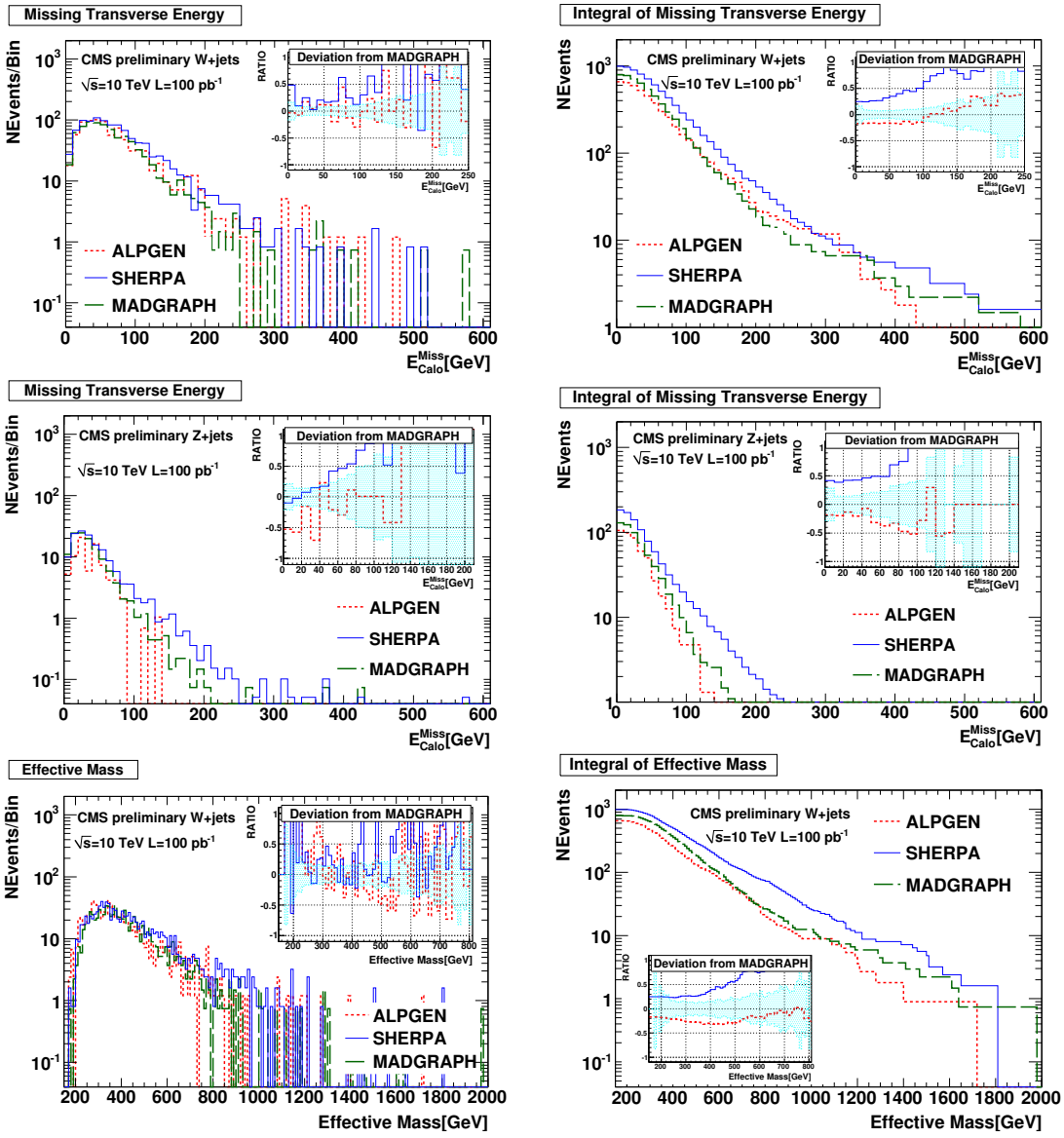


Figure 6.25: Differential (left) and integral (right) E_{Calo}^{Miss} and effective mass $M_{\text{eff}} = E_{Calo}^{Miss} + \sum_{i=5}^{n} P_T^{ji}$ distribution for W+jets and Z+jets (middle) production simulated with Alpgen, Sherpa and Madgraph after applying the single muon selection step. Ratio: relative differences of Alpgen and Sherpa with respect to Madgraph.

The multi-muon searches after the different selection steps are presented for VV+jets and A*+jets in Table 6.15. VV+jets and A*+jets are very important for tri-muon SUSY search (see Section 6.2.2). The number of events are generally in good agreement for all searches. There is good agreement for the tri-muon search after the final selection steps.

The fake muon production has been also checked for different channels using the various generators. Fake muons have been defined as all muons, which are not coming from SM bosons and gauginos (see details in Appendix A.1.2).

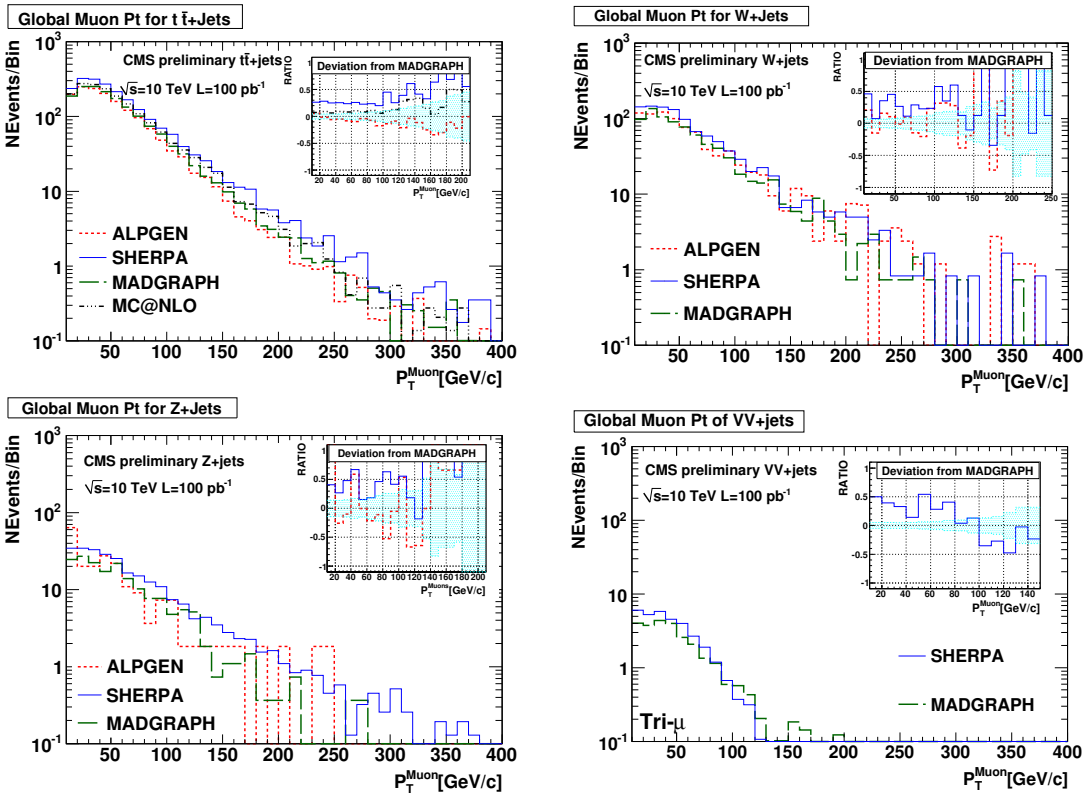


Figure 6.26: Global Muon P_T distribution with single muon SUSY search for $t\bar{t}$+jets (top-left), W+jets (top-right) and Z+jets (down-left). Global muon P_T distribution are presented for VV+Jets in Tri-muons SUSY search (down-right). Ratio: relative differences between of Alpgen and Sherpa with respect to Madgraph.

Table 6.15: Final number of predicted events after the different event selection steps of the single muon SUSY search for VV+jets production scaled to an integrated luminosity of $\mathcal{L} = 100pb^{-1}$. Sherpa and Madgraph are compared for different matching methods. All samples are normalized to the Madgraph cross section $\sigma = 20$pb.

Generators	Datasets	Single μ	Final	OSSF	Final	Tri μ	Final
SHERPA	A^*JETS	22.14	0.7	13.1	0.07	207	0.48
MADGRAPH	A^*JETS	34.5	0	13.1	0	230	1.04
SHERPA	VVJETS	3.72	1.21	1.14	0.09	13.1	3.31
MADGRAPH	VVJETS	3.17	1.29	0.45	0.12	8.69	3.89

Fig. A.2 shows the P_T distribution of fake muons after a loose isolation requirement for the Z+jets and $t\bar{t}$+jets production simulated with Alpgen, Sherpa, Madgraph and MC@NLO. It can be seen that fake muons produced in different channels with different generators are generally similar for Sherpa and Madgraph and slightly softer for Alpgen with respect to others.

Fake rate mostly depends on the heavy flavor content of the produced event, while the fake rate of QCD and electroweak boson production is an order of magnitude lower [44].

6.5.4 SUSY Signals for Multi-Muonic Searches

The SUSY signal samples have been analyzed for the multi-muonic SUSY searches at $\sqrt{s}$=10 TeV. The corresponding final number of events have been presented in Table 6.16 for all SUSY-LM benchmark points.

Table 6.16: Final number of predicted events at the final event selection step of the multi-muonic SUSY searches for SUSY-LM samples production scaled to an integrated luminosity of $\mathcal{L}$=100pb^{-1}. All samples are normalized to the corresponding cross section at $\sqrt{s}$=10 TeV.

Generators	Benchmark points	Single Muon	Di-Muon	Tri-Muon
SOFTSUSY + SUSYHIT + PYTHIA	LM0	**403**	**23.5**	8.95
	LM1	52	6.4	2.14
	LM2	11.2	0.5	0.09
	LM3	60	4.32	1.17
	LM4	33	2.33	0.7
	LM5	11.3	0.42	0.09
	LM6	9.48	0.85	0.33
	LM7	8.58	0.66	0.26
	LM8	27.1	2.2	0.62
	LM9	29.54	2.42	1.06
	LM9p	11.7	2.12	0.6
	LM9t175	30.96	2.59	1.05
	LM11	21.1	1.23	0.19

Bold font numbers represents points with a convenient number of selected events for a possible SUSY discovery at an integrated $\mathcal{L}$=100 pb^{-1}. The background estimations via various generators and settings for the multi-muonic SUSY searches with/without uncertainties, discussed in the appendix B.2.1, is discussed with SUSY discovery reaches at the end of this chapter. The final number of events presented in the table are used for the SUSY discovery reaches for multi-muonic SUSY searches. E^{Miss}_{Calo}, HT (sumEt), effective mass and the jet multiplicity are shown after the muon selection step of the single muon search in Figure 6.27.

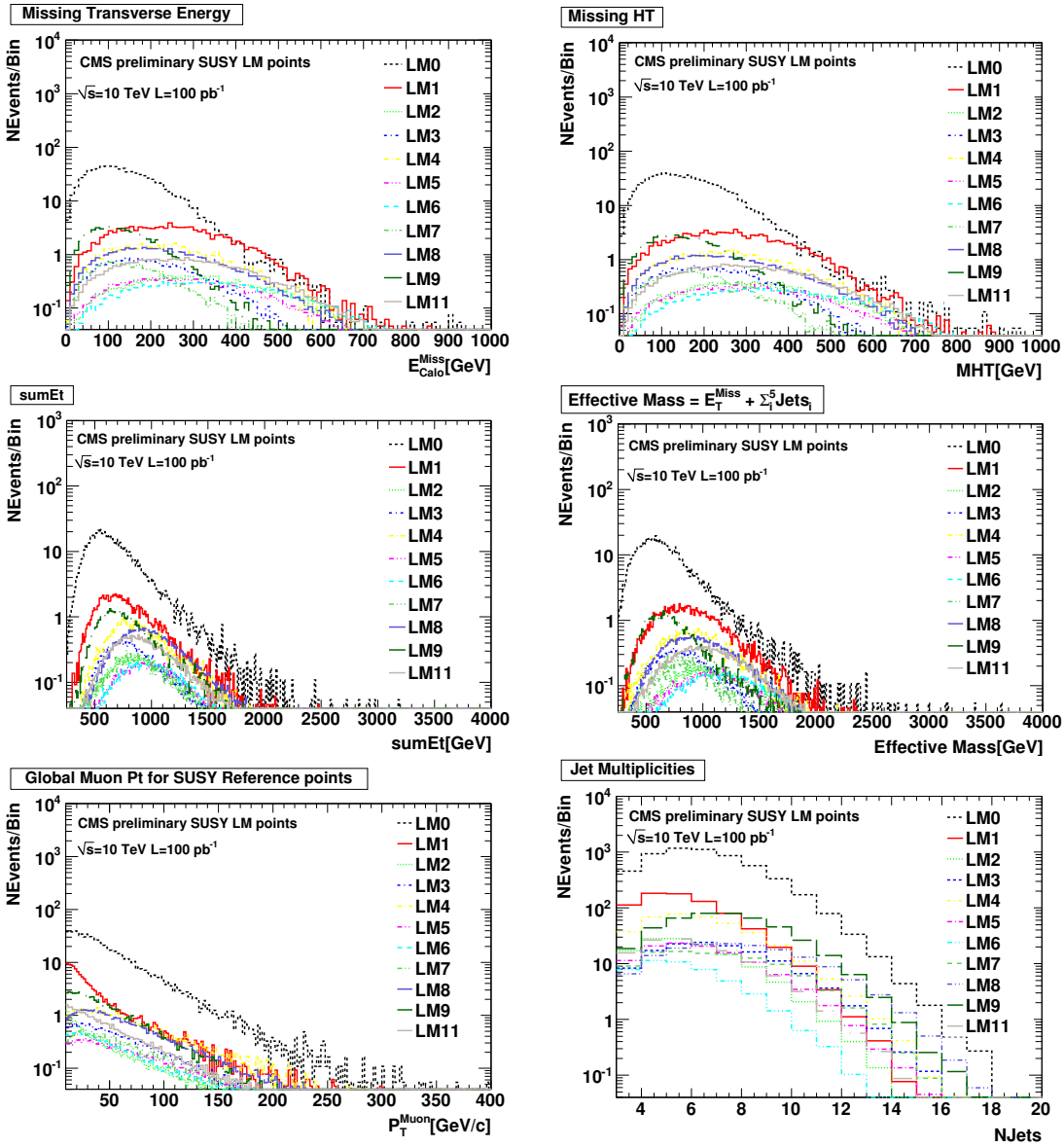


Figure 6.27: E_{Calo}^{Miss} (top-left), MHT (top-right), sumEt (middle-left), effective mass (middle-right), Muon P_T(down-left) and jet multiplicity (down-right) distribution for all SUSY-LM samples after applying the single muon selection step.

6.6 SUSY Discovery Reach

Generator differences of the SM background for hadronic and multi-leptonic (muons) SUSY searches have been analyzed at $\sqrt{s} = 10$ TeV for the four different generators: Sherpa, Alpgen, Madgraph and MC@NLO. Important SM processes have been generated including CMS detector simulations: Z/γ^*+jets, Zinvisible+jets, W+jets, $t\bar{t}$+jets, QCDjets and VV+jets. All SUSY LM benchmark points have been considered.

Two different systematics are considered for the final number of predicted events, internal and external. Internal systematics are defined as the systematics which arise from the variation of the matching, factorization and renormalization scales and the initial and final state radiations. The external systematics are defined as the difference between generators. The internal and external systematic uncertainties, obtained as the maximum deviation from the average, are summarized in Tables 6.17 and 6.18 for the hadronic and multi-leptonic SUSY searches, respectively. The following results have been observed: *For the exclusive and inclusive hadronic SUSY searches*, the internal systematics do not cause significant deviations ($\sim$ 10%) for the $t\bar{t}$+jets and Zinvisible+jets production for all generators. A significant deviation has been observed for the W+jets production. Although the W+jets samples show $\sim$ 30% uncertainties for the exclusive hadronic SUSY search, it reaches $\sim$ 50% for the inclusive search. However, these differences reach larger values for the external systematics. The external systematics for the exclusive and inclusive SUSY searches show maximum $\sim$ 30% differences for the $t\bar{t}$+jets and Zinvisible+jets production. For the W+jets production, it reaches $\sim$ 45% and $\sim$ 60% for the exclusive and inclusive hadronic SUSY searches, respectively. For the predicted number of final events, the internal systematics of total SM backgrounds are about $\sim$ 18% and $\sim$ 27%, the external systematics of total SM background are about $\sim$ 28% and $\sim$ 50% for the exclusive and inclusive searches, respectively. Thus, the inclusive hadronic SUSY search shows a larger dependence for the internal and external systematics with respect to the exclusive hadronic SUSY search.

For the multi-leptonic SUSY searches, the internal systematics show negligible deviations (around $\sim$ 15%) for the all production channels except for W+jets. The W+jets production gives $\sim$ 27% deviation for the single muon SUSY search. As expected, these differences reach larger values for external systematics. The $t\bar{t}$+jets, VV+jets and A*+jets production shows maximum $\sim$ 27% difference between generators. However, above $\sim$ 60% deviation has been observed for the Z+jets and W+jets production for the single- and di- muon SUSY searches. For the predicted number of final events, the internal systematics of total SM background are about $\sim$ 11%, $\sim$ 13% and $\sim$ 20% and the external systematics of total SM background are about $\sim$ 34%, $\sim$ 25% and $\sim$ 20% systematics for the single-, di- and tri-muon SUSY searches, respectively.

The impact of systematic uncertainties on the discovery reaches of the various SUSY searches is presented in Figure 6.28 for $\mathcal{L}$=100pb^{-1}. The discovery reach has been evaluated using the Profile Likelihood method (see details in Appendix B). The single muon and inclusive hadronic discovery reaches show the largest sensitivity to systematics since here the statistical errors are negligible compared to the systematical errors. The di-muons search shows the smallest sensitivity to systematic uncertainties. The tri-muons search needs a luminosity higher than $\mathcal{L}$=100pb^{-1} for SUSY discovery.

Table 6.17: Summary of systematic uncertainties for hadronic SUSY searches at $\sqrt{s}$=10 TeV. For each background the final prediction together with the internal, external and statistical uncertainties are shown. Additionally the SUSY expectation is presented for benchmark points LM0 and LM1 as well as the significance $\mathcal{S}$ with/without systematic uncertainties.

Background ± Internal ± External ± Statistical Uncertainties

$\mathcal{L}$=100pb^{-1}	Exclusive Njets+E_T^{Miss}	Reference	Inclusive Njets+E_T^{Miss}	Reference
$t\bar{t}$+jets	**39.5**±4.3±10.5±1.6	Table 6.3	**11.3**±1.9±4.1±0.6	Table 6.7
ZInvisible+Jets	**27**±4.1±7.1±2.2	Table 6.4	**10.8**±1.9±1.9±0.8	Table 6.8
W+jets	**43.5**±14.3±17.5±3.4	Table 6.5	**14**±7±9±2.24	Table 6.10
QCDjets	**10.75**$_*$±0.3	Table 6.6	**5.28**±6.1±0.58	Table 6.9
$\mathcal{S}_{Background}$	**120.8**±22.8±35.1±7.4		**41.4**±11.4±21.8±4.2	
$LM0$	623	Table 6.11	476	Table 6.11
$\mathcal{S}_{w/oSystematics}$	36.8		> 74	
$\mathcal{S}_{Internal/ExternalSyst.}$	23.2 / 19.5			
$LM1$	257	Table 6.11	249	Table 6.11
$\mathcal{S}_{w/oSyst.}$	18.6		25.1	
$\mathcal{S}_{Internal/ExternalSyst.}$	11.2 / 9.2		15.5 / 12.4	

Table 6.18: Summary of systematic uncertainties for multi-muonic SUSY searches at $\sqrt{s}$=10 TeV. For each background the final prediction together with the internal, external and statistical uncertainties are shown. Additionally the SUSY expectation is presented for benchmark points LM0 as well as the significance $\mathcal{S}$ with/without systematic uncertainties.

Background ± Internal ± External ± Statistical Uncertainties

$\mathcal{L}$=100pb^{-1}	Single μ+E_T^{Miss}	Di-(OS)μ+E_T^{Miss}	Tri-μ	Reference
$t\bar{t}$+jets	**239.4**±11.5±43±2.4	**7.8**±0.9±2.2±0.5	**0.2**±0.15±0.06	Table 6.12
W+jets	**117**±30±77.2±8.2	-	-	Table 6.14
Z+Jets	**6.7**±0.7±5.2±0.8	**0.5**±0.1±0.5±0.15	**1.02**±0.16±0.4±0.15	Table 6.13
VV+Jets	**1.3**±0.4±0.16	**0.12**±0.01±0.03	**3.9**±0.3±0.19	Table 6.15
A^*+Jets	-	-	**1.04**±0.28±0.7	Table 6.15
$\Sigma_{Background}$	**364.5**±42.5±125±11.6	**8.4**±1.1±2.4±0.65	**5.95**±1.13±1.1	
$LM0$	403	23.5	8.95	Table 6.16
$\mathcal{S}_{w/oSystematics}$	18.3	5.96	3	
$\mathcal{S}_{Int./Ext.Sys.}$	9 / 4.7	5.5 / 4.5	2.9 / 2.9	

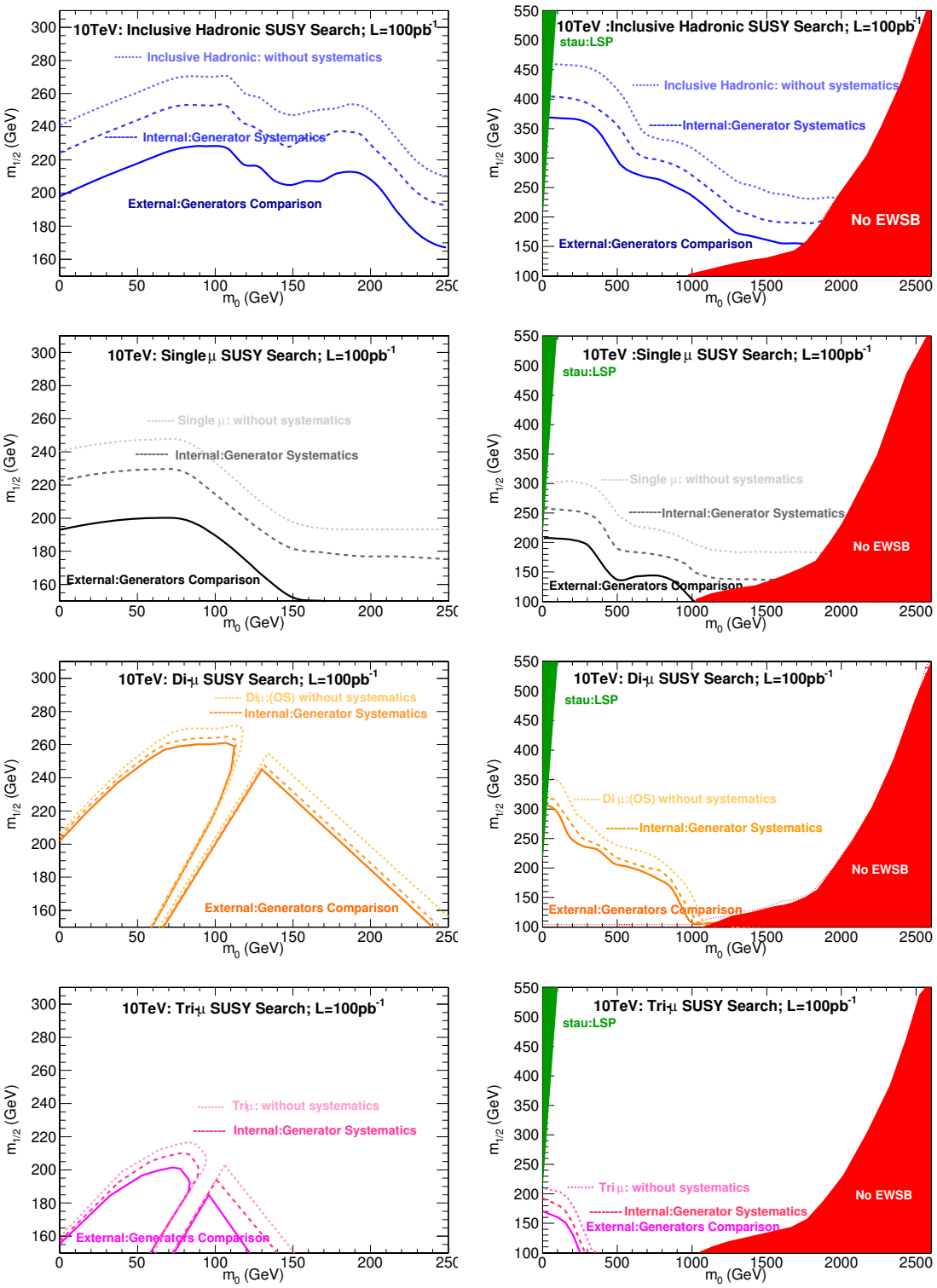


Figure 6.28: SUSY discovery reaches at a significance of $5\sigma_{\mathcal{S}}$ with tanβ=3 (left-plots) and tanβ=10 (right-plots) for hadronic and multi-muonic SUSY searches at $\sqrt{s}$ = 10 TeV. Systematic uncertainties are analyzed with correlated background estimations using the profile likelihood method (see Appendix B for details).

Chapter 7

Summary and Outlook

The Large Hadron Collider, which has started operation at the end of March 2010, is designed to probe the Standard Model and beyond the Standard Model (SM) particle physics in an energy range not accessible so far. The SM physics is a well established physical theory whose parameters have been measured by many experiments. Nevertheless, some of its features seem to be demanding a more comprehensive theory, i.e. the Hierarchy Problem between the Electroweak and the Planck scales and the radiative stability of the Higgs boson mass. Furthermore, according to recent cosmological precision measurements the universe is made from as yet unknown components, called Dark Energy and Dark Matter (DM). Supersymmetry (SUSY), in its minimal formulation, called the Minimal Supersymmetric Standard Model (MSSM), predicts a supersymmetric partner with different spin statistics for each SM particle. Besides providing a mechanism to radiatively stabilize the Higgs mass, it predicts also a massive stable neutral particle, which interacts very weakly, the so-called Lightest Supersymmetric Particle (LSP), which is a good DM candidate. In many models, the LSP is the lightest neutralino, which leads to a missing transverse energy signature in the multi-jet and/or lepton final states from the decays of heavy SUSY particles.

To discover physics beyond the SM it is necessary to simulate the SM events with Monte-Carlo (MC) methods. Primary workhorses, like Pythia and Herwig, use the leading-log approximation for generating parton showers. This approach resums the leading-logarithms of the soft and collinear divergences to all orders, but it is not good for reproducing large P_T jets and leptons. But heavy particle decays, as expected beyond the SM, lead to large transverse momenta of leptons and jets. Large P_T particles can be determined only by explicitly calculating the matrix element (ME) from Feynman diagrams. To avoid double counting of partons from the resummation via the leading-log approach and the explicitly calculated ME partons detailed *matching schemes* have been designed, like the MLM and CKKW methods. However, various MC generators, like Alpgen, Madgraph, Sherpa and MC@NLO which are based on matching ME calculations with a parton shower evolution (either Pythia or Herwig), lead to $40 - 50\%$ differences in the high P_T tails of the production of heavy gauge bosons and top production. This leads to signifi-

cant differences in the background estimation for the production of heavy particles beyond the SM. To determine these effects quantitatively the background for the CMS-SUSY benchmark points has been calculated with different generators, which leads to a mean background with an uncertainty taken to be the maximum difference from various generators to this mean background. This uncertainty is then propagated as a systematic uncertainty for the calculation of the discovery reach of the various SUSY search channels. As expected, for the channels with large cross-sections, the systematic uncertainty dominates over the statistical uncertainty and the effect of systematic uncertainties is large, while for search channels with a small cross-section, but also a small SM background, the background uncertainty place less of a role.

The discovery reaches with and without systematic uncertainties have been summarized in Figure 6.28 in Chapter 6 for the following search channels (assuming an integrated luminosity of 100 pb^{-1} and benchmark points LM0, LM1): all-hadronic, single- and di-muon plus missing transverse energy and tri-muon final states. As can be seen there, the systematic uncertainties are dominating for all-hadronic and single muon SUSY searches, which have the smallest statistical errors.

Outlook: in future, data on W, Z and top production will decide which MC generators describe data best. Especially by having clean tagged samples one can determine P_T and E_T^{Miss} distributions for SM processes accurately. Together with other data driven techniques one will be able to reduce the systematic errors of SM backgrounds significantly below the values determined from the difference between generators, thus paving the way for searches beyond the SM with higher sensitivity.

Appendix A

Lepton Selection in CMS reconstruction

The leptonic part of this study has attempted to search for SUSY using only muons for selection, since muons are considered to be the most robust physics objects in CMS. Missing transverse energy and jets, used in the hadronic searches, have been studied for a single and di-muon search with lepton veto step. In the tri-muon search those variables has been not considered for the SUSY discovery.

A.1 Muon Identification

The parameters used for the muon identification are motivated by the differences of prompt muons (Section A.1.1) and fake muons (Section A.1.2). The muon identification, which has been used for the muonic SUSY search, is based on three identifier:

- a **global muon** consisting of a reconstructed track in both, the silicon tracker and the muon system, with good muon track quality. The track quality has been determined by the normalized-χ^2 (χ^2/dof) of the global track and the number of valid hits N_{Hits} in the silicon tracker. The muon identification was thoroughly studied at CMS and the *GlobalMuonPromptTight* selection with $\chi^2/dof \leq 10$ was recommended with an additional requirement on the silicon track $N_{Hits} \geq 11$.

- **the relative isolation of the muon**. The isolation has been determined in a cone around the muon track defined by $0.01 < \Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.3$ and was calculated in the tracker ($isoTr$) by the sum of transverse momentum of all tracks and in the calorimeter ($isoCal$) by summing up the energy deposition in the ECAL and HCAL. The relative tracker isolation $isoTr/PT$ and the relative calorimeter isolation $isoCal/PT$ weights the isolation with the transverse momentum of the muon i.e. low PT muons were required to be better isolated than high PT muons. The requirement of isolated muons

(i.e. small values of $isoTr/PT$ and $isoCal/PT$) is a powerful discriminant to reject all kind of fake muons, since all fake muons are produced in jets.

- **the impact parameter significance** $Sdxy$ calculated as the impact parameter dxy divided by its uncertainty σ_{dxy}. The impact parameter has been determined as the distance in the x-y-plane of the muon track and the primary vertex of the event. Instead of the impact parameter the significance has been considered, since a selection on the significance automatically scales with the vertex resolution, so when at the beginning the vertex resolution is worse, the selection will be looser.

A.1.1 Prompt Muons

The characteristic properties of prompt muons correspond to the production mechanism. Prompt muons appear isolated in the detector, since they are produced in a clean decay without additional particles producing tracks in the tracker and energy deposit in the calorimeter around the muon track. The mother particle has a short lifetime and decays immediately after production. Therefore the track of the muon should be close to the primary vertex of the event and the fit of global track is expected to have a good quality.

There is a difference between SUSY muons and muons from SM boson decays. In the low mass SUSY region where the mass difference of gauginos is small, the production of the lightest neutralino χ_1^0 from decay of the gaugino results in a softer muon P_T distribution. The more heavy the initially produced particles are, the more central in the detector the muons are. Therefore muons produced in SUSY processes are more central than in SM processes [44].

A.1.2 Fake Muons

Fake muons have been defined as all muons, which are not coming from SM bosons and gauginos and can be categorized as follows.

- Muons from decays of heavy flavor hadrons consisting of b, c quarks produced in the hard process and in gluon splitting. Heavy flavor hadrons have a non-zero lifetime ($c\tau \sim$ 0.1-0.5 mm) [54] and can travel through the tracker before decaying. On the other hand the hadron decays of long living mesons, $K^{\pm}$, $\pi^{\pm}$, can produced muons from those decays.

- Escaping charged particles from the calorimeter can produce some tracks in the muon system and can be misidentified as muons, the so called punch-through muons.

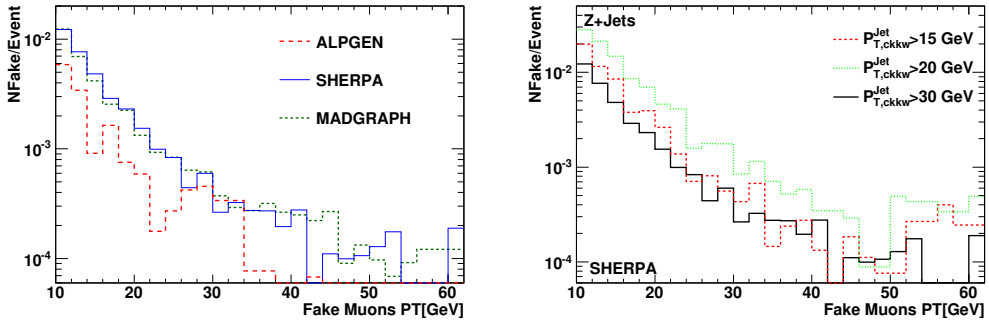


Figure A.1: P_T distribution of fake muons after the standard cuts (see section 5.1.1) for the Z+jets production simulated with MC event generators Sherpa, Madgraph and Alpgen(left) and in different matching scale (CKKW) systematics in Sherpa simulated(right). The distributions are normalized to the number of events. The fake muon production is prone to large theoretical uncertainties between Alpgen and other generators, the P_T distribution for fakes in different matching scale is rather similar (the $P^{Jet}_{m,ckkw} > 20$ GeV is full reconstructed samples compared to the others).

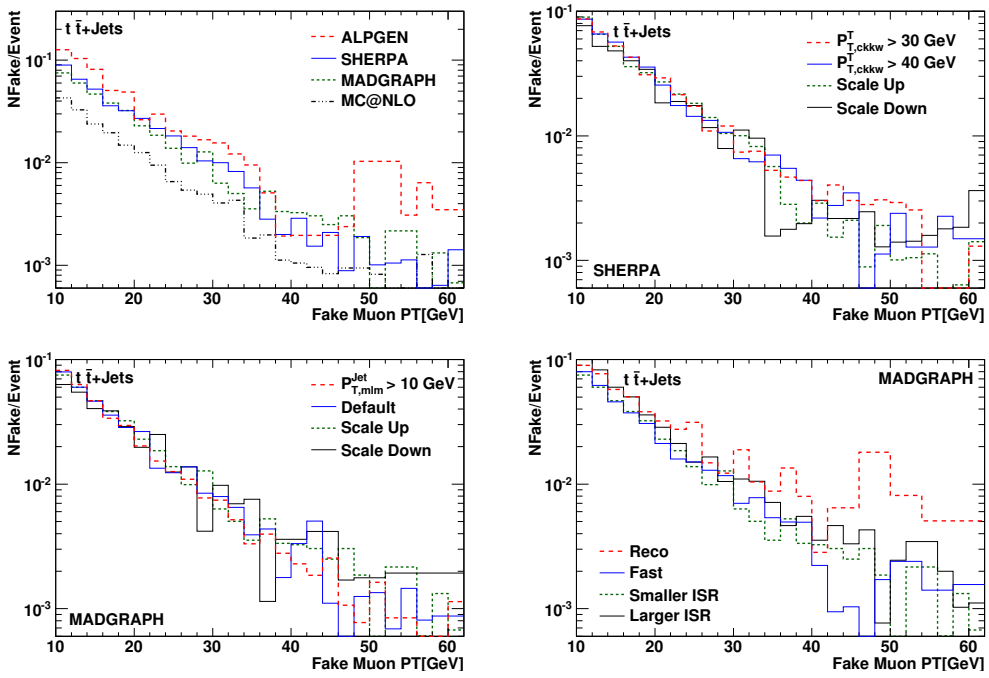


Figure A.2: P_T distribution of fake muons after the standard cuts (see section 5.1.1) on the $t\bar{t}$+jets production simulated with MC event generators Sherpa,MC@NLO, Madgraph and Alpgen(top-left) and in different matching scale (CKKW) systematics in Sherpa simulated(top-right). The distributions are normalized to the number of events. The fake muon production is prone to large theoretical uncertainties (L), the PT distribution for fakes in different channels is rather similar and the fake rate in the different channels depend on the heavy flavor content of the event (R).

The production rate of fake muons is prone to large systematic uncertainties. The contribution of fake muons from heavy flavor decays has large uncertainties from

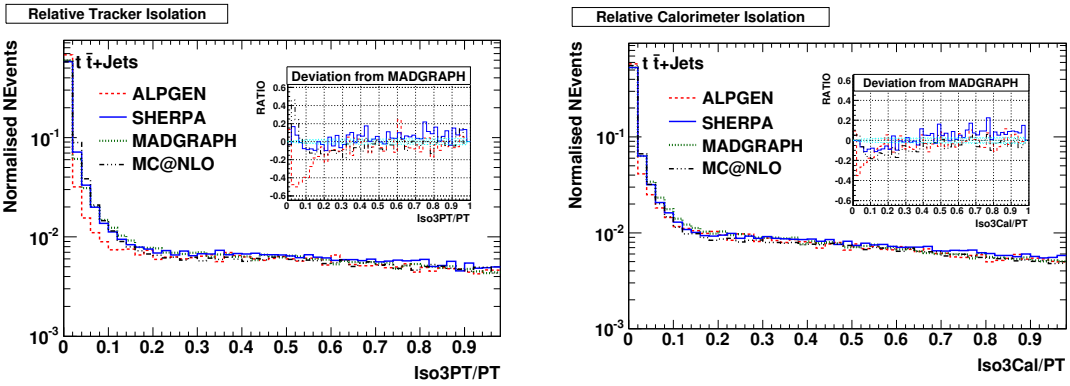


Figure A.3: Normalized relative isolations isoPT/PT, isoCal/PT between different MC truth samples of $t\bar{t} + jets$ for GlobalMuons. The standard cuts (see section 5.1.1) has been applied for the comparison.

quarks produced in initial and final state radiation in the parton shower evolution, since it is sensitive to the gluon parton distribution function and the splitting probability into heavy quarks. The sensitivity to MC details was checked with $t\bar{t}$+jets and Z+jets samples simulated with different MC generators.

A.2 Electron Identification

The electron part of this study has attempted to avoid the electrons for SUSY searches. Therefore the veto selection, used for the hadronic and multi-muonic searches, has been discussed with the corresponding the electrons definitions (see details in section 3.7.4).

The same isolation and impact parameter implementation have been simply used for the electron veto selection. `Loose Electrons` and `RobustTightElectrons` have been used for the hadronic and multi-muonic SUSY searches, respectively. The sequence used to reconstruct the electrons [141] is called `pixelMatchGsfElectrons`. Starting from an ECAL super-cluster (E_{SC}), a pixel track seed matching the super-cluster is searched for. If the seed found, the pattern recognition is performed with the Combinatorial Track Finder algorithm in a loose cut configuration while the final fit with the Gaussian Sum Filter(GSF) [142] . GSF is a fitting algorithm dedicated to electrons that accounts for the electron bremsstrahlung energy loss. As reconstructed pions are often misidentified based electron-id algorithm [143]. Electron isolation is also required: no tracks with $P_T > 1.5$ GeV have to lie a cone with $0.02 < \Delta R < 0.2$ around the electron: the ECAL deposit within a cone with ΔR=0.3 required to be $< 0.05 X E_{SC}$, while the HCAL deposit in a cone with $0.15 < \Delta R < 0.3$ has to be $< 0.2 X E_{SC}$. In our analysis the most simply configuration has been applied for the electron veto selection for hadronic and multi-muonic searches. The further details and explanation for the electron identification can be found in CMS PTDR2.

Appendix B

Statistical Methods

From a statistical point of view, any search for new physics is a Hypothesis test between the null Hypothesis H_0, which assumes, that there is no new physics present, i.e. the background only Hypothesis, and the Alternate Hypothesis H_1, which includes new physics contributions to the measured quantity. In order to claim discovery, the null Hypothesis has to be rejected at a given probability, the p-value. The general measure to express how good an observation can be explained by the null Hypothesis is the so called significance. It is the number of Gaussian standard deviations, beyond which the integral of one side of a Normal distribution is equal to the p-value of the null Hypothesis. For CMS statistical recommendations see [144]. The study has been performed with the RooStats Package included in ROOT package, which provides to describe the model and perform Hypotheses tests.

B.1 Model building

The first step has to be the correct determination of the two Hypotheses to test.
Each Hypothesis is in principle a probability density function *pdf* predicting the distribution of the observables of the experiment. If it is not completely determined, but contains free parameters, it is called a composite Hypothesis.
In the present case, the only observable is the number of events which survives certain cuts. So the basic pdf is a Poisson distribution, which has only one free parameter μ, the mean number of expected events.

$$P(n|\mu) = \frac{\mu^n e^{-\mu}}{n!} \tag{B.1}$$

where n is the number of observed events. In case of the null Hypothesis, the expected number of events will just be background b, $\mu = b$ whereas the mean of H_1 contains an additional contribution of new physics s, $\mu = s + b$.
If systematic uncertainties on the background are present, they are introduced via *nuisance parameters*. The most common assumption about the distribution of those

systematics is a Gaussian shape. So the mean of the Poisson distribution will be modified $b \rightarrow b\prod_i^N(1+\delta_i)$, where the δ_i are the nuisance parameters and there are N sources of systematics present. Each of these nuisance parameters is Gaussian distributed around 0 with the width corresponding to the relative uncertainty, it parametrizes. So the complete *pdf* including systematic uncertainties becomes:

$$P(n|s,\delta_1,...,\delta_N) = Poisson(n|s+b\prod_i^N(1+\delta_i))\prod_j^N Gauss(\delta_j|0,\sigma_j) \tag{B.2}$$

Sometimes, correlations between different nuisance parameters can be present. In this case, the one dimensional Gaussian distributions have to be replaced by a multivariate one including the correlations and widths of all nuisances in the covariance matrix.

B.2 Profile Likelihood method

Basically, testing the Hypotheses means defining a test statistic, a quantity which can be calculated from the observation and which allows for a good distinction between the two Hypotheses under consideration. This means, that the distribution of the test statistic under the different Hypotheses should be as different as possible. With a real value for the test statistic from a measurement, the p value can be calculated as the probability of the null Hypothesis to yield a result as extreme as observed or even more extreme, which in fact is the integral of the tail of the distribution. If an experiment is performed, it results in real measurements of the observables, whose distributions are predicted by the statistical model. By inserting the measurement into the model pdf, the so called likelihood function $L(DATA|Hypothesis)$ is built, which is the basic quantity for any statistical inference.
The test statistic used in the present analysis is the *Profile Likelihood ratio* which is the ratio

$$\lambda = \frac{L(n|s=0,\hat{\vec{\delta}})}{L(n|\hat{s},\hat{\vec{\delta}})} \tag{B.3}$$

It is the ratio of the Likelihoods for null and alternate Hypotheses, each maximized with respect to all free parameters. In the former case the only free parameters are the nuisances, in the latter the signal strength is an additional free parameter. Now the distribution of this quantity under the null Hypothesis has to be found out, which is in general done by simulating toys. Wilks's theorem [145] however states, that, in the limit of an infinite number of observations taken into account, the distribution of $-2ln\lambda$ under the null is a χ^2 distribution. The number of degrees of freedom is the difference in degrees of freedom between the alternate and the null hypothesis, so it is one in our case.
It has to be noted, that here, we use only one measurement to construct the likelihood, so the asymptotic assumption of Wilks's theorem is not fulfilled. It has

been shown however, that the Profile Likelihood method used this way also works astonishingly well for non asymptotic cases [146]. So with the value of $x = -2ln\lambda$ obtained from experiment, the p value is just

$$p = \int_x^\infty P_{\chi_1^2}(x')dx', \tag{B.4}$$

the integral of the χ^2 distribution with one degree of freedom from the observation to infinity. The significance Z can then be calculated from the Normal distribution $N(0,1)$ with $\mu = 0$ and $\sigma = 1$:

$$\int_Z^\infty N(x)dx \stackrel{!}{=} 1 \tag{B.5}$$

B.2.1 Specific model for this study

As this study is devoted to theoretical systematics from the Monte Carlo generation, instrumental uncertainties are not taken into account.
The background for each single SUSY search is a superposition of different standard model processes. As each of those can have different uncertainties, they are considered separately, i.e. $b = \sum_i^N b_i$, if there are N background processes present. For each of those processes, two systematic uncertainties are taken into account. One, σ_{theo}, is coming from different generator tunes, the comparison of different generators respectively. The second one, σ_{size}, is due to the limited size of Monte Carlo Samples which causes Poisson uncertainties on each b_i. All uncertainties are assumed uncorrelated, so each b_i receives two nuisance parameters, which are Gaussian distributed according to their relative uncertainty.

$$P\left(n|s,\vec{\delta}_{theo},\vec{\delta}_{size}\right) = Poisson\left(s + \sum_i^N b_i(1+\delta^i_{theo})(1+\delta^i_{size})\right)$$
$$\prod_j^N Gaussian(\delta^j_{theo}|0,\sigma^j_{theo}) \prod_k^N Gaussian(\delta^k_{size}|0,\sigma^k_{size}) \tag{B.6}$$

B.2.2 Datacard for the significance calculation in ROOTSTAT

```
//////////////input section, here you have to specify, your numbers
double signal=77; //signal expectation
//define different background expectations
double _b_1,(2,3,4,5)=11.32; //background (bg) expectations
int flag_1_1=1; //specifies, if bg has uncertainty.
double sigma_1_1=0.35; //specifies error (35%) on bg source
int flag_1_2=1; //specifies, if bg has uncertainty.
```

```
double sigma_1_2=0.1; //specifies error (10%) on bg source,
///////create the corresponding RooRealVars

RooRealVar n_obs("n_obs","n_obs",_signal + _b_1 + _b_2 + _b_3 + _b_4 + _b_5, 0, 5 *
(_signal + _b_1 + _b_2 + _b_3 + _b_4 + _b_5));
cout << "n_obs" << n_obs.getVal() << endl;
RooRealVar signal("signal","signal",_signal, 0, 5 * _signal);


RooRealVar flag_1_1, 2("flag_1_1, 2","flag_1_1, 2",_flag_1_1, 2);
RooRealVar flag_2_1, 2("flag_2_1, 2","flag_2_1, 2",_flag_2_1, 2);
RooRealVar flag_3_1, 2("flag_3_1, 2","flag_3_1, 2",_flag_3_1, 2);
RooRealVar flag_4_1, 2("flag_4_1, 2","flag_4_1, 2",_flag_4_1, 2);
RooRealVar flag_5_1, 2("flag_5_1, 2","flag_5_1, 2",_flag_5_1, 2);


RooRealVar b_1("b_1","b_1",_b_1);
RooRealVar b_2("b_2","b_2",_b_2);
RooRealVar b_3("b_3","b_3",_b_3);
RooRealVar b_4("b_4","b_4",_b_4);
RooRealVar b_5("b_5","b_5",_b_5);


RooRealVar delta_1_1, 2("delta_1_1, 2","delta_1_1, 2", 0, -1, 10);
RooRealVar delta_2_1, 2("delta_2_1, 2","delta_2_1, 2", 0, -1, 10);
RooRealVar delta_3_1, 2("delta_3_1, 2","delta_3_1, 2", 0, -1, 10);
RooRealVar delta_4_1, 2("delta_4_1, 2","delta_4_1, 2", 0, -1, 10);
RooRealVar delta_5_1, 2("delta_5_1, 2","delta_5_1, 2", 0, -1, 10);


RooRealVar sigma_1_1, 2("sigma_1_1, 2","sigma_1_1, 2",sigma_1_1, 2);
RooRealVar sigma_2_1, 2("sigma_2_1, 2","sigma_2_1, 2",sigma_2_1, 2);
RooRealVar sigma_3_1, 2("sigma_3_1, 2","sigma_3_1, 2",sigma_3_1, 2);
RooRealVar sigma_4_1, 2("sigma_4_1, 2","sigma_4_1, 2",sigma_4_1, 2);
RooRealVar sigma_5_1, 2("sigma_5_1, 2","sigma_5_1, 2",sigma_5_1, 2);


//Take this as the mean of all gaussinas
RooRealVar zero_var("zero_var","zero_var", 0);

//These shouldn't be fitted
n_obs.setConstant();
flag_1, (2, 3, 4, 5)_1, 2.setConstant();
b_1, (2, 3, 4, 5).setConstant();
sigma_1, (2, 3, 4, 5)_1, 2.setConstant();
zero_var.setConstant();


//create poisson for measurement
RooArgSet poisson_parameters(signal);
```

```
poisson_parameters.add(flag_1_1);
poisson_parameters.add(flag_1_2);
poisson_parameters.add(b_1,(2,3,4,5));
poisson_parameters.add(delta_1_1);
poisson_parameters.add(delta_1_2);
poisson_parameters.add(flag_2_1);
poisson_parameters.add(flag_2_2);


RooFormulaVar
poisson_mean("poisson_mean","signal + b_1 * (1 + delta_1_1 * flag_1_1) * (1 +
delta_1_2 * flag_1_2) + b_2 * (1 + delta_2_1 * flag_2_1) * (1 + delta_2_2 * flag_2_2) +
b_3 * (1 + delta_3_1 * flag_3_1) * (1 + delta_3_2 * flag_3_2) + b_4 * (1 + delta_4_1 *
flag_4_1)*(1+delta_4_2*flag_4_2)+b_5*(1+delta_5_1*flag_5_1)*(1+delta_5_2*
flag_5_2)",poisson_parameters);
// RooFormulaVar poisson_mean("poisson_mean","signal+b_1+b_2+b_3+b_4+
b_5",poisson_parameters);
cout << "mean" << poisson_mean.getVal() << endl;


//create gaussians for nuisance parameters
RooGaussian uncertainty_1(2,3,4,5)_1(2,3,4,5)("uncertainty_1(2,3,4,5)_1,2",
"uncertainty_1(2,3,4,5)_1,2",delta_1(2,3,4,5)_1,2,zero_var,sigma_1(2,3,4,5)_1,2);


//Create likelihood as a product of those
RooArgList pdfs(counting_model,uncertainty_1_1,uncertainty_2_1,
uncertainty_3_1,uncertainty_4_1,uncertainty_5_1);
pdfs.add(uncertainty_1_2);
pdfs.add(uncertainty_2_2);
pdfs.add(uncertainty_3_2);
pdfs.add(uncertainty_4_2);
pdfs.add(uncertainty_5_2);


RooProdPdf likelihood("likelihood","likelihood",pdfs);


//define Data
RooDataSet* data = new RooDataSet("Data", "Data", RooArgSet(n_obs));
data->add(RooArgSet(n_obs));


//define paramsOfInterest
RooArgSet *paramsOfInterest=new RooArgSet("paramsOfInterest");
paramsOfInterest->addClone(signal); //clone because we need s for complete
likelihood
```

```
//build likelihood ratio
ProfileLikelihoodCalculator plc(*data,likelihood, *paramsOfInterest);


//Get Result
// Get Confidence Interval
plc.SetParameters(*paramsOfInterest);//define the parameters which are
not nuisance


//get hypo test result
paramsOfInterest->setRealValue("signal",0);
plc.SetNullParameters(*paramsOfInterest);
HypoTestResult* lrhypo=plc.GetHypoTest();


cout<<"significance by hypotest:  "<<lrhypo->Significance()<<endl;
```

Appendix C

Data Cards of the simulated samples

Sherpa datacard for v1.1.3

```
(BEAM) !-- Beam parameters -----------------------------------
BEAM-1 = 2212 !  possible beam particles:  P+, P-, e+, e-
BEAM-ENERGY-1 = 5000.  !  in GeV
BEAM-POL-1 = 0. !  Polarization degree -1 ...  1
BEAM-SPECTRUM-1 = Monochromatic !  Monochromatic
K-PERP-MEAN-1 = 0.2
K-PERP-SIGMA-1 = 0.8
BEAM-2 = 2212 !  possible beam particles:  P+, P-, e+, e-
BEAM-ENERGY-2 = 5000.  !  in GeV
BEAM-POL-2 = 0. !  Polarization degree -1 ...  1
BEAM-SPECTRUM-2 = Monochromatic !  Monochromatic
K-PERP-MEAN-2 = 0.2
K-PERP-SIGMA-2 = 0.8
BEAM-SMIN = 1.e-10 !  Minimal fraction of nominal s after beam spectra
BEAM-SMAX = 1.0 !  Maximal fraction of nominal s after beam spectra
E-LASER-1 = 1.17e-9 !  Laser energy in GeV
P-LASER-1 = 0. !  Laser polarization +-1
E-LASER-2 = 1.17e-9 !  Laser energy in GeV
P-LASER-2 = 0. !  Laser polarization +-1
```

LASER-MODE = 0 ! 0 = all, 1,2,3 = individual components

LASER-ANGLES = Off ! On/Off

LASER-NONLINEARITY = On ! On/Off

(Initial State Radiation)

-- ISR parameters ----------

BUNCH-1 = 2212 ! possible beam particles: P+, P-, e+, e-

ISR-1 = On ! On/Off

BUNCH-2 = 2212 ! possible beam particles: P+, P-, e+, e-

ISR-2 = On ! On/Off

ISR-SMIN = $1.e{-}10$! Minimal fraction of nominal s for parton after ISR

ISR-SMAX = 1.0 ! Maximal fraction of nominal s for parton after ISR -

ISR-E-ORDER = 1 ! Perturbative order of electron structure function

ISR-E-SCHEME = 2 ! Beta-scheme : 0,1,2 , default = 2

PDF-SET = cteq6l1 ! the CTEQ6 PDFs cteq6m,cteq6d,cteq6l,cteq6l1

PDF-SET-VERSION = 1

PDF-GRID-PATH = CTEQ6Grid ! CTEQ6Grid,PDFsets,MRST99Grid

(Model) (model)

!-- Model parameters ----------

MODEL = SM ! Model

SM parameters

EW-SCHEME = 0 ! which parameters define the ew sector.

ALPHAS(MZ) = 0.118 ! strong coupling at scale M-Z

ALPHAS(default) = 0.0800 ! strong coupling

ORDER-ALPHAS = 1 ! NLO

1/ALPHAQED(0) = 137.036 ! inverse of alpha QED in the Thomson limit

1/ALPHAQED(default) = 132.51 ! inverse of alpha QED

SIN2THETAW = 0.2222 ! Weinberg angle at scale M-Z

VEV = 246. ! Higgs vev

LAMBDA = 0.47591 ! SM Higgs self coupling

CKMORDER = 0 ! order of expansion of CKM matrix in Cabibbo angle

```
CABIBBO = 0.22 !  Cabibbo angle (Wolfenstein parametrization)
A = 0.85 !  A (Wolfenstein parametrization)
RHO = 0.50 !  rho (Wolfenstein parametrization)
ETA = 0.50 !  eta (Wolfenstein parametrization)
MASS[6] = 172.4 !  top quark mass
MASS[5] = 4.8 !  b quark mass
SUSY parameters
SLHA-INPUT = LesHouches-SPS1A.dat
ADD parameters ...
STABLE[15]=0 !  set tau stable
-----------------------------------------------------------------
```

(Matrix Element) (ME)

```
!-- ME generators ------------------------------------
ME-SIGNAL-GENERATOR = Amegic !  Internal or Amegic
EVENT-GENERATION-MODE = Unweighted
COUPLING-SCHEME = Running-alpha-S ! Fixed (default value) or
!  Running (s for ME generators, pt for 2 - 2)
YUKAWA-MASSES = Fixed !  Fixed (polemass) or
!  Running (higgs mass)
YUKAWA-MASSES-FACTOR = 1.  !  Additional prefactor for yukawas
SCALE-SCHEME = CKKW ! see 'http://www.sherpa-mc.de/scales.html'
KFACTOR-SCHEME = 1 !  default=1
SUDAKOV-WEIGHT = 1 !  apply sudakov weight on single events
SCALE-FACTOR = 1.  !  factor the scale is multiplied with 0.5 or 2
-----------------------------------------------------------------
```

(Multiple Interactions)

```
!-- UE Underlying Event Setup file-------------------------------------
general parameters
MI-HANDLER = Amisic !  Amisic / None
hard underlying event parameters
CREATE-GRID 93 93 -> 93 93 !  processes to generate
PS-ERROR = 1.0e - 2 !  error for integration
```

```
REGULATE-XS = 0 !  regulate cross section
XS-REGULATION = 2.225 !  regulation parameter
SCALE-MIN = 2.225 !  minimum scale
RESCALE-EXPONENT = 0.16 !  rescaling exponent
REFERENCE-SCALE = 1800.0 !  reference energy scale
PROFILE-FUNCTION = Gaussian !  Gaussian / Double-Gaussian
PROFILE-PARAMETERS = 1.0 0.5 0.5 !  size (must be 1), coresize, matter fraction
-----------------------------------------------------------------
```

(Parton Shower)

```
!-- Parton showers -----------------------------------
SHOWER-GENERATOR = Apacic !
FSR-SHOWER = 1 !  1=On,0=Off
ISR-SHOWER = 1 !  1=On,0=Off
IS-PT2MIN = 4.  !  IS Shower cutoff
FS-PT2MIN = 1.  !  FS Shower cutoff
-----------------------------------------------------------------
```

(Fragmentation)

```
!-- Fragmentation parameters-------------------------
FRAGMENTATION = Ahadic !  Off, Lund (Pythia string fragmentation) or Ahadic
DECAYMODEL = Hadrons !  Lund or Hadrons
-----------------------------------------------------------------
```

(Processes)

```
!-- Processes to calculate ---------------------------
jet jet -> μ ν + n-0, 1, 2, 3 jets
Process :  93 93 -> 13 -14 93{3}
Order electroweak :  2
End Process
Process :  93 93 -> -13 14 93{3}
Order electroweak :  2
End Process
-----------------------------------------------------------------
```

(Selector)

```
!-- Parton level selectors ----------------------------
JetFinder sqr(20/E-CMS) 0.7
-----------------------------------------------------------------
```

(selector)

Alpgen datacard for v2.13

```
ih2 1 !  pp collision
ebeam 5000 !  beam energy in GeV
njets 3 !  jets multiplicities - exclusive
iqopt 1 !  factorization / renormalizations scales
ickkw 1 !
ptjmin 20 !  pt of the jet, in GeV
drjmin 0.7 !  ΔR
etajmax 5 !  η of the jet
iwdecmod 2 !  W decay to μ ν
ndns 9 !  choose pdf data set - cteq6l1
mc 0.  !  c quark mass
mb 4.8 !  b quarks mass
mt 172.4 !  t quark mass
Alpgen interface to the Pythia-MC in CMSSW.
```

Pythia datacard for the CMSSW

```
process.source = cms.Source("AlpgenSource",
GeneratorParameters = cms.PSet(
parameterSets = cms.vstring('generator'),
generator = cms.vstring('IXpar(2) = 1 !  inclus./exclus.  sample:  0/1',
'RXpar(1) = 20. !  ETCLUS : minET(CLUS)',
'RXpar(2) = 0.7!  RCLUS : deltaR(CLUS)')
),
pythiaHepMCVerbosity = cms.untracked.bool(False),
fileNames = cms.untracked.vstring('file:TEMPLATE'),
pythiaPylistVerbosity = cms.untracked.int32(0),
```

```
PythiaParameters = cms.PSet(
parameterSets = cms.vstring('pythia'),
pythia = cms.vstring('MSEL=0 !  (D = 1) ',
'MSTJ(11)=3 !  Choice of the fragmentation function',
'MSTP(143)=1 !  Call the matching routine in ALPGEN',
'MSTJ(11)=3 !  Choice of the fragmentation function',
'MSTJ(22)=2 !  Decay those unstable particles',
'PARJ(71)=10 !  for which ctau 10 mm',
'MSTP(2)=1 !  which order running alphaS',
'MSTP(33)=0 !  no K factors in hard cross sections',
'MSTP(51)=7 !  structure function chosen (external PDF CTEQ6L1)',
'MSTP(81)=1 !  multiple parton interactions 1 is Pythia default',
'MSTU(21)=1 !  Check on possible errors during program execution'
) ) )
```

MC@NLO-v3.3 (Herwig) datacard for the CMSSW

```
process.source = cms.Source("MCatNLOSource",
HerwigParameters = cms.PSet(
defaultHerwig = cms.vstring(),
parameterSets = cms.vstring('defaultHerwig')
),
doMPInteraction = cms.untracked.bool(True),
processNumber = cms.untracked.int32(-1362),
herwigHepMCVerbosity = cms.untracked.bool(False),
doHardEvents = cms.untracked.bool(True),
herwigVerbosity = cms.untracked.int32(0),
comEnergy = cms.untracked.double(10000.),
MCatNLOParameters = cms.PSet(
mcatnloReadin = cms.vstring('TWIDTH=1.4',
'ECM=10000', 'FREN=1', 'FFACT=1', 'HVQMASS=172.4', 'WMASS=80.419',
'WWIDTH=2.124', 'ZMASS=91.17', 'ZWIDTH=2.495', 'HGGMASS=165', 'HGGWIDTH=1.04',
'IBORNHGG=1', 'V1GAMMAX=30', 'V1MASSINF=0', 'V1MASSSUP=0', 'V2GAMMAX=30',
'V2MASSINF=0', 'V2MASSSUP=0', 'HGAMMAX=30', 'HMASSINF=0', 'HMASSSUP=0',
```

```
'UMASS=0.32', 'DMASS=0.32', 'SMASS=0.5', 'CMASS=1.55', 'BMASS=4.95',
'GMASS=0.75', 'VUD=0.9748', 'VUS=0.2225', 'VUB=0.0036', 'VCD=0.2225',
'VCS=0.9740', 'VCB=0.041', 'VTD=0.009', 'VTS=0.0405', 'VTB=0.9992',
'AEMRUN=YES', 'IVCODE=1', 'IL1CODE=1', 'IL2CODE=1', 'PDFGROUP=LHAPDF',
'PDFSET=20060', 'LAMBDAFIVE=1', 'SCHEMEOFPDF=MS', 'LAMBDAHERW=-1',
'FPREFIX=zmumu', 'EVPREFIX=zmumu', 'WGTTYPE=1', 'PDFLIBRARY=HWLHAPDF',
'HERPDF=EXTPDF', 'LHAOFL=FREEZE'),
parameterSets = cms.vstring('mcatnloReadin')
),
mcatnloVerbosity = cms.untracked.int32(0),
lhapdfSetPath = cms.untracked.string(''),
printCards = cms.untracked.bool(False),
numHardEvents = cms.untracked.int32(10000),
maxEventsToPrint = cms.untracked.int32(0),
stringFileName = cms.untracked.string('stringInput.txt'),
useJimmy = cms.untracked.bool(True) )
```

Appendix D

MC Samples in CMS Database

All Madgraph samples and some of Sherpa samples are available in the CMS database[1].

Table D.1: Madgraph data samples in DBS

Samples	cs[pb]	N events	Datasets	Settings
MadGraph-v4:				
$t\bar{t}$+jets	317	1000K	$t\bar{t}$JetsFall08	$P_{T,mlm}^{jets} \geq 30$ GeV
		660K	$t\bar{t}$Jets-Winter09	$P_{T,ckkw}^{jets} \geq 10$ GeV
		730K	$t\bar{t}$Jets-Winter09	$P_{T,ckkw}^{jets} \geq 30$ GeV
		552K	$t\bar{t}$Jets-Winter09	Scales Up (x2)
		630K	$t\bar{t}$Jets-Winter09	Scales Down (x1/2)
		935K	$t\bar{t}$Jets-Winter09	Larger ISR
		720K	$t\bar{t}$Jets-Winter09	Smaller ISR
W+jets	40000	5430K	WJets-Summer08	$P_{T,mlm}^{jets} \geq 30$ GeV
		2900K	WJets-Winter09	Scales Up (x2)
		6000K	WJets-Winter09	Scales Down (x1/2)
Z+jets	3700	1010K	ZJets-Summer08	$P_{T,mlm}^{jets} \geq 30$ GeV
		5420K	ZJets-Winter08	$P_{T,mlm}^{jets} \geq 30$ GeV
		995K	ZJets-Winter08	Scales Down (x1/2)
		200K	ZJets-Winter08	Scales Up (x2)
ZInvisible+Jets	2000	1010K	ZInvJets-Summer08	Default
QCD+jets	15+E06	1000K	QCDJets-Summer08	ht100to250
	400000	500K	QCDJets-Summer08	ht250to500
	14000	400K	QCDJets-Summer08	pt500to1000
	370	250K	QCDJets-Summer08	pt1000toInfinity
A^*+Jets	6000	1200K	A^*Jets-Summer08	$P_{T,mlm}^{jets} \geq 30$
VV+Jets	20	900K	VVJets-Summer08	$P_{T,mlm}^{jets} \geq 30$

[1]https://cmsweb.cern.ch/*dbs_discovery/_advanced*?userMode=user

Table D.2: Sherpa, Alpgen and MC@NLO data samples in DBS

Samples	cs[pb]	N events	Datasets	Settings
Alpgen-v2.13:				
W+jets	43000	2000K	WJets-Winter09	$P^{jets}_{T,mlm} \geq 30$ GeV
Z+jets	4000	1000K	ZJets-Winter09	$P^{jets}_{T,mlm} \geq 30$ GeV
$t\bar{t}$+jets	210	500K	ZJets-Winter09	$P^{jets}_{T,mlm} \geq 30$ GeV
Sherpa-v1.1.3:				
$t\bar{t}$+jets	204	500K	$t\bar{t}$Jets-Winter09	$P^{jets}_{T,ckkw} \geq 10$ GeV
	206	800K	$t\bar{t}$Jets-Winter09	$P^{jets}_{T,ckkw} \geq 30$ GeV
	206	800K	$t\bar{t}$Jets-Winter09	$P^{jets}_{T,ckkw} \geq 40$ GeV
	186	770K	$t\bar{t}$Jets-Winter09	Scales Up (x2)
	227	770K	$t\bar{t}$Jets-Winter09	Scales Down (x1/2)
	206	770K	$t\bar{t}$Jets-Winter09	Lund Fragmentation
W+jets	42000	5200K	WJets-Summer08	Default
		4900K	WJets-Winter09	$P^{jets}_{T,ckkw} \geq 30$ GeV
		4900K	WJets-Winter09	1. Order EWeak Corr.
Z+jets	4000	900K	ZJets-Summer08	Default
		1000K	ZJets-Winter09	$P^{jets}_{T,ckkw} \geq 30$ GeV
		1420K	ZJets-Winter09	$P^{jets}_{T,ckkw} \geq 15$ GeV
		1000K	ZJets-Winter09	Scales Up (x2)
		1100K	ZJets-Winter09	Scales Down (x1/2)
ZInvisible+Jets	7000	3600K	ZInvJets-Winter09	Default
		1500K	ZInvJets-Winter09	Lund Fragmentation
		860K	ZInvJets-Winter09	Scales Up (x2)
		920K	ZInvJets-Winter09	Scales Down (x1/2)
QCD+jets	11.1+E06	1000K	QCDJets-Winter09	pt40to120
	31145	500K	QCDJets-Winter09	pt120to280
	181	400K	QCDJets-Winter09	pt280to500
	5.2	250K	QCDJets-Winter09	pt500to5000
A^*+Jets	6000	1200K	A^*Jets-Winter09	$P^{jets}_{T,ckkw} \geq 30$
VV+Jets	14.5	900K	VVJets-Winter09	$P^{jets}_{T,ckkw} \geq 30$
MC@NLO-v3.3:				
$t\bar{t}$+jets	366	460K	$t\bar{t}$Jets-Winter09	$P^{jets}_{T,NLO} \geq 30$

List of Figures

List of Tables

Bibliography

[1] T. N. P. Commitee, "From the Nobel Committee's Scientific Background". http://nobelprize.org.

[2] B. Schwarzschild, Phys. Today 61N12, 16–20 (2008).

[3] D. Griffiths, "Introduction to elementary particles". Weinheim, Germany: Wiley-VCH (2008) 454 p.

[4] M. E. Peskin and D. V. Schroeder, "An Introduction to quantum field theory". Reading, USA: Addison-Wesley (1995) 842 p.

[5] H. E. Haber, G. L. Kane, and T. Sterling, Nucl. Phys. B161, 493 (1979).

[6] G. L. Kane, W. W. Repko, and W. B. Rolnick, Phys. Lett. B148, 367–372 (1984).

[7] R. Barate et al., Phys. Lett. B565, 61–75 (2003).

[8] H. E. Haber and G. L. Kane, Phys. Rept. 117, 75–263 (1985).

[9] W. de Boer, "Vorlesung Introduction to Supersymmetry (SS2010)". Karlsruhe Institute of Technology (KIT).

[10] W. de Boer, Prog. Part. Nucl. Phys. 33, 201–302 (1994).

[11] W. de Boer, Nucl. Phys. Proc. Suppl. 173, s3–s7 (2009).

[12] S. P. Martin, "A Supersymmetry Primer". 1997. hep-ph/9709356.

[13] D. I. Kazakov, "Beyond the standard model (in search of supersymmetry)". 2000. hep-ph/0012288.

[14] P. Abreu et al., Nucl. Phys. B367, 511–574 (1991).

[15] R. Brandelik et al., Phys. Lett. B76, 361 (1978).

[16] E. Halkiadakis, Nuovo Cim. 032C, 255–263 (2009).

[17] D. E. Acosta et al., Phys. Rev. Lett. 94, 091803 (2005).

[18] F. Krauss, A. Schalicke, S. Schumann, and G. Soff, Phys. Rev. D70, 114009 (2004).

[19] S. Schael et al., Eur. Phys. J. C47, 547–587 (2006).

[20] **CDF** Collaboration, A. Patwa, "W and Z production and standard model Higgs search from the Fermilab Tevatron". 2006.

[21] V. Buscher, AIP Conf. Proc. 1200, 3–11 (2010).

[22] M. D'Onofrio, AIP Conf. Proc. 1200, 12–21 (2010).

[23] M. H. Seymour, Comp. Phys. Commun. 90, 95–101 (1995).

[24] G. Corcella and M. H. Seymour, Phys. Lett. B442, 417–426 (1998).

[25] G. Corcella, M. L. Mangano, and M. H. Seymour, JHEP 07, 004 (2000).

[26] M. Dobbs, Phys. Rev. D64, 034016 (2001).

[27] S. Frixione and B. R. Webber, JHEP 06, 029 (2002).

[28] **D0** Collaboration, V. M. Abazov et al., "Measurement of differential $Z/\gamma^* +$ jet $+ X$ cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$-TeV", volume B669. 2008.

[29] **CDF** Collaboration, A. Messina, "Measurement of the W + jet cross section at CDF", volume 37. 2007.

[30] **D0** Collaboration, M. Zielinski, "D0 measurement of the dijet azimuthal decorrelations", volume 842. 2006.

[31] V. M. Abazov et al., Phys. Lett. B678, 45–54 (2009).

[32] Wikipedia, "The Standard Model". PBS NOVA, Fermilab, Office of Science, United States Department of Energy.

[33] T. P. Cheng and L. F. Li, "Gauge theory of elementary particle physics: Problems and solutions". Oxford, UK: Clarendon (2000) 306 p.

[34] S. L. Glashow, "Partial Symmetries of Weak Interactions", volume 22. 1961.

[35] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. B47, 365–368 (1973).

[36] D. J. Gross and F. Wilczek, Phys. Rev. D9, 980–993 (1974).

[37] S. Weinberg, Phys. Rev. Lett. 31, 494–497 (1973).

[38] P. W. Higgs, Phys. Rev. 145, 1156–1163 (1966).

[39] H. W. Nilsen, "Studying Z/γ^*+jets Production in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV". 2009.

[40] J. Goldstone, Nuovo Cim. 19, 154–164 (1961).

[41] N. Cabibbo, Phys. Rev. Lett. 10, 531–533 (1963).

[42] H. Baer and X. Tata, "Weak scale supersymmetry: From superfields to scattering events". Cambridge, UK: Univ. Pr. (2006) 537 p.

[43] C. Sander, "Interpretation des Ueberschusses in diffuser galaktischer Gamma-Strahlung oberhalb 1GeV als Annihilationssignal dunkler Materie". 2005.

[44] M. Niegel, "Search for Supersymmetry in Tri-Muon Final States with the CMS Detector". 2009.

[45] R. N. Mohapatra, "Unification and Supersymmetry: The Frontiers of Quark - Lepton Physics". Berlin, Germany: Springer (1986) 309 P.

[46] U. Amaldi, W. de Boer, and H. Fuerstenau, Phys. Lett. B260, 447–455 (1991).

[47] J. Wess and B. Zumino, Phys. Lett. B49, 52 (1974).

[48] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, "THE HIGGS HUNTER'S GUIDE". SCIPP-89/13.

[49] I. Simonsen, "A Review of minimal supersymmetric electroweak theory". 1995. hep-ph/9506369.

[50] L. Alvarez-Gaume, J. Polchinski, and M. B. Wise, Nucl. Phys. B221, 495 (1983).

[51] G. L. Bayatian et al., J. Phys. G34, 995–1579 (2007).

[52] H. Baer, C. Balazs, A. Belyaev, T. Krupovnickas, and X. Tata, JHEP 06, 054 (2003).

[53] G. W. Bennett et al., Phys. Rev. Lett. 92, 161802 (2004).

[54] C. Amsler et al., Phys. Lett. B667 (2008).

[55] http://bigscience.web.cern.ch/bigscience/en/lhc/lhc2.html.

[56] http://press.web.cern.ch/press/PressReleases/Releases2010/PR13.10E.html.

[57] **ALICE** Collaboration, "ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC". LHC Tech. Proposal. CERN, 1995.

[58] **LHCb** Collaboration, "LHCb : Technical Proposal". Tech. Proposal. CERN, 1998.

[59] **ATLAS** Collaboration, "ATLAS: Technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN". CERN-LHCC-94-43.

[60] **CMS** Collaboration, "CMS physics". Technical Design Report CMS. CERN, 2006.

[61] R. Brun et al., "The CERN Large Hadron Collider: Accelerator and Experiments vol.i I, II". CERN and JINS. CERN-2009.

[62] http://cms.web.cern.ch/cms/Detector/FullDetector/index.html.

[63] R. Adolphi et al., JINST 0803, S08004 (2008).

[64] D. Kotlinski, Nucl. Instrum. Methods Phys. Res., A 477, 446–50 (2002), no. 1-3,.

[65] **CMS** Collaboration, "The CMS tracker system project". Technical Design Report CMS. CERN, 1997.

[66] **CMS** Collaboration, "The CMS electromagnetic calorimeter project". Technical Design Report CMS. CERN, 1997.

[67] **CMS** Collaboration, V. Adzic et al., "Energy Resolution of the Barrel of the CMS Electromagnetic Calorimeter," Technical Report CMS-NOTE-2006-148, CERN, 2006.

[68] **CMS** Collaboration, "The Hadron Calorimeter". Technical Design Report CMS. CERN, 1997.

[69] C. Kerem et al., "CMS HCAL installation and commissioning". Conf. Ser.160 012055 (8pp). J. Phys., 2009.

[70] **CMS** Collaboration, "The Magnet Project". Technical Design Report CMS. CERN, 1997.

[71] **CMS** Collaboration, "The CMS muon project". Technical Design Report CMS. CERN, 1997.

[72] **CMS** Collaboration, "CMS TriDAS project: The trigger systems". Technical Design Report CMS. CERN, 1997.

[73] V. Buege, "Virtualisation of Grid Resources and Prospects of the Measurement of Z Boson Production in Association with Jets at the LHC". Vdm Verlag, 2008.

[74] **CMS** Collaboration, "CMS computing". Technical Design Report CMS. CERN, 2005.

[75] C. Jones et al., "The new CMS Event Data Model and Framework," in *Proceedings of the CHEP06 Conference*. CERN, Mumbia, India, 2006.

[76] R. Brun et al., "ROOT: An Object-Oriented Data Analysis Framework: Users Guide 5.16". 2007.

[77] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP 05, 026 (2006).

[78] T. Sjostrand, "PYTHIA 8 Status Report". 2008. hep-ph/0809.0303.

[79] J. Yarba, "The PythiaInterface of the CMS Software Framework". https://twiki.cern.ch/twiki/bin/view/CMS/PyhiaInterface.

[80] G. Corcella et al., "HERWIG 6.5 release note". 2002. hep-ph/0210213.

[81] M. Bahr et al., "Herwig++ 2.3 Release Note". 2008. hep-ph/0812.0529.

[82] F. Stoeckli, "The HerwigInterface of the CMS Software Framework". https://twiki.cern.ch/twiki/bin/view/CMS/HerwigInterface.

[83] S. Frixione and B. R. Webber, "The MC@NLO 3.3 event generator". 2006. hep-ph/0612272.

[84] F. Maltoni and T. Stelzer, JHEP 02, 027 (2003).

[85] F. Caravaglios, M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. B539, 215–232 (1999).

[86] M. Mangano et al., JHEP 07, 001 (2003).

[87] J. Alwall et al., Comput. Phys. Commun. 176, 300–304 (2007).

[88] T. Gleisberg et al., JHEP 02, 056 (2004).

[89] F. Krauss, R. Kuhn, and G. Soff, JHEP 02, 044 (2002).

[90] F. Krauss et al., Comput. Phys. Commun. 174, 876–902 (2006).

[91] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, JHEP 11, 063 (2001).

[92] J.-C. Winter, F. Krauss, and G. Soff, Eur. Phys. J. C36, 381–395 (2004).

[93] M. Merschmeyer and M. Niegel, A. Cakir, "The SherpaInterface of the CMS Software Framework". https://twiki.cern.ch/twiki/bin/view/CMS/SherpaInterface.

[94] A. Djouadi and M. Muhlleitner, M.and Spira, Acta Phys. Pol. B 38, 635–644. 8 p (2006), no. hep-ph/0609292. CERN-PH-TH-2006-200,.

[95] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. 108, 56–74 (1998).

[96] M. Muhlleitner, Acta Phys. Polon. B35, 2753–2766 (2004).

[97] B. Allanach, Comput. Phys. Commun. 143, 305–331 (2002).

[98] K. Amako et al., Nuclear Physics B - Proceedings Supplements 150, 44–49 (2006).

[99] A. Heisterss et al., "Measurement of Jets with the CMS Detector at the LHC". CMS NOTE 2006/036. 2006.

[100] G. C. Blazey et al., "Run II jet physics". 2000. hep-ex/0005012.

[101] G. P. Salam and G. Soyez, JHEP 05, 086 (2007).

[102] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. B406, 187–224 (1993).

[103] T. C. Collaboration, "Performance of Jet Algorithms in CMS". CMS PAS JME-07-003. 2007.

[104] R. Kalman, Journal of Basic Engineering 82, 35–45 (1960).

[105] S. Baffioni et al., Eur. Phys. J. C49, 1099–1116 (2007).

[106] L. Moneta, I. Antcheva, R. Brun, and A. Kreshuk.

[107] S. D. Drell and T.-M. Yan, Ann. Phys. 66, 578 (1971).

[108] J. C. Collins and D. E. Soper, Ann. Rev. Nucl. Part. Sci. 37, 383–409 (1987).

[109] J. M. Campbell, J. W. Huston, and W. J. Stirling, Rept. Prog. Phys. 70, 89 (2007).

[110] M. A. Dobbs et al., "Les Houches guidebook to Monte Carlo generators for hadron collider physics". 2004. hep-ph/0403045.

[111] F. Krauss, JHEP 08, 015 (2002).

[112] F. Caravaglios and M. Moretti, Phys. Lett. B358, 332–338 (1995).

[113] H. Murayama, I. Watanabe, and K. Hagiwara, "HELAS: HELicity Amplitude Subroutines for Feynman diagram evaluations". KEK-91-11.

[114] A. Denner, H. Eck, O. Hahn, and J. Kublbeck, Nucl. Phys. B387, 467–484 (1992).

[115] G. Altarelli and G. Parisi, Nucl. Phys. B126, 298 (1977).

[116] S. Hoche et al., "Matching parton showers and matrix elements". 2006. hep-ph/0602031.

[117] M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. B632, 343–362 (2002).

[118] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. B269, 432–438 (1991).

[119] S. Catani, Y. L. Dokshitzer, and B. R. Webber, Phys. Lett. B285, 291–299 (1992).

[120] S. Hoeche, "Perturbative QCD in Event Generation". 2009.

[121] S. Schumann, "Simulation of signal and background processes for collider experiment". 2008.

[122] T. Plehn, "LHC Phenomenology for Physics Hunters". 2008. hep-ph/0810.2281.

[123] J. Alwall et al., Eur. Phys. J. C53, 473–500 (2008).

[124] T. Gleisberg et al., Czech. J. Phys. 55, B529–B536 (2005).

[125] F. Krauss, A. Schalicke, S. Schumann, and G. Soff, Phys. Rev. D72, 054017 (2005).

[126] "Monte Carlo Sample Production." https://twiki.cern.ch/twiki/bin/view/CMS/GeneratorProduction.

[127] "Generator Interfaces." https://twiki.cern.ch/twiki/bin/view/CMS/GeneratorTaskList.

[128] "Event Generation Offline Guide." https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideEventGeneration.

[129] "The Physics Analysis Toolkit." https://twiki.cern.ch/twiki/bin/view/CMS/SWGuidePhysicsTools.

[130] W. de Boer and et. al., "Search for SUSY with one, two and three muons at 10 TeV". CMS AN2010/001. 2010.

[131] W. de Boer and et. al., "Side Band measurement in SUSY search with muons". CMS AN2010/075. 2010.

[132] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, JHEP 09, 127 (2008).

[133] A. V. Gladyshev and D. I. Kazakov, Phys. Atom. Nucl. 70, 1553–1567 (2007).

[134] T. A. Collaboration, "Data Driven Determination of *tbart* Background to Supersymmetry Searches in ATLAS". ATL-PHYS-PUB-2009-083. 2009.

[135] "SUSY RA2 Project Tables: Inclusive N-Jet and MET Signature." https://twiki.cern.ch/twiki/bin/view/CMS/SusyRA2InclusiveNJetProjectTable.

[136] U. Blumenschein, "Suche nach assoziierter Chargino-Neutralino Produktion in Proton-Antiproton Kollisionen bei 1.96TeV". 2005.

[137] L. Randall and D. Tucker-Smith, Phys. Rev. Lett. 101, 221803 (2008).

[138] H. Flaecher et al., "SUSY search with Dijet Events". CMS AN2008/071. 2008.

[139] T. Whyntie et al., "Extending the early SUSY search with all-hadronic dijet events to n-jet topologies". CMS AN2008/114. 2008.

[140] "SUSY RA1 Project Tables: Exclusive N-Jet and MET Signature." https://twiki.cern.ch/twiki/bin/view/CMS/SusyRA1ExclusiveNJetProjectTable.

[141] S. Baffioni and et al., "Electron reconstruction in CMS". CMS NOTE2006/040. 2006.

[142] Y. Sirois and et al., "Reconstruction of electron tracks using gaussian sim filter in CMS". CMS AN2005/011. 2005.

[143] J. Branson and et al., "A cut based method for electron identification in CMS". CMS AN2008/082. 2008.

[144] "Search Procedures in CMS." https://twiki.cern.ch/twiki/bin/viewauth/CMS/SearchProcedures.

[145] S. S. Wilks, Ann. Math. Statist. 9 60-62 (1938).

[146] R. D. e. a. Cousins, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 595 (2) (2008).

Acknowledgements

There is no perfect work which can be done without any help. This thesis is the consequence of a hard period of study evolved by the contribution of many people and now I would like to express my gratitude to all the people supporting me from all the aspects for the period of my thesis.

First, I would like to thank my supervisor Prof. Dr. Wim de Boer for giving me the opportunity to come to Karlsruhe and perform my PhD studies as a member of his group. Thank you for always being available with solutions and motivations on short notice whenever I got stuck.

Thank you to Prof. Dr. Günther Quast for agreeing to take part in the Ph.D. committee, and to the *Koreferent* for evaluating the thesis.

I am most grateful to Dr. Valery Zhukov and Dr. Stefan Gieseke for all the advice, help and many interesting discussions.

I would like to thank my friend and colleague, Dr. Martin Niegel. I could not have imagined having a better mentor for my Ph.D., and without him commonsense, knowledge, perceptiveness, extremely good guidance and inspiring suggestions during my hard times and the preparation of this thesis. I would like to express my deep and sincere gratitude to him. I want to say a lot of things to express my feelings but in short, he will always be a special person for me.

I would like to thank my colleagues who created a friendly and and productive atmosphere Karlsruhe Institute of Technology (KIT), Faculty of Physics, IEKP room 8-15: Daniel Tröndle and Eva Ziebarth. I am also thankful to my colleagues Tim Hanisch and Markus Bonsch who provided valuable comments and developments for different parts of this thesis.

Finally, i would like to express my gratitude to my wife Özlem Cakir, to my parents, to my brother for their constant moral support. Özlem, I want especially to thank you for your love and understanding, and for encouraging and supporting me whenever I needed it.

Special thanks are owed to Adnan Hatipoglu, Fatih Ok and Orhan Arslan for contributing to an inspiring, friendly, and respectful environment in Karlsruhe.

I am grateful to IEKP, KIT for giving me a financial support during my thesis.