Universiteit Antwerpen
Doctoral school in Physics

ADD Extra Dimension searches with Monojet events and Resonance searches with Dijet events with the CMS detector

Ph.D. Thesis

Ph.D. student:
Dott. Marco Cardaci

Advisors:
Prof. P. Van Mechelen
Prof. A. De Roeck

Contents

Chapter 1

Preface

The study of the nature of Elementary Particles finds its deeper roots in the philosophical doctrine of Atomism, studied by ancient Greek philosophers such as Leucippus, Democritus and Epicurus in the fifth century BC. These atomists theorized that the natural world consists of two fundamental and opposite, indivisible bodies - atoms and void. Atoms were uncreatible, unchangeable and eternal. The different possible packings and scatterings of the atoms within the void were responsible of the shifting outlines and bulk of the objects that organisms feel, see, eat, hear, smell, and taste. Although the ideas of the atomists were not far from the actual understanding of the reality, since then, many advances have been made in the quest of a conclusive and comprehensive model of the fundamental constituents of matter. During the 20th century, the work on fundamental interactions and Quantum Mechanics, and the discovery of a large "zoo" of particles at accelerator laboratories all over the world, have culminated in the formulation of the so called Standard Model (SM) of Particle Physics.

The SM is a spectacular scientific achievement, summa of knowledge about the interactions between the particles which make up all the visible matter in the Universe, describing the properties of quarks and leptons in terms of basic principles arising from the nature of the electromagnetic, weak and strong interactions, tested at an high level of precision by many accelerator- and non accelerator-based experiments, which have made it one of the major cornerstones of physics. Despite the beautifulness of its principles and its experimental reliability, the SM is an effective theory valid up to a cut-off energy, affected by a problem known as hierarchy problem, a puzzling difference between the electroweak ($\sim 10^2$ GeV) and the Planck ($\sim 10^{19}$ GeV) scales, which suggests extensions of the model capable to stabilize the too sensitive behavior of the theory at energies of the order of the cut-off induced by the quadratic divergences in the scalar sector. Furthermore, it leaves open questions like: Why are there 6 quark flavors? Why are quarks arranged in generations? Why are there so many different forces?

How do we unify gravity with the other forces? Why is gravity so weak? Many are the proposed models capable to solve the hierarchy problem and to answer the several questions arising from the SM: Technicolor, Grand Unified Theory, Superstrings, Compositeness, Extra Dimensions, Extra Color, and Contact Interactions, just to mention a few. For instance, Grand Unified Theories postulate that the electroweak and strong forces come from a single interaction at much higher energies. These theories generally predict an extra heavy Z boson, and occasionally predict an extra heavy W boson.

The European Organization for Nuclear Research (CERN) has been playing a leading role in the search for new physics. After the discovery of neutral currents (1973) and the discovery of the W and Z bosons (1983), it turned to the search for the Higgs boson, Supersymmetry and theories of new physics such as Extra Dimensions. In 1989, the Large Electron Positron Collider (LEP) was brought into operation. During 11 years, scientists have looked at the SM, validating its predictions to a large extent, but were unable to give a conclusive answer on the nature of the spontaneous symmetry breaking mechanism. CERN has now concluded the construction of an accelerator and two general purpose detectors of the last generation: the Large Hadron Collider (LHC) and the ATLAS (A Toroidal LHC ApparatuS) and the CMS (Compact Muon Solenoid) detectors. They will, in conjunction, scrutinize an unprecedentedly explored collision energy range. The designed LHC energy of 7 TeV per beam and luminosity of 10^{34} cm^{-2} s^{-1} will be reached only after several stages of machine upgrade. Currently, it is understood that the machine will operate at progressively increasing energies, such as 3.5, 5 and finally 7 TeV per beam.

In the present dissertation we will discuss with emphasis techniques envisaged by CMS to search for resonances decaying into partons at nominal LHC energy and luminosity. The motivation on the study of the so called dijet resonances is two-fold: not only theoretical, which from the number of models mentioned appears quite clear, but also experimental. LHC is a parton-parton collider in a previously unexplored energy region. If new parton-parton resonances exist then the LHC will produce them copiously.

A solution to the hierarchy problem is given by the ADD (Arkani-Hamed, Dimopoulos, Dvali) model of Extra Dimensions. The ADD model of Extra Dimensions solves the hierarchy problem introducing n extra spatial dimensions, which in the simplest scenario are compactified over a torus and all have the same radius R. In this model the Planck scale become an effective scale related to the more fundamental scale M_D via the relation: $M_D^{n+2} R^n \sim M_{Pl}^2$. Consequences of the ADD model are that the extra dimensions can be 'macroscopic' for TeV-ish M_D scales and be tested at colliders such as Tevatron and LHC. Gravitons interact 'weakly' with the ordinary SM particles and escape detection. Therefore a clear ADD signature of direct graviton production, which will be

probed with the CMS detector, is the jet + missing transverse energy signature, also called Monojet signature. Here we will give a description of the techniques which will be used by CMS for the search in the Monojet channel at 7 and 5 TeV energy per LHC beam.

This thesis is divided into five parts. In Chapter 1, the SM and the theoretical framework of Extra Dimensions will be presented. Emphasis will be given to the description of the hierarchy problem and of the ADD model solution to it. A brief review of the detector characteristics, together with the CMS software components for simulation and analysis, will be given in Chapter 2. In Chapter 3, there will be a description of 2 different techniques for CMS dijet searches: the inclusive jet rate versus jet p_T in the search of Contact Interactions and the dijet rate versus dijet mass in the search of resonances. In Chapter 4 and 5, a description of the CMS plans for search in the Monojet channel at 14 and 10 TeV respectively, will be given. In Chapter 5 there will be an overview of the results obtained in the study of the Monojet channel, which must be seen as an improvement of the results presented in Chapter 4.

Chapter 2

Theoretical Introduction

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of Particle Physics is a Quantum Field Theory summa of the knowledge about the interactions between the particles which make up all the visible matter in the universe. All of these particles can be distinguished in fermions and bosons. Fermions are particles with half integer spin, which obey Fermi-Dirac statistics. In contrast to fermions, bosons are particles with integer spin, obeying the Bose-Einstein statistics.

The SM of Particle Physics is a gauge theory of the Strong and ElectroWeak (EW) interactions, with the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$. The fermions are gauge charged particles and they interact via the force mediating particles of the SM, the gauge bosons.

The $SU(3)_c$ group is the gauge group of the Strong interaction and its gauge bosons are eight gluons (g). The $SU(2)_L \times U(1)_Y$ group is the gauge group of the EW interaction and its gauge bosons are the photon, the W's and the Z bosons.

According to what charges they carry, the fermions can be classified in quarks and leptons. The fermions carrying color charge of the Strong interaction are the six quarks: down (d), up (u), strange (s), charm (c), bottom (b), top (t). The fermions not carrying the color charge are the six leptons: electron (e), muon (μ), tauon (τ), and corresponding neutrinos (ν).

In the SM, the lagrangian and the action play a central role. The most general aspect of an action in $3+1$ dimensions is:

$$S\left(\Phi, \tau_1, \tau_2\right) = \int_{\tau_1}^{\tau_2} \mathcal{L}\left(\Phi, \partial_\mu\right) d^4x \tag{2.1}$$

where $\mathcal{L}$ is the lagrangian, d^4x is the metric in the Minkowski space, Φ is a set of local fields and τ_1 and τ_2 are the boundaries of integration.

A lagrangian can be invariant under a symmetry group transformation. According to Noether's theorem, it is always possible to associate to an invariance under a symmetry group transformation, a conservation law. For instance, to the invariance of the lagrangian under rotation is associate the conservation of the angular momentum. More in general, to the invariance under a Poincaré group transformation corresponds conservation of energy, momentum and angular momentum.

Among the symmetries which can interest a physics system of particular importance are:

- Charge conjugation, C: exchange of particle in antiparticle.
- Parity conjugation, P: change of sign of spatial coordinates.
- Time reversal, T: change of sign of time coordinate.

Strong and electromagnetic interactions conserve separately C, P and T, while the weak interaction violates separately C and P, and the combined CP.

Gauge symmetries play a central role in the SM. An example of gauge symmetry is the electromagnetic gauge symmetry of the group $U(1)$. The QED lagrangian is defined as following:

$$\mathcal{L}_{\mathcal{QED}} = \frac{1}{4} F_{\mu\nu}(x) F^{\mu\nu}(x) + \sum_{l=e,\mu,\tau} \overline{\psi}^{l}(x) \left(i D_\mu \gamma^\mu - m_l \right) \psi^{l}(x) \tag{2.2}$$

where $F_{\mu\nu}$ is the electromagnetic tensor, ψ_l is the Dirac leptonic field and D_μ is the covariant derivative, defined as $\partial_\mu + ieA_\mu$. This lagrangian is invariant under a local phase transformation of the leptonic field and a corresponding photonic transformation:

$$\text{LOCAL} \quad U(1) \begin{cases} \psi_l(x) \to \psi_l'(x) = e^{-i\alpha(x)} \psi_l(x) \\ A_\mu(x) \to A_\mu' = A_\mu(x) + \frac{1}{e} \partial_\mu A_\mu(x) \end{cases} \tag{2.3}$$

Spin-one fields can appear in an interacting theory if they are coupled via the gauge principle.

In this section, we write down specifically what the field content of the SM is. The interactions will then follow as the most general set compatible with that field content.

Let us now build the SM model lagrangian. Let us start from the free leptonic massless lagrangian:

$$\mathcal{L}^{\mathcal{L}} = i \left(\overline{\psi}_l \not{\partial} \psi_l + \overline{\psi}_{\nu_l} \not{\partial} \psi_{\nu_l} \right) \tag{2.4}$$

where the Feynman notation $\not{\partial} = \gamma^\mu \partial_\mu$ is used. We can rewrite the lagrangian in LEFT-handed and RIGHT-handed components of the field, by means of the projector operators:

$$\begin{cases} \psi^L_{l,\nu_l} = \frac{1}{2}(1-\gamma_5)\,\psi_{l,\nu_l} \\ \psi^R_{l,\nu_l} = \frac{1}{2}(1+\gamma_5)\,\psi_{l,\nu_l} \end{cases} \tag{2.5}$$

The lagrangian can be rewritten in this way:

$$\mathcal{L}^{\mathcal{L}} = i\left(\overline{\psi}^L_l \not{\partial} \psi^L_l + \overline{\psi}^R_l \not{\partial} \psi^R_l + \overline{\psi}^L_{\nu_l} \not{\partial} \psi^L_{\nu_l} + \overline{\psi}^R_{\nu_l} \not{\partial} \psi^R_{\nu_l}\right) \tag{2.6}$$

In the following we will neglect the term on the RIGHT-handed neutrinos, because they are sterile, which means that they do not couple to the other particles because they do not carry any charge. We can define a singlet and a doublet in the following way:

$$\begin{cases} \Psi^L_l \equiv \begin{pmatrix} \psi^L_{\nu_l} \\ \psi^L_l \end{pmatrix} \quad ; \quad \overline{\Psi}^L_l = \begin{pmatrix} \overline{\psi}^L_{\nu_l} \overline{\psi}^L_l \end{pmatrix} \\ \Psi^R_l \equiv \psi^R_l \quad ; \quad \overline{\Psi}^R_l = \overline{\psi}^R_l \end{cases} \tag{2.7}$$

(we will use Ψ^R_l and ψ^R_l without distinction in the text) and rewrite the lagrangian:

$$\mathcal{L}^{\mathcal{L}} = i\left(\overline{\Psi}^L_l \not{\partial} \Psi^L_l + \overline{\Psi}^R_l \not{\partial} \Psi^R_l\right) \tag{2.8}$$

We can add to the free non interacting leptonic lagrangian, an interaction term:

$$\mathcal{L}^{\mathcal{L}} + \mathcal{L}^{\mathcal{LB}} = \mathcal{L}^{\mathcal{L}} - g J^\mu_i W_{i\mu} - g' J^\mu_Y B_\mu = i\left(\overline{\Psi}^L_l \not{D} \Psi^L_l + \overline{\Psi}^R_l \not{D} \Psi^R_l\right) \tag{2.9}$$

where D_μ is the covariant derivative:

$$D_\mu = \partial_\mu + ig\frac{\tau_i}{2} W_{i\mu} + ig' Y B_\mu \tag{2.10}$$

where $W_{i\mu}$ (with i = 1, 2, 3) and B_μ are the bosonic fields. We can write a similar lagrangian also for the quark interactions in the EW sector:

$$\mathcal{L}^{\mathcal{Q}} + \mathcal{L}^{\mathcal{QB}} = i\left(\overline{\Psi}^L_q \not{D} \Psi^L_q + \overline{\Psi}^R_q \not{D} \Psi^R_q\right) \tag{2.11}$$

We can finally introduce the self-interacting term of the lagrangian for the gauge bosons:

$$\mathcal{L}^{\mathcal{B}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} G_{i\mu\nu} G^{\mu\nu}_i \tag{2.12}$$

where the various tensors are defined in this way:

$$\begin{cases} B^{\mu\nu} = \partial^\nu B^\mu - \partial^\mu B^\nu \\ G_i^{\mu\nu} = F_i^{\mu\nu} + g\epsilon_{ijk} W_j^\mu W_k^\nu \\ F_i^{\mu\nu} = \partial^\nu W_i^\mu - \partial^\mu W_i^\nu \end{cases} \tag{2.13}$$

It is worth emphasizing at this point why certain terms do not appear among the lagrangian terms we wrote up to now, particularly mass terms for the fermionic fields. The reason is that only terms which are singlets under $SU(3)_c \times SU(2)_L \times U(1)_Y$ can appear in the Lagrangian, otherwise they would not respect gauge invariance, changing under a gauge transformation.

The vanishing of the masses is a consequence of the gauge invariance of the theory and can be evaded only if this symmetry is spontaneously broken. We can see that this can be obtained by adding a single complex scalar doublet, called the Higgs field:

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.14}$$

We can now write down the lagrangian which rules the dynamic of the field:

$$\mathcal{L}^{\mathcal{H}} = -\left(D^\mu \Phi\right)^\dagger \left(D_\mu \Phi\right) - V\left(\Phi\right) = \left(D^\mu \Phi\right)^\dagger \left(D_\mu \Phi\right) - \mu^2 \Phi^\dagger \Phi - \lambda \left(\Phi^\dagger \Phi\right)^2 \tag{2.15}$$

where the covariant derivative acts on the Higgs field as follow:

$$D^\mu \Phi = \left(\partial^\mu + ig \frac{\tau_i}{W_i^\mu} + ig' B^\mu\right) \Phi \tag{2.16}$$

where the τ_i are the Pauli matrices. We can now write the Yukawa terms which give rise to the mass of the fermions. Let us do it for the leptonic case.

$$\mathcal{L}^{\mathcal{LH}} = -g_l \left(\overline{\Psi}_l^L \Phi \psi_l^R + \overline{\psi}_l^R \Phi^\dagger \Psi_l^L\right) - g_{\nu_l} \left(\overline{\Psi}_l^L \Phi \psi_{\nu_l}^R + \overline{\psi}_{\nu_l}^R \tilde{\Phi}^\dagger \Psi_l^L\right) + h.c. \tag{2.17}$$

where $\tilde{\Phi} = i\sigma_2 \Phi^*$ is the charge conjugate of the Higgs field. A similar lagrangian $\mathcal{L}^{\mathcal{QH}}$ can be written for the quarks.

In order to read off the particle masses we must identify the part of the SM lagrangian which is quadratic in the fluctuations. The expansion of $\mathcal{L}^{\mathcal{L}} + \mathcal{L}^{\mathcal{LB}} + \mathcal{L}^{\mathcal{B}}$ is trivial and just contributes with the spin-half and spin-one kinetic terms. Everything else comes from the expansion of $\mathcal{L}^{\mathcal{H}}$.

We can rewrite the Higgs field in the unitary gauge:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta_1(x) + i\eta_2(x) \\ v + H(x) + i\eta_3(x) \end{pmatrix} = e^{i\frac{\tau_i}{2}\theta_i} \begin{pmatrix} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{pmatrix} \tag{2.18}$$

Using the following result:

$$D_\mu \Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \partial_\mu H \end{pmatrix} - \frac{i}{2\sqrt{2}} \begin{pmatrix} gW_{3\mu} + g'B_\mu & gW_{1\mu} - igW_{2\mu} \\ gW_{1\mu} + igW_{2\mu} & -gW_{3\mu} + ig'B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{2.19}$$

the expansion of the scalar-field kinetic term becomes:

$$\begin{aligned} -\left(D_\mu \Phi\right)^\dagger \left(D^\mu \Phi\right) = \tfrac{1}{2}\partial_\mu H \partial^\mu H - \tfrac{1}{8}(v+H)^2 g^2 \left(W_{1\mu} - iW_{2\mu}\right)\left(W_1^\mu - iW_2^\mu\right) \\ -\tfrac{1}{8}(v+H)^2 g^2 \left(-gW_3^\mu + g'B^\mu\right)\left(-gW_{3\mu} + g'B_\mu\right) \end{aligned} \tag{2.20}$$

The expansion of the Higgs lagrangian will be:

$$\begin{aligned} \mathcal{L}^{\mathcal{H}} = -\tfrac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \tfrac{\lambda}{4} H^4 \\ -\tfrac{1}{8} g^2 \left(v + H\right)^2 \left|W_{1\mu} - iW_{2\mu}\right|^2 \\ -\tfrac{1}{8}\left(v + H\right)^2 \left(gW_{3\mu} + g'B_\mu\right)^2 \\ +... \end{aligned} \tag{2.21}$$

which explicitly express the mass content of the SM, in particular the Higgs mass.

The SM requires a non-vanishing Vacuum Expectation Value (VEV) for Φ at the minimum of the potential. This will occur if $\lambda > 0$ and $\mu^2 < 0$, resulting in $\langle\Phi\rangle = \sqrt{-\mu^2}/2\lambda$. Since we know experimentally that $\langle\Phi\rangle$ is approximately 174 GeV, from measurements of the properties of the weak interactions, it must be that μ^2 is very roughly of order $-(100\text{ GeV})^2$. The problem is that μ^2 receives enormous quantum corrections from the virtual effects of every particle that couples, directly or indirectly, to the Higgs field. We will see it more in detail in section 2.2.3.

We can finally add up all terms collected and write the SM lagrangian in the EW sector:

$$\sum_{i=1}^{3} \mathcal{L}^{\mathcal{L}}{}_i + \mathcal{L}^{\mathcal{LB}}{}_i + \mathcal{L}^{\mathcal{Q}}{}_i + \mathcal{L}^{\mathcal{QB}}{}_i + \mathcal{L}^{\mathcal{B}}{}_i + \mathcal{L}^{\mathcal{H}}{}_i + \mathcal{L}^{\mathcal{LH}}{}_i + \mathcal{L}^{\mathcal{QH}}{}_i \tag{2.22}$$

where the index i run over the 3 generations of quarks and leptons.

2.2 Extra dimensions

In the physics community is widely agreed that some form of new physics must exist beyond the SM. One possibility is that of extra spatial dimensions, although only a few years ago not many physicists would have thought this even a remote possibility. The study of the physics of TeV-scale extra dimensions has its origins in the ground breaking work of Arkani-Hamed, Dimopoulos and Dvali(ADD) [1].

Since then, extra dimensions has evolved from a single idea to a new paradigm with many authors employing extra dimensions as a tool to address the large number of issues that remain unanswerable in the SM context. This in turn has lead to other phenomenological implications which should be testable at colliders and elsewhere. A list of some of these ideas includes, e.g.:

- Addressing the hierarchy problem.
- Producing electroweak symmetry breaking without a Higgs boson.
- The generation of the ordinary fermion and neutrino mass hierarchy, the CKM matrix and new sources of CP violation.
- TeV scale grand unification or unification without SUSY while suppressing proton decay.
- New Dark Matter candidates and a new cosmological perspective.
- Black hole production at future colliders as a window on quantum gravity.

This list hardly does justice to the wide range of issues that have been considered in the extra dimension context.

Extra dimensions can be considered as a set of tools which can be realized through various models, as much as the Quantum Field Theory can be realized through the SM. In the following we will concentrate on those models which can appear at the TeV scale, hence those models which can be probed at the LHC (See i.e. [2]).

For every model it is possible to define a length scale which characterizes the size of the extra dimensions. There are two equivalent ways to depict an extra-dimensional model depending on whether the description is at a scale small or large compared to the length scale. At a scale large or comparable to the length scale a 4-dimensional description is appropriate. At a large scale it is not possible to resolve the extra-dimensional world and only the 4-dimensional effective world will be observable. At this stage the concept introduced by Kaluza Klein enters to play a role, see below.

At a scale small compared to the length scale a higher dimensional language is more appropriate. This language puts an emphasis on the higher dimensional space structure and allows to treat the effects of all Kaluza Klein (KK) modes.

In the following only the 4-dimensional approach will be used in order to draw consequences of colliders experiments and relate them to known results of the SM.

Let us consider a field Φ depending on the ordinary 4 dimension and n extra dimensions:

$$\Phi(x^\mu, y^i)(\mu = 1, 2, 3, 4; i = 1, ..., n) \tag{2.23}$$

where y^i parametrize a compact space. The basic idea of Kaluza Klein is to move to the Fourier space of the extra-dimensional space. This is done by expanding the field in a basis depending only on the extra dimensions:

$$\Phi(x^\mu, y^i) = \frac{1}{\sqrt{V}} \sum_k \phi^{(k)}(x^\mu) f_k(y^i) \tag{2.24}$$

Here the fields $\phi^{(k)}(x^\mu)$, depending only on the ordinary space, are the so called KK modes and Φ can be seen as a tower of KK modes.
Let us now consider the lagrangian of N-dimensional theory (N $= 4 + n$), limiting ourself to the free part of the theory. The lagrangian will be made of a kinetic term and a mass term, such as in the following expression:

$$\mathcal{L} = \partial_N \Phi \partial^N \Phi - M^2 \Phi^2 \tag{2.25}$$

Under certain boundary condition assumptions we can integrate by parts the action of this lagrangian, as in the following:

$$S = \int d^4x d^k y \frac{1}{2}(\partial_N \Phi \partial^N \Phi - M^2 \Phi^2) = -\int d^4x d^k y \frac{1}{2} \left(\Phi(\Box + M^2)\Phi\right) \tag{2.26}$$

and replacing the expanded expression of the field we obtain:

$$= -\frac{1}{V} \int d^4x d^k y \sum_{k,k'} \frac{1}{2} \phi^{(k)} f_k \left[f_{k'} \partial_\mu \partial^\mu \phi^{(k')} + \phi^{(k')} (\partial_i \partial^i + M^2) f_{k'} \right] \tag{2.27}$$

$$= -\sum_{k,k'} \left(\frac{1}{V} \int d^k y f_k f_{k'} \right) \int d^4x \frac{1}{2} \phi^{(k)} (\partial_\mu \partial^\mu + m_{k'}^2) \phi^{k'} \tag{2.28}$$

$$= -\sum_k \int d^4x \frac{1}{2} \phi^{(k)} (\partial_\mu \partial^\mu + m_{k'}^2) \phi^k \tag{2.29}$$

where $m_{k'}$ are given by the linear partial differential equation:

$$\left(\partial_i \partial^i + M^2\right) f_k = m_k^2 f_k \tag{2.30}$$

which define a self-adjoint problem, solved by finding the eigenvalues of the Klein-Gordon operator. The solutions of the problem implies orthogonal eigenfunctions:

$$\langle f_k | f_{k'} \rangle = \frac{1}{V} \int d^k y f_k f_{k'} = \delta_{k,k'} \tag{2.31}$$

The physical upshot of this discussion is that a free 4+n-dimensional field is equivalent to an infinite number of 4-dimensional scalars with masses m_k.

Now we can come back to the issue concerning the boundary conditions. In two points of the previous discussion the boundary conditions play a role: the vanishing of the surface integral obtained by integrating by part the action and the possibility to solve the Klein-Gordon equation. Any choice of boundary conditions will have to satisfy these two needs. Let us have a look to some possible boundary conditions:

- Boundary conditions on a circle S^1 for 1 extra dimension.

  $$\Phi(y + 2\pi R) = \Phi(y) \tag{2.32}$$

  the new quantum number k is simply the 5th component of the momentum along the new dimension. If the higher dimensional field is massless, than the Kaluza-Klein masses will describe the spectra:

  $$m_k = \frac{k}{R} \tag{2.33}$$

  The fact that the masses can be expressed in units of $1/R$ means that an integer number of wave lengths can lie along the extra-dimensional direction (see 2.1).

- Boundary conditions on an "interval" for 1 extra dimension (see Fig. 2.2).

  Dirichlet or Neumann boundary conditions or a mix of them at $y = 0$ and $y = \pi$.

- Boundary conditions on torus for 2 extra dimensions (see Fig. 2.3).

  $$\Phi(x^4 + 2\phi, x^5) = \Phi(x^4, x^5) \tag{2.34}$$
  $$\Phi(x^4, x^5 + 2\phi) = \Phi(x^4, x^5) \tag{2.35}$$

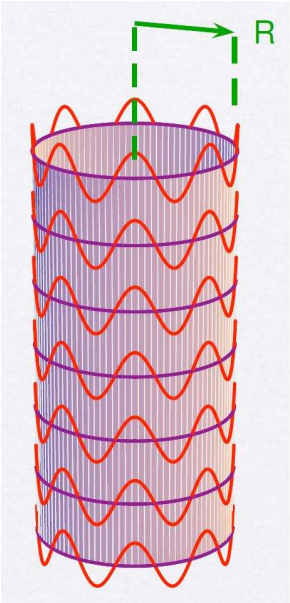


Figure 2.1: Graphical representation of boundary conditions on a circle S^1. Only an integer number of wave lengths can lie along the extra-dimensional direction.

it is the analogous of the compactification on the circle for 2 extra dimensions. Also in this case periodic boundary conditions are imposed (see fig. 2.3).

- Chiral square boundary conditions for 2 extra dimensions (see Fig. 2.4).

$$\Phi(y,0) = e^{ik\pi/2}\Phi(0,y) \tag{2.36}$$
$$\Phi(y,\pi R) = e^{ik\pi/2}\Phi(\pi R,y) \tag{2.37}$$
$$\partial_5\Phi_{(x^4,x^5)=(y,0)} = -e^{ik\pi/2}\partial_5\Phi_{(x^4,x^5)=(0,Y)} \tag{2.38}$$
$$\partial_5\Phi_{(x^4,x^5)=(y,\pi R)} = -e^{ik\pi/2}\partial_5\Phi_{(x^4,x^5)=(\pi R,Y)} \tag{2.39}$$

Let us go back to the Kaluza-Klein decomposition. We have seen that in order to decompose the higher dimensional field in Kaluza-Klein modes, the solutions to the eigenvalue problem is needed. These solutions are called solutions on the bulk and can lead to Kaluza-Klein modes with $m_k = 0$, i.e. $m_0 = 0$, which are called zero modes. These modes are of particular interest. The reason is that a particle of mass zero should have been already observed because it can be produced at low energies. Let us see two cases of solutions on the bulk:

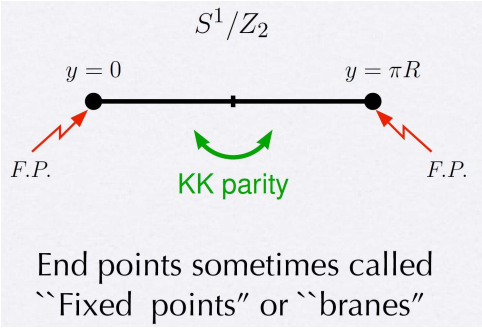


Figure 2.2: Graphical representation of interval compactification.

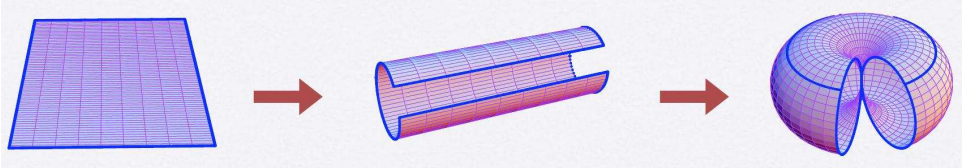


Figure 2.3: Sketch of compactification on a torus for scenarios with 2 extra dimensions.

- Gauge field in a flat space and 1 extra dimension.

  In this case the eigenvalue problem (see 2.30) assume the form:

$$f_k''(y) + 2f_k'(y) = m_k^2 f_k(y) \tag{2.40}$$

  which is solved by:

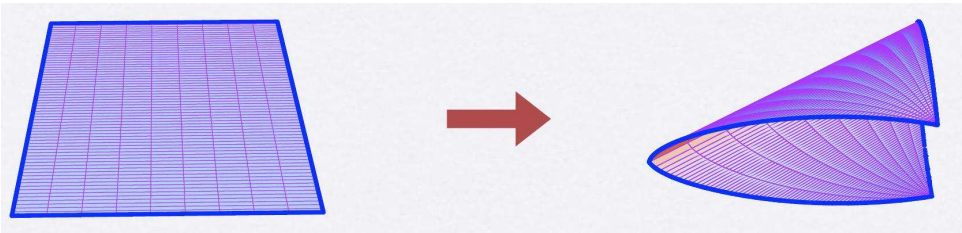


Figure 2.4: Chiral compactification.

$$m_0^2 = 0 \tag{2.41}$$
$$f_0(y) = 1 \tag{2.42}$$

- Fermion field in a flat space and 1 extra dimension.

  In this case the Kaluza-Klein decomposition has to be performed for the two chiralities of the fermion; therefore we end up to the two following differential equations:

$$f'_{k,L} - M f_{k,L} = m_k f_{k,R} \tag{2.43}$$
$$f'_{k,R} - M f_{k,R} = m_k f_{k,L} \tag{2.44}$$

  which are solved by:

$$m_0 = 0; f_0^{L,R}(y) = \sqrt{\frac{1 - e^{-2ML}}{2ML}} e^{\pm My} \tag{2.45}$$

Anyway, these kind of solutions may or may not be allowed depending on the boundary conditions. For the fermions for instance, the boundary conditions in a circle and in a torus allow both chiralities, while the interval and chiral square compactifications allow only one of the two chiralities, giving rise to a chiral theory. In this direction, extending the notion of zero modes to those modes which have small mass compared to the other KK modes and identifying them with the known particles, namely those of the SM, one opens the way to the building up of an extra-dimensional theory which is a realistic extension of the SM.

Let us now introduce an interaction to the free theory. As far as the interaction will be treated as a perturbation to the free lagrangian, everything can be expressed in terms of KK modes. In this case not only the mass spectrum but also the wavefunctions will become observable, since they determine the details of the interaction among KK modes. For instance, let us consider a Yukawa term of interaction between two fermions and a scalar higher dimensional fields:

$$\int d^4x d^k y \lambda_k \overline{\Psi} \Psi \Phi \rightarrow \sum_{k_1,k_2,k_3} \lambda_{k_1,k_2,k_3} \int d^4x \overline{\psi}^{(k_1)} \psi^{(k_2)} \phi^{(k_3)} \tag{2.46}$$

$$\tag{2.47}$$

where:

$$\lambda_{k_1,k_2,k_3} = \frac{\lambda_k}{V\sqrt{V}} \int d^k y f_{k_1} f_{k_2} f_{k_3} \tag{2.48}$$

and the approximation is valid only for λ_k small enough. This integral sometimes has very interesting selection rules. For instance, for the case of 5-D and compactification in the interval, this is not vanishing when $k_1 \pm k_2 \pm k_3 = 0$. Hence, at tree level, no KK mode can decay into a 0 mode. This opens the possibility of having in the theory new stable particles. Obviously the rule that no KK mode can decay into a 0 mode is only valid at tree level and can be violated at loop level considering a KK 2 mode which, by means of a loop, can decay in two 0 modes (see Fig. 2.5). Nevertheless, also including loop corrections the KK 1 mode cannot further decay, becoming a candidate particle for the explanation of Dark Matter.

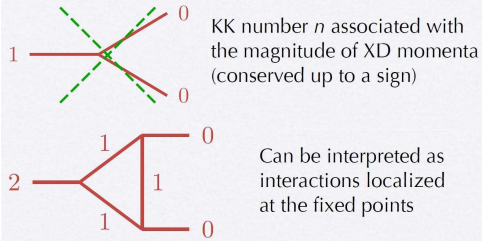


Figure 2.5: At tree level, no KK mode can decay into a 0 mode. This rule can be violated at loop level considering a KK 2 mode which, by means of a loop, can decay in two 0 modes.

Now we can generalize the previous discussion to warped spaces in the case of 5 dimensions. For this aim we define a metric:

$$ds^2 = a^2(y)\eta_{\mu\nu}dx^\mu dy^\nu - dy^2 \tag{2.49}$$

where $a(y)$ is the warping factor along the 5th dimension. For $a(y)$ constant, the metric reduces to a flat metric. In this scenario, the scalar fields, the equations of motion and their solutions are:

$$\Phi(x^\mu, y) = \frac{a(y)^{-1}}{\sqrt{L}} \sum_k \phi^k(x^\mu) f_k(y) \tag{2.50}$$

$$\begin{cases} \text{Eq. of motion :} \\ f_k'' + 2\frac{a'}{a} f_k' - \left[\frac{a''}{a} + 2\frac{a'^2}{a^2} + M^2\right] f_k = -m_k^2 a^{-2} f_k \\ \text{Solution for } a(y) = e^{-ky} \text{ and squared bulk mass } M^2 = \left[c_s^2 - c_s - \frac{15}{4}\right] k^2 : \\ f_k(y) = N_k e^{ky} \left[J_{|c_s+1/2|}(m_k e^{ky}/k) + b_k) b_k Y_{|c_s+1/2|}(m_k e^{ky}/k)\right] \end{cases} \tag{2.51}$$

and for the fermions:

$$\Psi_{L,R}(x^\mu, y) = \frac{a(y)^{-3/2}}{\sqrt{L}} \sum_k \psi^{L,R}(x^\mu) f_k^{L,R}(y) \tag{2.52}$$

$$\begin{cases} \text{Eq. of motion :} \\ f_{k,L}' - (c_f - 1/2)\frac{a'}{a} f_{k,L} = -m_k a^{-1} f_{k,R} \\ -f_{k,R}' - (c_f - 1/2)\frac{a'}{a} f_{k,R} = -m_k^* a^{-1} f_{k,L} \\ \text{Solution for } a(y) = e^{-ky} \text{ and bulk mass } M = c_f k : \\ f_k(y) = N_k e^{ky} \left[J_{|c_s+1/2|}(m_k e^{ky}/k) + b_k) b_k Y_{|c_s+1/2|}(m_k e^{ky}/k)\right] \end{cases} \tag{2.53}$$

while for the gauge fields, with gauge fixing $\frac{1}{2\xi}\left\{\eta^{\mu\nu}\partial_\mu A_\nu - \xi\partial_y\left[a(y)^2 A_5\right]\right\}^2$, we obtain:

$$A_\mu(x^\mu, y) = \frac{1}{\sqrt{L}} \sum_k A_\mu^{(k)}(x^\mu) f_k(y) \tag{2.54}$$

$$\begin{cases} \text{Eq. of motion :} \\ f_k''(y) + 2\frac{a'}{a} f_k'(y) = -m_k^2 a^{-2} f_y \\ \text{Solution for } a(y) = e^{-ky} : \\ f_k(y) = N_k e^{ky} \left[J_1(m_k e^{ky}/k) + b_k) b_k Y_1(m_k e^{ky}/k)\right] \end{cases} \tag{2.55}$$

Here all the wavefunctions are normalized according to:

$$\frac{1}{L}\int d^k y f_k f_{k'} = \delta_{k,k'} \tag{2.56}$$

In equations 2.51 and 2.53 the mass of the higher dimensional field M is not related to the mass m_k of the KK modes. Even in the presence of a non zero M is possible to have zero modes. Instead, it regulates the exponential warping factor. Also in the warped 5 dimensions scenario the boundary conditions determine the mass spectrum of the KK modes.

The zero modes have an associated profile along the 5th dimension. The profile controls the way how the fields couple to each other. The coupling between the various KK modes is proportional to the overlap integrals of the profiles.

Exponentially small numbers and the hierarchy among the particles of the SM are in this way explained by means of fundamental parameters which expresses the relative localization of the fields along the additional dimension.

Also the scalars can be localized. In most of the extra-dimensional theories a fine tuning is needed in order to get a light scalar. Nevertheless it is possible to obtain light scalars in other ways. Let us consider a gauge field in 5-D and impose Dirichlet boundary conditions for the first 4 components of the ordinary space and Neumann boundary conditions for the 5th component of the field. In a warped background the equation of motion for the 5th component of the 0 mode is:

$$\partial_y^2(a^2 f_0) \to f_0(y) = N_0 a^{-2}(y) \to \sqrt{\frac{2kL}{e^2kL-1}} e^{2ky} \tag{2.57}$$

where $\to$ is the limit to the so called Randall-Sundrum scenario, where the warp factor is an exponential: $a(y) = e^{-ky}$. In this case the gauge field happen to be localized at $y = L$, near the InfraRed (IR) brane (or higher brane, in opposition to the UltraViolet (UV) brane), where the warp factor has the smallest value (see Fig. 2.6). Interestingly enough, the localization mechanism does not need adjustable parameters because the localization happens dynamically.

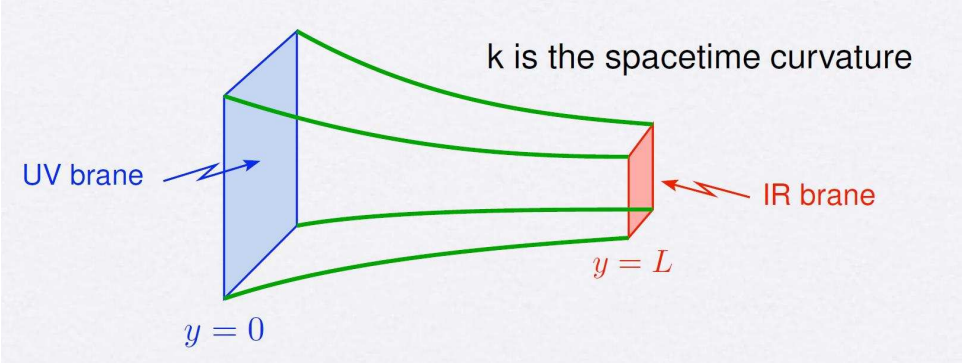


Figure 2.6: Schematic view of Randall Sundrum extra dimensions bulk. The UV brane and the IR brane are the two extremes of the bulk where the warped factor has the largest and the smallest value, respectively.

Another way to obtain localization for a scalar is in the case when we have color charged fermions localized toward the IR brane. They can couple with gluons and be able to form bound states which can give rise to scalars which are localized too.

2.2.1 Large extra dimensions

There is an important difference between the weak scale and the Planck scale, at the origin of the hierarchy problem. While the EW interactions have been probed at distances $\sim m_{EW}^{-1}$, the gravitational interaction has only been measured in the ~ 100 μm range. Our interpretation of M_{Pl} as a fundamental energy scale is then based on the assumption that gravity is unmodified over 33 orders of magnitude. Given the crucial way in which the fundamental role attributed to M_{Pl} affects our current thinking, it is worthwhile questioning this extrapolation and seeking new alternatives to the standard picture of physics beyond the SM. The philosophy of the ADD model [1] is that m_{EW} is the only fundamental short distance scale in nature, even setting the scale for the strength of the gravitational interaction. This can be realized by introducing a number n of extra compact spatial dimensions, which in the simplest scenario are compactified over a torus and all have the same radius R. The Planck scale M_D of this $4+n$-dimensional theory is taken to be ~ 1 TeV. Gauss law tells us what is the form of the gravitational potential of a mass m as a function of a radius r:

$$V(r) = \frac{1}{M_D^{n+2}} \frac{m}{r^{n+1}}, \quad r \ll R \tag{2.58}$$

This expression of the potential is valid only when the distances are quite small compared to the compactification radius. When observed from a distance which is much larger than R, the gravitational flux lines can not continue to penetrate in the extra dimensions and the potential is reduced to the effective potential:

$$V(r) = \frac{1}{M_D^{n+2} R^n} \frac{m}{r}, \quad r \gg R \tag{2.59}$$

which expresses the usual $1/r$ dependence. The now effective 4 dimensional M_{Pl} is related to M_D by the following relation:

$$M_{Pl}^2 \sim M_D^{n+2} R^n \tag{2.60}$$

We can therefore estimate the required R for a given number of extra dimensions:

$$R \sim 10^{\frac{30}{n}-17}\text{cm} \times \left(\frac{1 \text{ TeV}}{m_{EW}}\right)^{1+\frac{2}{n}} \tag{2.61}$$

$R \sim 10^{13}$ cm corresponding to $n = 1$, is not viable because it implies deviations of gravity over solar system distances. For all $n \geq 2$ the modification of gravity only becomes noticeable at distances smaller than those currently probed by experiments, i.e. to $n = 2$ corresponds R from hundreds of μm to ~ 1 mm.

While gravity has not been probed at distances smaller to $55\mu m$ [4], the SM gauge forces have certainly been accurately measured at weak scale distances. Therefore, the SM particles cannot freely propagate in the extra dimension, but, thanks to some mechanism, must be localized to a 4-dimensional sub-manifold, while 4+n-dimensional gravitons can freely travel through the 4+n-dimensional space.

Of course, the localization of the SM fields is a non-trivial task in any explicit realization of this framework. A number of ideas for such localizations have been proposed in the literature, i.e. in the context of trapping zero modes on topological defects [5]. There certainly are other possible ways for realizing the proposal of large extra dimensions, without affecting its most important consequences.

The theory above m_{EW} is unknown, being whatever gives a sensible quantum theory of gravity in the higher dimensional space. The assumption is that dangerous higher dimension operators (suppressed only by m_{EW}) which lead to proton decay, neutral meson mixing etc., are not induced. Any extension of the SM at the weak scale must also not give dangerously large corrections to precision EW observables. There could be unknown contributions from the physics above m_{EW}. However, at least the purely gravitational corrections do not introduce any new EW breaking beyond the W, Z masses, and therefore should decouple as loop factor $\times(m_{W,Z}/m_{EW})^2$, which is already quite small even for $m_{EW} \sim$ TeV.

Independently of any specific realization, there are a number of dramatic experimental consequences of the ADD framework. First, as already mentioned, gravity becomes comparable in strength to the gauge interactions at energies $m_{EW} \sim$ TeV. The LHC would then not only probe the mechanism of EW symmetry breaking, it would probe the true quantum theory of gravity.

Second, for the case of 2 extra dimensions, the gravitational force law should change from $1/r^2$ to $1/r^4$ on distances between 100 μm and 1 mm, and this deviation could be observed by experiments measuring gravity at sub-millimeter distance, such as at Tevatron and LHC accelerators which can reach energies of the order of the TeV.

2.2.2 Graviton-SM interaction in large extra dimensions

In the large extra dimensions theories we have seen that light SM particles have to be localized (up to some extent) to four dimensions, while the graviton (and, as we will see, other particles) can propagate in more dimensions [6].

In the technical jargon, the surface along which some of the particles are localized is called "brane", which stands for a membrane that could have more spatial dimensions than the usual 2-dimensional membrane. In our case the most important object will be a 3-brane, which has 3 spatial dimensions just like our observed world, which could be embedded into more dimensions.

In a low-energy effective field theory approach is implicit the assumption that the theory will be valid only up to some cutoff scale, above which the dynamics that generates the brane has to be taken into account.

Let us set up some notation. The 3-brane will be assumed to be described by a flat 4-dimensional space-time while the extra dimensions compactified over an n-dimensional torus.

The coordinates in the bulk are denoted by X^M, with M = 0, 1, . . . , 3 + n, the coordinates on the brane are denoted by x_μ, with $\mu = 0, 1, 2, 3$, while the coordinates along the extra dimensions only are denoted by x_m, with m = 4, . . . , 3 + n.

For now we will concentrate on the bosonic degrees of freedom. What are these degrees of freedom that we need to discuss in a low-energy effective theory? Since we need to discuss the physics of higher dimensional gravity, this should include the metric in the 4 + n dimensions $G_{MN}(X)$, and also the position of the brane in the extra dimensions $Y^M(x)$. Note, that the metric is a function of the bulk coordinates X^M, while the position in the brane is a function of the coordinates along the brane x_μ.

In addition, we would like to take into account the fields that are localized to live along the brane. These could be some scalar field $\Phi(x)$, gauge field $A_\mu(x)$ or fermionic field $\Psi(x)$. These fields are also functions of the coordinates along the brane x_μ.

The effective theory that we are trying to build up should describe small fluctuations of the the fields around the vacuum state. So we have to specify what we actually mean by the vacuum. We will assume that we have a flat brane embedded into flat space. The corresponding choice of vacuum is then given by:

$$G_{MN}(X) = \eta_{MN} \tag{2.62}$$

$$Y^M(x) = \delta^M_\mu x^\mu \tag{2.63}$$

with by convention $\eta_{MN} = \text{diag}\,(+, -, -, ..., -)$. The bulk action will just be a generalization to more dimensions of the usual Einstein-Hilbert action, plus perhaps the cosmological constant:

$$S = \int d^{4+n}X \sqrt{|G|}(M_D^{n+2} R_{4+n} + \Lambda) \tag{2.64}$$

where R_{4+n} is the $4 + n$ dimensional scalar curvature. In order to be able to write down an effective action for the brane localized fields one has to first find the induced metric on the brane. This is the metric that should be used to contract Lorentz indexes of the brane field. To get the induced metric, we need to find what the distance between two points on the branes is, x and $x + dx$:

$$ds^2 = G_{MN} dY(x)^M dY(x)^N = G_{MN} \frac{\partial Y^M}{\partial x^\mu} dx^\mu \frac{\partial Y^N}{\partial x^\nu} dx^\nu \tag{2.65}$$

therefore the metric will be:

$$g_{\mu\nu} = G_{MN}\left(Y(x)\right) \partial_\mu Y^M \partial_\nu Y^N \tag{2.66}$$

We can write the brane induced part of the action, which has to be invariant both under the general coordinates transformation of the bulk coordinates X and under the general coordinate transformations of x. It is clear that the invariance under the general coordinates transformation of the bulk coordinates just corresponds to the usual general covariance of a higher dimensional gravitational theory. The additional requirement that the action also be invariant under the coordinates transformations of the brane coordinates x is an expression of the fact that x is just one possible parametrization of the surface (brane), which itself can not have a physical significance, and any different choice of parametrization has to give the same physics. This will ensure the usual 4-dimensional Lorentz invariance of the brane induced action. Thus there are two separate coordinate transformations that the action has to be invariant under, that corresponding to X and to x. Practically this means that one has to contract bulk indexes with bulk indexes, and brane indexes with brane indexes.
Thus the general form of the brane action would be:

$$S = \int d^{4+n}\sqrt{|g|}\left[-f^4 - R^{(4)} + \frac{g^{\mu\nu}}{2} D_\mu \Phi D_\nu \Phi - V(\Phi) - \frac{g^{\mu\nu} g^{\rho\sigma}}{4} F_{\mu\rho} F_{\nu\sigma} + ...\right] \tag{2.67}$$

Here the possible constant piece f corresponds to the energy density of the brane, called the brane-tension. This brane tension has to be small (in the units of the fundamental Planck scale) in order to be able to neglect its back-reaction on the gravitational background.

As discussed above, in addition to the usual bulk coordinates invariance, there is also a 4-dimensional parametrization invariance, which is corresponding to the fact that a different parametrization of the surface describing the brane would yield the same physics. Thus one needs an additional gauge fixing condition, which will eliminate the non-physical components from Y^M. There are four coordinates, so one needs four conditions, which can be picked as:

$$Y^\mu(x) = x^\mu \tag{2.68}$$

which is a complete gauge fixing. Thus out of the 4 + n components of Y^M only the components along the extra dimension $Y^m(x)$, where m can be 4, 5, .

. . , 3 + n, correspond to physical degrees of freedom. These n physical fields correspond to the position of the brane within the bulk.

Let us now concentrate on the modes of the bulk graviton, and for now set the fluctuations of the brane to zero, which translated in mathematical language means:

$$Y^M = \delta^M_\mu x^\mu; Y^m = 0 \tag{2.69}$$

In this case the metric will be of the form:

$$g_{\mu\nu}(x) = G_{MN}(x_\mu, x^m = 0) \tag{2.70}$$

the action is given by:

$$S = \int d^4x \mathcal{L}_{SM} \sqrt{g}(g_{\mu\nu}, \Phi, \Psi, A, ...) \tag{2.71}$$

and the definition of the energy-momentum tensor is:

$$\sqrt{g} T^{\mu\nu} = \frac{\delta S_{SM}}{\delta g_{\mu\nu}} \tag{2.72}$$

Expanding in the fluctuation around flat space:

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{h_{\mu\nu}(x_\mu)}{M_D^{\frac{n}{2}+1}} \tag{2.73}$$

at linear order the interaction between the SM matter and the graviton field is given by:

$$S_{int} = \int d^4x T^{\mu\nu} \frac{h_{\mu\nu}(x_\mu)}{M_D^{\frac{n}{2}+1}} \tag{2.74}$$

Thus generically the graviton couples linearly to the energy-momentum tensor of the matter. Note, that in the above expression what appears is the graviton field at the position of the brane, which is not a mass eigenstate field from the 4-dimensional point of view, but rather a superposition of all KK modes. We can therefore use the KK expansion of the graviton field:

$$h_{MN}(x, y) = \sum_{k_1=-\infty}^{\infty} ... \sum_{k_n=-\infty}^{\infty} \frac{h^{\vec{k}}_{MN}(x)}{\sqrt{(V_n)}} e^{i\frac{\vec{k}\cdot\vec{y}}{R}} \tag{2.75}$$

which plugged in the 2.75:

$$\sum_{\vec{k}} \int d^4x T^{\mu\nu} \frac{1}{M_D^{\frac{n}{2}+1}} \frac{h_{\mu\nu}^{\vec{k}}}{\sqrt{V_n}} = \sum_{\vec{k}} \int d^4x \frac{1}{M_{Pl}} T^{\mu\nu} h_{\mu\nu}^{\vec{k}} \tag{2.76}$$

allows to infer the coupling of the SM fields to the individual KK modes. Thus we can see that an individual KK mode couples with strength $1/M_{Pl}$ to the SM fields. However, since there are many of them, the total coupling in terms of the field at the brane sums up to a coupling proportional to $1/M_D$.

Next, let us discuss what the different modes contained in the bulk graviton field are. Clearly, the graviton is a D by D symmetric tensor, where D = n+4 is the total number of dimensions. Therefore this tensor has in principle D(D + 1)/2 components. However, we know that general relativity has a large gauge symmetry, D-dimensional general coordinate invariance. Therefore we can impose D separate conditions to fix the gauge, for example using the harmonic gauge:

$$\partial_M h_N^M = \frac{1}{2} \partial_N h_M^M \tag{2.77}$$

However, just like in the ordinary Lorentz gauge for gauge theories, this is not yet a complete gauge fixing. Gauge transformations which satisfy the equation $\Box \epsilon_M = 0$ are still allowed, and this means that another D conditions can be imposed. This means that generically a graviton has D(D + 1)/2 - 2D = D(D - 3)/2 independent degrees of freedom. For D = 4 this gives the usual 2 helicity states for a massless spin two particle, however in D = 5 we get 5 components, in D = 6 we get 9 components, etc. This means that from the 4-dimensional point of view a higher dimensional graviton will contain particles other than just the ordinary 4-dimensional graviton. This is quite clear, since the higher dimensional graviton has more components, and thus will have to contain more fields. The question is what these fields are and how many degrees of freedom they contain. The total number of degrees of freedom of the bulk graviton are: 5 (graviton) + 3(n - 1) (vectors) + (n^2 - n - 2)/2 (scalars) +1 (radion) = (4 + n)(1 + n)/2 = D(D - 3)/2.

Let us now derive the Feynman rules necessary to compute the rate for graviton-emission processes. At low energy and small curvature the equations of motion of the effective theory reduce to the Einstein equation in D = 4 + n dimensions:

$$\mathcal{G}_{AB} \equiv \mathcal{R}_{AB} - \frac{1}{2} g_{AB} \mathcal{R} = -\frac{T_{AB}}{M_D^{2+n}} \tag{2.78}$$

Expanding the metric g_{AB} around its Minkowski value η_{AB}, we obtain:

$$M_D^{1+n/2}\mathcal{G}_{AB} = \Box h_{AB} - \partial_A\partial^C h_{CB} - \partial_B\partial^C h_{CA} - \partial_A\partial_B h_C^C \\ -\eta_{AB}\Box h_C^C + \eta_{AB}\partial^C\partial^D h_{CD} = -M_D^{-1-n/2}T_{AB} \tag{2.79}$$

where only the first power of h is kept, indexes are raised or lowered using the flat-space metric and summation over repeated indexes is understood. To this equation corresponds the lagrangian:

$$\mathcal{L} = \frac{1}{2}h^{AB}\Box h_{AB} + \frac{1}{2}h_A^A\Box h_B^B - h^{AB}\partial_A\partial_B h_C^C + h^{AB}\partial_A\partial_C h_B^C - \frac{1}{M_D^{1+n/2}}h^{AB}T_{AB} \tag{2.80}$$

We can choose the field parametrization:

$$G_{\mu\nu}^{(n)} \equiv h_{\mu\nu}^{(n)} + \frac{\kappa}{3}\left(\eta_{\mu\nu} + \frac{\partial_\mu\partial_\nu}{\hat{n}^2}\right)H^{(\overrightarrow{n})} - \partial_\mu\partial_\nu P^{(n)} + \partial_\mu Q_\nu^{(n)} + \partial_\nu Q_\mu^{(n)} \tag{2.81}$$

$$V_{\mu j}^{(n)} \equiv \frac{1}{\sqrt{2}}\left[ih_{\mu j}^{(n)} - \partial_\mu P_j^{(n)} - \overrightarrow{n}_j Q_\mu^{(n)}\right] \tag{2.82}$$

$$S_{jk}^{(n)} \equiv h_{jk}^{(n)} - \frac{\kappa}{n-1}\left(n_{jk} + \frac{\hat{n}_j\hat{n}_k}{\hat{n}^2}\right)H^{\overrightarrow{n}} + \hat{n}_j P_k^{(n)} + \hat{n}_k P_j^{(n)} - \hat{n}_j\hat{n}_k P^{(n)} \tag{2.83}$$

$$H^{(\overrightarrow{n})} \equiv \frac{1}{\kappa}\left[h_j^{(n)j} + \hat{n}^2 P^{(n)}\right] \tag{2.84}$$

$$Q_\mu^{(n)} \equiv -i\frac{\hat{n}_j}{\hat{n}^2}h_\mu^{(n)j} \tag{2.85}$$

$$P_j^{(n)} \equiv \frac{\hat{n}_k}{\hat{n}^2}h_j^{(n)k} + \hat{n}_j P^{(n)} \tag{2.86}$$

$$P_j^{(n)} \equiv \frac{\hat{n}^j\hat{n}^k}{\hat{n}^4}h_{jk}^{(n)} \tag{2.87}$$

where κ is choosen:

$$\kappa = \sqrt{\frac{3(\delta-1)}{\delta+2}} \tag{2.88}$$

in order to have canonical parametrization of the field $H^{(\overrightarrow{n})}$. And by choosing the unitary gauge $Q_\mu^{(n)} = 0$, $P_j^{(n)} = 0$ and $P^{(n)} = 0$, which eliminate the non-physical degrees of freedom absorbed by the massive fields, we can rewrite the lagrangian:

$$\begin{aligned}\mathcal{L} = &\textstyle\sum_{all\,\overrightarrow{n}} -\frac{1}{2}G^{(-\overrightarrow{n})\mu\nu}\left(\Box + m_n^2\right)G_{\mu\nu}^{(\overrightarrow{n})} + \frac{1}{2}G_{\mu}^{(-\overrightarrow{n})\mu}\left(\Box + m^2\right)G_{\nu}^{(\overrightarrow{n})\nu} \\ &-G^{(-\overrightarrow{n})\mu\nu}\partial_\mu\partial_\nu G_{\lambda}^{(\overrightarrow{n})\lambda} + G^{(-\overrightarrow{n})\mu\nu}\partial_\mu\partial_\lambda G_{\nu}^{(\overrightarrow{n})\lambda} - \tfrac{1}{4}|\partial_\mu V_{\nu j}^{(\overrightarrow{n})} - \partial_\nu V_{\mu j}^{(\overrightarrow{n})}|^2 \\ &+\tfrac{m^2}{2}V^{(-\overrightarrow{n})\mu j}V_{\mu j}^{(\overrightarrow{n})} - \tfrac{1}{2}S^{(-\overrightarrow{n})jk}(\Box + m^2)S_{jk}^{(\overrightarrow{n})} - \tfrac{1}{2}H^{(-\overrightarrow{n})}(\Box + m^2)H^{(\overrightarrow{n})} \\ &-\tfrac{1}{M_{Pl}}\left[G^{(\overrightarrow{n})\mu\nu} - \tfrac{\kappa}{3}\eta^{\mu\nu}H^{(\overrightarrow{n})}\right]T_{\mu\nu}\end{aligned} \tag{2.89}$$

where $m_n = n/R$ is the Kaluza-Klein graviton squared mass.

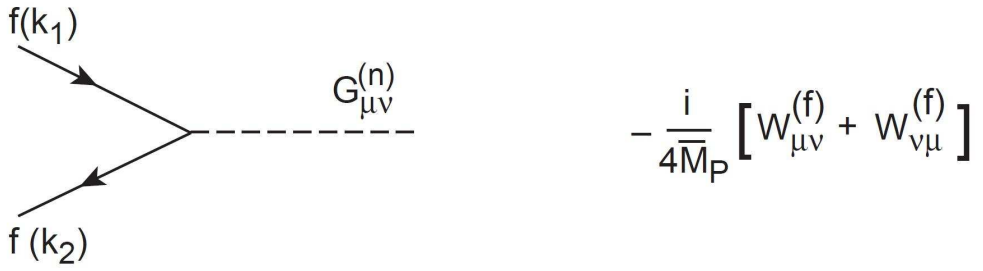


Figure 2.7: Difermion-Graviton interaction vertex and corresponding Feynman rule, where $W_{\mu\nu}^{(f)} = (k_1 + k_2)_\mu\gamma_\nu$.

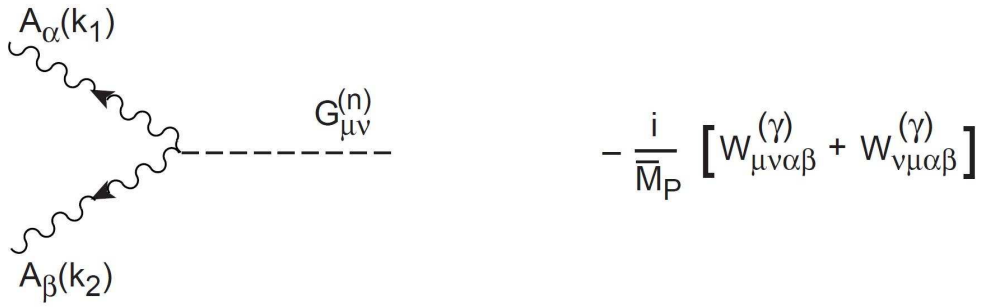


Figure 2.8: Diboson-Graviton interaction vertex and corresponding Feynman rule, where $W_{\mu\nu\alpha\beta}^{(\gamma)} = \frac{1}{2}\eta_{\mu\nu}(k_{1\beta}k_{2\alpha} - k_1\cdot k_2\eta_{\alpha\beta}) + \eta_{\alpha\beta}k_{1\mu}k_{2\nu} + \eta_{\mu\alpha}(k_1\cdot k_2\eta_{\nu\beta}) - k_{1\beta}k_{2\nu}) - \eta_{\mu\beta}k_{1\nu}k_{2\alpha}$.

The graviton interaction lagrangian is given by:

$$\mathcal{L} = -\frac{1}{M_{Pl}}G_{\mu\nu}^{(n)}T^{\mu\nu} \tag{2.90}$$

The graviton will interact with the SM particles and in particular with the QED sector. The lagrangian of QED coupled with quantum gravity is given by (see [10]):

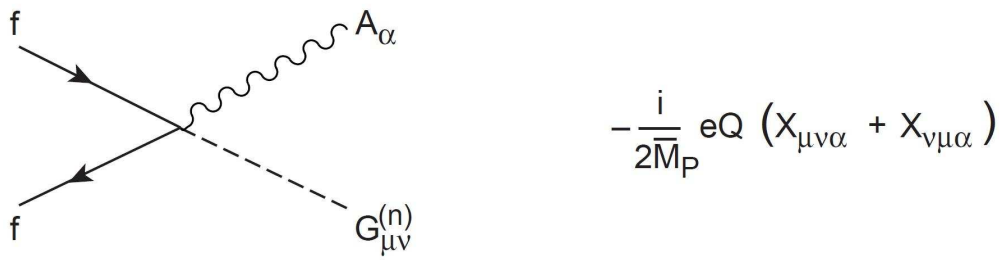


Figure 2.9: Difermion-Boson-Graviton vertex and corresponding Feynman rule, where $X_{\mu\nu\alpha} = \gamma_\mu \eta_{\nu\alpha}$.

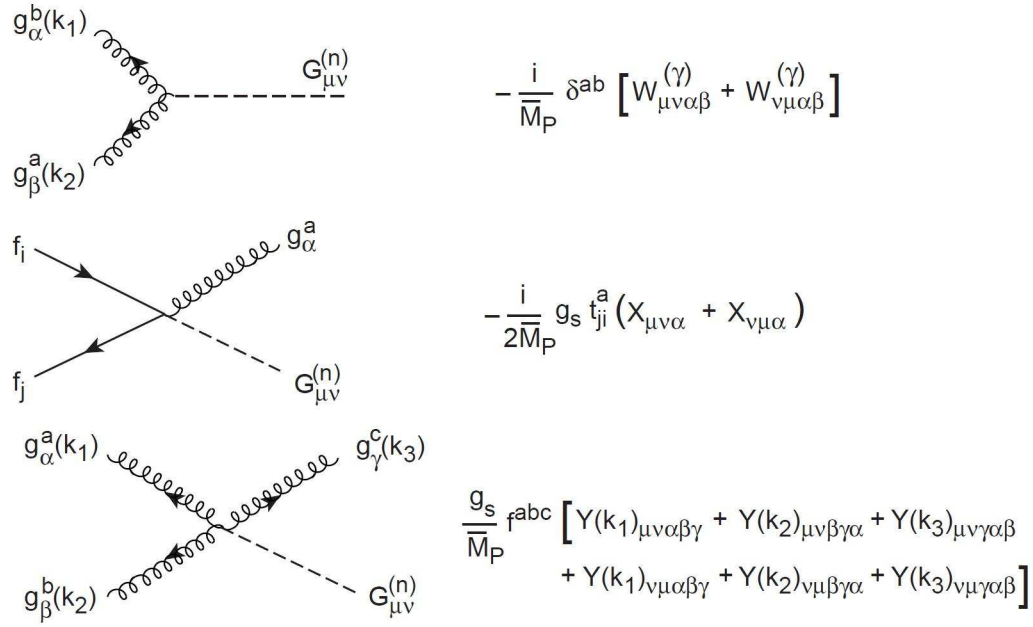


Figure 2.10: QCD-Graviton vertexes and corresponding Feynman rules, where $Y(k) = k_\mu(\eta_{\nu\beta}\eta_{\alpha\gamma} - \eta_{\nu\gamma}\eta_{\alpha\beta}) + k_\beta\left(\eta_{\mu\alpha}\eta_{\nu\beta} - \frac{1}{2}\eta_{\mu\nu}\eta_{\alpha\gamma}\right) + k_\gamma\left(\eta_{\mu\alpha}\eta_{\nu\beta} - \frac{1}{2}\eta_{\mu\nu}\eta_{\alpha\beta}\right)$.

$$\mathcal{L}_{QED} = \sqrt{-g}\left(i\overline{\psi}\gamma^a D_a\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right) \tag{2.91}$$

where the covariant derivative is:

$$D_a = e^\mu_a\left(\partial_\mu - ieQA_\mu + \frac{1}{2}\sigma^{bc}e^\nu_b\partial_\mu e_{c\nu}\right) \tag{2.92}$$

where $\sigma^{ab} = \left(\gamma^a\gamma^b - \gamma^b\gamma^a\right)$, e_a^μ are the vierbein fields and Q is the electric charge of the fermion ψ. Greek indexes refer to general coordinate transformations and Latin indexes to Lorentz transformations. By varying the Lagrangian with respect to the vierbein, we obtain:

$$T_{\mu\nu} = \frac{i}{4}\overline{\psi}(\gamma_\mu\partial_\nu + \gamma_\nu\partial_\mu)\psi - \frac{i}{4}(\partial_\nu\overline{\psi}\gamma_\nu + \partial_\nu\overline{\psi}\gamma_\mu)\psi \tag{2.93}$$

$$+\frac{1}{2}eQ\overline{\psi}(\gamma_\mu A_\nu + \gamma_\nu A_\mu)\psi + F_{\mu\lambda}F_\nu^\lambda + \frac{1}{4}\eta_{\mu\nu}F^{\lambda\rho F_{\lambda\rho}} \tag{2.94}$$

From equation 2.90 we derive the Feynman rules in figures 2.7, 2.8 and 2.9.
The generalization of these results to the case of QCD coupled to quantum gravity is straightforward. The energy-momentum tensor for QCD has the same form as in the QED case with the replacements: $eQA_\mu \to g_s A_\mu^a t^a$ and $F_{\mu\nu} \to G_{\mu\nu}^a$. We obtain the Feynman rules in figure 2.10.

We will see in section 2.2.4 the phenomenological consequences of ADD extra dimensions at hadron colliders.

2.2.3 Hierarchy problem in extra dimensions

Quite confidently we could define particle physics as the physics of the SM, its problems and the study of the possible solutions.
Since the end of the LEP accelerator as running experiment in 2000, there has been no large accelerator running at CERN, until September of 2008, when in LHC beams were made circulating several times counterclockwise and clockwise.
At the end of 2009 LHC will resume operation providing, among other things, a tool for investigation of:

- the physical mechanism that provides the breaking of the EW symmetry, which is at the root of most of the problems of the SM.
- the physics at the TeV scale.
- identification of the particles which make up the Dark Matter.

The EW lagrangian has a symmetric part and a symmetry breaking part. The former has been tested with an extreme accuracy, at SLC, LEP, Tevatron and many other accelerators all over the world and it is dictated by the gauge symmetry $SU(2) \times U(1)$. The latter part is completely unspecified and it is a mere conjecture. It is simply the simplest addition to the model which allows to do computations. We know little on that part of the lagrangian:

- No Higgs field has been observed at LEP2 and a limit at 95% C.L. of the Higgs mass at 114.4 GeV has been established.
- Radiative corrections constrain the mass to less than 185 GeV (at 95% C.L., this including LEP2 results).
- A limit at 95% C.L. of the Higgs mass in the range $160 \div 170$ GeV and at 90% C.L. in the range $157 \div 180$ GeV has been set at Tevatron from direct searches (see 2.11).
- From measurement of the W mass and the Fermi coupling, the V.E.V is determined to be 174 GeV.
- The Higgs has to transform as a doublet under $SU(2) \times U(1)$.

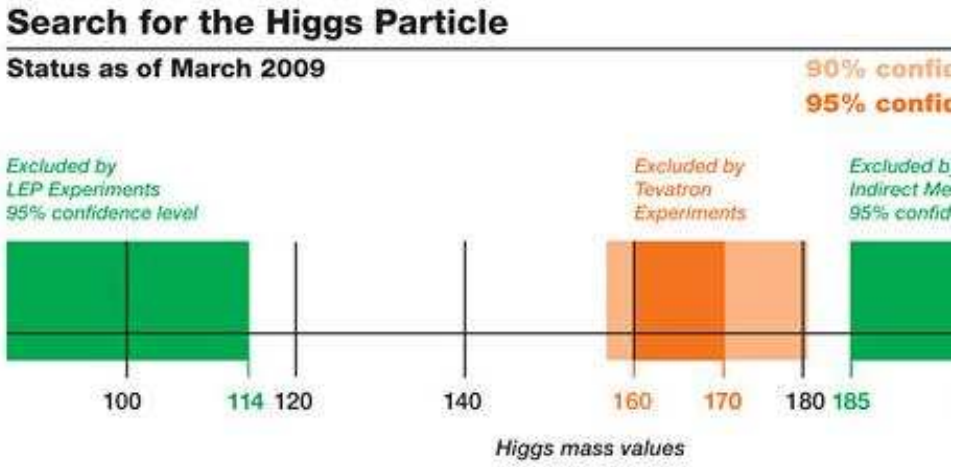


Figure 2.11: The Higgs mass theoretical and experimental limits as of March 2009.

The couplings are very symmetric, while the symmetry is badly broken in the Higgs sector and the bosons and fermions are extremely massive, i.e. in the case of the Z, W and top. Furthermore, the fact that the mass of the top and the bottom are different means that even the global $SU(2)$ symmetry is broken. The symmetry is perfectly respected in the currents and explicitly broken in the mass spectra. In a gauge theory the fact that the symmetry is broken but that also the mass spectra are broken, is the defining feature of spontaneous symmetry breaking.

Fig. 2.12 shows the Higgs mass theoretical bounds. If we define Λ as the scale up to which the SM is valid, then the upper bound is obtained by requiring no

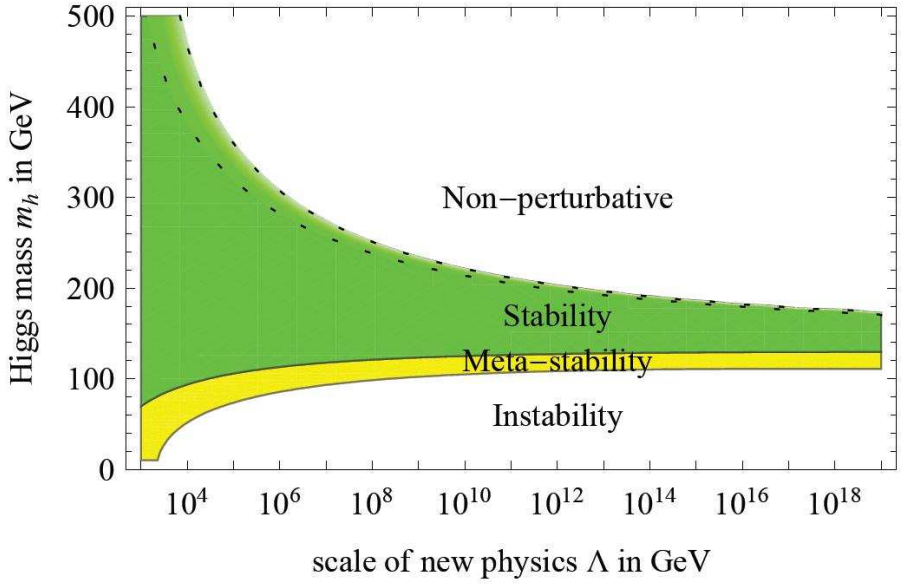


Figure 2.12: The Higgs mass theoretical limits. The upper bound is obtained by requiring no Landau Pole up to Λ, the lower bound comes by requiring vacuum stability or metastability.

Landau Pole up to Λ and the lower bound comes by requiring vacuum stability or metastability. The higher the value of the scale Λ up to which the SM is required to be valid, the more severe the constraint on the SM Higgs mass. Requiring validity up to the M_{GUT} or M_{Pl} scale, requires the Higgs mass to be constrained in the range 128 GeV $< m_H <$ 180 GeV (where these values depend on the actual value of the top mass). Instead, if the SM is only valid up to $\sim$ 1 TeV, the Higgs must still be well below 1 TeV. Therefore the LHC was designed to cover the whole range of possible Higgs masses.

Let us write the most general lagrangian:

$$\mathcal{L} = o(\Lambda^2)\mathcal{L}_2 + o(\Lambda)\mathcal{L}_3 + o(1)\mathcal{L}_4 + o(1/\Lambda)\mathcal{L}_5 + o(1/\Lambda^2)\mathcal{L}_6 + ... \tag{2.95}$$

and let us assume that there must be some more fundamental theory above Λ, the SM will be some low energy limit of this more fundamental theory. Then the lagrangian can be written as an expansion on $1/\Lambda$. In absence of some dynamical reason the coefficients of the operators should be dictated by dimension. The non renormalizable part of the lagrangian constituted by the operator of dimension 5

is suppressed by powers of Λ, while the renormalizable part should represent the SM lagrangian. In this part of the lagrangian the fermions masses are protected by chiral symmetry and gauge symmetry, namely Λ is replaced by $m \log \Lambda$, while the Higgs sector is not protected by symmetries and in the limit of μ^2 to zero there are no additional chiral symmetries, therefore the mass of the Higgs and of the bosons must be of the order of Λ and then Λ must be of the order of the EW scale. In particular the loop corrections to the Higgs mass, coming mostly from heavy particle like W, Z and top, are of the order of Λ. For instance, the loop correction coming from the top loop, which is the one more largely contributing, is given by:

$$\delta m^2_{Higgs}(top) = -\frac{3G_F}{2\pi\sqrt{2}} m_t^2 \Lambda^2 \sim -(0.2\Lambda)^2 \tag{2.96}$$

In general, any particle having a mass of the order of the scale of new physics Λ gives a contribution to the Higgs mass parameter (see Fig. 2.13). For scales Λ of the order of the Planck scale the contribution are too high and loop contribution should be cut off near the weak scale.

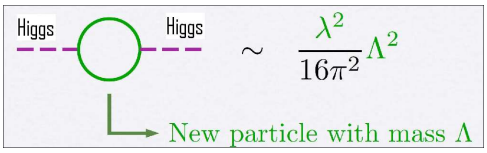


Figure 2.13: A particle having a mass of the order of the scale Λ gives a contribution to the Higgs mass parameter through loop contribution.

Theories with more than 4 dimensions are intrinsically non renormalizable. This is not a problem as far as the theory is valid below a cut off scale Λ, also called KK scale. For a flat space in 5 dimensions the scale is:

$$\Lambda \propto \frac{24\pi^3}{N_c \pi R} \tag{2.97}$$

while for the RS scenario the scale is:

$$\tilde{\Lambda} \propto \frac{24\pi^3}{N_c k L} \tilde{k} \tag{2.98}$$

The scale of new physics Λ being close to the EW scale allows to solve the hierarchy problem, leaving open only the question why the KK scale and the EW scale are so close.

The Higgs field enter to play in the extra dimensions by means of a Gauge field. The fifth component of the polarization of a Gauge field is a scalar and

it can be identified with the Higgs field. The zero mode of such a field has the following profile:

$$f_h(y) = \sqrt{\frac{2kL}{e^2kL - 1}} e^{2ky} \tag{2.99}$$

therefore is localized in the IR brane.

The potential of the Higgs field is then determined by the 5-dimensional gauge invariance. In some cases the VEV of the Higgs field is zero. In other cases the potential has a non trivial minimum. These are the cases when the symmetry is broken and so are the most viable. In 5-dimensional case the Higgs potential at loop level can be written as following:

$$V(h) = \sum_r \pm \frac{N_r}{4\pi^2} \int_0^{infty} dpp^3 log\rho(-p^2) \tag{2.100}$$

which is the Coleman-Weinberg potential, where:

$$\rho_{top}(z^2) = 1 + F_1(z^2) sin^2\left(\frac{h}{\sqrt{2}f_h}\right) + F_2(z^2) sin^4\left(\frac{h}{\sqrt{2}f_h}\right) \tag{2.101}$$

with $f_h \propto \tilde{k}/\sqrt{kL}$. This means that there is a cut off at $\tilde{k}$ and this is related to the fact that the Higgs field is localized on the higher brane where the scale is the order of Λ.

2.2.4 Remarks on the collider phenomenology of ADD Extra Dimensions

In the ADD model the graviton Kaluza-Klein modes have masses equal to $|n|/R$, and therefore the different excitations have mass splittings:

$$\Delta m \sim \frac{1}{R} = M_D \left(\frac{M_D}{M_{Pl}}\right)^{2/n} \tag{2.102}$$

If we take $M_D = 1$ TeV and $n = 4$, 6, 8, Δm is equal to 20 keV, 7 MeV, and 0.1 GeV, respectively. Only for a large number of extra dimensions does the mass splitting becomes comparable with the experimental energy resolution.

For experimental applications, it is convenient to express the graviton-production rate in terms of inclusive cross sections, where the contributions of the different Kaluza-Klein modes have been summed up. For not too large n, the mass splitting Δm is so small that the sum over the different Kaluza-Klein states can be replaced by a continuous integration. The number of modes with Kaluza-Klein index between $|k|$ and $|k| + dk$ is given by:

$$dN = S_{n-1}|k|^{n-1}dk, S_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)} \tag{2.103}$$

where S_{n-1} is the surface of a unit-radius sphere in n dimensions. We can express the differential cross section for inclusive graviton production as:

$$\frac{d^2\sigma}{dtdm} = S_{n-1}\frac{M_{Pl}^2}{M_D^{n+2}}m^{n-1}\frac{d\sigma_m}{dt} \tag{2.104}$$

The cross sections for the parton processes relevant to graviton plus jet production in hadron collisions are:

$$\frac{d\sigma_m}{dm}(q\overline{q} \to gG) = \frac{\alpha_s}{36}\frac{1}{sM_{Pl}^2}F_1(t/s, m^2/s) \tag{2.105}$$

$$\frac{d\sigma_m}{dm}(qg \to qG) = \frac{\alpha_s}{96}\frac{1}{sM_{Pl}^2}F_2(t/s, m^2/s) \tag{2.106}$$

$$\frac{d\sigma_m}{dm}(gg \to gG) = \frac{3\alpha_s}{16}\frac{1}{sM_{Pl}^2}F_3(t/s, m^2/s) \tag{2.107}$$

where t and s are the Maldestam variables and:

$$\begin{aligned} &F_1(x,y) = \tfrac{1}{x(y-1-x)}\left[-4x(1+x)(1+2x+2x^2)\right. \\ &\left.+y(1+6x+18x^2+16x^3) - 6y^2x(1+2x) + y^3(1+4x)\right] \end{aligned} \tag{2.108}$$

$$F_2(x,y) = -y(y-1-x)F_1(\frac{x}{y-1-x}, \frac{y}{y-1-x}) \tag{2.109}$$

$$F_3(x,y) = \frac{1}{x(y-1-x)}\left[1+2x+3x^2+2x^3+x^4\right. \tag{2.110}$$

$$\left.-2y(1+x^3)+3y^2(1+x^2)-2y^3(1+x)+y^4\right] \tag{2.111}$$

as derived in [10]. We want to stress again that the calculation of graviton emission is based on an effective low-energy theory, valid below the scale M_D.

Gravitons couple to matter only gravitationally, but the $1/M_P^2$ suppression present in their production cross section can be compensated by the large multiplicity of the Kaluza-Klein modes or, in other words, by the D-dimensional phase-space factor. However the $1/M_P^2$ suppression in the graviton decay rate into ordinary matter is not compensated by phase space and therefore its lifetime is $\tau_G \sim M_P^2/m^3 \sim (TeV/m)^3 10^3$ seconds.

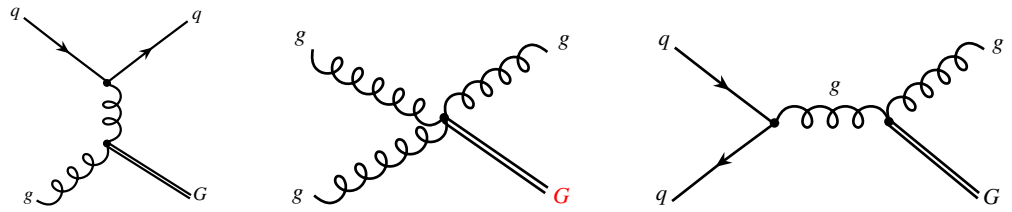


Figure 2.14: Three different production modes for real gravitons G in the ADD extra dimensions model.

The $1/M_P^2$ suppression factor can be also interpreted as the small probability that a graviton propagating in the D-dimensional space crosses the 3-dimensional brane. For experimental purposes, this means that the Kaluza-Klein graviton behaves like a massive, non-interacting, stable particle and its collider signature is imbalance in final state momenta and missing mass. Since the relevant observables for graviton production are described only by inclusive cross sections, the graviton has a continuous distribution in mass. This mass distribution corresponds to the probability of emitting gravitons with different momenta in the extra dimensions. Notice that this is a peculiarity of the graviton signal with respect to other new-physics processes. For instance, SUSY with conserved R-parity also can yield an excess of missing energy events, but these correspond to a fixed invisible-particle mass. Moreover, in the case of SUSY, the missing energy signal is in some cases accompanied by a variety of leptons, photons, and hadronic activity coming from the decay of heavier particles. For graviton-production in the perturbative regime, each extra particle in the final state is associated with an extra suppression factor. The emission of a single graviton in the extra dimensions violates momentum conservation along the directions transverse to the brane. This is not surprising, since translational invariance in the D-dimensional space is broken by the presence of the brane. In other words, the brane can radiate gravitons into the extra dimensions conserving the total energy (since time invariance is preserved), but absorbing any arbitrary transverse momentum smaller than the energy tension. From a 4-dimensional point of view, energy and momentum are conserved, but the Kaluza-Klein gravitons can have any arbitrary mass smaller than about M_D, the approximate cutoff on the validity of our description.

The leading experimental signal of graviton production at the LHC is pp $\rightarrow$ jet + missing transverse energy (MET) coming from the subprocess $qg \rightarrow qG$ (which gives the largest contribution), $q\overline{q} \rightarrow gG$, and $gg \rightarrow gG$ (see fig. 2.14). This is the signature investigated in this thesis.

A different signal for graviton production comes from events with a photon and missing energy in the final state.

From the phenomenological point of view the key issue is that when the fundamental scale M_D is of the order of TeV, gravity can become stronger than in ordinary space and light Kaluza-Klein gravitons can be directly produced.

Chapter 3

Experimental apparatus

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) (see Fig. 3.1) is the proton-proton collider under construction at CERN. It will collide protons with a center of mass energy $\sqrt{s} = 14$ TeV with a design luminosity $\mathcal{L}$=10^{34} cm^{-2} s^{-1}.

One of the critical aspects in accelerating the protons up to an energy of 7 TeV is the required bending magnetic field which, for the LHC bending radius ($R \sim$ 2780 m) is about 8.4 T. 1232 LHC superconducting 14.2 m long dipole magnets will create this magnetic field; they are placed in the eight curved sections which connect the straight sections of the LHC ring. The super-conducting magnets use a Ni-Ti conductor, cooled down to 1.9 K by means of super-fluid Helium. The choice of a collider obliges to install two separate magnetic chambers which will lay in the same mechanical structure and cryostat for economical reasons.

The high luminosity of the LHC is obtained by a high frequency bunch crossing and by a high number of protons per bunch: two beams of protons with an energy of 7 TeV (3.5 - 5 TeV in the initial physics runs), circulating in two different vacuum chambers, will contain each 2808 bunches filled with about 1.15×10^{11} protons. The main machine parameters are summarized in Table 3.1.

The operating conditions at the LHC are extremely challenging for the experiments. The $p-p$ total inelastic cross section at $\sqrt{s}$= 14 TeV is about 80 mb, several orders of magnitude larger than the typical cross section for events with large momentum transfer. Most of the inelastic events consist of soft $p-p$ interactions characterized by outgoing particles with a low transverse momentum. These events are referred to as *minimum bias*. It is expected that each bunch crossing will produce about 20 minimum bias events in the high luminosity phase and 4 minimum bias events in the low luminosity phase. Hence, each interesting events will be readout entangled with a large number of minimum bias events, which constitute the *pile-up*. The high interaction rate ($\sim 10^9$ events/s) and the

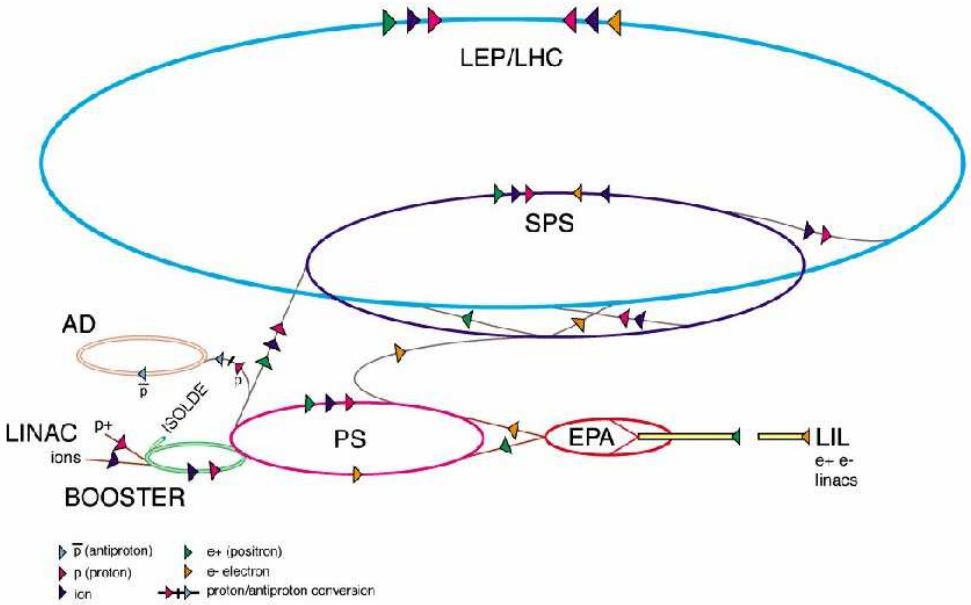


Figure 3.1: Scheme of the LHC injection chain.

high bunch crossing frequency impose stringent requirements on the data acquisition and trigger systems and on the detectors. The trigger has to provide an high rejection factor, maintaining at the same time a high efficiency in selecting the interesting events. The detectors have a fast response time (25-50 ns) and a fine granularity (and therefore a large number of readout channels) in order to minimize the effect of the pile-up. Furthermore the high flux of particles coming from the $p-p$ interactions implies that each component of the detector, including the read-out electronics, has to be radiation resistant.

3.2 The CMS detector

The CMS (*Compact Muon Solenoid*) detector is one of the two general purpose experiment that will take data at the LHC. The CMS structure is a typical one for experiments at colliders: a cylindrical central section (the *barrel*) closed at its end by two caps (the *endcaps*), as sketched in Fig. 3.2. The coordinates system in CMS are chosen with the z axis along the beam direction, the x axis directed toward the center of the LHC ring and the y axis directed upward, orthogonally to the z and x axes. Given the cylindrical structure of CMS, a convenient and commonly used coordinate system is r, ϕ, η, where r is the distance from the z

Beam parameters	
Beam energy	7 TeV
Maximum luminosity	$10^{34}\text{cm}^{-2}\text{s}^{-1}$
Time between collisions	25 ns
Bunch length	7.7 cm
RMS beam radius at the interaction point	16.7 μm
Technical parameters	
Ring length	26658.9 m
Radiofrequency	400.8 MHz
Number of bunches	2808
Number of dipoles	1232
Dipole magnetic field	8.33 T

Table 3.1: *The relevant LHC parameters for $p-p$ collisions.*

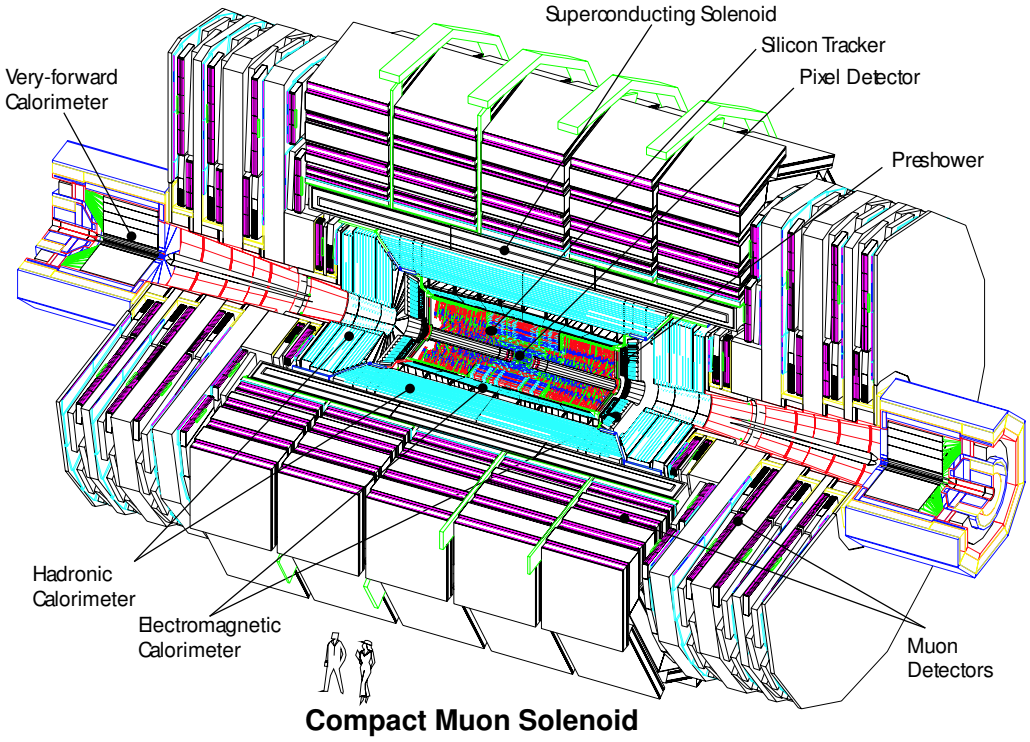


Figure 3.2: CMS overview.

axis, ϕ is the azimuthal angle in the xy plane and η is the pseudorapidity defined

as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \quad , \tag{3.1}$$

where θ is the angle with respect to the beam axis. CMS is characterized by high hermeticity with a full coverage in ϕ and up to $\eta = 5$ in pseudorapidity. The detector consists of different subdetectors which have different tasks. Starting from the innermost region to the outermost they are: the inner tracking system; the electromagnetic calorimeter (ECAL); the hadron calorimeter (HCAL) and the muon chambers. In addition several subdetectors with coverage beyond $|\eta| = 5$ are planned (CASTOR, ZDC sampling calorimeters, FP420 proton taggers and fast timing detectors). A characteristic feature of CMS is that it has a compact design.

3.2.1 The tracking system

CMS tracker [11] is the subdetector closest to the interaction point, placed in the 3.8 T magnetic field of the superconductive solenoid; it is designed to determine the interaction vertex, measure with good accuracy the momentum of the charged particles and identify the presence of secondary vertices. The tracker must be able to operate without degrading its performances in the hard radiation environment of LHC and it has to comply with severe constraints on the material budget, the material of the tracker through which all particles will pass (see Fig. 3.3).. The amount of material needs to be kept to a minimum in order to limit effects like multiple scattering, bremsstrahlung, photon conversions and nuclear interactions and not to degrade the excellent energy resolution of the electromagnetic calorimeter.

The CMS collaboration has adopted silicon technology for the whole tracker. The pixel detector consists of three barrel layers and two endcap disks at each side (fig. 3.4). The barrel layers are located at 4.4 cm, 7.3 cm and 10.2 cm are 53 cm long. The two end disks, extending from 6 to 15 cm in radius, are placed on each side at $|z|$= 34.5 cm and 46.5 cm. This design allows to obtain at least two points per track in the $|\eta| < 2.2$ region for tracks originating within 2 σ_z from the central interaction point. The total number of channels is about 48 millions in the Barrel and 17 millions in the Endcap, organized in about 16000 modules of 52 columns and 80 rows. The total active area is close to 0.92 m^2. The presence of high magnetic field causes a noticeable drift of the electrons (and a smaller drift for the holes) from the ionization point along the track with a Lorentz angle of about 32°. This leads to a charge sharing between pixels which, using an analog readout, can be exploited to considerably improve the resolution down to about 10 μm. In the endcap the modules of the detector are arranged in a turbine-like shape with a 20° tilt, again in order to enhance the charge sharing.

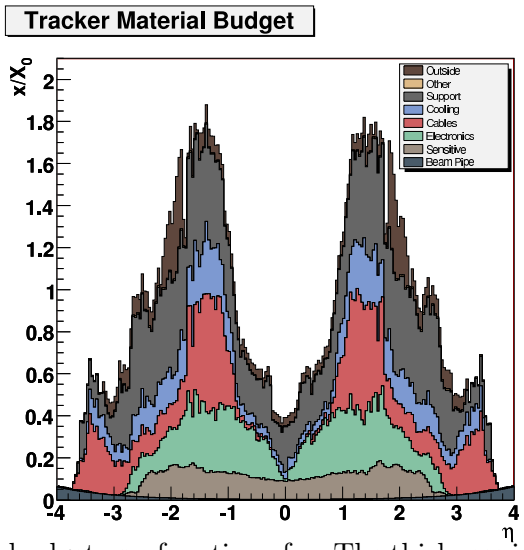


Figure 3.3: Material budget as a function of η. The thickness is expressed in terms of radiation length (X_0). The peak around η=1.5 corresponds to the cables and services of the tracker.

The inner and outer tracker detector are based on silicon strips. They are p^+ strips on a n-type bulk whose thickness is close to 300 and 500 μm respectively in the inner and outer tracker. In the barrel the strips are parallel to the beam axis while for the endcaps they have a radial orientation. The inner tracker is made of 4 barrel layers, the two innermost are double sided and the endcaps count 3 disks each. The outer tracker consists of 6 layers in the barrel (the two innermost are double sided) while the endcaps are made of 9 layers (the first, the second and the fifth are double sided). On the whole the siliconstrips tracker is made of about 10 million channels for an active area close to 198 m^2.

3.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) measures the energy of the electrons and photons. The design of the CMS ECAL [12] was driven by the requirements imposed by the search of the Higgs boson in the channel $H \to \gamma\gamma$, where a peak in the diphoton invariant mass at the Higgs mass has to be distinguished from a continuous background. A good resolution and a fine granularity are therefore required: both of them improve the invariant mass resolution on the diphoton system by improving respectively the energy and angle measurement of the two γs. The fine granularity also helps to obtain a good π^0/γ separation.
ECAL is a hermetic, homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals, 61200 crystals mounted in the central barrel part, and 7324 crystals

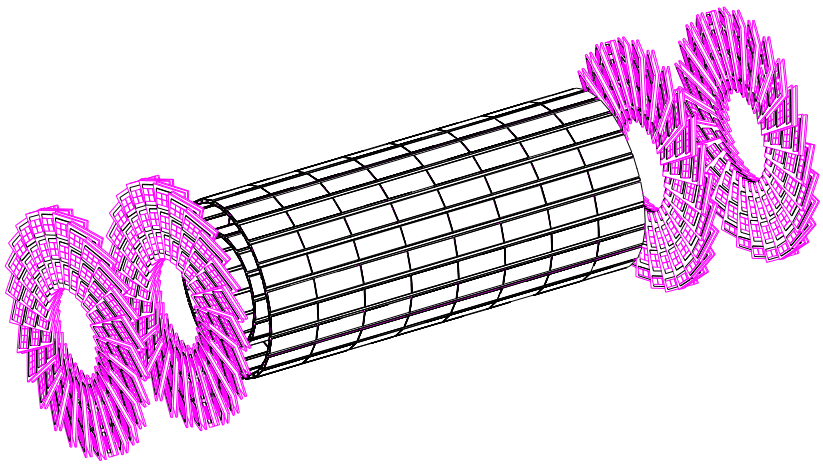


Figure 3.4: The inner pixel detector. The three barrel section and the two disks of the endcap with blades disposed in a turbine-like shape are visible. The configuration shown here corresponds to the one for the high luminosity running.

in each endcap (fig. 3.5). The choice of lead tungstate scintillating crystals was driven by the characteristics of these crystals: they have a short radiation length ($X_0 = 0.89$ cm) and a small Moliere radius ($R_M = 2.2$ cm); they are fast, as 80% of the scintillation light is emitted within 25 ns and radiation hard. The use of $PbWO_4$ crystals has thus allowed the design of a compact calorimeter to be placed inside the solenoid, and that is fast, with fine granularity and radiation resistant. However the relative low light yield (30 γ/MeV) requires the use of photodetectors with intrinsic gain that can operate in a magnetic field. In the barrel, silicon avalanche photodiodes (APDs) are used as photodetectors, while vacuum phototriodes (VPTs) have been chosen for the endcaps. In addition, the sensitivity of both the crystals and the APDs response to temperature changes requires temperature stability. A water cooling system guarantees a long term stability within 0.1°C in order to preserve the ECAL energy resolution performances.

The barrel region has a pseudorapidity coverage up to $|\eta| < 1.479$. It has an inner radius of 129 cm and is structured in 36 supermodules, each containing 1700 crystals and covering half the barrel length and 20° angle in ϕ. Each supermodule is divided along η into four modules which in turn are made of submodules, the basic assembling alveolar units. The barrel crystals have a front face cross-section of $\sim 22\times$ 22 mm^2 and have a length of 230 mm, corresponding to $25.8X_0$. The crystal axes are oriented with a 3° tilt with respect to the pointing geometry, as can be seen in Fig. 3.6 to avoid that the particles can directly escape into the dead regions between the crystals. The granularity of the barrel is $\Delta\phi \times \Delta\eta =$

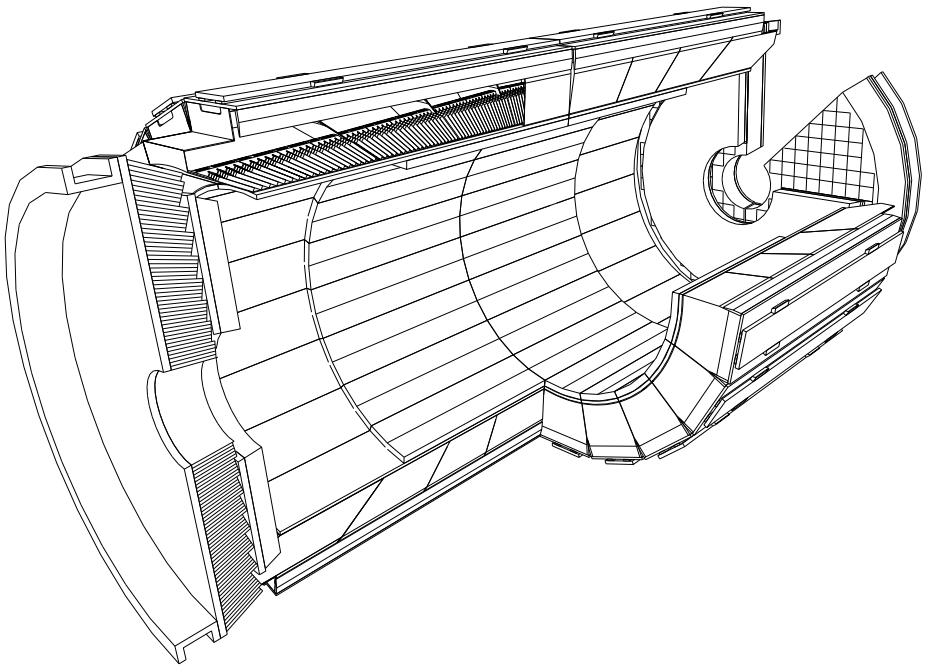


Figure 3.5: Scheme of the barrel and of the endcaps of the CMS ECAL.

0.0175×0.0175 and the crystals are grouped, from the readout point of view, into 5×5 arrays corresponding to the *trigger towers*.

The endcaps cover the pseudorapidity region $1.48 < |\eta| < 3.0$, ensuring precision measurements up to $\eta < 2.5$. The endcap crystals have dimensions of $28.6 \times 28.6 \times 220$ mm^2. Each endcap is structured in two "Dees" consisting of semi-circular aluminum plates from which are cantilevered structural units of 5×5 crystals, known as "supercrystals".
A preshower device, whose principal aim is to identify neutral pions in the endcaps within $1.653 < |\eta| < 2.6$, is placed in front of the crystal calorimeter. The active elements are two planes of silicon strip detectors which lie behind disks of lead adsorber at depths of $2X_0$ and $3X_0$.

ECAL targets and performances

One of the relevant issue in evaluating the performances of the electromagnetic calorimeter is its energy resolution. In the relevant energy range between 25 GeV and 500 GeV, the energy resolution is usually parametrized as the sum in quadra-

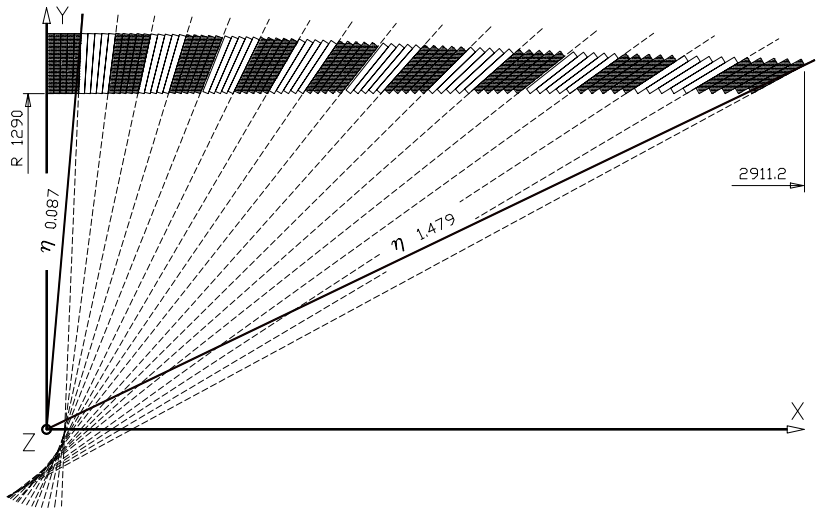


Figure 3.6: Cross section view of the crystals in the calorimeter. The groups of five crystals are highlightened by the changing of color. The shape of the crystals is different between the groups. The 3° tilt with respect to the pointing geometry is clearly visible

ture of three different terms:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad , \tag{3.2}$$

where a, b and c are named respectively stochastic, noise and constant term, and E is the energy expressed in GeV.
The target values for CMS are 2.7% for a, 200 MeV when adding the signal of 5×5 crystals for b, and 0.5% for c. Their relative contributions are reported in the plot of Fig. 3.7. It can be noticed that above 50 GeV the resolution is dominated by the constant term. Different effects contribute to the different terms in equation 3.2:

- the stochastic term a receives a contribution from the fluctuations in the number of electrons which reaches the preamplifier (n_e). These fluctuations are proportional to $\sqrt{n_e}$ and therefore proportional to the square root of the deposited energy. Contributions to this term come from the light yield of the crystals, from the efficiency in collecting light onto the photodetector surface and from the quantum efficiency of the photodetector. The fluctuations in the electrons multiplication process also contribute to this term, as an excess noise term F.

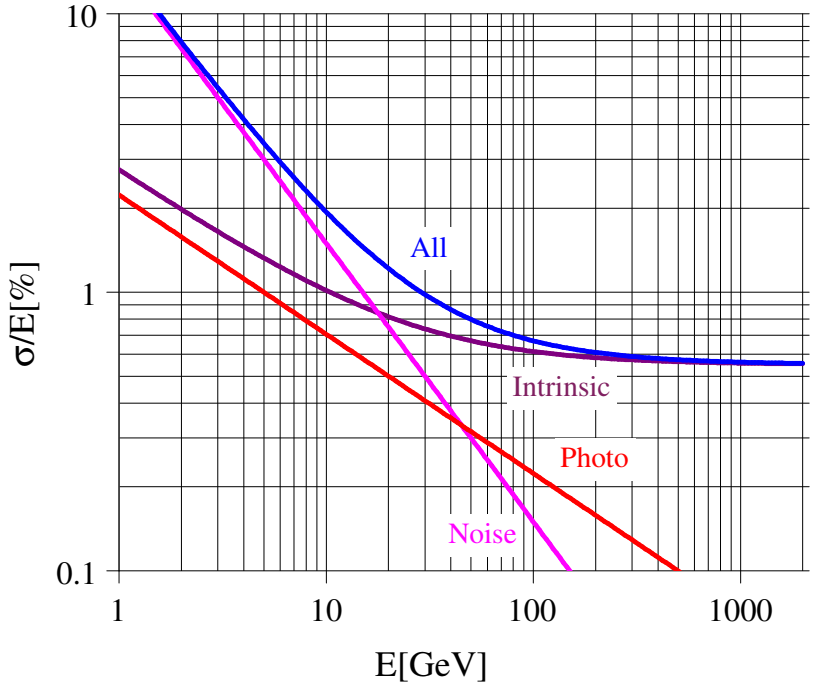


Figure 3.7: The expected ECAL energy resolution versus the energy of the impacting electron. The different contributions are superimposed separately. The term "Intrinsic" includes the shower containment and a constant term (0.55%).

- The noise term b accounts for all the effects that can alter the measurements of the energy deposit independently of the energy itself. This term receives contributions from the electronic noise and from the pile-up events, whose contribution are different in the barrel and in the endcaps and can vary with the luminosity of LHC. The target values for the barrel (at $\eta = 0$) and the endcaps (at $\eta = 2$) in the low luminosity running are respectively 155 MeV and 205 MeV while in the high luminosity running they are 210 MeV and 245 MeV.

- The constant term determines the energy resolution at high energy. Many different effects contribute to this term: the stability of the operating conditions such as the temperature and the high voltage of the photodetectors; the rear and lateral leakage of the electromagnetic shower and the presence of the dead material of the supporting structure between the crystals; the light collection uniformity along the crystal axis; the intercalibration between the channels which contributes almost directly to the overall energy resolution since the most of the energy is contained into few crystals; the

radiation damage of the $PbWO_4$ crystals [1].

3.2.3 The hadron calorimeter

The hadron calorimeter (HCAL) [13], placed just outside the electromagnetic calorimeter, plays a major role in the reconstruction of jets and missing energy. Its resolution must guarantee a good reconstruction of the dijets invariant mass and an efficient measurement of the missing energy which represent an effective signature in many channels of physics beyond the SM.

HCAL resolution was measured. Portions of each of the four HCAL subsystems were placed in a test beam to measure the characteristics and obtain a reference calibration. The electromagnetic calorimeter (ECAL) was also included in the test beam setup. When the hadronic energy resolution is parameterized as $\sigma/E = a/\sqrt{E} \oplus b$ where E is in GeV, a is the stochastic and b the constant term, for the barrel HCAL and ECAL combination, one measures $a = 84.7 \pm 1.6\%$ and $b = 7.4 \pm 0.8\%$ [14]. The energy resolution of the endcap is similar to the barrel. The corresponding values for HF are a = 198% and b = 9% [15]. The forward jets have typically very high energies, so the stochastic term of the resolution can be higher for the HF than for the other calorimeters while still providing the required energy resolution [15].

Similarly to the other subdetectors, HCAL has to provide a good hermeticity, which is critical for determining the missing energy, and a quite fine granularity to allow a clear separation of dijets from resonance decays and improve the resolution in the invariant mass of the dijets. Moreover it has to provide a number of interaction lengths sufficient to contain the energetic particles from high transverse momentum jets. The dynamic range has to be large to detect signals ranging from the signal of a single minimum ionizing muon up to an energy of 3 TeV.

The pseudorapidity region $|\eta| < 3$ is covered by the Hadron Barrel (HB, up to $|\eta| < 1.74$) and the two Hadron Endcaps (HE) calorimeters. HCAL is composed of brass layers as absorbers interleaved by plastic scintillator layers, 4 mm thick, used as an active medium. The absorber layer thickness is between 60 mm in the barrel and 80 mm in the endcaps, while the scintillators layers are 4 mm thick. In terms of interaction lengths λ, the barrel ranges from 5.46 λ at $|\eta| = 0$ up to 10.82 λ at $|\eta| = 1.3$; the endcap corresponds on average to 11 λ. The scintillator in each layer is divided into tiles with a granularity matching the granularity of the ECAL trigger towers ($\Delta\eta \times \Delta\phi = 0.0875 \times 0.0875$) and the light is collected

[1] Changes in the crystal transparency caused by irradiation and subsequent annealing lead to a variation in the crystal response to a given deposited energy; this effect develops on a short term scale ($\sim$ hours) and need to be tracked and corrected properly using a laser monitoring system.

by wavelength shifters.

The hadron calorimeter barrel is radially restricted between the outer extent of the electromagnetic calorimeter (R = 1.77 m) and the inner extent of the magnet coil (R = 2.95 m). This constrains the total amount of material which can be used to absorb the hadronic shower. Therefore, an outer hadron calorimeter or tail catcher (HO) is placed outside the solenoid complementing the barrel calorimeter. The HO will utilize the solenoid coil as an additional absorber equal to $1.4 / \sin\theta$ interaction lengths and will be used to identify late starting showers and to measure the shower energy deposited after HB, the minimal absorber depth. The central ring (ring 0) has two layers of HO scintillators on either side of a 19.5 cm thick piece of iron (the tail catcher iron) at radial distances of 3.82 m and 4.07 m, respectively. All other rings have a single HO layer at a radial distance of 4.07 m. The HO is constrained by the geometry of the muon system. The effect of leakage is visible at $\eta = 0$ (ring 0) from 70 GeV and above increasing with energy. The mean fraction of energy in HO increases from 0.38% for 10 GeV pions to 4.3% for 300 GeV pions. The effect of shower leakage has a direct consequence on the measurement of MET in an event.

The two hadronic forward (HF) calorimeters improve the HCAL hermeticity, covering the pseudorapidity region $3 < |\eta| < 5$. It is placed at 11.15 m from the interaction point outside the magnetic field. Due to the extremely harsh radiation environment a different detection technique is used: a grid of quartz (radiation hard) fibers is embedded in a iron absorber.

3.2.4 The magnetic field

An important aspect of the CMS experiment is its axial high magnetic field. The magnet system of CMS is composed of three main parts: the superconducting solenoid, the barrel return yoke and the endcap return yoke. The return yoke is made of iron and contains the muon detectors. It is a 12-sided cylindrical structure, with a total length is about 11 m and it is divided into five rings of about 2.5 m each. It has an outer diameter of 14 m and a total weight of about 7000 tons. Each ring is divided into three iron layers where the muon detectors are inserted. The thickness of the border layers is 630 mm and the middle layer is 295 mm thick. Each endcap is composed by three independent disks, the outermost is 300 mm thick and the others are 600 mm thick.

The superconductive coil is housed into a vacuum tank and kept at the temperature of the liquid helium. The vacuum tank is supported only by the central barrel ring of the yoke and in its turn supports the calorimeter system (ECAL and HCAL) and the tracker.

3.2.5 The muon system

The muon system [16] has 3 functions: muon identification, momentum measurement, and triggering. Good muon momentum resolution and trigger capability are enabled by the high-field solenoidal magnet and its flux-return yoke. The latter also serves as a hadron absorber for the identification of muons. In CMS the muon detectors are placed beyond the calorimeters and the solenoid. Because the muon system consists of about 25 000 m^2 of detection planes, the muon chambers had to be inexpensive, reliable, and robust.

The barrel drift tube (DT) chambers cover the pseudorapidity region $|\eta| < 1.2$ and are organized into 4 stations interspersed among the layers of the flux return plates. The number of chambers in each station and their orientation were chosen to provide good efficiency for linking together muon hits from different stations into a single muon track and for rejecting background hits. In the two endcap regions of CMS, where the muon rates and background levels are high and the magnetic field is large and non-uniform, the muon system uses cathode strip chambers (CSC). The CSCs identify muons between $0.9 < |\eta| < 2.4$. There are 4 stations of CSCs in each endcap, with chambers positioned perpendicular to the beam line and interspersed between the flux return plates. A crucial characteristic of the DT and CSC subsystems is that they can each trigger on the p_T of muons with good efficiency and high background rejection, independent of the rest of the detector.

Because of the uncertainty in the ability of the muon system to measure the correct beam-crossing time when the LHC reaches full luminosity, a complementary, dedicated trigger system consisting of resistive plate chambers (RPC) was added in both the barrel and endcap regions. The RPCs provide a fast, independent, and highly-segmented trigger with a sharp p_T threshold over a large portion of the rapidity range ($|\eta| < 1.6$) of the muon system. The RPCs are double-gap chambers, operated in avalanche mode to ensure good operation at high rates.

3.2.6 The trigger system

At the nominal LHC luminosity, given that the beam crossing interval is 40 MHz, the expected event rate is about 10^9 Hz. Given the typical size of a raw event (~ 1 MB) it is not possible to record all the information for all the events. Indeed, the event rate is largely dominated by soft $p - p$ interactions with particles of low transverse momentum. The triggering system must have a large reduction factor and an high efficiency on the potential interesting events, reducing the rate down to 200 Hz, which is the maximum sustainable rate for storing events. The trigger system consists of two main steps: a Level 1 (L1) Trigger and a High Level Trigger (HLT). L1 consists of custom-designed, largely programmable electronics, whereas the HLT is a software system implemented in a filter farm of more than

three thousands commercial processors. The rate reduction capability is designed to be at least a factor of 10^5 for the combined L1 and HLT.

The Level 1 trigger

The L1 trigger [17] reduces the rate of selected events down to 50 (75) kHz for the low (high) luminosity running. The full data are stored in pipelines of processing elements, while waiting for the trigger decision. The L1 decision about taking or discarding data from a particular bunch crossing has to be taken in 3.2 μs. If the L1 accepts the event, the data are moved to be processed by the HLT. To deal with the 25 ns bunch crossing rate, the L1 trigger has to take a decision in a time too short to read data from the whole detector, therefore it employs the calorimetric and muons information only, since the tracker algorithms are too slow for this purpose. The Level-1 trigger is organized into a Calorimeter Trigger and a Muon trigger whose information is transferred to the Global Trigger which takes the accept-reject decision.
The Calorimeter Trigger is based on trigger towers, arrays of 5 crystals in ECAL , which match the granularity of the HCAL towers. The trigger towers are grouped in calorimetric region of 4×4 trigger towers. The Calorimeter Trigger identifies, from the calorimetric region information, the best four candidates of each of the following classes: electrons and photons, central jets, forward jets and τ-jets identified from the shape of the deposited energy. The information of these objects is passed to the Global Trigger, together with the measured MET. The Muon trigger is performed separately for each muon detector (see Sec. 3.2.5). The information is then merged and the best four muon candidates are transferred to the Global Trigger.
The Global Trigger takes the accept-reject decision exploiting both the characteristic of the single objects and of combination of them.

The High Level Trigger

The HLT [17] reduces the output rate down to 200 Hz. The idea of the HLT trigger software is the regional reconstruction on demand, that is only those objects in the useful regions are reconstructed and the uninteresting events are rejected as soon as possible. This leads to the development of three "virtual trigger" levels: at the first level only the full information of the muon system and of the calorimeters is used, in the second level the information of the tracker pixels is added and in the third and final level the full event information is available.

3.3 CMS software components

The high-level goals of the CMS software are to process and select events inside a HLT Farm, to deliver the processed results to experimenters within the CMS Collaboration, and to provide tools for them to analyze the processed information in order to produce physics results. The overall collection of software, now referred to as CMSSW (and in the past called ORCA, acronym of Object Oriented Reconstruction for Cms Analysis), is built around a Framework, an Event Data Model, Services needed by the simulation, calibration and alignment and reconstruction modules that process event data so that physicists can perform analysis. The physics and utility modules are written by detector groups. The modules can be plugged into the application framework at run time, independently of the computing environment. The software should be developed keeping in mind not only performance but also modularity, flexibility, maintainability, quality assurance and documentation. CMS has adopted an object-oriented development methodology, based primarily on the C++ programming language.

The primary goal of the CMS Framework and Event Data Model (EDM) is to facilitate the development and deployment of reconstruction and analysis software. The event data model is centered around the Event. The Event holds all data that was taken during a triggered physics event as well as all data derived from the taken data.

The detailed CMS detector and physics simulation is implemented for all CMS detectors in both the central region (Tracker, Calorimeters and Muon Systems) and in the forward regions, including the field map from the 3.8 T solenoid [18]. In addition, several test-beam prototypes and layouts have been simulated. The full simulation program implements the sensitive detector behavior, track selection mechanisms, hit collection and digitization (i.e. detector response). The detailed simulation workflow is as follows:

- A physics group configures an appropriate Monte Carlo event generator (several are used) to produce the data samples of interest;
- The production team/system runs the generator software to produce generator event data;
- The physics group validates the generator data samples and selects a configuration for the Geant 4 simulation (detector configuration, physics cuts, etc.);
- The production team/system runs the GEANT 4-based simulation of CMS, with generator events as input, to produce (using the standard CMS framework) persistent hits in the sensitive detectors;

- The physics group validates this hit data which are then used as input to the subsequent digitization step, allowing for pile-up. This step converts hits into digitization (also known as "digis") which correspond to the output of the CMS electronics.

The full simulation relies on the GEANT 4 toolkit. GEANT 4 provides a rich set of physics processes describing electromagnetic and hadronic interactions in detail. It also provides tools for modeling the full CMS detector geometry and the interfaces required for retrieving information from particle tracking through these detectors and the magnetic field.

The digitization step, following the hit creation step, constitutes the simulation of the electronic readout used to acquire data by the detector and DAQ systems. It starts from the hit positions and simulated energy losses in the sensitive detectors, and produces an output that needs to be as close as possible to real data coming from CMS. Information from the generation stage (e.g. particle type and momentum) is preserved in the digitization step.

The CMS Fast Simulation framework [19] has been developed in view of doing physics analyses, develop and tune reconstruction algorithms, design detector upgrades, without being penalized by CPU time considerations while still benefiting from an accurate simulation of the detector effects. It is an object-oriented system for which C++ has been chosen as programming language. The Fast Simulation does not depend on the GEANT software at all, and this allows an execution time at least two orders of magnitude faster than the full simulation. To achieve this performance, a number of simplifying assumptions were made, a number of dedicated parameterizations were used, and some optimized reconstruction algorithms were developed (e.g., for tracking). The Fast Simulation is validated and tuned to the full simulation.

Chapter 4

Dijets

4.1 Introduction

The SM describes the properties of quarks and leptons in terms of basic principles arising from the nature of the electromagnetic, weak and strong interactions. Despite the beautifulness of these principles, the theory is incomplete, leaving open questions like: Why are there 6 quark flavors? Why are quarks arranged in generations? Why are there so many different forces? How do we unify gravity with the other forces? Why is gravity so weak?

Technicolor [20], Grand Unified Theory [21], Superstrings [22], Compositeness [23], Extra Dimensions [24] and Extra Color [25] [26], and Contact Interactions [27] [28] (CI's) are theories of new physics which could give an answer to one or more of these fundamental questions.

Since the first day of data taking experiments like CMS will look for the presence of new particles in one of the most promising channels for discovery of these particles: the high mass dijet resonance channel. In these processes all the final and initial state particles are partons (quarks, antiquarks and gluons). Narrow resonances decaying in two jets are produced by all the models already mentioned.

The Large Hadron Collider (LHC) at CERN will produce many events with two energetic jets resulting from proton-proton collisions at various energies up to 14 TeV.

The main background of a signal interaction producing two final state partons is constituted by the QCD $2 \rightarrow 2$ scattering processes. These dijet events result from parton scattering, produced by the strong interaction of quarks (q) and gluons (g) inside the protons.

In this thesis is discussed how the dijets at $\sqrt{s} = 14$ TeV will be used in the search for two signals of new physics: CI's and resonances decaying into dijets,

with an emphasis on the latter signal. This generic search is applied to two models of quark Compositeness, that are used as benchmarks for the sensitivity to new physics. The first model is a contact interaction [29] among left-handed quarks at an energy scale Λ^+ in the process $qq \to qq$, modeled with the effective Lagrangian $L_{qq} = (\pm 2\pi/\Lambda^2)(\overline{q}_L\gamma^\mu q_L)(\overline{q}_L\gamma_\mu q_L)$ with + chosen for the sign. The second model is an excited quark (q^*) [30] in the process $qg \to q* \to qg$, detectable as a dijet resonance. All processes presented here have been simulated using PYTHIA version 6.4 [31].

4.2 Dijet Measurement and Cleanup

This analysis uses a CMS cone algorithm with a cone size R = 0.5 to reconstruct jets (for a detailed overview of the jet algorithms used in CMS have a look at [32]). Three kinds of jets are discussed: GenJets are from the Monte Carlo particles, CaloJets are observed at the calorimeter level after a full CMS detector simulation, and corrected CaloJets have Lorentz vectors that have been scaled to equal the GenJet on average (for a more detailed overview of the jet reconstruction, correction and resolution, have a look at [33]). Dijets are the two jets with largest transverse momentum, p_T, in the event. All jets used in this analysis satisfy the requirement $|\eta| < 1.3$, which is the optimal pseudorapidity cut, as we will see. Fig. 4.1 demonstrates that jet response in this η region is expected to vary smoothly and can be measured reliably in data via dijet p_T balance [34].

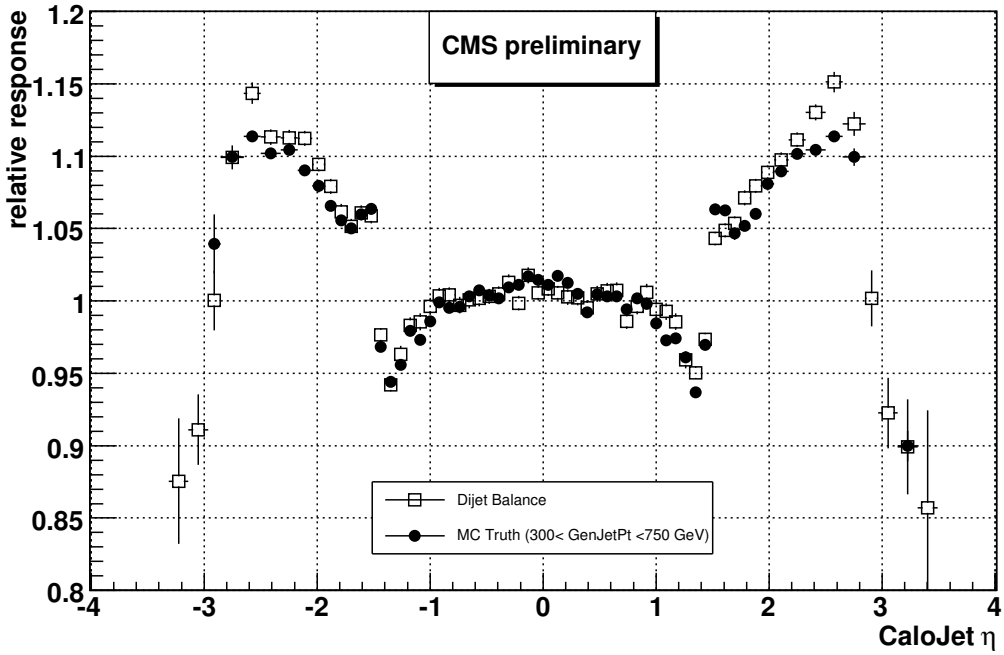
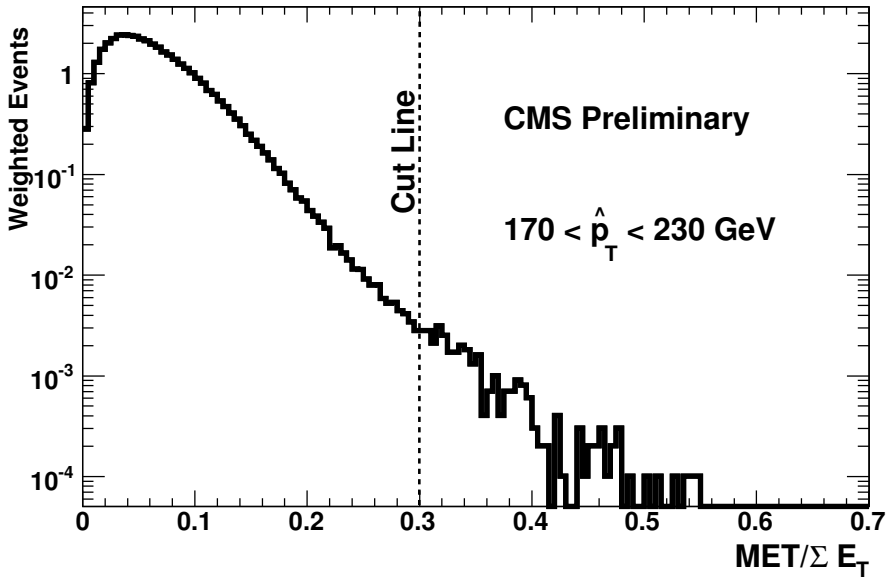


Figure 4.1: Relative jet response vs. η from MC truth compared to simulations of dijet p_T balance that are planned for collider data.

Backgrounds from cosmic rays, beam halo, and detector noise are expected to occasionally produce events with large or unbalanced energy depositions. They will be removed by requiring $\not{E}_T / \sum E_T < 0.3$ and $\sum E_T < 14$ TeV, where $\not{E}_T$



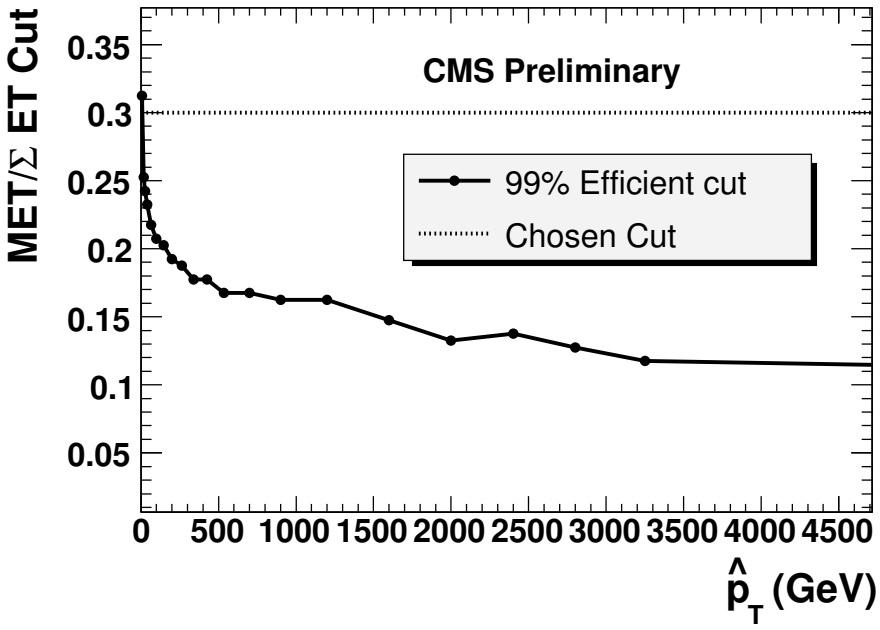


Figure 4.2: Top) MET/ΣE_T for a bin in $\hat{p}_T$ (170 GeV $< \hat{p}_T <$ 230 GeV) and a cut to remove unphysical background. Bottom) A cut that is 99% efficient compared to our cut.

($\sum E_T$) is the magnitude of the vector (scalar) sum of the transverse energies measured by all calorimeter towers in the event. This cut is estimated to be more than 99% efficient for both QCD jet events and the signals of new physics considered (Fig. 4.2). In the high p_T region relevant for this search, jet reconstruction is fully efficient.

4.3 Dijet Resonances in the Dijet Mass distribution

It is natural to search for resonances on the dijet invariant mass spectrum through simple bump hunting. This approach preserves the generality which allows to discover particles expected by one or another particular model as well as to give useful inputs to the building of new models.

The work presented in this chapter is a natural continuation of a previous analysis [35] which described the CMS plans to search for high mass dijet resonances and was published in the Physics TDR [36]. That study discussed the triggers we will use to do the search which were explored in detail previously [37].

The first motivation for our work is to validate the simulation of dijet resonance signals and backgrounds in CMSSW by comparing it to ORCA. An agreement between the two simulations would allow us to be confident on the studies performed in ORCA and to go further in our analysis with CMSSW based simulation. We will compare the dijet resonance invariant mass shapes and the QCD differential cross section in the two simulations.

The second motivation is to produce estimates that demonstrate our sensitivity for three nominal resonances masses at integrated luminosities of 10 pb^{-1}, 100 pb^{-1} and 1 fb^{-1}. We note that detailed studies were previously performed [35] to estimate our sensitivity for 100 pb^{-1}, 1 fb^{-1}. The third and most important motivation is to improve the dijet resonances analysis by optimizing the $|\eta|$ cut.

Jets are reconstructed using both the iterative and midpoint cone algorithms, with indistinguishable results for this analysis, which we will not show here. Below we will discuss three types of jets: reconstructed, corrected and generated. The reconstructed jet energy, E, is defined as the scalar sum of the calorimeter tower energies inside a cone of radius $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.5$, centered on the jet axis. The jet momentum, $\vec{p}$, is the corresponding vector sum: $\vec{p} = \sum E_i \hat{u}_i$ with $\hat{u}_i$ being the unit vector pointing from the origin to the energy deposition E_i inside the cone. The jet transverse momentum, p_T, is the component of $\vec{p}$ in the transverse plane. The E and $\vec{p}$ of a reconstructed jet are then corrected for the non-linear response of the calorimeter to a generated jet. Generated jets come from applying the same jet algorithm to the Lorentz vectors of stable generated particles before detector simulation. On average, the p_T of a corrected jet is equal to the p_T of the corresponding generated jet. The corrections estimated from a GEANT [39] simulation of the CMS detector increase the average jet p_T by roughly 50% (10%) for 70 GeV (3 TeV) jets in the region $|\eta| < 1.3$. Further details on jet reconstruction and jet energy corrections can be found elsewhere [40, 41]. The jet measurements presented here are within the region $|\eta| < 1.3$, where the sensitivity to new physics is expected to be the highest. No HO was used in the simulation.

The dijet system is composed of the two jets with the highest p_T in an event

(leading jets), and the dijet mass is given by $m = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2}$. The estimated dijet mass reconstructed width varies from 9% at a dijet mass of 0.7 TeV to 4.5% at 5 TeV.

CMS will record events that pass a first level trigger followed by a high level trigger. For an instantaneous luminosity of 10^{32} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, we consider three event samples collected by requiring at least one jet in the high level trigger with corrected transverse energy above 60, 120 and 250 GeV, prescaled by factors of 2000, 40 and 1, respectively. For an integrated luminosity of 100 pb^{-1}, the three event samples will effectively correspond to 0.05, 2.5, and 100 pb^{-1}. The first event sample will be used to measure the trigger efficiency of the second sample. The second and third event samples will be used to study dijets of mass above 330 and 670 GeV, respectively, for which the trigger efficiencies are expected to be higher than 99% [42].

4.3.1 DataSets

Mass (GeV)		700	2000	5000
Cross section for $\|\eta\| < 1.0$	Q* (pb)	0.7954E+03	0.9011E+01	0.1819E-01
	Axigluon (pb)	0.3225E+03	0.5792E+01	0.1548E-01
	E_6 Diquark (pb)	0.8113E+02	0.4197E+0	0.4648E-01
	Z' (pb)	0.8858E+01	0.1813E+00	0.5501E-03
	W' (pb)	0.1461E+02	0.3494E+00	0.8717E-03
	RS Graviton (pb)	0.3573E+02	0.1828E+00	0.2637E-03

Table 4.1: Lowest order cross section times branching ration times acceptance for dijet resonances with $|\eta| < 1.0$ from various models. See [35].

The DataSets used in our analysis are the CMSSW Monte Carlo Samples at 14 TeV produced by the LPC MC production group at the FermiLab Tier1. For the aim of the study of the response of the CMS detector to the dijet resonances signal, we have used:

- 4 K events Z'(700 GeV) to dijets,
- 3 K events Z'(2000 GeV) to dijets,
- 3 K events Z'(5000 GeV) to dijet;

produced and reconstructed with CMSSW_1_2_0, where the Z' has been generated with PYTHIA, and the cross section calculated from the Sequential Standard Model variety, and used to mimic a generic spin 1 resonance. From these events

sample i	$\hat{p_T}$ (GeV)	Events N_i	σ_i (mb)
1	0 - 15	25277	5.52E+01
2	15 - 20	50100	1.46E+00
3	20 - 30	99465	6.32E-01
4	30 - 50	98420	1.63E-01
5	50 - 80	100590	2.16E-02
6	80 - 120	48558	3.08E-03
7	120 - 170	50195	4.94E-04
8	170 - 230	49352	1.01E-04
9	230 - 300	49253	2.45E-05
10	300 - 380	30551	6.24E-06
11	380 - 470	50251	1.78E-06
12	470 - 600	50079	6.83E-07
13	600 - 800	29748	2.04E-07
14	800 - 1000	30193	3.51E-08
15	1000 - 1400	27441	1.09E-08
16	1400 - 1800	30116	1.06E-09
17	1800 - 2200	27125	1.45E-10
18	2200 - 2600	30178	2.38E-11
19	2600 - 3000	30321	4.29E-12
20	3000 - 3500	29920	8.44E-13
21	3500 - 7000	21364	1.08E-13

Table 4.2: QCD Dataset from CMSSW_1_2_0. For each sample number i we list the $\hat{p_T}$ range, the number of events generated, and the cross section. See text for how these numbers are used to calculate the QCD cross section.

a generic dijet invariant mass shape is obtained and if needed normalized to the cross sections anticipated by all the models considered in our analysis. All the expected cross sections are listed in table 4.1. Our background is mainly given by the 2 to 2 QCD scattering of partons interaction. To study the background we use 958,497 events divided in different $\hat{p_T}$ ranges as in table 4.2.

4.3.2 Jet Algorithms

The algorithms, midpoint cone with cone size $R = 0.5$ and iterative with cone size $R = 0.5$, give very similar results in terms of the dijet analysis, so we chose to only show Iterative cone jets in this chapter for the simplest comparison with ORCA.

The inputs were either CaloTowers (CaloJets), or HepMCparticles (GenJets). For the cone algorithms, CMSSW_1_2_0 jet corrections from MCJet were applied

to form corrected CaloJets (CorJets). In all cases we form the dijet invariant mass from the first two jets in the collection, the leading jets for p_T ordered jet collections, and we require each leading jet has $|\eta| < 1.0$ unless otherwise stated.

4.3.3 Dijet Mass from QCD and Resonances

The dijet mass distribution will be used to search for dijet resonances. We can use a fit or a prediction for the QCD background. Strongly produced resonances have large enough rate to be seen above the background. We expect a convincing signal for an excited quark (q*) with 100 pb^{-1}, well beyond the Tevatron exclusion of $M < 0.78$ TeV in the dijet channel [43, 44].

As it can seen in Fig. 4.12, the QCD cross section rises significantly with the $|\eta|$ cut due to the large forward scattering amplitude. The dijet resonance signal only gradually increases with the cut. Therefore we expect an optimal $|\eta|$ cut at relatively low η values, somewhere within the central region of the barrel.

4.3.4 Dijet Mass Reconstructed Width for Resonances

Due to calorimetric response, for resonances the dijet mass from CaloJets is shifted to lower values than GenJets (Fig. 4.5). Correcting the CaloJets allows us to recover the GenJet response. Corrected CaloJets peak at the same dijet mass as GenJets. In Fig. 4.3 we show a fit to the Gaussian core of the dijet mass distribution for a 2 TeV Z′. Similar fits are done for a 0.7 and 5.0 TeV resonance to obtain the reconstructed width vs. resonance mass in Fig. 4.4.

Fig. 4.4 shows that the reconstructed width for corrected CaloJets varies from 9% at $M = 700$ GeV to 4.5% at $M = 5000$ GeV, and is well fit by the parameterization $\sigma/M = 0.038 + 38/M$ where M is the true resonance mass in GeV. This supersedes prior results [35]. In the previous analysis the reconstructed width was given by the simple functional form:

$$\frac{\sigma}{M} = 0.04548 + \frac{1.306}{\sqrt{M}} \tag{4.1}$$

This estimate was subsequently used to form the bin width of the dijet invariant mass plots. The width used for each bin was chosen to be of size 1 σ to optimally resolve a resonance bump on the spectra. For best comparison with the previous analysis, we have used that same bin width in this analysis.

Fig. 4.4 also compares this expected reconstructed width to the GenJet level width and the natural width of the input resonance.

We can ideally decompose the shape of the dijet invariant mass distribution for narrow dijet resonances in two parts: a first Gaussian part, due to intrinsic jet energy resolution and a second part constituted low mass Breit-Wigner tail part

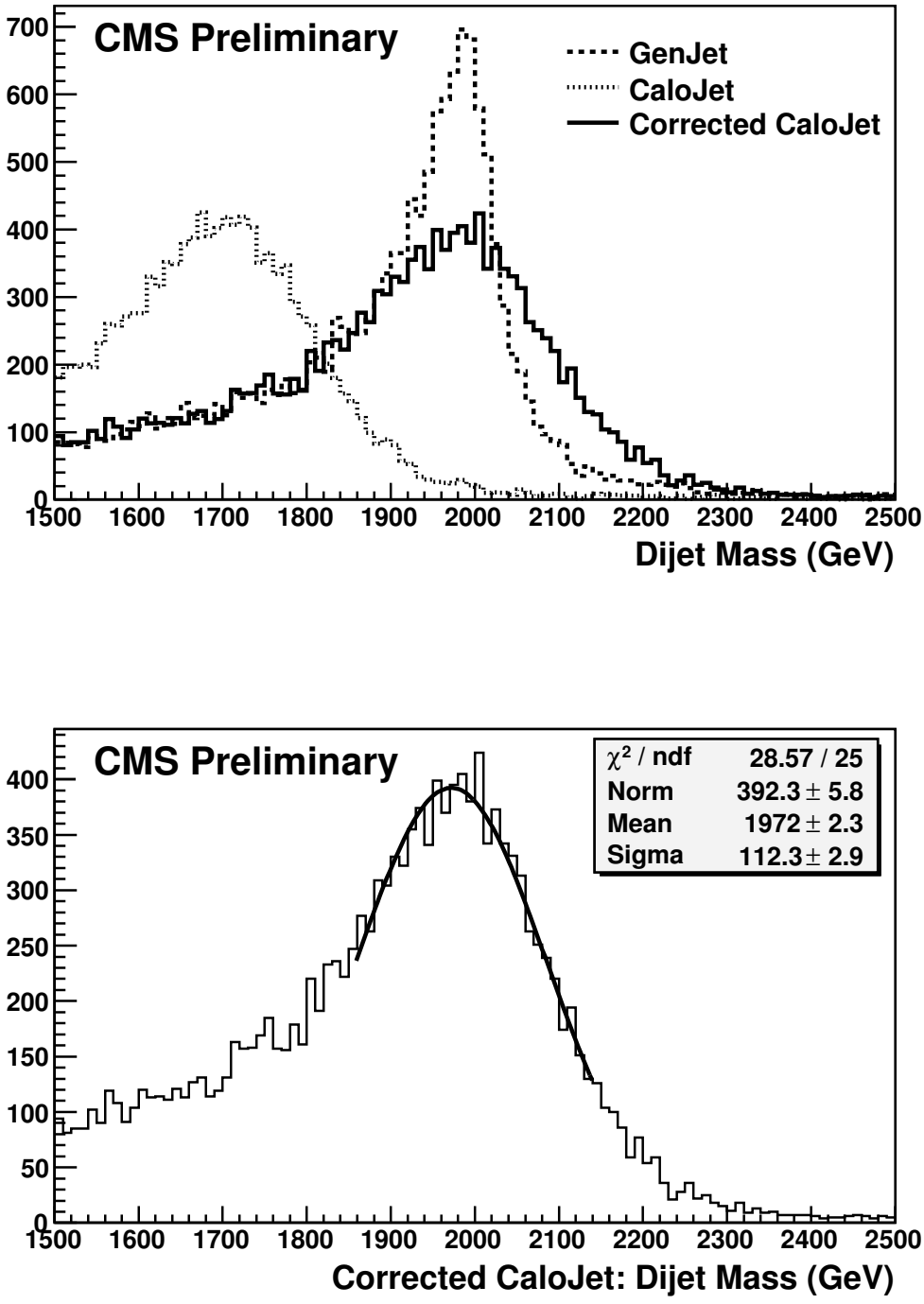


Figure 4.3: Top) Dijet invariant mass plot for a 2 TeV Z' resonance at Gen, Calo and corrected Jet levels. Bottom) Fit of the Gaussian core of the mass peak to estimate the dijet mass reconstructed width.

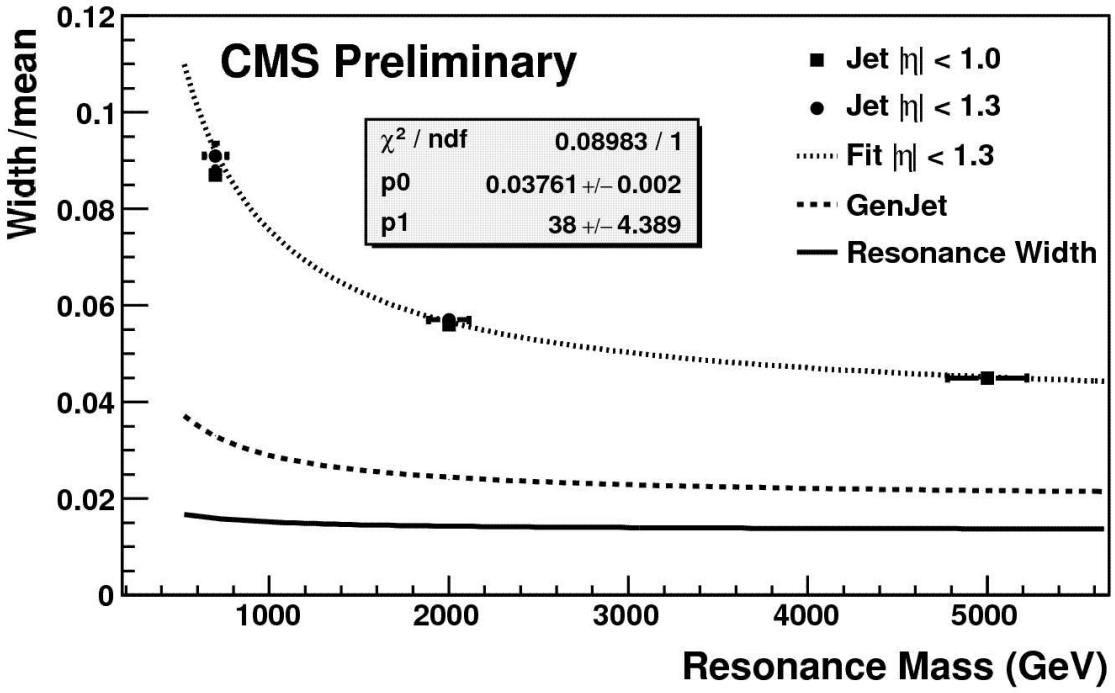


Figure 4.4: Width / mean for dijet resonances at various levels vs. input resonance mass.

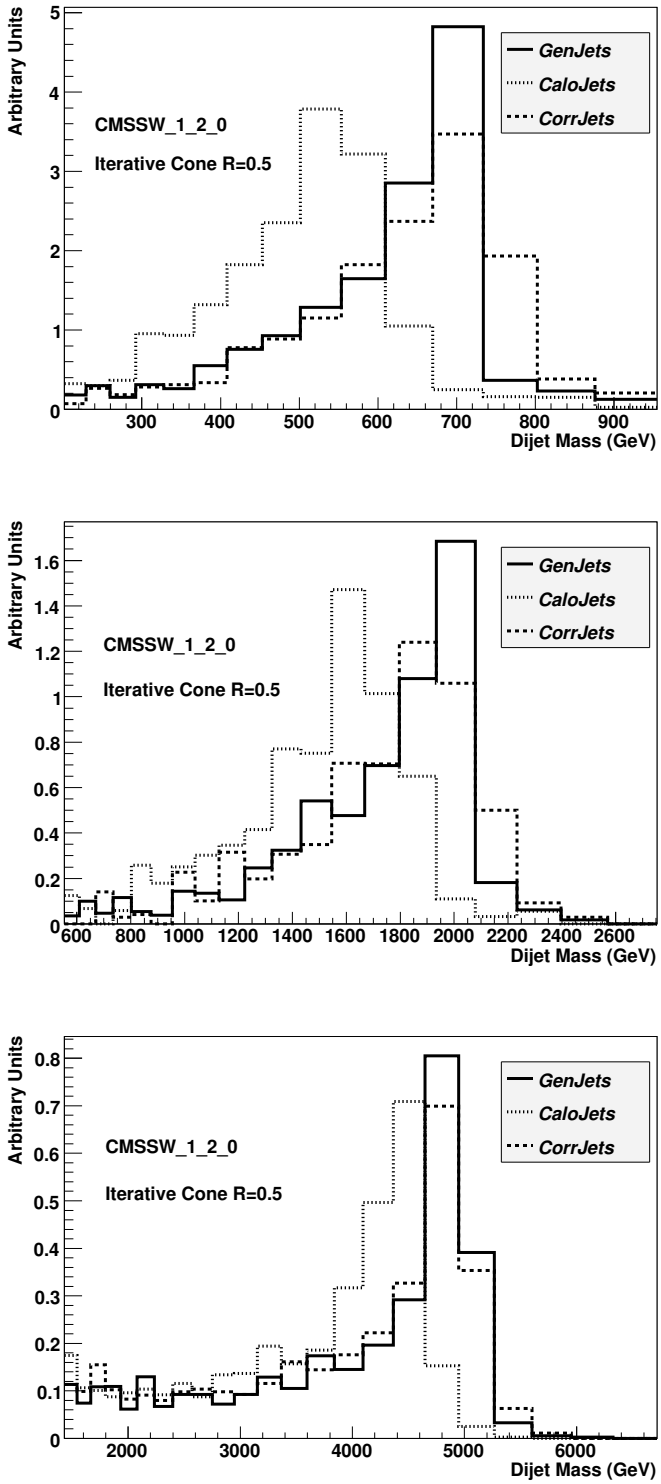


Figure 4.5: Dijet invariant mass from a Z' of mass 700 GeV (top plot), 2000 GeV (middle plot), and 5000 GeV (bottom plot). Histograms are shown for GenJets, CaloJets, and corrected CaloJets. The LHC energy is 14 TeV.

due mainly to radiation of the final state partons. For a 5 TeV Z' there is also a contribution on the tail from the parton distributions. This is demonstrated in figure 4.5 where we can see a very long low mass tail at generator level for the 5 TeV Z' resonance but a smaller one for the 2 TeV and 700 GeV Z' resonances. This affects both the particle level and calorimeter level jets. The size of this effect depends on the width of the narrow resonance and the kind of partons in the initial state.

4.3.5 QCD background

The differential cross section for the QCD background is determined by combining the samples in table 4.2 appropriately weighted by the cross section and events listed. In detail, the differential cross section, $d\sigma/dm$, in a bin of dijet mass is calculated from the contributions of all the $\hat{p_T}$ samples by the equation

$$\frac{d\sigma}{dm} = \sum_{i=1}^{21} \frac{n_i \sigma_i}{N_i \Delta m} \tag{4.2}$$

where n_i is the number of events in the bin of dijet mass originating from $\hat{p_T}$ sample i, σ_i is the cross section for $\hat{p_T}$ sample i, N_i is the total number of events generated in $\hat{p_T}$ sample i, and Δm is the width of the dijet mass bin.

We use the same mass bins as the previous analysis [35]. The width of the bin is the mass reconstructed width for a Z', discussed in the previous section, and given by Equation 4.1.

In figure 4.6 we show the QCD differential cross section for GenJets, CaloJets and corrected CaloJets. The bin edges, cross section values, and MC statistical errors are listed in tables 4.3, 4.4 and 4.5 for future reference. The cross section for CaloJets is much less than the cross section for GenJets, because the CaloJet energy is much lower due to CMS calorimeter response, and the entire CaloJet distribution is essentially shifted to lower mass with respect to the GenJet distribution. The jet corrections fix most of the calorimeter response problem. The cross section for corrected CaloJets is similar to the cross section for GenJets.

An important test of the cross section weight procedure is whether the resulting dijet cross section is smoothly falling. In figure 4.7 we show that we are able to fit the distribution with a four parameter function between a mass of 0.3 and 7 TeV and get a reasonable χ^2 of 39 for 36 degrees of freedom, an acceptable 32% probability.

The fit function used
$\frac{d\sigma}{dm} = p_0 \frac{(1-\frac{m}{\sqrt{s}}+p_3(\frac{m}{\sqrt{s}})^2)^{p_1}}{(m/\sqrt{s})^{p_2}}$
is identical to that employed by the Tevatron [45] The values of the the parameters of the fitted function can be found in figure 4.7.

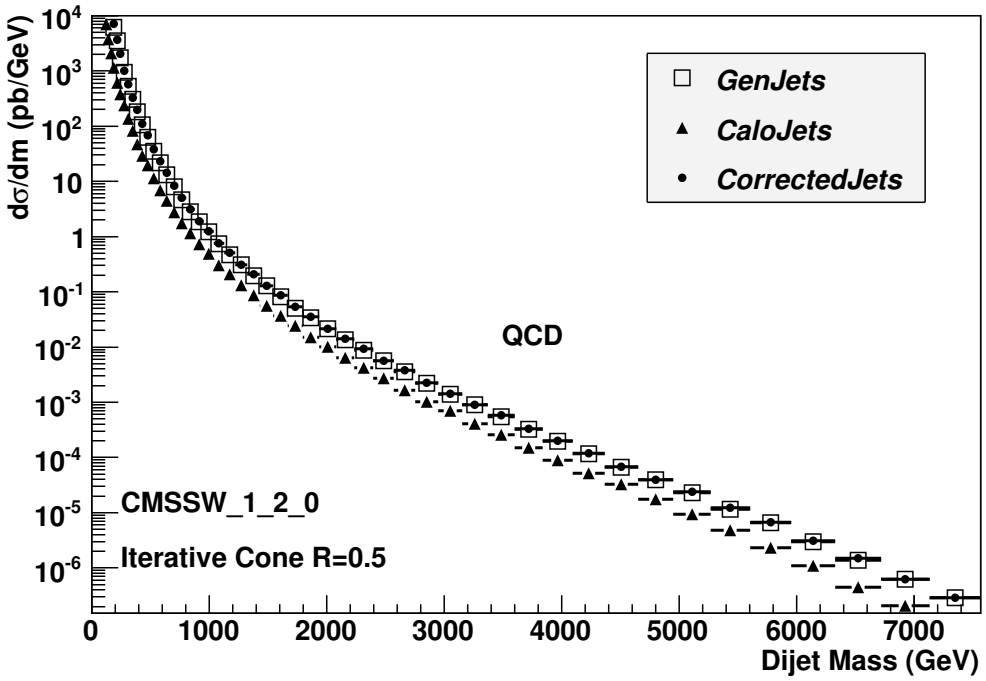


Figure 4.6: QCD differential cross section as a function of dijet invariant mass are shown for GenJets, CaloJets, and corrected CaloJets. LHC energy at 14 TeV.

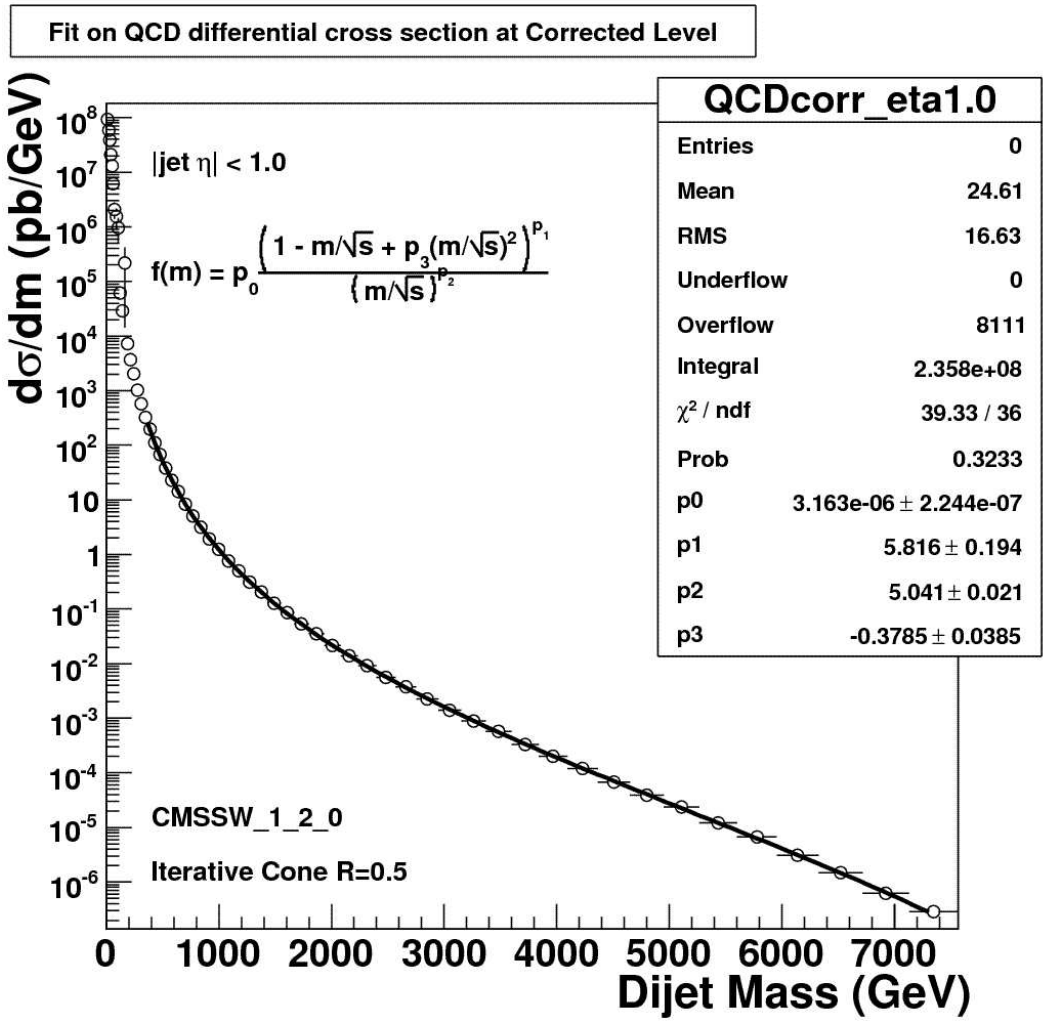


Figure 4.7: The QCD differential cross section fit to a smooth parameterization with 4 parameters. LHC energy at 14 TeV.

4.3.6 CMSSW - ORCA comparison

In this section we will present a comparison of the dijet resonances and of the QCD background in CMSSW and ORCA. This will include a direct comparison of the dijet resonance shapes for 3 mass points at generator and corrected level and of the QCD differential cross sections at generator and corrected level.

In figure 4.8 we can see the dijet invariant mass plots made with CMSSW and ORCA. The plots coming from the different CMS software are normalized to 1 and superimposed to allow an easier comparison. The agreement between the two simulations is quite good and allow us to be confident on the studies performed in ORCA and to go further in our analysis with CMSSW based simulation.

In figure 4.9 we can see the fractional difference between QCD cross sections in CMSSW and ORCA for corrected CaloJets and for GenJets. The differences between the CMSSW and ORCA simulations for Corrected CaloJets are similar to the differences for GenJets. Known differences between the two CMSSW and ORCA simulations include the following effects which are responsible:

- GenJets from CMSSW include neutrinos and muons while GenJets from ORCA did not. This makes the GenJet energy slightly larger and the cross section at a fixed GenJet dijet mass larger.
- Underlying event tune in CMSSW is different than in ORCA.
- Corrected calorimeter response in CMSSW may be somewhat different than ORCA.

The two QCD distributions are close enough for our current sensitivity study.

4.3.7 Signal significance and Background Statistical error

In figure 4.10 we compare the cross section for an excited quark dijet resonance signal to the statistical uncertainties expected on the QCD dijet background for three luminosity scenarios: 10 pb^{-1}, 100 pb^{-1} and 1 fb^{-1}. The normalization of the excited quark signal came from our lowest order calculation in table 4.1 and the shape of the excited quark signal comes from the CMSSW simulation of dijet resonances for corrected CaloJets in figure 4.5. The statistical uncertainties on the QCD background was obtained for the expected rates of corrected CaloJets, coming from the cross sections in fig. 4.6 and table 4.5 multiplied by the luminosities and trigger prescales discussed previously [35].

Figure 4.10 shows that we will be sensitive to an excited quark signal: with only 10 pb^{-1} we can see at 2 TeV excited quark signal beginning to emerge above our statistical precision with a total significance of 4.1 σ, neglecting systematic uncertainties. With 100 pb^{-1} the same 2 TeV signal has a convincing significance

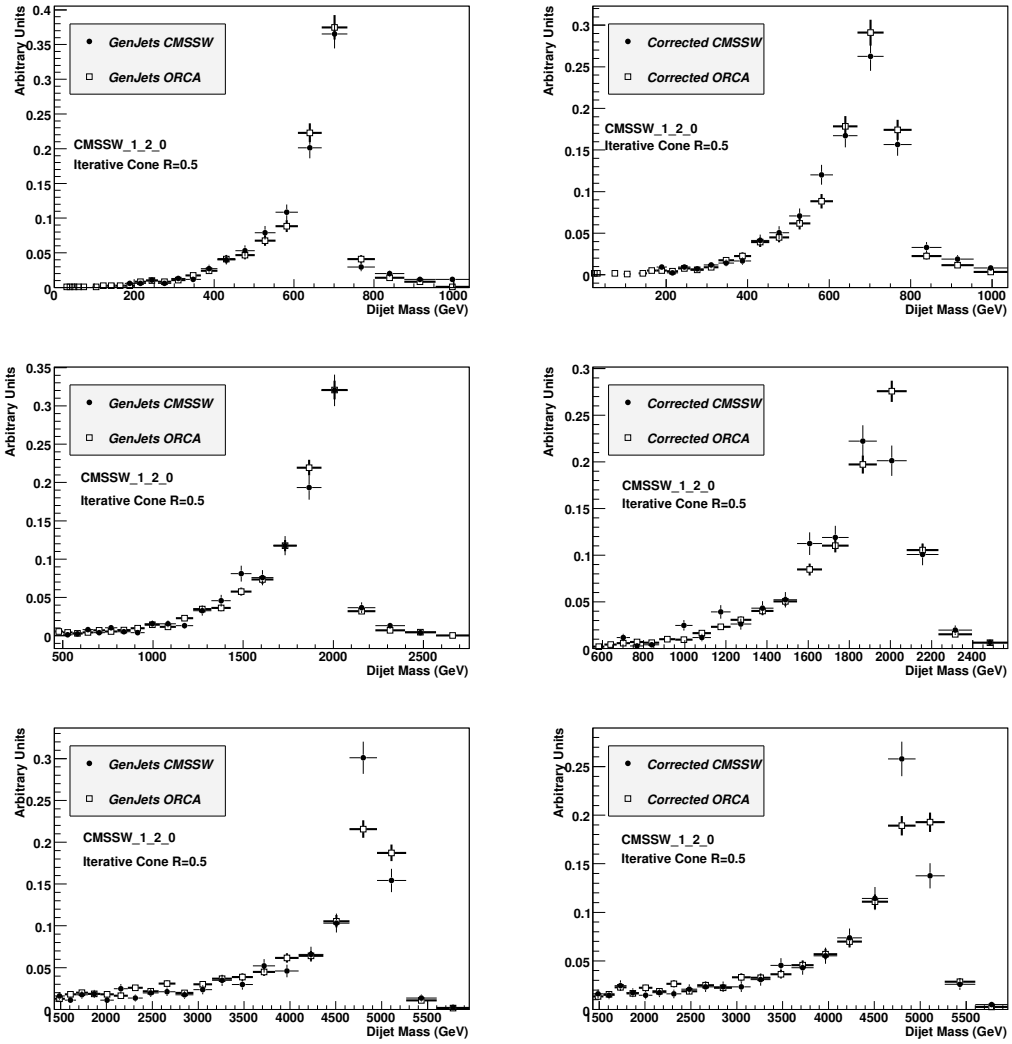


Figure 4.8: Dijet invariant mass from CMSSW and ORCA is compared for a Z' of mass 700 GeV (top plots), 2000 GeV (middle plots), and 5000 GeV (bottom plots), for GenJets (left plots) and corrected CaloJets (right plots).

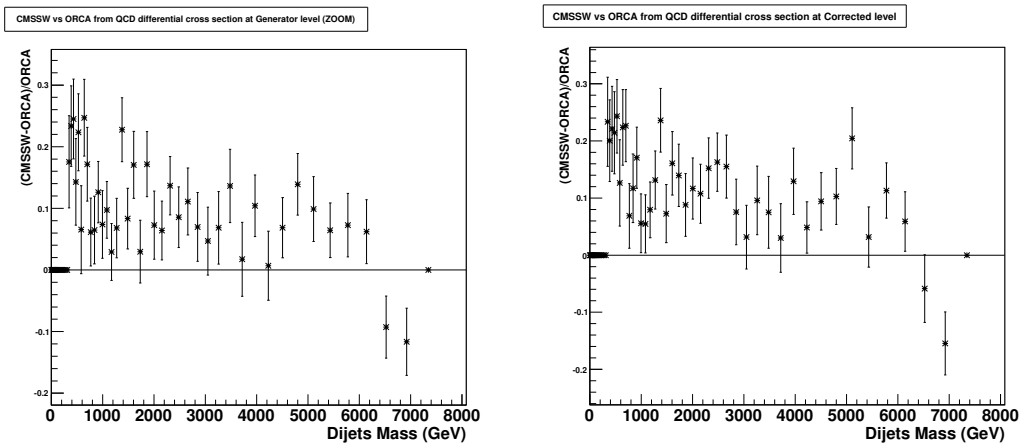


Figure 4.9: Fractional difference between QCD cross sections in CMSSW and ORCA for GenJets (left plot) and corrected CaloJets (right plot)

of 13 σ. In Figure 4.10 we list the total statistical significance for an excited quark signal at each resonance mass next to the mass peak on each plot. This total significance comes from summing the bin-by-bin significances in quadrature. For example in fig. 4.11 we show the bin-by-bin significance of the excited quark signal for an integrated luminosity of 1 fb^{-1}. The total significance for the signal is then found by summing the bin-by-bin significance in quadrature for the complete resonance. We will use this simple and robust technique for estimating statistical significance in the next section to optimize our $|\eta|$ cut.

4.3.8 $|\eta|$ cut optimization

The dijet resonance analysis presented so far required each leading jet to be in the region $|\eta| < 1.0$, just like in the Physics TDR II. This cut was inherited from the Tevatron, but it was never optimized for CMS. Here we will estimate the optimal value of this cut for statistical sensitivity to dijet resonance signals.

In Figure 4.12 we show the differential cross section as a function of dijet mass for the QCD background and a dijet resonance signal for 36 values of the $|\eta|$ cut, varying from $|\eta| < 0.5$ to $|\eta| < 4.0$ in steps of 0.1. The signal shape is from figure 4.5 and the signal normalization is from our lowest order calculation for an excited quark for the value $|\eta| < 1$ multiplied by the relative signal acceptance, R, of the given $|\eta|$ cut compared to the $|\eta| < 1$ cut in the dijet resonance simulation

$$R = \frac{N(|\eta| < \text{cut})}{N(|\eta| < 1)} \tag{4.3}$$

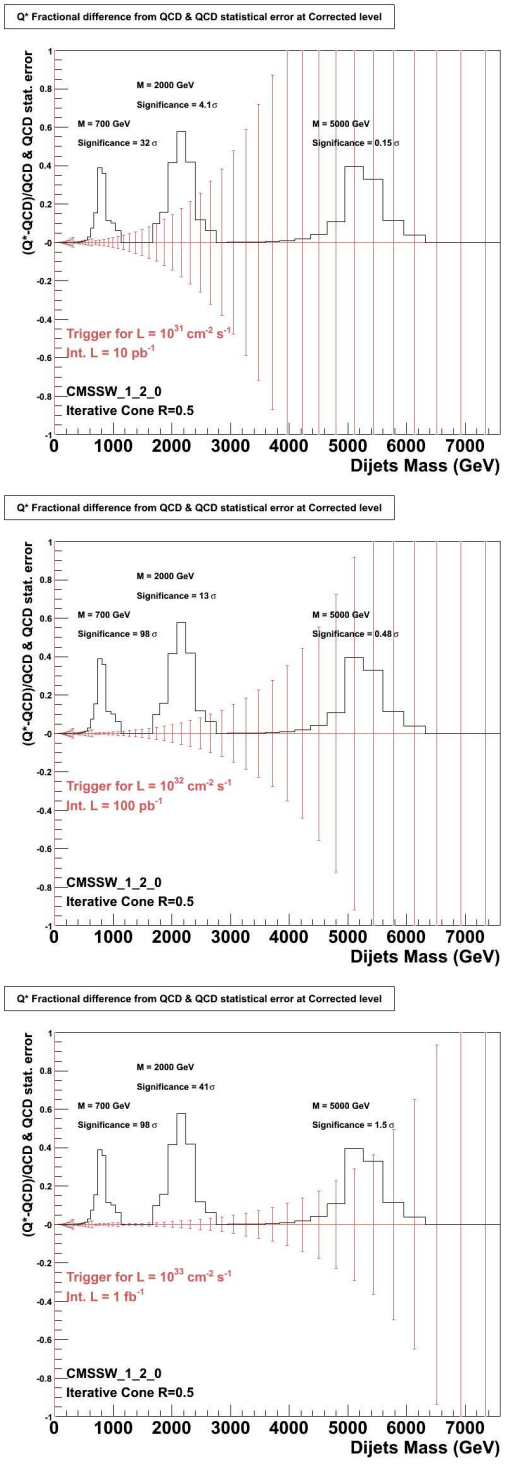


Figure 4.10: Fractional difference between an excited quark signal and the QCD background compared to the QCD statistical error for the 3 luminosity scenarios and 14 TeV LHC energy.

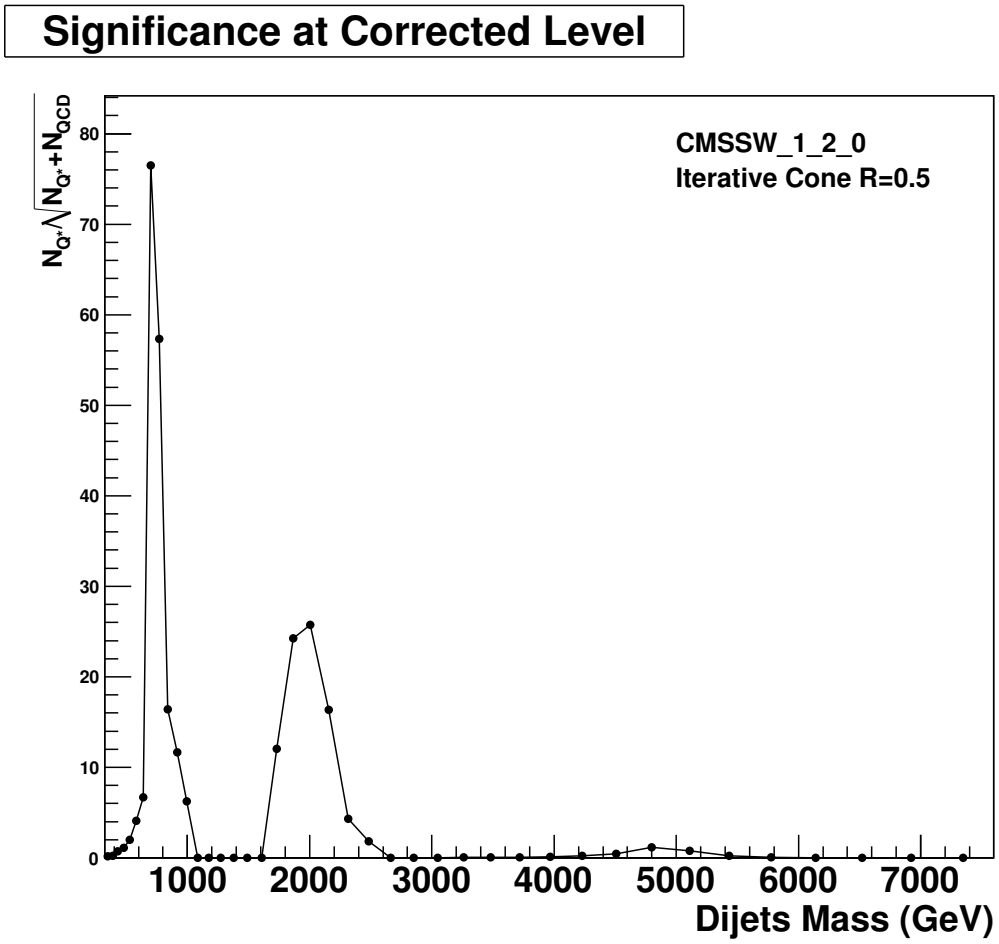


Figure 4.11: The bin by bin statistical significance of an excited quark signal compared to the QCD background as a function of dijet mass in units of standard deviations (σ) for 1 fb^{-1} of integrated luminosity.

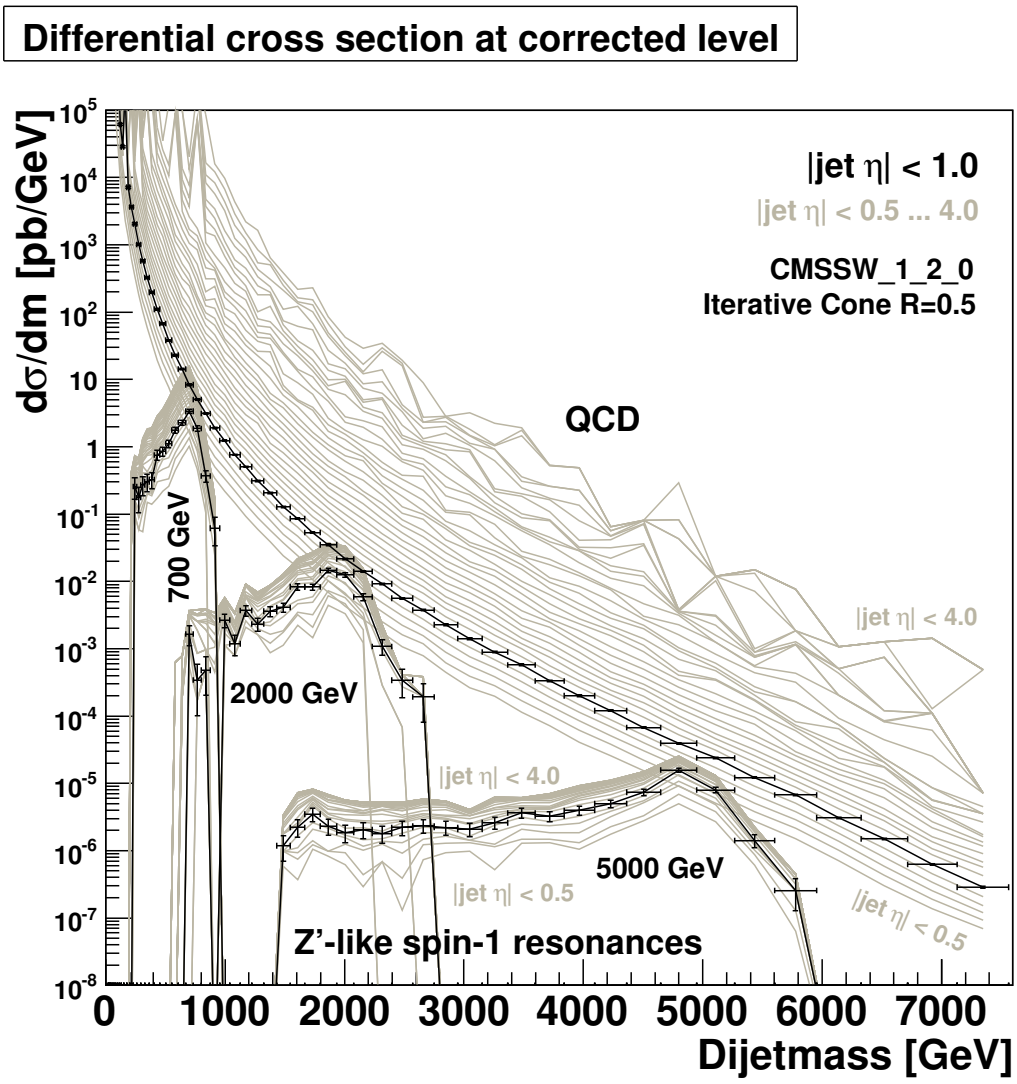


Figure 4.12: The cross section for a dijet resonance signal (see text) and the QCD background as a function of dijet mass is shown for the cut $|\eta| < 1$ (points and dark curves) and for various values of the $|\eta|$ cut between 0.5 and 4.0 (light curves).

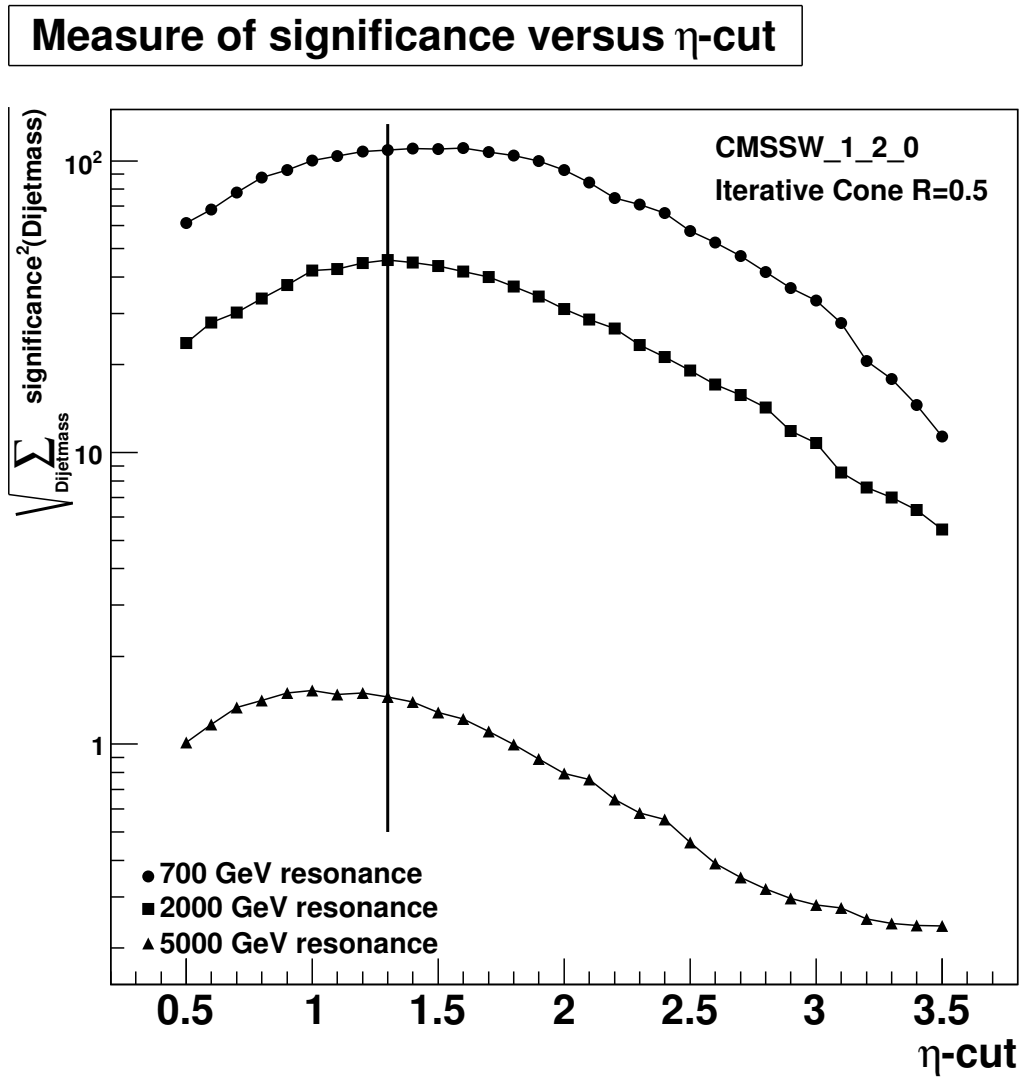


Figure 4.13: The statistical significance of a dijet resonance signal (see text) as a function of the $|\eta|$ cut for 1 fb^{-1} of integrated luminosity is shown for a resonance of mass 700 GeV (triangles), 2000 GeV (squares) and 5000 GeV (circles).

where N is the number of events measured in the resonance simulation with the particular $|\eta|$ cut. The background shape and normalization is from the QCD simulation with the given $|\eta|$ cuts. As we relax the eta cut from the default value of $|\eta| < 1.0$ (inherited from Tevatron) to higher values, up to $|\eta| < 4.0$, the QCD background cross section increases by nearly three orders of magnitude. This is because QCD is dominated by t-channel scattering, simple Rutherford-like scattering, which peaks in the forward direction. In contrast the dijet resonance signal increases only slightly as we relax the $|\eta|$ cut from 1 to 4, because dijet resonance production is an s-channel process, which is always more isotropic and hence more centrally produced than a t-channel process. Thus if the $|\eta|$ cut is completely relaxed we expect the significance of the signal over the QCD background to decrease dramatically. Conversely if the $|\eta|$ cut is tightened to very small values then we will have very few signal events and we again expect the significance of the signal to be low. Somewhere in between we expect an optimal value.

We note that the precise rate as a function of $|\eta|$ for the resonances will depend on the detailed helicity of the resonance decay to dijets: the resonance spin, and whether the resonance decays to quarks (fermions) or gluons (bosons). However, compared to the t-channel pole of QCD, which increases the rate dramatically at high $|\eta|$, all s-channel resonances are roughly isotropic in rate vs. $|\eta|$. Here we have done the optimization for the angular distribution of a spin 1 particle (Z') decaying to quarks. We expect the exact value for the optimization to be slightly different for other kinds of dijet resonances from spin arguments alone.

In Figure 4.13 we show the total statistical significance of the signal compared to the background as a function of the $|\eta|$ cut for an integrated luminosity of 1 fb^{-1}. The significance is calculated by summing in quadrature the bin-by-bin statistical significance over all the bins of the resonance. The significance is calculated independently for each of three resonance masses, 0.7, 2, and 5 TeV. An $|\eta|$ cut of 1.3 shown by a vertical line in Figure 4.13, maximizes the significance for a 2 TeV resonance, and has a very similar significance as the maximum significance cut for a 0.7 and 5 TeV resonance as well. It is likely the best single value for an $|\eta|$ cut for this analysis.

While we recommend a single cut at $|\eta| < 1.3$, we also note from Figure 4.13 that as the resonance mass increases the $|\eta|$ cut which maximizes the significance decreases from $|\eta| < 1.5$ at $M = 0.7$ TeV to $|\eta| < 1.0$ at $M = 5$ TeV.

Although a sliding cut appears to be optimal from a statistical point of view, and appears to be well motivated from the physics, we do not recommend a sliding cut for the first resonance searches at CMS for the following reasons. A sliding cut might introduce a mass dependence into the analysis which would increase the chances of generating spurious signals, particularly if the eta cut allowed jets outside the Barrel for one mass region but required them all to be

inside the barrel for another mass region. The simpler $|\eta| < 1.3$ cut will provide a single uniform region of calorimeter response in the barrel and better control the systematic uncertainties. Further, at this point we do not want to to tune the analysis too much on this single resonance hypothesis. We want to keep the search generic, and we are just using this resonance model to guide us toward what is the general region of the calorimeter we should use for the analysis. We note that in the future it may be possible to increase our sensitivity slightly by going to a sliding cut, assuming the systematics are well understood, but at this point we can only recommend a single cut at $|\eta| < 1.3$.

LowEdge	BinWidth	Xsec	Error
176.208	25.35	6230.54	165.606
201.558	27.709	3520.52	103.593
229.267	30.202	1745.16	55.7613
259.469	32.837	946.724	30.1296
292.306	35.623	530.676	17.8358
327.929	38.564	308.194	9.52522
366.493	41.671	182.429	5.0275
408.164	44.948	106.265	2.8561
453.112	48.408	62.3932	2.10753
501.52	52.056	34.9452	0.904957
553.576	55.905	21.6299	0.533114
609.481	59.961	13.2723	0.407858
669.442	64.237	8.04574	0.208832
733.679	68.743	4.65967	0.128876
802.422	73.489	2.82766	0.0745046
875.911	78.489	1.8693	0.0443728
954.4	83.75	1.22003	0.0366235
1038.15	89.2999	0.735903	0.015389
1127.45	95.13	0.464107	0.0090216
1222.58	101.26	0.311264	0.00710543
1323.84	107.73	0.194982	0.00415676
1431.57	114.52	0.128931	0.00296396
1546.09	121.67	0.0816044	0.00205769
1667.76	129.19	0.0498288	0.0012685
1796.95	137.08	0.0338663	0.000859646
1934.03	145.4	0.0212371	0.000563268
2079.43	154.12	0.0135589	0.000342808
2233.55	163.31	0.0087951	0.000226432
2396.86	172.95	0.0055904	0.000147664
2569.81	183.08	0.00356572	9.98124e-05
2752.89	193.72	0.00219261	6.1252e-05
2946.61	204.91	0.00138634	3.94669e-05
3151.52	216.64	0.000892097	2.60676e-05
3368.16	228.98	0.000551004	1.68574e-05
3597.14	241.93	0.000326153	9.57556e-06
3839.07	255.52	0.000197919	5.2081e-06
4094.59	269.79	0.000116352	2.97595e-06
4364.38	284.77	6.71595e-05	1.64709e-06
4649.15	300.49	3.97971e-05	1.02349e-06
4949.64	317	2.27664e-05	5.99011e-07
5266.64	334.3	1.15444e-05	2.72508e-07
5600.94	352.47	6.63961e-06	1.86045e-07
5953.41	371.53	3.00695e-06	7.99895e-08
6324.94	391.53	1.38764e-06	3.70904e-08
6716.47	412.49	6.23733e-07	1.77658e-08
7128.96	434.04	2.89385e-07	8.76233e-09

Table 4.3: QCD Cross Section: For each bin of dijet mass, the lower bin edge, bin width, cross section and statistical error are shown for GenJets. This table requires each leading jet have $|\eta| < 1.0$.

LowEdge	BinWidth	Xsec	Error
176.208	25.35	1135.88	39.7205
201.558	27.709	607.806	19.1209
229.267	30.202	374.299	11.5163
259.469	32.837	238.055	8.26914
292.306	35.623	135.6	3.74459
327.929	38.564	82.6966	2.81829
366.493	41.671	47.4103	1.25388
408.164	44.948	29.2787	0.806616
453.112	48.408	19.6777	0.529353
501.52	52.056	11.2761	0.299
553.576	55.905	6.84988	0.187001
609.481	59.961	4.42505	0.122353
669.442	64.237	2.75349	0.0706298
733.679	68.743	1.74294	0.041263
802.422	73.489	1.13647	0.0254853
875.911	78.489	0.727467	0.0154331
954.4	83.75	0.485326	0.0104687
1038.15	89.2999	0.304899	0.00667743
1127.45	95.13	0.207286	0.00451574
1222.58	101.26	0.130636	0.00315487
1323.84	107.73	0.0863515	0.00221818
1431.57	114.52	0.0557825	0.00144018
1546.09	121.67	0.0368266	0.000947805
1667.76	129.19	0.0240253	0.000634211
1796.95	137.08	0.0151729	0.00038339
1934.03	145.4	0.0101689	0.000270564
2079.43	154.12	0.0063538	0.000167948
2233.55	163.31	0.00421239	0.000114512
2396.86	172.95	0.00269539	7.54368e-05
2569.81	183.08	0.00165136	4.76605e-05
2752.89	193.72	0.00103095	3.07408e-05
2946.61	204.91	0.0006948	2.10842e-05
3151.52	216.64	0.000410783	1.18865e-05
3368.16	228.98	0.000258583	7.3132e-06
3597.14	241.93	0.000149933	3.27718e-06
3839.07	255.52	8.84026e-05	2.12847e-06
4094.59	269.79	5.2033e-05	1.25522e-06
4364.38	284.77	3.2433e-05	8.70349e-07
4649.15	300.49	1.73912e-05	4.55174e-07
4949.64	317	9.29762e-06	2.16129e-07
5266.64	334.3	4.84632e-06	1.22788e-07
5600.94	352.47	2.2938e-06	6.20523e-08
5953.41	371.53	1.09457e-06	3.05291e-08
6324.94	391.53	4.50902e-07	1.31544e-08
6716.47	412.49	2.06498e-07	6.19593e-09
7128.96	434.04	8.26654e-08	3.29172e-09

Table 4.4: QCD Cross Section: For each bin of dijet mass, the lower bin edge, bin width, cross section and statistical error are shown for CaloJets. This table requires each leading jet have $|\eta| < 1.0$.

LowEdge	BinWidth	Xsec	Error
176.208	25.35	7095.88	234.455
201.558	27.709	3628.44	110.304
229.267	30.202	2057.76	66.9814
259.469	32.837	1013.92	35.7433
292.306	35.623	575.302	20.2388
327.929	38.564	325.685	10.0813
366.493	41.671	197.084	7.07505
408.164	44.948	110.164	3.31332
453.112	48.408	67.7506	2.36817
501.52	52.056	38.1676	1.03521
553.576	55.905	22.9012	0.635682
609.481	59.961	14.2555	0.398322
669.442	64.237	8.35325	0.230471
733.679	68.743	5.05407	0.140335
802.422	73.489	3.12702	0.0885729
875.911	78.489	1.90581	0.0486402
954.4	83.75	1.23545	0.0297051
1038.15	89.2999	0.763632	0.0166184
1127.45	95.13	0.505249	0.0109348
1222.58	101.26	0.311324	0.00688451
1323.84	107.73	0.206268	0.00476263
1431.57	114.52	0.127736	0.00300978
1546.09	121.67	0.0856802	0.00219033
1667.76	129.19	0.053064	0.00136085
1796.95	137.08	0.0350776	0.000909819
1934.03	145.4	0.0216121	0.000568529
2079.43	154.12	0.0140688	0.000374959
2233.55	163.31	0.00921438	0.000252592
2396.86	172.95	0.0056144	0.000149737
2569.81	183.08	0.00376091	0.000103029
2752.89	193.72	0.00226809	6.45224e-05
2946.61	204.91	0.00141281	4.08883e-05
3151.52	216.64	0.000893011	2.66554e-05
3368.16	228.98	0.000575467	1.82693e-05
3597.14	241.93	0.000333011	9.67375e-06
3839.07	255.52	0.0002008	6.05387e-06
4094.59	269.79	0.000120327	2.73859e-06
4364.38	284.77	6.74338e-05	1.68441e-06
4649.15	300.49	3.94103e-05	9.6517e-07
4949.64	317	2.394e-05	7.11255e-07
5266.64	334.3	1.21597e-05	3.2555e-07
5600.94	352.47	6.73573e-06	1.70832e-07
5953.41	371.53	3.07937e-06	8.5573e-08
6324.94	391.53	1.49843e-06	4.52914e-08
6716.47	412.49	6.27062e-07	1.82275e-08
7128.96	434.04	2.86332e-07	9.27392e-09

Table 4.5: QCD Cross Section: For each bin of dijet mass, the lower bin edge, bin width, cross section and statistical error are shown for corrected CaloJets. This table requires each leading jet have $|\eta| < 1.0$.

4.4 Inclusive Jet p_T from QCD and Contact Interactions

Inclusive jet p_T is a QCD measurement that is sensitive to new physics. For the cross section in Fig. 4.14 we count all jets inside a p_T bin and η interval and divide by bin width and luminosity. CaloJets are shifted to lower p_T than GenJets. Corrected CaloJets agree fairly well with GenJets. Fig. 4.14 shows the ratio of the jet rates for corrected CaloJets and GenJets, the "resolution smearing", which is close to 1 at high p_T. A simple extra correction for this resolution smearing in real data would be to divide the observed jet rate by this Monte Carlo ratio.

Fig. 4.15 shows that CI's [27] create a large rate at high p_T, and discovery is possible with only 10 pb^{-1}. Errors are expected to be dominated by jet energy scale uncertainty ($\sim$10%) in early running (10 pb^{-1}). Fig. 4.15 shows a contact scale $\Lambda^+ = 3$ TeV would be observable despite jet energy errors for $p_T > 1$ TeV. Statistics and PDFs give smaller uncertainties. For 10 pb^{-1} we can discover a CI beyond the Tevatron exclusion of $\Lambda^+ < 2.7$ TeV [46].

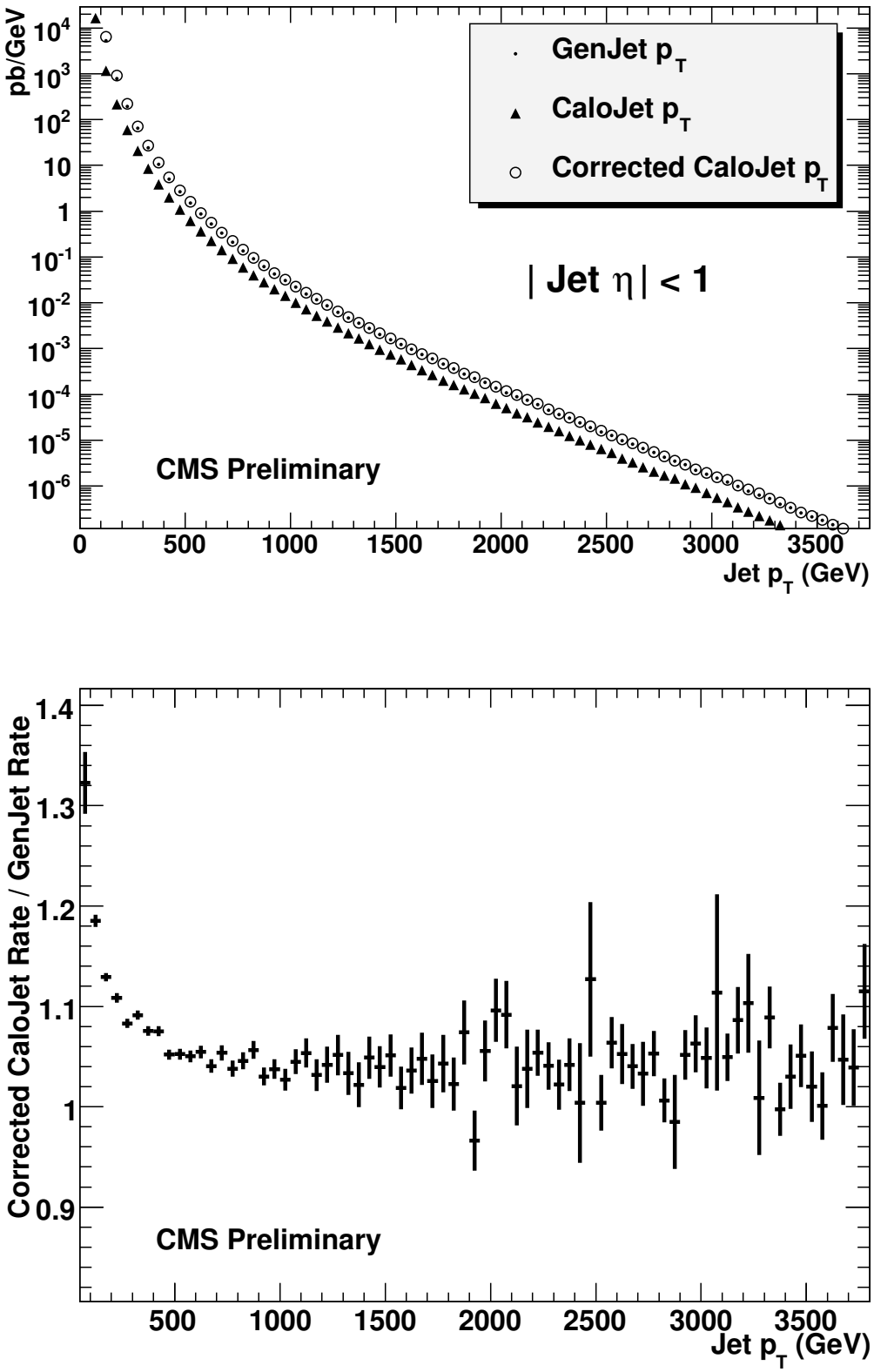


Figure 4.14: Top) Inclusive jet p_T spectra of the QCD multijets sample. Bottom) The ratio of the rates of corrected CaloJets and GenJets.

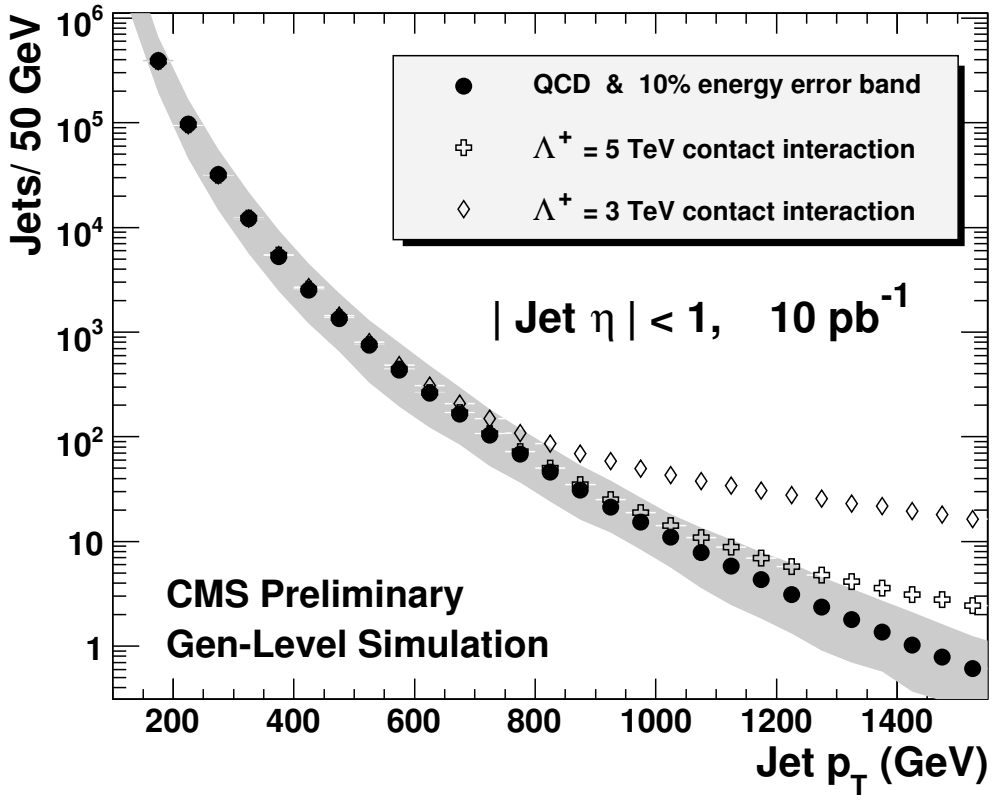


Figure 4.15: CI's and systematic uncertainties in jet cross section for 10 pb^{-1}.

4.5 Conclusions

We have repeated the basic plots in the dijet resonance analysis of Physics TDR II using CMSSW 1_2_0, and we have compared some basic distributions between ORCA and CMSSW 1_2_0. We find reasonable levels of agreement.

Figure 4.6 shows the differential cross section versus dijet mass, where both leading jets have $|\eta| < 1$, and the mass bins have a width roughly equal to the dijet mass reconstructed width. Considering first the QCD processes, the cross section for corrected jets agrees with the QCD prediction from generated jets. To determine the background shape either the Monte Carlo prediction or a parameterized fit to the data can be used. Fig. 4.10 shows a simulation of narrow dijet resonances with a q^* production cross section. This is compared to the statistical uncertainties in the QCD prediction, including trigger prescaling. This comparison shows that with an integrated luminosity of 100 pb^{-1} a q^* dijet resonance with a mass of 2 TeV would produce a convincing signal above the statistical uncertainties from the QCD background. For comparison, a Tevatron search has excluded q^* dijet resonances with mass, M, below 0.87 TeV [49]. The heaviest dijet resonances that CMS can discover (at five standard deviations) with 100 pb^{-1} of integrated luminosity, using this search technique and including the expected systematic uncertainties [50, 35], are: 2.5 TeV for q^*, 2.2 TeV for Axigluons [51] or Colorons [52], 2.0 TeV for E_6 Diquarks [53], and 1.5 TeV for Color Octet Technirhos [54]. Studies of the jet η cut have concluded that the optimal sensitivity to new physics is achieved with $|\eta| < 1.3$ for a 2 TeV spin 1 dijet resonance decaying to $q\bar{q}$ [55].

In conclusion, CMS plans to use measurements of rate as a function of jet p_T and dijet mass to search for new physics in the data sample collected during early LHC running. With integrated luminosity samples in the range 10-100 pb^{-1}, CMS will be sensitive to CI's and dijet resonances beyond those currently excluded by the Tevatron.

Chapter 5

Monojet at 14 TeV

5.1 Introduction

Here we describe the procedures which will be used for the search for evidence of large ADD extra dimensions in the MET plus a single jet (MET +1 jet) channel, using the CMS detector. The assumptions concerning the detector performance and the simulated samples of events are those corresponding to early stages of the data taking, for an integrated luminosity of up to $100\,\mathrm{pb}^{-1}$ and 7 TeV per beam.

5.1.1 Previous studies and current limits

The most recent result of direct Newtonian potential measurements reports $r_C < 55\,\mu\mathrm{m}$ as a limit for the $n = 2$ case [56], while the value of M_D depends on the compactification mechanism.
Astrophysical limits are derived mostly from photon annihilation into Kaluza-Klein gravitons in supernova cores. One of the strongest bound comes from the graviton emission from SN1987A. Even though there are several uncertainties such as the lack of precise knowledge of temperatures in the core, a conservative choice of the core parameters has led to a lower limit $M_D = 1.6\,\mathrm{TeV}$ for $n = 3$ [57].
The production of gravitons in hadron collisions can be probed via events where the final-state hadronic jet or photon recoils against the graviton, which is not detected and results in missing energy. Searches in both the jet+MET and the γ+MET have been performed by CDF ($1.1\,\mathrm{fb}^{-1}$ and $2.0\,\mathrm{fb}^{-1}$ of data respectively, [58]) and DØ ([59, 60]), where a good agreement with SM expectations has been observed. Depending on n, the sensitivity to large extra dimensions at the Tevatron is comparable to (or better than) the one from LEP experiments [61]. The lower limits on M_D as a function of the number of extra dimensions in the model

	LEP	DØ		CDF		
n	γ+MET	jet+MET	γ+MET	jet+MET	γ+MET	combined
2	1.600	0.99	0.921	1.310	1.080	1.400
3	1.200	0.80	0.877	1.080	1.000	1.150
4	0.940	0.73	0.848	0.980	0.970	1.040
5	0.770	0.66	0.821	0.910	0.930	0.980
6	0.660	0.65	0.810	0.880	0.900	0.940

Table 5.1: 95% confidence limits on M_D from DØ [59, 60], CDF [58], and LEP [61] in the γ+MET and jet+MET signatures (Table from Ref. [62]). Values are in TeV.

are reported in Table 5.1.

It is worth to notice that astrophysical and direct measurements are sensitive to the infra-red KK spectrum of gravitons, while collider experiments reach the ultraviolet part. The model can always be modified by the introduction of a cut-off parameter μ to fix the lower bound of the graviton mass spectrum [10]. If $\mu > 50\,\text{MeV}$, astrophysical signatures like supernovae photon emission would not be observable anymore and the lightest graviton excitation will be accessible only at colliders.
Expectations for the ATLAS experiment for the maximum reach in M_D for low luminosity running and $30\,\text{fb}^{-1}$ are 7.7, 6.2, 5.2 TeV corresponding to $n = 2$, 3 and 4. For high luminosity running and $100\,\text{fb}^{-1}$ these results can be extended up to 9.1, 7.0 and 6.0 TeV respectively [63]. This chapter is an estimate of these reaches for the early stage of CMS data-taking, namely $100\,\text{pb}^{-1}$.
In Sec. 5.2 the software and algorithms used for the signal and background generation and simulation are presented. This is followed by a discussion on the trigger (Sec. 5.3), selection cuts and their optimization in Sec. 5.4. In Sec. 5.5, procedures for estimating the most relevant backgrounds using data-driven methods are outlined, and in Sec. 5.6 the impact of the major systematic effects on signal and analysis is discussed. Finally a discussion on the discovery potential is given in (Sec. 5.7).

5.2 Samples generation and reconstruction at 14 TeV

The new physics process under study is simple and has the following relevant features:

- A high-transverse momentum (p_T) ($> 300 \div 400\,\text{GeV}$) jet in the central region of the detector ($|\eta| < 1.7$), plus possible less energetic jets due to

initial-final state radiation. No other energetic jets are produced, neither central nor in the very forward direction, creating a typical 'mono-jet' signature;

- a large MET (same order of magnitude of p_T(jet)) which is almost back-to-back to the leading jet in the transverse plane.

Several SM processes are prone to mimic such a topology. In this study the most relevant ones are addressed:

- Z+jets, with the Z decaying into two neutrinos. This channel leads to invisible energy recoiling against one (or more) jets and is described by the same signature as the signal. The contribution from this "irreducible" background needs to be estimated and subtracted;
- W+jets followed by a leptonic W decay. When the lepton (e, μ, τ) is not reconstructed by the detector, the signal is quite similar to the Z+jets one.
- QCD dijets, when one or more jets are mismeasured and/or a significant amount of MET is produced in the hadron decays (high-p_T neutrinos and high-p_T particles outside the detector acceptance);
- top quark production, both $t\bar{t}$ and single-top, especially for events with few or collimated jets or events where leptons are not tagged.

In addition, a fraction of $\gamma + jet$ production might have an undetected photon and sneak in the signal region. The strategy for the analysis adopted in the work is that of a standard counting experiment, i.e. no special event kinematics is used. Once the SM background has been properly modeled, the MET+1 jet signal should be visible as an excess on the MET spectrum on top of what expected from the SM. The excess is evaluated and used to estimate the discovery potential of the experiment. In such a counting experiment it is very important to be able to validate the estimation of the SM contributions: this will be accomplished by using control regions where the SM background is expected to be large and the signal small. The full hermeticity of the apparatus makes CMS an ideal detector to improve the missing energy measurements, while the excellent performance of tracker and electromagnetic calorimeter are instrumental to reject the leptons contamination in jets.

In the following, details about the signal and background production are given.

5.2.1 Signal production

The ADD-model signal has been produced with the SHERPA generator [64] version 1.0.11 for the three different production modes (see Fig. 5.1). SHERPA is a multi-purpose event generation framework that is capable of simulating lepton-lepton,

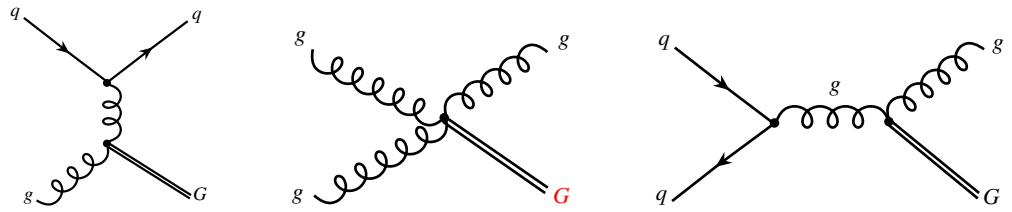


Figure 5.1: Three different production modes for real gravitons G in the ADD extra dimensions model.

lepton-photon, photon-photon and fully hadronic collisions, such as proton-proton reactions.

In order to explore the sensitivity in a wide energy spectrum, 18 different samples with M_D ranging from 2 to 7 TeV and $n = 2$, 3, 4 have been produced. Since the cross sections in [10] are computed in an effective theory approach, the cut prescription $\sqrt{\hat{s}} < M_D$ has been introduced in the generation step. A p_T cut-off on the parton recoiling against the graviton was also introduced, by requiring $\hat{p}_T > 200$ GeV. Parton density functions from the CTEQ5L model were used. With these production parameters, signal cross sections (LO) were evaluated by SHERPA for the 18 samples. They are reported along with their errors from the generation stage (phase space sampling) in Tab. 5.2 (pb).

M_D	$n = 2$	$n = 3$	$n = 4$
2 TeV	49.246 ± 0.056	29.147 ± 0.033	18.914 ± 0.022
3 TeV	12.462 ± 0.015	6.392 ± 0.007	3.874 ± 0.005
4 TeV	4.253 ± 0.005	1.844 ± 0.002	0.998 ± 0.001
5 TeV	1.783 ± 0.002	0.650 ± 0.001	0.308 ± 0.001
6 TeV	0.862 ± 0.001	0.266 ± 0.001	0.109 ± 0.001
7 TeV	0.466 ± 0.001	0.124 ± 0.001	0.044 ± 0.001

Table 5.2: ADD cross sections (and errors from generation stage) as evaluated by the SHERPA program with the generation parameters detailed in the text. All values are in pb.

For each subsample with a given (M_D, n) a statistics of about 10^5 events was produced. As a preliminary check at generator level, a few variables are compared in Fig. 5.2 for $n = 2$ and variable M_D. Although the graviton often gets a larger boost when moving to higher energies, the transverse momentum and jet multiplicity are found not to show any striking dependence neither on n nor on M_D and the topologies are fairly similar. All the ADD signals have most of gravitons with $p_T > 200$ GeV (thus the sum of jets p_T of the same order), one or more jets from initial or final gluon radiation accompanying the mono-jet from

hard interaction, and the angle in the transverse plane between the graviton and the jets (here the leading p_T one is displayed) is greater than $\pi/2$.

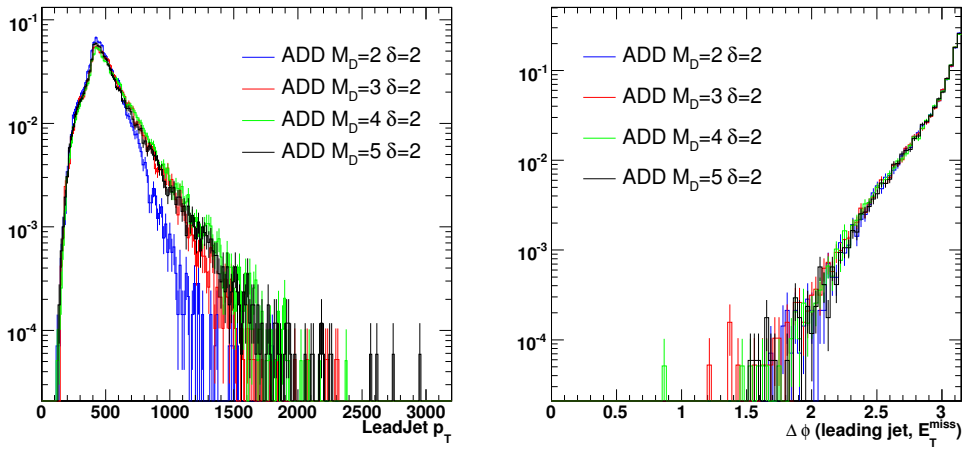


Figure 5.2: The p_T of generated graviton (left) and the azimuthal angular difference between the leading jet and the MET (right) at generator level, for $\delta = 2$ (where δ is the number of extra dimension, referred as n in the text) and different M_D.

Each of these features will play a key role in isolating the signal from the different background sources. No different selections are required for the analysis of the different signal samples.

5.2.2 Background production

All the boson+jets samples have been produced with ALPGEN 2.12 [65] after slicing the outgoing parton in $\hat{p}_T$ bins. Since the availability of high-energy jets is essential, additional background samples with $0 < \hat{p}_T < 3200$ GeV for W+jets and $0 < \hat{p}_T < 3200$ GeV for Z+jets were generated.

ALPGEN has been used also for the inclusive $t\bar{t}$ production, and the large multijet QCD backgrounds were generated by PYTHIA 6.409 [66], using both the JetMet (for the analysis in the signal region) and the Muon (for the QCD estimation in a W+jet sample, Sec. 5.5) Primary Datasets (PDs). Here PD is referred to a "chunck" of data skimmed according to a physics trigger and apportioned to a Tier1. The idea behind the PD concept is to have flexibility on reconstructing and transferring the data according to the analysis needs.

All the background samples were produced during the 'Computing, Software and Analysis challenge' performed in the Fall 2007 (CSA07) [67] with the version 1_4_6 of CMSSW and 1_6_7 for reconstruction.

SM Channel	Notes	σ·BR (pb)	Evs. ($\times 10^3$)
Z+jets	$0 < \hat{p}_T < 3200\,\text{GeV}$	2792	567.5
$Z(\to \nu\nu)$+jets	$0 < \hat{p}_T < 3200\,\text{GeV}$	239.1	265.1
QCD dijets PDJetMet	from Gumbo	$2.531 \cdot 10^6$	6322
QCD dijets PDMuon	from Gumbo	$2.531 \cdot 10^6$	152.6
$W(\to e\nu)$+jets	$0 < \hat{p}_T < 3200\,\text{GeV}$	14.767	579.1
$W(\to \mu\nu)$+jets	$0 < \hat{p}_T < 3200\,\text{GeV}$	14.767	579.1
$W(\to \tau\nu)$+jets	$0 < \hat{p}_T < 3200\,\text{GeV}$	14.767	579.1
$t\bar{t}$	from Chowder	$447 \times 1.85^{(*)}$	19,696
single-t (incl. e)	inclusive e channel	27	52.81
single-t (incl. μ)	inclusive μ channel	27	48.24

(*) k-factor

Table 5.3: Overview of the background statistics exploited in the analysis. All samples are produced in the CSA07 and cross sections are from Ref. [69].

The so-called 'Gumbo' part of CSA07 inclusive sample ('soup') was used for non-diffractive dijet production and minimum-bias background events, and the inclusive $t\bar{t}$ production was sorted among the 'Chowder' component of the soup. The Gumbo sample is known to be affected by a production bug occurred at the simulation step, which resulted in an incorrect decay of some heavy particles as c, b, τ and gave rise to an excess of large MET. This effect is predominantly in the high-MET region and, in the present study, is induced basically by b-quarks. Since this wrong manipulation affects decay chains where there is a decay length bigger than 10 cm (essentially from boosted particles with heavy quarks), the bug has been worked around at generator level, by requiring a decay length between two vertexes bigger than 10 cm [68]. This introduce a bias which removes about 1.6% of total multijet events.
Inclusive single-top production was generated with the `MC@NLO` generator [70].
The generated events per process are given in Table 6.2, along with the cross sections scaled with the branching ratio and generator cuts. When a binning in p_T has been used, the number of events is integrated over the full range.

Studies are currently ongoing to address the machine induced background ('beam halo') and the effect of cosmic rays. On the base of the experience acquired at Tevatron, we foresee that the use of pointing tracks and out-of-time energy deposits can be efficient on rejecting these contributions.

The CSA07 production targeted for $1\,\text{fb}^{-1}$ of statistics and therefore the data samples have been reconstructed with calibration and alignment constants based on the understanding of the detector obtained with $100\,\text{pb}^{-1}$ of data. Miscalibrations were not taken into account at the trigger level.

ADD signal samples have been generated using Fast Simulation with CMSSW version 1_6_9. The reconstruction and the analysis step which will be discussed in next sections have been entirely performed using the Physics Analysis Toolkit (PAT) [71].

5.3 Trigger optimization studies at 14 TeV

A detailed description of the trigger system in CMS can be found in [72]. This study optimizes a trigger stream based on the following two quantities:

- the sum of module of transverse momenta $p_T(j)$ of all jets above a threshold p_T^0:
$$HT = \sum_{p_T(j)>p_T^0} |\vec{p}_T(j)| \,; \tag{5.1}$$
- the module of the vectorial sum of jets transverse momenta above a threshold p_T^0:
$$MHT = \left| \sum_{p_T(j)>p_T^0} \vec{p}_T(j) \right| . \tag{5.2}$$

We considered the trigger setup adopted in SUperSYmmetry (SUSY) searches, which shares many of the backgrounds with the present analysis (in particular the QCD dijet events).
The trigger envisioned for SUSY consists in HT $>$ 200 GeV and a threshold p_T^0 = 10 GeV at first level trigger (L1), followed by HT $>$ 250 GeV, MHT $>$ 100 GeV and p_T^0 = 20 GeV at High Level Trigger (HLT). It is reasonable to assume that such a HT + MHT trigger will be viable for the analysis under study, especially in the early stages when MHT is believed to be a quantity more reliable than MET. The QCD rate at HLT from this trigger is expected to amount to 3 Hz for an instantaneous luminosity of 10^{+31}; the efficiency for the signal under study was found to be $(65.9 \pm 1.2)\%$. Although the efficiency is not particularly high, it has been proven that 100% of the signal events passing the analysis selections will pass also the trigger selections. Hence, we consider the trigger suitable for this analysis and we will use it in the following.
The trigger has been object of an threshold variation study, whose results are shown in Figures 6.1. The signal efficiency at HLT as a function of the p_T^0 is reported on the left-hand for different sets of (HT, MHT) thresholds. The study demonstrates that it is eventually possible to significantly increase the signal efficiency at HLT by changing the HT threshold to HT $>$ 200 GeV, while it is insensitive to variations on the MHT threshold. For a HT $>$ 200 GeV, the

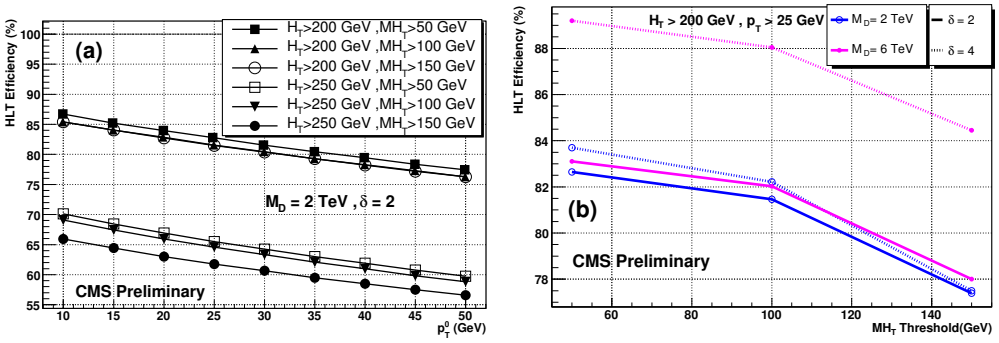


Figure 5.3: *a*) Efficiencies for the HT + MHT trigger for a benchmark signal. Here the MHT, HT thresholds are fixed and the p_T^0 cuts are varied. *b*) Efficiencies for different benchmark signals, as a function of MHT. In the plots δ is the number of extra dimension, referred as n in the text.

background rate would increase above the 10 Hz for $p_T^0 < 20$ GeV. So, an alternative HLT stream would be HT > 200 GeV , MHT > 100 GeV and $p_T^0 = 20$ GeV which can increase the signal efficiency up to $\sim 84\%$ (Fig. 5.3*b*)) maintaining the background rate below the 10 Hz. Although this HLT trigger would be optimal in terms of number of signal events kept for further analysis, it will not be supported. This is not a problem in terms of the foreseen analysis cuts, as discussed before, but it could be an optional solution in case of different analysis approaches.

5.4 Signal and background analysis at 14 TeV

The cut-based analysis procedure is described in this section. After a quick reminder about how the relevant objects are reconstructed, the selection variables are presented along with their motivation. Threshold cuts on the quantities that have been demonstrated to be most sensitive to SM suppression have been optimized, with the aim of maximizing the signal over background ratio in the relevant kinematic region.

5.4.1 Definitions of variables

In all this work, standard reconstruction algorithms and methods as provided by the PAT are exploited. PAT objects that are relevant for the signal selection and background estimation are muons, photons, jets and MET. In addition, tracks with $p_T > 15$ GeV are reconstructed in a framework independent from PAT.
Jets were obtained using an iterative cone (IC) algorithm) with $\Delta R = 0.5$ [32]. Since a cleaning of jets from electrons will be performed in the off-line analysis,

no electron-jet separation was applied in reconstructing jet objects. In the study raw jets were used without any Monte Carlo correction, which is a more suitable approach for the LHC very early stage.
The MET was calculated from the vectorial sum of the transverse energy of all calorimeter towers with E >0.8 GeV and E_T >0.5 GeV and also in this case the raw quantity was used.

5.4.2 Cascade selection and efficiencies

In order to reduce the impact of hard gluon radiation in the selection, the analysis exploited a collection of signal and background samples having jets with p_T >40 GeV and $|\eta| < 3$. A long tail in the missing energy spectrum is a distinctive signal feature, hence a cut MET > 400 GeV was imposed early at the preselection level. These simple cuts have already an important effect on the boson+jets (where the next-to-leading jets are usually soft), single-top (distinguished by high-$|\eta|$ jets) and QCD (whose contribution dominates at moderate missing energy). Nevertheless, the long missing energy tail from multijet could be the most (reducible) important background and subsequent cuts are designed to reduce it.

Kinematic cuts and QCD cleaning at 14 TeV

Jets are ordered with decreasing p_T (leading, secondary, tertiary etc.). The leading jet is required to have $p_T(\text{jet}\,1) > 350$ GeV and $|\eta(\text{jet}\,1)| < 1.7$. While decreasing the amount of $Z(\nu\nu)$+jets, QCD and $t\bar{t}$ only by a $20 \div 25\%$ of the original anlyzed samples, single-top processes become negligible already at this level (less than 1 event in the signal region).
Fig. 5.4 displays the distribution of the leading jet p_T for all the generated backgrounds except single-top. Since most of the dominant SM processes have the same p_T shape as the signal, the reductions of signal acceptances are comparable. The selection does not allow to enhance the signal/background ratio, but is meant to define a kinematic region where further cuts can take place.

Not surprisingly, the number of selected jet is one of the most effective variables to disentangle the signal (with only one parton outgoing from the hard interaction) from the QCD multijet and $t\bar{t}$ background. As explained in Sec. 5.2, the signal jet multiplicity is peaked around 2 and steeply decreasing for higher number of jets (Fig. 5.5). A veto against events with more than two jets has been found as an optimal choice: the signal efficiency reduces not more than 20% while the rejection factors for QCD and $t\bar{t}$ are 2.5 and 3.4 respectively. SHERPA generator uses the CKKW (Catani, Khun, Krauss, Webber) approach for the matching of Matrix Element to Parton Shower. There are many different approaches of matching on the market. Without entering in the details of the

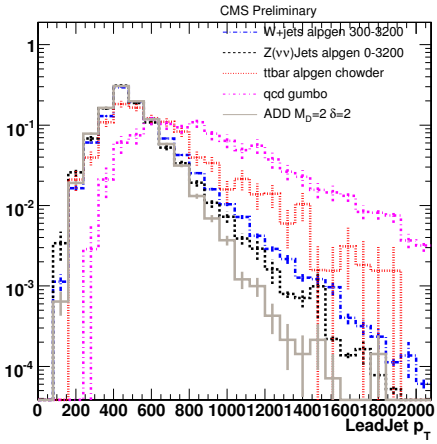


Figure 5.4: Transverse momentum of leading jet for ADD signal ($M_D = 2\,\text{TeV}$, $n = 2$) and relevant backgrounds. Event numbers are normalized to $100\,\text{pb}^{-1}$. In the plot δ is the number of extra dimension, referred as n in the text.

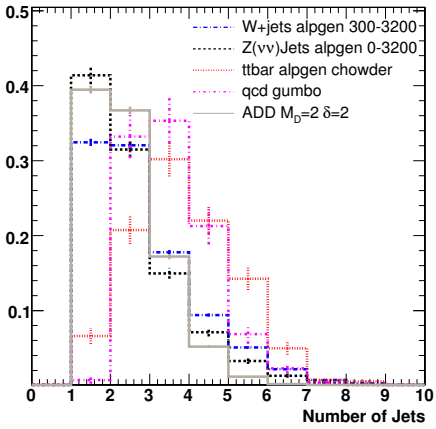


Figure 5.5: Jet number distribution for signal and relevant backgrounds. A veto against three or more jet events turns out to be the optimal cut to minimize the QCD contribution. In the plots δ is the number of extra dimension, referred as n in the text.

matching problem, it is clear that any estimation of the jet multiplicity will be dependent on the matching approach. On real data, it will be therefore necessary to variate the cut on the multiplicity and verify that the result obtained does not depend on the multiplicity cut.

Topological Constraints

In order to reduce the background further, it is worth to profit of the back-to-back topology of the signal event.
When events with an angular difference in the transverse plane $\Delta\phi$(jet 1,MET) > 2.8 are selected, a fraction of the processes where missing energy does not recoil with jets is rejected. The rejection is largely enhanced by requiring $\Delta\phi$(jet 2,MET) > 0.5, that further reduces $t\bar{t}$ by a factor 2.1 and QCD by a factor 42. Moreover, the angle $\Delta\phi$(jet 2,MET) (Fig. 5.6) has proved to be an effective variable to discriminate against W+jets process. The effect can be ascribed to the fact that in such processes the secondary jet comes often from the same object that causes MET (as τ-jets or electron cluster misidentified as jets). On the other hand, the secondary jet in signal is often produced from primary jet gluon radiation, hence it is almost collinear with the primary jet.
Although the irreducible $Z(\nu\nu)$+jets background is reduced by less than 12%, most of the multijet contamination is successfully suppressed and the signal to background ratio is effectively improved.

Charged leptons cleaning cuts

After the jet selections and the topological cuts, a significant fraction of background events can still mimic the signal when leptons fake jets. Therefore, it is important to clean the events from both isolated lepton contaminations (as those from the $W(l\nu)$+jets channel) and electrons and photons reconstructed as jets. Two variables may be used to accomplish the tasks:

- Jet Electromagnetic Fraction (JEMF), defined as the fraction of jet energy collected by the electromagnetic calorimeter over the total energy of the jet in both the hadronic and electromagnetic calorimeter. Electrons and photons reconstructed as jets are characterized by a JEMF close to 1: therefore, it is possible to reject high-energy electrons in the final state imposing that the JEMF of the two most energetic jets is below 0.9. A cut on both primary and secondary jets was successfully in removing a significant fraction of background events from QCD, $t\bar{t}$ and $W(e\nu)$+jets;
- Track Isolation Veto (TIV). A hollow cone $0.02 < \Delta R < 0.35$ is drawn around a track with $p_T > 15\,\text{GeV}$. The sum of tracks inside the cone is performed and the TIV defined as:

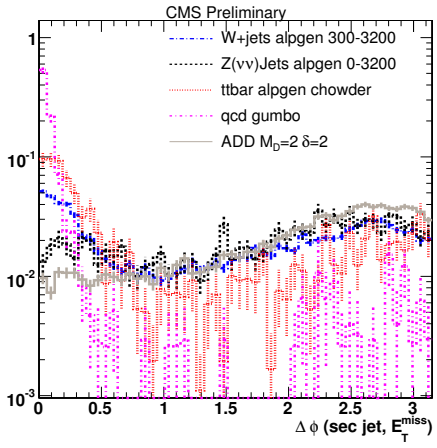


Figure 5.6: Angle in the transverse plane between the MET and the secondary jet. Choosing events with $\Delta\phi(\text{jet}\,2,\text{MET}) > 0.5$ can exclude a large part of QCD, $t\bar{t}$ and W+jet backgrounds. In the plot δ is the number of extra dimension, referred as n in the text.

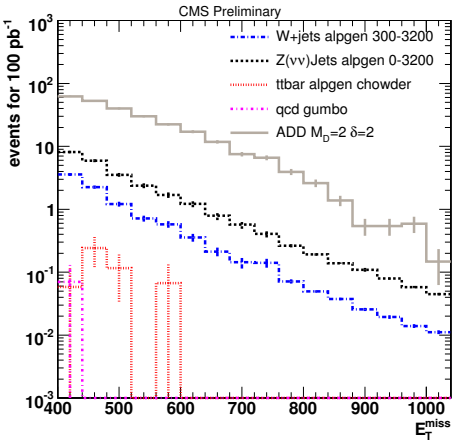


Figure 5.7: MET distribution after all selections are applied. Histograms are overlaid, not stacked. In the plot δ is the number of extra dimension, referred as n in the text.

$$TIV = \frac{1}{p_T(\mathrm{tk}\,1)} \sum_{R \in \Delta R} p_T^j \quad , \tag{5.3}$$

where $p_T(\mathrm{tk}\,1)$ is the transverse momentum of the leading track. A small value of TIV is typical of well-isolated leptons and can be effectively used to clean the sample. It has been found that requiring TIV < 0.1 results in a reduction of $W(\mu\nu)$+jets, $W(\tau\nu)$+jets and $t\bar{t}$ by a factor 15.3, 2.3 and 7.1 respectively.

The combined usage of these two variable is commonly referred to as ''Indirect Lepton Veto" [73].

Missing energy distributions for signal and background are shown in Fig. 5.7. The signal shows up as an excess of events in addition to the $Z(\nu\nu)$+jets dominant background. In Sec. 5.7 the reach limits for this selection will be discussed.

5.4.3 Summary of selection cuts

The present analysis relies on a simple series of cuts applied to uncorrected MET and uncorrected jets. The set of cuts presented above are summarized:

1. *Trigger* : Customized trigger with HT $>$ 200 GeV at L1 and HT $>$ 200 GeV , MHT $>$ 100 GeV at HLT;
2. *Preselection* : MET $>$ 400 GeV , $p_T(\mathrm{jet}) > 40\,\mathrm{GeV}$, $|\eta(\mathrm{jet})| < 3$;
3. *Kinematic cuts* : $p_T(\mathrm{jet}\,1) > 350\,\mathrm{GeV}$, $|\eta(\mathrm{jet}\,1)| < 1.7$, $|\eta(\mathrm{jet}\,2)| < 3$;
4. *Jet cleaning* : veto against 3 or more jets in the event;
5. *Angular cuts* : $\Delta\phi(\mathrm{jet}\,1,\mathrm{MET}) > 2.8$, $\Delta\phi(\mathrm{jet}\,2,\mathrm{MET}) > 0.5$;
6. *Indirect Lepton Veto* : JEMF < 0.9, TIV < 0.1

The effect of each group of cuts is reported in Tab. 5.4 for all the SM processes and in Tab. 5.5 for some benchmark signals. Quoted are the relative efficiency of the cuts along with the absolute expected number of survived events (assuming $100\,\mathrm{pb}^{-1}$ of data). Absolute efficiencies and statistical errors can be computed from these numbers in a straightforward way.

Table 5.5 confirms that both the kinematic and geometric features of the reconstructed ADD signal are uniform for M_D from 2 to 6 TeV and n from 2 to 4. Selection efficiencies are consistent and the number of events are only scaled by the cross section.

	$t\bar{t}$	$Z(\nu\nu)$+jets	QCD	$W(e\nu)$+jets
Trigger	3860 (4.88)	1276 (0.457)	492329 (0.252)	1199 (60.8)
MET $>$ 400 GeV	36.6 (0.948)	54.8 (4.29)	17.9 ($3.6 \cdot 10^{-3}$)	19.5 (5.47)
JEMF < 0.9	32.0 (87.5)	52.4 (95.6)	17.2 (96.0)	8.76 (53.0)
TIV < 0.1	12.2 (38.1)	46.3 (88.3)	14.2 (82.8)	4.27 (48.8)
$p_T(\mathrm{j}\,1) > 350\,\mathrm{GeV}$, $\lvert\eta(\mathrm{j}\,1)\rvert < 1.7$	9.8 (79.9)	36.6 (78.9)	11.8 (83.0)	3.27 (77.3)
Number of jets < 3	2.2 (22.6)	28.9 (79.0)	4.63 (39.1)	2.33 (71.4)
$\Delta\phi$(j 1,MET) > 2.8, $\Delta\phi$(j 2,MET) > 0.5	0.48 (21.8)	25.7 (88.9)	< 0.701 (< 1.51)	1.97 (84.2)
	$W(\mu\nu)$+j	$W(\tau\nu)$+j	single-t(e)	single-t(μ)
Trigger	1617 (85.5)	1488 (80.6)	29.8 (1.10)	20.7 (0.767)
MET $>$ 400 GeV	64.7 (15.1)	36.3 (9.21)	0.767 (2.58)	1.18 (5.67)
JEMF < 0.9	60.6 (95.1)	32.0 (88.0)	0.307 (40.0)	1.12 (95.2)
TIV > 0.1	5.87 (9.68)	13.0 (40.7)	0.153 (50.0)	0.336 (30.0)
$p_T(\mathrm{j}\,1) > 350\,\mathrm{GeV}$, $\lvert\eta(\mathrm{j}\,1)\rvert < 1.7$	4.46 (75.8)	9.92 (76.2)	0.102 (66.7)	0.280 (83.3)
Number of jets < 3	2.84 (63.7)	6.93 (69.9)	0.102 (100.0)	0.280 (100.0)
$\Delta\phi$(j 1,MET) > 2.8, $\Delta\phi$(j 2,MET) > 0.5	1.99 (69.8)	5.50 (83.4)	0.0511 (50.0)	0.112 (50.0)

Table 5.4: Number of selected events for each group of cuts in the relevant background samples, normalized to $100\,\mathrm{pb}^{-1}$. In parenthesis the partial efficiencies are quoted in percent.

	$n = 2$		$n = 4$	
$M_D =$	2 TeV	6 TeV	2 TeV	6 TeV
Trigger	3060	54.4	1190	7.98
MET > 400 GeV	691 (23)	12 (22)	245 (21)	3.0 (38)
JEMF < 0.9	659 (95)	11 (96)	232 (95)	2.9 (95)
TIV > 0.1	539 (82)	9.5 (82)	185 (80)	2.2 (76)
$p_T(\text{jet } 1) > 350\,\text{GeV}$, $\lvert\eta(\text{jet } 1)\rvert < 1.7$	343 (64)	6.5 (68)	117 (63)	1.6 (73)
Number of jets < 3	287 (84)	5.4 (83)	98.3 (84)	1.2 (75)
$\Delta\phi(\text{jet } 1,\text{MET}) > 2.8$, $\Delta\phi(\text{jet } 2,\text{MET}) > 0.5$	261 (91)	4.9 (91)	90.1 (92)	1.1 (92)
Total Efficiency (%)	8.1 ± 0.5	8.5 ± 3.8	7.1 ± 0.7	13.2 ± 13.2

Table 5.5: Number of selected events for each group of cuts in four signal subsamples, normalized to $100\,\text{pb}^{-1}$. Uncertainties on efficiencies are statistical only. In parenthesis are the partial efficiencies are quoted in percent.

While an estimation of SM background using data-driven techniques will be performed in the next section, one can anticipate that a significant sensitivity to ADD should be reached for $M_D \sim 4\,\text{TeV}$ for $n = 2$ and $M_D \sim 3\,\text{TeV}$ for $n = 4$. This is already an excellent improvement with respect to the most recent Tevatron results (Tab. 5.1), and could be reached in a very early stage of the LHC data taking.

It is worth to underline that significant achievement could already be obtained with $10\,\text{pb}^{-1}$. In order to compensate for the lack of statistics, the kinematic cuts can be relaxed to MET > 250 GeV and $p_T(\text{jet } 1) > 250\,\text{GeV}$. Although the signal over background ratio is expected to be reduced be a factor 2, a signal could already be observed and exclusion limits around $M_D \sim 2\,\text{TeV}$ for $n = 2$ could be reached.

5.5 Data-driven background estimation at 14 TeV

Since the present analysis is tailored to a very early data-taking, evaluating the background from the early data sample instead of Monte Carlo is important. In the following, procedures are proposed to determine the irreducible background of $Z(\nu\nu)$+jets (here referred also as "invisible Z" background) and $W(e/\mu/\tau\nu)+$ jets. Control regions which are easy to be identified in early data are defined, and the rescaling of the background into the signal region, presented in the previous section, is discussed.

5.5.1 $Z(\nu\nu)$+jets background estimation

In the CMS PTDR [73, 74], the invisible Z background is estimated by use of $Z \to \mu\mu$ production. The basic idea is to remove the leptons from the event and calculate the total missing energy. This allows to model the MET spectrum for $Z(\nu\nu)$+jets after correcting for relative branching fractions, efficiencies and acceptances. Unfortunately few $Z \to \mu\mu$ events satisfying all the search criteria are expected in data samples of less than $1\,\mathrm{fb}^{-1}$; thus different strategies are needed.

In Ref. [75] it has been shown how the Z invisible background can be derived from samples of events containing a high-p_T $W(\to l\nu)$ boson, as well as a photon produced with high-energy jets. The MET spectrum is obtained by removing the identified lepton or photon and correcting for residual differences between these events and invisible Z events. The higher statistics of these samples allows the application of all search criteria. The number of jets in these events are not easily predicted, but to good approximation do not depend upon whether the boson is a Z, W or photon. Compared to using these techniques in SUSY searches, we do not expect any background of the signal in the control regions.

The use of the W+jets with a subsequent $W \to l\nu$, can be used to predict the missing energy spectrum for invisible Z at high-p_T and QCD appears strongly suppressed. The usage of this method is discussed in the following.

Estimate with the $W(\mu\nu)$+ jets channel

A control region with maximal purity in $W(\mu\nu)$+jets events has a quite high statistics at the LHC and one can profit of the excellent performance of the CMS muon system.
The event selection defining the control region has been maintained as close as possible to that for MET+1 jet signal, except for the muon requirement. The strategy proposed consists of different steps:

- the event is triggered by a HLT for a single (not isolated) muon;
- exactly one muon with $p_T > 20\,\mathrm{GeV}$ is required;
- the muon must fulfill the tracker isolation requirements. The muon isolation variable is defined as the the sum of the transverse momenta of the tracks in a cone of $\Delta R = 0.3$ around the muon track position, excluding the muon track: $\mu_{Iso} = \sum_{trk}^{\Delta R} p_T(\mathrm{trk})$. It is considered isolated when $\mu_{Iso} < 1\,\mathrm{GeV}$. Jets with an isolated muon closer than ΔR =0.5 are discarded;
- jets that are closer than a $\Delta R = 0.5$ cone from the muon direction are vetoed;

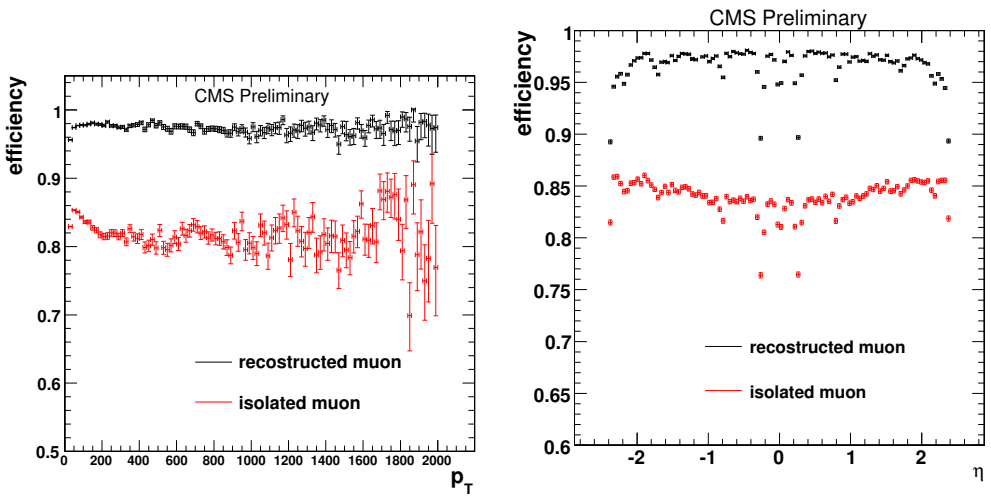


Figure 5.8: Muon selection efficiency as a function of its transverse momentum (left) and pseudorapidity (right). Efficiencies after the trigger and after isolation requirements are displayed and error bars are statistical only.

- the resulting sample undergoes the signal selections, with the isolated muon track excluded from the Track Indirect Veto.

Once the contamination from other SM processes has been controlled, the number of Z background events can be obtained by introducing only the cross section ratio and muon trigger and isolation efficiency. As discussed later, systematic uncertainties on these factors are small and can be controlled in a data-driven way themselves.
Figures 6.5 show the reconstruction and isolation efficiency for the W+jets sample.

Different backgrounds with at least one well-isolated muon and jet are prone to pass the selection. The most significant backgrounds are found to be $W(\tau\nu)$+ jets, $t\bar{t}$ and multijet events. The composition of the control region is reported in Tab. 5.6 for the $W(\mu\nu)$+jets channel and the background at different stages of the event selection. All values are quoted for $100\,\mathrm{pb}^{-1}$.

Since the requirement of an isolated muon undermines the effect of lepton cleaning algorithm, important contaminations from $t\bar{t}$ and $W(\tau\nu)$+jets affect the region and have to be accounted in order to use it for the invisible Z background measurement. Since many uncertainties on QCD events can modify this tail, a data-driven procedure is considered for this source too.

The latter task was accomplished by looking at a kinematic region where the QCD events are dominant. Therefore, the multijet distribution has been acquired

Selection	$W(\mu\nu)$+jets	$W(\tau\nu)$+jets	$t\bar{t}$	QCD
Single isolated μ	294	40.0	7924	394033
Kinematic cuts	36.0	6.32	7.12	0.0961
Jet cleaning	28.7	4.93	3.47	0.0961
Angular cuts	25.9	4.47	2.61	0.0907
Indirect Lepton Veto	20.0	3.32	2.00	0.0906

Table 5.6: Number of events survived selections for the $W(\mu\nu)+$ jets control region. All normalizations refer to $100\,\mathrm{pb}^{-1}$. QCD events are cleaned from the bugged events and rescaled.

with the following steps:

- To enrich the QCD statistics, a set of relaxed cuts were put in place: soft muon with $10 < \mu_{Iso} < 20\,\mathrm{GeV}$ isolation criteria, leading jet p_T and the $\Delta\phi$(jet 2,MET) cut removed. The number of the QCD events was derived for the control region MET > 400 GeV and for a a low-MET region MET < 100 GeV ;
- the efficiency ratio between relaxed and tight isolation cuts is estimated.
- normalization and efficiency ratio allow to infer the number of the multijet events in the high missing energy tail.

QCD events are the bulk of background in the MET < 100 GeV region with relaxed cuts (68,324 events). With the same cuts, 8.810 events are found for MET > 400 GeV . No events from $t\bar{t}$ and $W(\tau\nu)+$ jets survive to the anti-isolation cuts and contribution from $W(\mu\nu)+$ jets (42 events) is negligible. When tight isolation cuts were applied to the low-MET region, the remaining QCD events amount to 4.47 events. With this method, the QCD contribution to the $W(\mu\nu)+$ jets control region can be evaluated to be $8.810/68324 \cdot 4.466 = 5.76 \cdot 10^{-5}$ events, which is pretty coherent with the smallness of the QCD background.

Concerning the other contaminations, $W(\tau\nu)+$jet has a significant impact and has to be subtracted. On the other hand, selection efficiencies of the last three cuts, that are defining the event topology (Tab. 5.6), are consistent between $W(\tau\nu)+$jet and $W(\mu\nu)+$jet within $2 \div 3\%$. This indicates that the shapes of the missing energy distribution is quite similar for the two processes, so the ratio indicated by Monte Carlo could be used to retrieve the number $W(\tau\nu)+$jet from $W(\mu\nu)+$jet, with little systematic uncertainty.
As a cross-check of the method, it has been observed that all $W(\tau\nu)$ events entering the region have a muon from τ decay. Therefore, the composition of the control region can be assumed to be:

$$N^{Contr} = N(W(\mu\nu) + \text{jets})^{Contr} + N(W(\tau\nu) + \text{jets})^{Contr} + N(t\bar{t})^{Contr} = \quad (5.4)$$
$$= [1 + Br(\tau \to \mu)]\, N(W(\mu\nu) + \text{jets})^{Contr} + N(t\bar{t})^{Contr} \quad .$$

With $Br(\tau \to \mu) = 0.1736 \pm 0.0005$ [76], the method described produces the number of events in the control region $N(W(\mu\nu) + \text{jets})^{Contr} = 19.9 \pm 4.5\,(\text{stat}) \pm 1.6\,(\text{syst})$, which is consistent with the Monte Carlo result. Subtracting this value from the control region leads to a $N(W(\tau\nu) + \text{jets})^{Contr} = 3.45 \pm 0.77\,(\text{stat}) \pm 0.27\,(\text{syst})$, to be compared with the result in Tab. 6.7.

To reproduce the number of $Z(\nu\nu)$+jets invisible background, the amount of selected $W(\mu\nu)$+jets has to be rescaled for the following factors:

- Ratio between $Z(\nu\nu)$+jets and $W(\mu\nu)$+jets production cross sections. With the mono-jet topology under study, comparison can be restricted to $W(\mu\nu)$+1 jet and $Z(\to \nu\nu)$+1 jet;
- muon reconstruction and isolation efficiency;
- Trigger efficiency for the single-muon trigger stream. In order to reduce the simulation dependence, the efficiency can be measured with a standard 'Tag and Probe' method, exploiting the $Z \to \mu\mu$ channel and having a triggered isolated muon as a tag. Statistic uncertainty has been calculated *e.g.* in Ref. [77] and is about 2%.

The theoretical cross section ratio was obtained with the ALPGEN Monte Carlo: the two spectra and their ratio are showed in Fig. 5.9 and Fig. 5.10. The ratio has been found to be constant with a 0.92% approximation, and affected by PDF uncertainty by about 1%. The quadratic sum of these effects can be assumed as a systematic uncertainty.
The efficiency ratio also can be studied with Monte Carlo and data using again 'Tag and Probe' method on the muons from the Z decay. In the present Monte Carlo study, this efficiency was estimated as 81% and the expected uncertainity at $100\,\text{pb}^{-1}$ is 4.9%(calculated with the method exploited in [75]).
Applying all the correction factors with their uncertainty, the number of invisible Z events in the signal region is found to be $N(Z(\nu\nu)+\text{jets})^{Sig} = 21.9 \pm 4.9\,(\text{stat}) \pm 2.1\,(\text{syst})$. In Fig. 5.11 the Z background MET distributions from Monte Carlo and from data-driven method are displayed, along with their ratio in Fig. 5.12.

The two shapes are consistent and confirm that the $W(\mu\nu)+$ jets process can do a good job in reproducing the $Z(\nu\nu)+$ jets channel.
It is worth to underline that the procedure is a little different from what reported in Ref. [75], since here the raw MET in the control region is compared with the

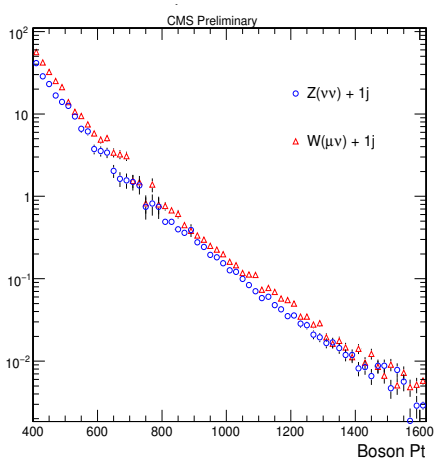


Figure 5.9: $(Z(\nu\nu), W(\mu\nu))$+1 jet spectra produced with ALPGEN .

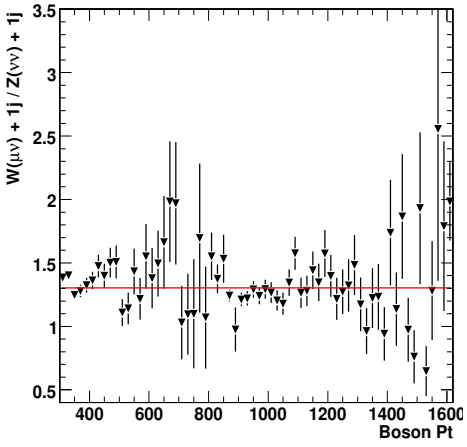


Figure 5.10: Ratio of $(Z(\nu\nu), W(\mu\nu))$+1 jet spectra produced with ALPGEN .

signal region, instead of a sum of muon p_T and raw MET. As soon as that will become a standard procedure, muon deposits will be removed and this estimate is expected to improve.

Estimate with the $Z(\mu\mu)$+jets channel

A possible cross-check for the procedure reported above is the usage of the $Z(\mu\mu)$+jets events as a standard candle. As a feasibility study, an outline for this method and results on efficiency are reported here.
The procedure to define the $Z(\mu\mu)$+jets control region is quite simple:

- Selection of two and only two isolated muons with $p_T > 7\,\text{GeV}$. Isolation requirements are the same of the previous selection on W;
- muons are deleted from the jet list, with a veto cone $\Delta R = 0.5$ around the track;
- a cut on muon invariant mass is applied $|M(\mu\mu) - 91.19| < 20\,\text{GeV}$;
- same signal selection is applied to the obtained sample, taking care to exclude the two muons from the lepton cleaning.

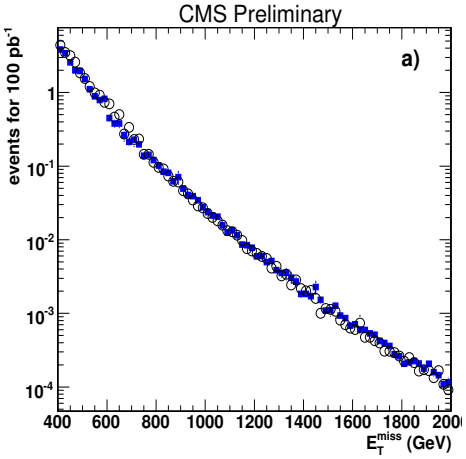


Figure 5.11: Selected $Z(\nu\nu)$+jets events (black circles) and estimate from $W(\mu\nu)$+jets procedure (blue squares). Error bars come from Monte Carlo.

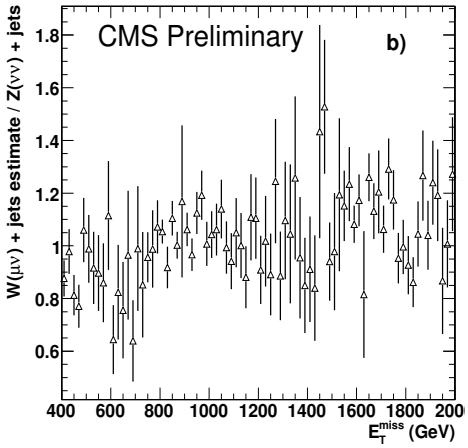


Figure 5.12: Ratio between selected $Z(\nu\nu)$+jets events and estimate from $W(\mu\nu)$+jets procedure.

The well-known payback of the method is statistic, that is a factor ~ 10 smaller than that from the W method. Table 5.7 quotes the number of events resulting from that region, after $1\,\mathrm{fb}^{-1}$ of integrated data. Contribution from $t\bar{t}$ and $W(\mu/\tau\nu)+$jets are well controlled by the double-muon requirement and in the Z mass window are reduced below percent level.

Selection	Events number
Muon $p_T > 100\,\mathrm{GeV}$	560.4
Single isolated muon	386.5
Z mass window	343.8
Kinematic cuts	33.75
Jet cleaning	28.71
Angular cuts	26.12
Indirect Lepton Veto	22.59

Table 5.7: Selection efficiencies for a control region with the $Z(\mu\mu)$+jets channel. Statistic is here after $1\,\mathrm{fb}^{-1}$ of integrated luminosity.

The $Z(\mu\mu)$ control region has to be rescaled to the $Z(\nu\nu)$ control one by correcting for the $Z \to \mu\mu$ branching ratio, muon reconstruction and isolation efficiency.
From Monte Carlo, we measured a muon isolation efficiency of 81% ($\sim$ 4% systematic error) and an acceptance of 95%. Z mass reconstruction is estimated to be 90% efficient. Scaling back to the $100\,\mathrm{pb}^{-1}$ stage, one obtains $N(Z(\nu\nu) + \mathrm{jets})^{Sig} = 24.1 \pm 16\,(\mathrm{stat}) \pm 1,(\mathrm{syst})$, that is consistent with the previous estimate. The quoted systematic error is retrieved from the 5

Once a larger statistics will be available in LHC data, the latter method could be used in its full swing. Since the topology of $Z(\mu\mu)+$jets closely resembles $Z(\nu\nu)+$jets and MET+1 jet signal, it is the primary tool to be used to estimate signal efficiency from data. Cross-checks between the two independent methods will strengthen the evaluation of systematic effects and can improve all the data-driven estimates.

5.5.2 QCD estimate

Even though the amount of QCD background was shown to be small, the limited statistics in this region may induce large uncertainties. Therefore, data-driven methods to control the multi-hadronic processes have been inquired.
As a preliminary approach, the feasibility of the 'ABCD' technique has been tested. The method consists in selecting two variables with discriminating power to create a 2-parameters space, that is partitioned in 4 regions by the variables cuts. If variables are uncorrelated and cuts are such that region C is dominated

	MET $< E_T^0$	MET $> E_T^0$
$\Delta\phi$(jet 2,MET) > 0.5	B	C
$\Delta\phi$(jet 2,MET) < 0.5	A	D

Table 5.8: Definition of A, B, C, D regions from the cuts of the two MET and $\Delta\phi$(jet 2,MET) variables. Here E_T^0 =100 GeV .

by signal and A, B, D only background, the background events into the signal region can be estimated by the formula: $N_C = N_D \times N_B/N_A$. In the framework of hadronic SUSY searches [78], the dependence of the ratio N_B/N_A is fitted as a function of MET and N_C is estimated from a bin-to-bin rescaling of the N_D distribution.
In this study the parameter space was divided by the MET and $\Delta\phi$(jet 2,MET) cuts as in Tab. 5.8. A relaxed missing energy threshold $E_T^0 = 100$ GeV here is instrumental both to minimize the signal contamination in the A background region, and to enhance the QCD contribution in D.

The distributions of the QCD processes and the ADD signal with $M_D = 2$, $n = 2$ in the parameter space are displayed in Fig. 5.13. The signal events lay uniformly on the whole angular range and, if not properly subtracted from region D, lead to a background overestimate.

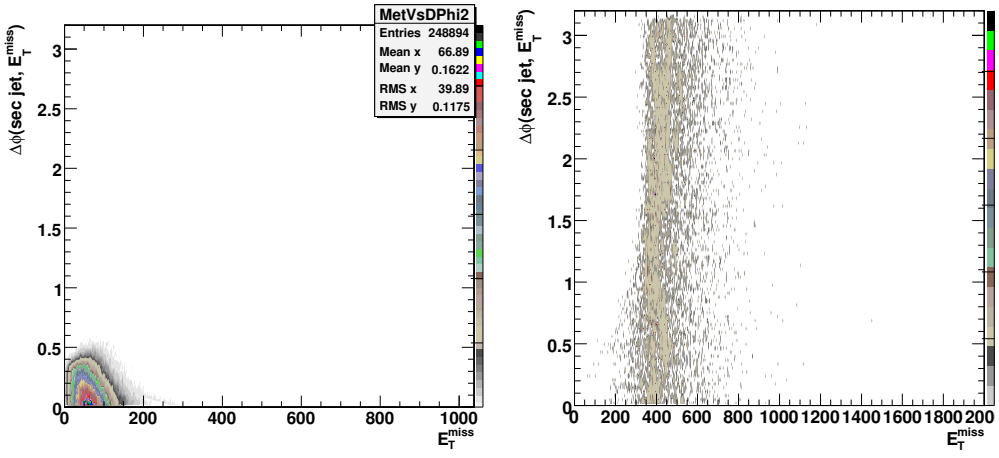


Figure 5.13: Signal (right) and QCD background (left) 2-d distributions in the parameter space. While the latter is well clustered within the (200 GeV , 0.5 rad) region, the signal spans both the D and C regions.

Numerically, the Monte Carlo prediction for the four regions are displayed in Tab. 5.9 for signal and QCD and $E_T^0 = 100$ GeV or 400 GeV .

QCD(Sign)	MET < 100 GeV	MET > 100 GeV	MET < 400 GeV	MET > 400 GeV
$\Delta\phi 1 > 0.5$	90.0 (0)	32.0 (327)	93.7 (101)	0.289 (225)
$\Delta\phi 2 < 0.5$	19755 (0.148)	3842 (66.6)	19965 (30.6)	12.9 (36.1)

Table 5.9: Number of QCD (signal) events expected in the four regions from Monte Carlo predictions, evaluated with the MET cut at 100 and 400 GeV . The regions are distributed in the table as in Tab. 5.8.

The method has been firstly checked assuming an 'only-background' scenario, to test its ability to predict QCD in the C region. The horizontal axis is sliced in 20 GeV bins and the ratio N_B/N_A for the QCD is plotted bin per bin (Fig. 5.14). A polynomial fit was applied in the $0 <$ MET < 100 GeV range and results extrapolated to the C region are 39.9(0.178) for $E_T^0 = 100(400)$ GeV . If signal contamination is taken into account in the range, this produces 42.4(1.79) for $E_T^0 = 100(400)$ GeV . Although reducing the fit range to $20 <$ MET < 100 GeV makes this estimates increase by 44%, comparison with Tab. 5.9 indicates that the 'closure test' has a satisfactory predictivity power.

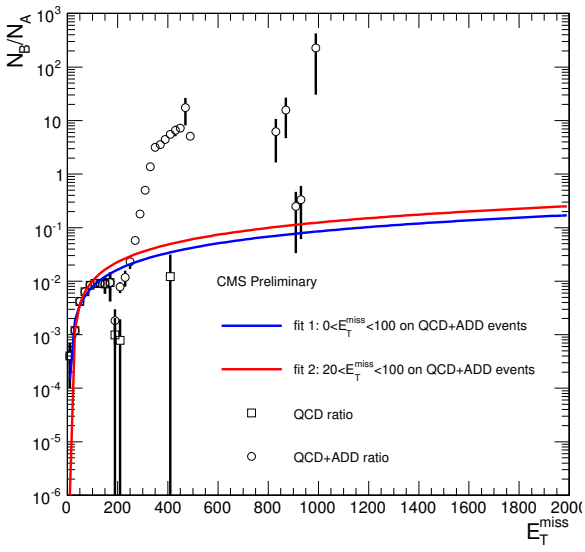


Figure 5.14: Comparison between the two fit regions (0,100) GeV and (20,100) GeV applied on QCD only and QCD+ADD events. On the y axis, N_B/N_A is plotted for MET < 100 GeV and N_C/N_D is plotted for MET < 400 GeV for QCD+ADD events.

When the procedure was repeated on the ratios from 'signal+background' events, the fit results was not significantly affected . Nevertheless, the ADD contamination in the D region induces a QCD overestimation, equal to 1.79 events with the first fit and 2.58 with the second one (for $E_T^0 = 400\,\text{GeV}$). The effect can be ascribed to a lack of statistics in high-MET region (devised selections are too good) and studies for a more detailed optimization of both fit function and fit region are in plan.
In all, the method confirms that QCD background is negligible and is expected to work in a real data scenario. Until optimization/stability tests, a 100% systematic error can be conservatively introduced.

5.5.3 Other background sources

The Monte Carlo results reported in Tab. 5.4 indicate that the $W+$jets production followed by a $W \to \tau\nu$ decay can account for $\sim 16\%$ of the total background. These events can be evaluated in a data-driven way by using again the control region addressed in Sec. 5.5.1, where the QCD component has been suppressed to the 1% level. Beside the irreducible contribution, that region is composed of $W+$jets with the W decaying to $\mu\nu$ (79%) or $\tau\nu$ (13%) final states and $t\bar{t}$ (8%). The latter component can be ascribed to very boosted top pair originating large missing energy tail from W and b-jet decays. Since this kind of processes typically give rise to "fat" jets where W and b quark decay products are closely bundled, b-tagging discriminating techniques are not introduced at this stage. The contribution of $W(\tau\nu)$+jets in the signal region can be derived from the data-driven estimate in Eq. 5.5 after rescaling for the muon trigger, reconstruction and isolation efficiencies as done for $Z(\nu\nu)$+jets. These factors have been quoted in Sec. 5.5.1 and correspond to about 80% with a 2% of systematic uncertainty. This results in $N(W(\tau\nu) + \text{jets})^{Sign} = 4.89 \pm 1.09\,(\text{stat}) \pm 0.46\,(\text{syst})$ with $100\,\text{pb}^{-1}$.
To evaluate the number of survived events for the other W channels, the ratio between them predicted by Monte Carlo is took into account. The procedure has been validated by CDF collaboration in the search for the same channel [79] and result barely affected from systematics. It produces $N(W(\mu\nu)+\text{jets})^{Sign} = 1.76 \pm 0.39\,(\text{stat}) \pm 0.17\,(\text{syst})$ and $N(W(e\nu)+\text{jets})^{Sign} = 1.75 \pm 0.39\,(\text{stat}) \pm 0.16\,(\text{syst})$. All these estimates are consistent with the direct Monte Carlo estimates reported in Tab. 5.4.

5.6 Impact of systematic effects at 14 TeV

Systematic uncertainties play an important role in this analysis where no strong mass peak is expected. The use of Monte Carlo predictions to estimate signal

and background leads to introduce two types of systematic uncertainties: experimental and theoretical, both of which may affect normalization and shapes for signal and backgrounds.

5.6.1 Systematic effects from theory

Since strategies have been defined for the normalization of the most relevant background sources, theoretical uncertainties on cross sections can be ignored for similar reasons. Nevertheless, the $t\bar{t}$ process is still not controlled in a data-driven manner, so its theoretical uncertainty should be considered. Taking into account a $\Delta\sigma(t\bar{t}) = 83\,(\text{PDF}) \pm 50\,(\text{scale})\,\text{pb}$ [80], this can affect the background estimation by 0.86%. Since the cross section of top production is planned to be measured from the first data with well-established techniques up to 8% [81], no account for this uncertainty has been included.

Theoretical errors on ADD signal cross sections produced from SHERPA specified in Tab. 6.1 are tiny in all the benchmark points, but those errors have been computed for a specific energy scale $Q = \sqrt{\hat{s}}$. To estimate the cross section sensitivity to the renormalization and factorization scale, the scale has been varied from $Q/2$ to $2Q$ in the SHERPA generation step for different PDF choices. Results indicates that a $+11\%/-13\%$ uncertainty has to be considered, expected to improve with next orders computations.

A relevant source of systematic uncertainty can be associated to the modeling of background from different generators. In the early days of CMS this will be assessed by running different physics models, *e.g.* ALPGEN vs. PYTHIA vs. `MC@NLO`. Since the current analysis deals with a low jet multiplicity final state, we do not expect generator discrepancies to be large. As already pointed out, ALPGEN and PYTHIA have been compared for all the variables involved in the analysis. Although the absolute numbers of Monte Carlo events can vary about 20% from generators, only efficiency ratios are important provided the background is measured from control regions. They were found to be consistent within $\sim 5\%$ and the value is taken as a conservative estimate of the effect.

Initial and Final State Radiation can in principle distort the multiplicity of jets and modify the event topology.
Since in the analysis no more than two jets are required, effect should be visible only on a selected secondary jet radiated from the primary. Such a jet must be hard ($p_T > 40\,\text{GeV}$) and fulfill very tight geometrical cuts as $\Delta\phi(\text{jet}\,2,\text{MET}) > 0.5$, thus is very unlikely to come from a final state radiation. In addition, the high-p_T overall content demanded from the chosen trigger leaves few chances to a radiated jet to pass the selection. Therefore, we assume that gluon radiation are not an issue for the analysis.

5.6.2 Instrumental systematic effects

The uncertainty associated with the instantaneous luminosity can turn to be the largest one among the instrumental systematic effects. It affects all the processes by the same amount. In the early data taking stage, it can be assumed to have a $\pm 10\%$ fluctuation and is supposed to decrease with a better beam control and machine parameters knowledge. The measure of the luminosity will be done thanks to the forwarded detectors.

One of the major systematics on the acceptance is the uncertainty on the jet energy scale. It has been emulated in a standard way, by shifting the jet 4-vector with a common $(1 \pm \alpha)$ factor and repeating the analysis.
For this early LHC stage, $\alpha = 10\%$ fluctuation can be conservatively assumed irrespective on the jet energy. Nevertheless, jet calibration will be improved since the CMS startup and better energy knowledge will be likely to be achieved soon by diverse data-driven approaches. To take into account these improvements, the effects for 4% and 7% jet energy uncertainty have been computed.
For a lower jet energy, the number of events passing the p_T cut decreases consistently. Increasing the energy scale results in more jets above the threshold, but most of them are rejected by the jet multiplicity veto: as a consequence, higher energy produce a slight loss of events.

The performance of missing energy reconstruction at the startup are expected to be very uncertain, so the modeling of MET with a data-driven technique is worth. Nevertheless, MET has been evaluated from calorimeter towers, so uncertainty on energy deposits can have an additional effect on the number of events passing the 400 GeV cut on MET. Even in this case, the effect has been simulated by rescaling uncorrected MET by a $(1 \pm \alpha)$ factor and the analysis is repeated. Again, three different miscalibration phases have been tested and results are in Table 5.10. Since the missing energy has the largest discriminating power in such analysis, large effects on the events number are expected. An improvement of calorimeter calibration at the level of 4% can contain the uncertainty within the 10%.

Systematic uncertainties due to MET and jet energy resolution have found to be negligible: after a Gaussian smearing of energy (by 10%) and ϕ angle (by 0.1 rad), the maximum effect is a 3% on signal efficiency while the S/B is quite unaffected.

The robustness of the angular cuts have been checked by shifting the azimuthal variables thresholds in the selection. The $\delta\phi$ jet resolution has resulted to be quite flat with the jet p_T in the observed kinematic region and it can be estimated to be about 0.02 rad [32]. When combined with the MET angular resolution, a conservative estimate $\Delta(\delta\phi) \sim 0.05$ rad can be assumed. It was observed to induce fluctuations in the background about 1%.

Source	Notes	Effect on number of signal events (%)
Hard process scale	$\sqrt{Q}/2 \div 2\sqrt{Q}$	$^{+11}_{-13}$
Background modeling	ALPGEN / PYTHIA comparison	5.0
Selection cuts	on $\Delta\phi$(jet 2,MET) thresh.	1
PDF	CTEQ6M param. + CTEQ5L/CTEQ6L diff.	$^{+8.7}_{-6.7}$
Jet energy scale	$\alpha = 4\%$	$^{-0.2}_{-1.0}$
Jet energy scale	$\alpha = 7\%$	$^{-0.2}_{-1.6}$
Jet energy scale	$\alpha = 10\%$	$^{-0.8}_{-4.0}$
MET energy scale	$\alpha = 4\%$	$^{+10.7}_{-10.7}$
MET energy scale	$\alpha = 7\%$	$^{+18.8}_{-18.9}$
MET energy scale	$\alpha = 10\%$	$^{+25.4}_{-26.8}$
Tot Theor. uncert. on S		$^{+14.1}_{-14.6}$
Tot Instr. uncert. on S	$\alpha = 4\%$	$^{+10.6}_{-11.9}$
Tot Instr. uncert. on S	$\alpha = 7\%$	$^{+18.6}_{-21.6}$
Tot Instr. uncert. on S	$\alpha = 10\%$	$^{+24.6}_{-31.8}$
Luminosity	after $100\,\text{pb}^{-1}$	10.0

Table 5.10: Overview of the effect from systematic uncertainties considered in the analysis. Header/footer in the last column correspond to the +/− variation imposed on parameters indicated in the second column.

5.6.3 Total impact of systematic effects

An overview of the effects induced by theoretical and instrumental uncertainties is summarized in Table 5.10. A significant correlation is expected between the jet and the MET energy scale, as the calorimeter towers are the input for the jet calculation. On the other hand, additional variables affect the jet reconstruction as clusterization algorithms, towers p_T, cluster cone etc. Systematic effects affecting the background have been quadratically summed to systematic uncertainties from data-driven methods, discussed in the previous section.

The simplicity of the analysis, based on very few object types reconstructed with very basic criteria, is robust against many systematics, which is a major advantage.
The effect of energy uncertainties on the calorimeter deposits is important on transverse energy, but would be significantly lowered once MET corrections will be better understood from data.
The lepton selection will be measured using the tag-and-probe method from the large sample of Z decays. As discussed above, with a data sample of $100\,\text{pb}^{-1}$ enough Zs should be collected to estimate the lepton efficiencies to the percent

level.

5.7 Discovery potential and exclusion limits at 14 TeV

The mono-jet discovery reach can be produced by considering all the relevant background sources, the ADD signal efficiency, and the impact of systematic effects. Combining together the results from the previous section, the total background can be estimated as:

$$N_B = 33.2 \pm 5.1\,(\text{stat}) \pm 3.7\,(\text{syst})$$

for $100\,\text{pb}^{-1}$of integrated luminosity. This number has to be compared with the signal number of events N_S obtained from the selection (Tab. 5.5) together with instrumental and theoretical systematic errors.
As a powerful significance estimator the S_{12} from Ref. [82] is chosen, that is tailored for low background, Gaussian-distributed statistic and allows a straightforward implementation of statistic and systematic errors:

$$S_{12} = 2\left(\sqrt{N_S + N_B} - \sqrt{N_B + \Delta_B^{Syst}}\right) \times \sqrt{\frac{N_B + \Delta_B^{Syst}}{N_B + \Delta_B^{Syst} + \left(\Delta_B^{Stat}\right)^2}} \quad , \quad (5.5)$$

whereΔ_B^{Syst}, Δ_B^{Stat} are the instrumental uncertainties on the background expectation.

Using S_{12} and the results from the analysis on different M_D, n samples, a discovery plot can be sketched as a function of the fundamental mass (Fig. 6.10). Results for the two different energy miscalibration stages are represented for the early startup phase and systematic uncertainties on signal cross section are included. As a first outlook to the analysis potential, the signal and background sensitivities have been extrapolated to the $1\,\text{fb}^{-1}$ stage: here a better knowledge of instantaneous luminosity (5%) and energy scale (4%) is assumed, while leaving the theoretical errors to the same magnitude.

Similarly, the exclusion limit can be set by assuming that no signal is present and there are only fluctuations from systematic effects on top of background. The 95% C.L. limit was computed for all the generated signal samples with different M_D, n and plotted as a function of the luminosity. Including the luminosity uncertainty and the theoretical and instrumental systematic effect quoted above, the exclusion plot in Fig. 5.16 is obtained. It can be used to indicate the fundamental mass exclusion limits as a function of integrated luminosity. Sensitivities are displayed for two different extra dimensions and quoted in Tab. 5.11 (extrapolations are needed here to go beyond the 7 TeV and below 2 TeV limits).

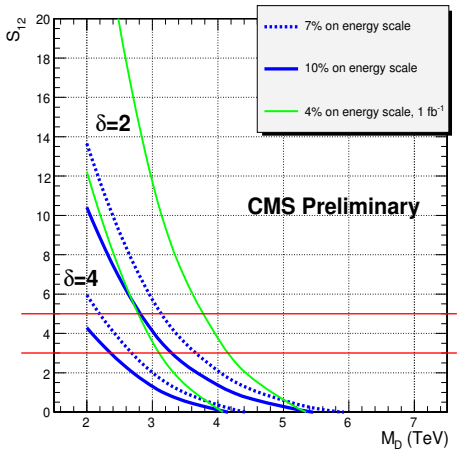


Figure 5.15: Discovery potential of the analysis as a function of M_D and n. Significance estimator S_{12} is defined in the text and the assumed integrated luminosity is $100\,\mathrm{pb}^{-1}$ (blue lines). Discovery potential with an improved energy knowledge and $1\,\mathrm{fb}^{-1}$ is plotted as a reference (green lines). In the plot δ is the number of extra dimension, referred as n in the text.

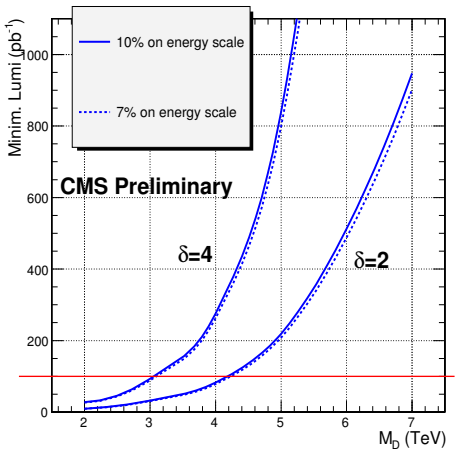


Figure 5.16: Exclusion plot at 95% C.L., representing the minimum statistic necessary to produce a given mass exclusion limit. In both cases, two different energy calibration uncertainties are displayed and sensitivity is plotted for two different extra dimension scenarios. In the plot δ is the number of extra dimension, referred as n in the text.

	L=100 pb^{-1}		L=1 fb^{-1}	
	$n=2$	$n=4$	$n=2$	$n=4$
5σ limit	2.82	1.83	3.77	2.78
3σ limit	3.30	2.37	4.15	3.11
95% C.L.	4.18	3.04	7.18	5.15

Table 5.11: Fundamental mass physics reach of the analysis (TeV), for $n = 2$, 4 and two different luminosities. Significance estimator S_{12} and scaling of the efficiencies with different masses and n have been used.

5.8 Conclusions for 14 TeV case

A simulation study on the ADD model in the G+jet channel with a Monojet plus MET signature has been performed with the CMS detector, focusing on the conditions expected with $100\,\mathrm{pb}^{-1}$ of data.

Simple selection cuts based on the signal events topology, MET and jet energy have been used to reduce the SM backgrounds (mainly QCD, $t\bar{t}$ and $W(l\nu)+$ jets), along with a lepton veto algorithm against the $W(\mu/\tau\nu)+$ jets rejection. All cuts thresholds have been tuned in order to obtain the better signal over reducible background ratio. An HT and MHT based trigger stream has been adopted.
Strategies to infer the irreducible $Z(\nu\nu)+$ jets, $W(l\nu)+$ jets and QCD background processes have been deployed. They give estimates consistent with Monte Carlo and are effective even at the early data-taking stage.
An ADD parameter scan has been performed in order to calculate the CMS sensitivity to the studied model. A first evidence for a MET+1 jet signal can be obtained for values of the fundamental scale M_D lower than 3.30(2.37) TeV for n = 2(4), while 95% C.L. exclusion limits are expected to be 4.18(3.04) TeV for n = 2(4). Current limits from Tevatron are exceeded by almost a factor 3, leaving large room to explore gravity scale with the early LHC.

Chapter 6

Monojet at 10 TeV

6.1 Introduction

The Monojet analysis was repeated for a machine scenario for an integrated luminosity of up to $200\,\mathrm{pb}^{-1}$ and 5 TeV proton beams, according to the directive given during the Chamonix Workshop, which was held in order to decide the machine start-up energy.

We profited of the opportunity to repeat the analysis in order to add some improvement to it.

The new analysis exploits corrected jets and MHT, the vectorial sum of the jets above a certain threshold, which replace the MET in all the analysis. This variable is expected to be less affected by instrumental effects than MET.

6.2 Signal production at 10 TeV

The ADD-model signal has been produced with the SHERPA Monte Carlo generator [64] version 1.1.2.

Different samples with M_D ranging from 1 to 3 TeV and n = 2,4,6 have been produced with detector full simulation of CMSSW. Also a Fast Simulation [19]. was adopted for cross-checking purpose, with n spanning from 2 to 4. All the relevant quantities produced in this simulation are know to be consistent with those from full-simulation [83].

A p_T cut-off on the parton recoiling, $\hat{p}_T > 150\,\mathrm{GeV}$, was required. The CTEQ61L Parton Density Functions (PDF) [84] have been used.
With these production parameters, the signal cross sections at leading order were calculated by SHERPA for the 15 different samples and are given in Tab. 6.1.

For each subsample with a given (M_D, n), $2 \cdot 10^4$ events were produced. As already demonstrated in Ref. [85], the transverse momentum, jet multiplicity,

	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
$M_D = 1\,\text{TeV}$	279.11	171.79	109.98	70.50	44.45
$M_D = 2\,\text{TeV}$	33.03	17.41	10.64	6.92	4.58
$M_D = 3\,\text{TeV}$	7.28	3.02	1.57	0.93	0.58

Table 6.1: ADD cross sections as evaluated by the SHERPA program with $\sqrt{s}$ =10 TeV. All values are in pb.

SM Channel	σ·BR (pb)	Gen. evs. ($\times 10^3$)	Int. Lumin. (pb^{-1})
$Z(\to \nu\nu)$+jets	3700	10^3	270
QCD $\hat{p}_T > 80\,\text{GeV}$	1,934,639	$3 \cdot 10^3$	1.5
QCD $\hat{p}_T > 170\,\text{GeV}$	62,563	$3 \cdot 10^3$	48
QCD $\hat{p}_T > 300\,\text{GeV}$	3665	$3 \cdot 10^3$	820
QCD $\hat{p}_T > 470\,\text{GeV}$	316	$3 \cdot 10^3$	9500
$W(\to l\nu)$+jets	35,550	10^4	281
$t\bar{t}$	317	10^3	3150
single-t (tW, s-ch., t-ch.)	93	550	5910

Table 6.2: Overview of the background statistics exploited in the analysis, together with the corresponding integrated luminosity. All samples are produced with the full simulation and cross sections are from Ref. [86].

and event shape do not show any striking dependence from n; a larger M_D will only result in more boosted jets.

6.3 Background production at 10 TeV

The set of background processes has been generated, with a sample size corresponding to an integrated luminosity larger than $200\,\text{pb}^{-1}$ (with the exception of $\hat{p}_T > 80\,\text{GeV}$ and $\hat{p}_T > 170\,\text{GeV}$ QCD samples).
Most of the background samples were sorted by the 10 TeV Monte Carlo production performed in the Summer 2008 with a full simulation of the CMS detector.

The leptons and jets have been processed with the Physics Analysis Toolkit (PAT) [71] which applies the standard energy corrections to the jets and computes isolation parameters for the leptons. Jets were obtained using an iterative cone (IC) algorithm with $\Delta R = 0.5$ [32], but the results do not depend significantly on the clusterization details. The jet energy scale is corrected by applying a Monte Carlo based energy correction. This includes the relative corrections which produce a uniform response along η and the absolute corrections, which correct the jet energy back to the generator particle level [33].

6.4 Signal and background analysis at 10 TeV

After a brief description of the motivation and results for the trigger, the cut-based analysis procedure is outlined in this section. The strategy will closely resemble the 14 TeV one. In this analysis we opted for a more robust trigger, and replaced the usage of MHT and HT with a more standard jet trigger. Following the beam energy reduction, some of the p_T cuts are lowered. The choice of thresholds does not result from a sophisticated significance maximization, but was determined by looking for a high $N_S/\sqrt{N_B}$ ratio (with $N_S(N_B)$ number of signal(background) events after the selection). The relative improvements obtained by further adjusting the cuts were found to be negligible.

6.4.1 Trigger selection

The trigger choice has been based on the usage of the most simple quantities, that are capable to select the signal with the maximal efficiency. The MET or MHT objects are no longer used in the online selection: the former assumes an optimal control of missing energy (which is unlikely to be reached in an early data taking stage), the latter results in a too low acceptance for the mono-jet signal. As a consequence, one of the simple single jet streams designed for a phase with a $10^{31}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ luminosity is proposed here. It requires:

- at Level 1 (L1), at least 1 jet with $p_T > 70\,\mathrm{GeV}$;
- at High Level Trigger, at least 1 jet with $p_T > 110\,\mathrm{GeV}$.

The jet energies are those of corrected jets. Both the L1 and HLT triggers are intended to be unprescaled throughout the duration of the initial CMS run, with a multijets rate at the HLT predicted to amount to $8.1 \pm 1.1\,\mathrm{Hz}$.

The High Level Trigger response has been reproduced off-line using reconstructed jets. The efficiency of this trigger path is shown in Fig. 6.1 as a function of the MHT lower cut. As it will be shown in the following, these variable demonstrates to be the most effective in enriching the signal contribution. For events with MHT $> 250\,\mathrm{GeV}$ the trigger efficiency of the signal is close to 100%.

The effect of a harder cut for HLT was also explored, rising the p_T threshold to 180 GeV . Reduction on acceptance in the background (about 80%) is larger than in signal (about 40%), but this gain is compensated by a much less effectiveness of off-line selections, especially the MHT cut explained below. The `HLT_Jet110` path is thus considered an optimal choice.

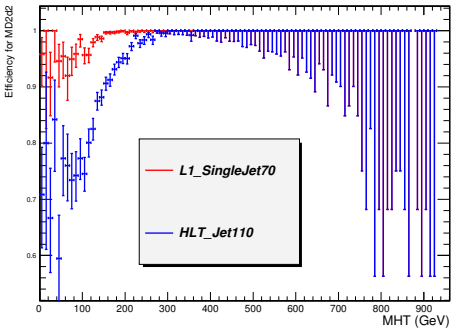


Figure 6.1: Efficiencies for the single-jet trigger for a signal with $M_D = 2\,\text{TeV}$ and $n = 2$ as a function of MHT.

6.4.2 Selection and efficiencies

To reduce the impact of objects not coming from hard interaction, only jets with transverse momenta larger than 50 GeV within the hadronic calorimeter acceptance $|\eta| < 3$ are considered. As shown in Fig. 6.2(a), the signal leads to a long tail in the MHT distribution, hence a cut MHT $> 250\,\text{GeV}$ was imposed at the preselection level. These simple cuts have already an important impact on the multi-jet background, for which the events have typically a few jets balanced in E_T, possibly accompanied by more softer jets, resulting in a quite low MHT value.

Charged leptons cleaning cuts

Before further jet selections, it is important to clean the events from contaminations with isolated leptons (as those from the $W(l\nu)$+jets channel) and electrons and photons identified as jets. Here the same procedure was used as in the previous analysis at 14 TeV ("Indirect Lepton Veto"), where two variables were exploited the JEMF and the TIV.

In order to suppress cosmic background, at least one vertex coming from interaction point and at least two $p_T > 5\,\text{GeV}$ tracks inside leading jet cone were requested.

In Fig. 6.3(b) we can see some of the distributions associated to the variables we will use to select the signal and reject the backgrounds.

Factorization method for QCD

After the complete set of selections has been applied, only 1 Monte Carlo event from QCD background have been found, corresponding to 0.2 events after nor-

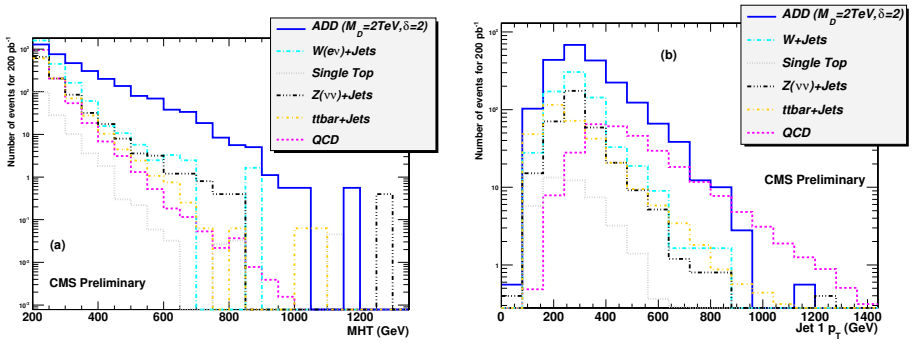


Figure 6.2: (*a*) Vectorial sum of transverse energy (MHT) for ADD signal ($M_D = 2\,\text{TeV}$, $n = 2$) and relevant backgrounds before any selection, after $200\,\text{pb}^{-1}$. (*b*) Transverse momentum of the leading jet, after the selection cut on MHT. Histograms are overlaid and normalized to the same area, in order to highlight the different shapes.

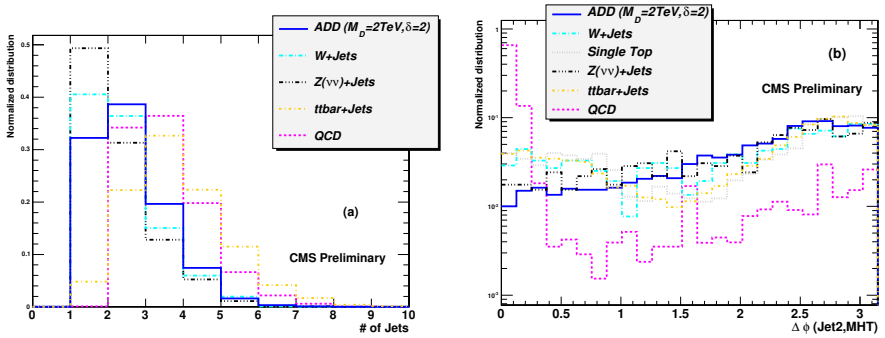


Figure 6.3: (*a*) Number of jets for signal and relevant backgrounds, for MHT $>$ 250 GeV and jets with transverse momenta larger than 50 GeV for the others and $|\eta| < 3$. A veto against three or more jet events rejects the hadronic objects in the signal that are not related to the hard interaction. Histograms are overlaid and normalized to the same area, in order to highlight the different shapes. (*b*) Angle in the transverse plane between the MHT and the secondary jet, after the preselection cuts.

QCD $\hat{p}_T$	$> 80\,\text{GeV}$	$> 170\,\text{GeV}$	$> 300\,\text{GeV}$	$> 470\,\text{GeV}$
N. of evts after ϵ_1	< 0.6	16 ± 8	97 ± 5	98 ± 2
Eff. after ϵ_2	$(1.06 \pm 0.02) \cdot 10^{-3}$	$(5.02 \pm 0.04) \cdot 10^{-3}$	$(7.77 \pm 0.18) \cdot 10^{-4}$	$(7.2 \pm 0.6) \cdot 10^{-5}$
QCD_{fact}	$< 0.65 \cdot 10^{-3}$	$(80 \pm 40) \cdot 10^{-3}$	$(754 \pm 43) \cdot 10^{-4}$	$(706 \pm 60) \cdot 10^{-5}$

Table 6.3: Number of events (second row) and efficiency (third row) for the two sets of cuts explained in the text, when measured on four QCD bins. The two set have been checked to be not correlated to a good approximation, so the total number of selected events can be estimated as the product QCD_{fact} in the fourth row.

malization to integrated luminosity. Since even an underestimation by e.g. a factor 3 of this contribution would affect the total background by less than 0.5%, it has been neglected in the total background.

The value of the QCD rejection factor has been further checked with a factorization approach. The signal region has been divided in two sets of cuts:

1. events with MHT $> 250\,\text{GeV}$, leptons removed with the Indirect Lepton Veto, and cut on the leading jet pseudorapidity;

2. events with p_T cut on the leading jet, veto for more than 2 jets, and azimuthal cuts as quoted above. To have a number of multi-jet events sufficient to compute efficiencies, leading jet momentum cut has been lowered to $150\,\text{GeV}$: therefore, the final estimate should be an upper bound.

The linear correlation coefficient ρ among all the jet-related variables has been evaluated and these two sets were demonstrated to be the ones with small correlation ($\rho < 0.15$, mostly due to correlation between leading jet and MHT). Therefore, the total efficiency for QCD in the control region can be reproduced by the product of the two efficiencies to a good approximation. The number of multi-jet events surviving the first set (ϵ_1), together with the efficiency of the second set (ϵ_2), is reported in Tab. 6.3. Combining together all the $\hat{p}_T$ samples, number of multi-jet events after $200\,\text{pb}^{-1}$ (QCD_{fact}) is foreseen not to be larger than 0.2 events, confirming that QCD multi-jet processes do not introduce additional uncertainties in background knowledge and can be ignored in the total background.

6.4.3 Summary of selection cuts

The set of cuts presented above are summarized in the following:

	$t\bar{t}$	$Z(\nu\nu)$+	QCD	$W(e\nu)$	$W(\mu\nu)$	$W(\tau\nu)$	single-t
Trigger	28,970	11,390	143 M	31,320	19,320	20,600	4460
	(45.7)	(2.8)	(0.1)	(1.2)	(0.8)	(0.1)	(24.1)
Pres.	318	358	288	90	391	230	44
	(1.1)	(3.1)	$(2 \cdot 10^{-4})$	(0.3)	(2.0)	(1.1)	(1.0)
ILV	52.5	305	214	31.9	38.5	90.9	7.2
	(16.5)	(85.4)	(74.3)	(35.4)	(9.8)	(39.5)	(16.5)
Kin.	37.4	245	187	24.6	24.6	72.1	4.5
	(64.7)	(80.5)	(87.8)	(77.1)	(63.9)	(79.3)	(62.5)
Mult.	8.2	205.6	70.9	18.8	22.9	59.8	2.8
	(24.1)	(83.8)	(37.9)	(76.7)	(93.3)	(83.0)	(61.7)
Ang.	6.4	182.5	0.2	17.2	19.7	46.7	2.3
	(78.6)	(88.8)	(0.28)	(91.5)	(86.0)	(78.1)	(82.1)

Table 6.4: Number of selected events for each group of cuts in the relevant background samples, normalized to $200\,\mathrm{pb}^{-1}$. In parenthesis the partial efficiencies are quoted in percent (For shortness' sake $Z(\nu\nu)$++jets and $W(l\nu)$+jets are simply indicated as $Z(\nu\nu)$+ and $W(l\nu)$).

1. *Trigger* : Standard single-jet trigger with $p_T > 70\,\mathrm{GeV}$ at L1 and $p_T > 110\,\mathrm{GeV}$ at HLT;

2. *Preselection* : MHT $> 250\,\mathrm{GeV}$, $p_T(\mathrm{jet}) > 50\,\mathrm{GeV}$, $|\eta(\mathrm{jet})| < 3$;

3. *Indirect Lepton Veto* : $0.1 <$ JEMF < 0.9, TIV > 0.1;

4. *Kinematic cuts* : $p_T(\mathrm{jet}\,1) > 200\,\mathrm{GeV}$, $|\eta(\mathrm{jet}\,1)| < 1.7$;

5. *Jet multiplicity* : veto against 3 or more jets in the event;

6. *Angular cuts* : $\Delta\phi(\mathrm{jet}\,1,\mathrm{MHT}) > 2.8$, $\Delta\phi(\mathrm{jet}\,2,\mathrm{MHT}) > 0.5$.

In Tab. 6.4 is reported the effect of each group of cuts, for the SM processes and in Tab. 6.5 for some signal samples. The numbers given in the tables are the absolute expected number of selected events (assuming $200\,\mathrm{pb}^{-1}$ of data) and the relative efficiency of the cuts.

Beside the irreducible background, the only important events are those with a muon or an electron from W (or from τ from W) decay; the W can be directly produced or coming from a top quark. These leptons are not identified by the Indirect Lepton Veto either because they are outside the pseudorapidity acceptance, or too close to a high p_T jet.

	$n = 2$			$n = 4$	
$M_D =$	1 TeV	2 TeV	3 TeV	2 TeV	3 TeV
Trigger	51,000	6180	1370	2010	301
Pres.	11,140	2123	498	753	133
	(21.8)	(34.3)	(36.5)	(37.3)	(44.1)
ILV	9572	1825	426	641	113
	(85.9)	(86.0)	(85.5)	(85.2)	(85.4)
Kin. cuts	6785	1368	314	487	88.4
	(70.9)	(75.0)	(73.7)	(76.0)	(78.0)
Jet mult.	5605	1044	401	374	64.4
	(82.6)	(76.3)	(77.3)	(76.8)	(72.9)
Ang. cuts	4934	906	206	322	55.8
	(88.0)	(86.8)	(86.8)	(86.1)	(86.6)
Total Eff. (%)	8.8 ± 0.1	13.7 ± 0.4	14.1 ± 0.4	13.2 ± 0.4	17.7 ± 0.4

Table 6.5: Number of selected events for each group of cuts in five signal subsamples, normalized to $200\,\mathrm{pb}^{-1}$. In parenthesis are the partial efficiencies (after the trigger) quoted in percent. Uncertainties on efficiencies are statistical only.

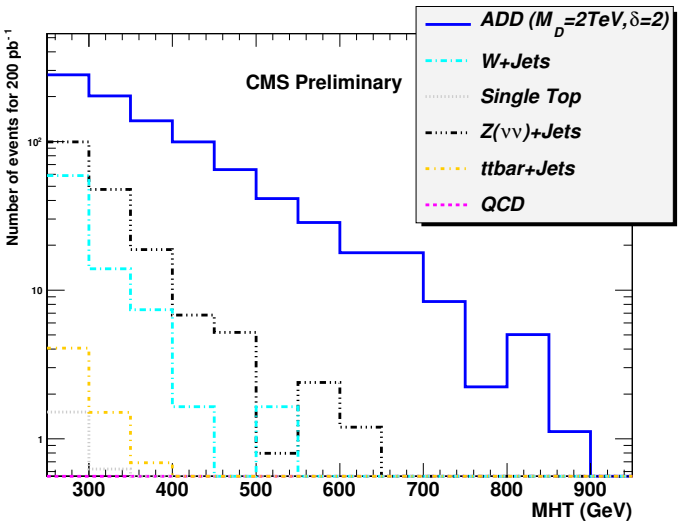


Figure 6.4: MHT distribution after all selections are applied. Histograms are overlaid and number of events correspond to $200\,\mathrm{pb}^{-1}$.

The MHT distributions for signal and background after the cuts are shown in Fig. 6.4. The signal shows up as an excess of events in addition to the dominant background from $Z(\nu\nu)$+jets.

The behavior of simulated $ZZ/WW/ZW$+jets processes is not shown here, because their impact on final selection is negligible. In fact, MHT preselection cut has a rejection factor larger than 200, that results in a contamination about 2% in the control region.

The effects of the machine-induced background and cosmic rays have not been included, but we are confident that the devised set of cut reduces this kind of effects to a negligible level; in particular, the beam halo is expected to be removed by the jet electromagnetic fraction lower cut. As an additional check, the number of tracks closest by $\Delta R < 0.5$ to the jet axis have been considered for a sub-sample of signal events. When the signal is also asked to have at least one primary vertex and 2 or more jet tracks with $p_T > 5\,\text{GeV}$, less than 1% of events are rejected. Such a cross-check demonstrates that the proposed procedure allows to select jets coming from genuine collision events, while objects that are displaced from the interaction point (as cosmic rays) do not contaminate the sample.

6.5 Impact of systematic effects at 10 TeV

In the next section, it will be shown how the relevant background processes can be estimated in control regions that will be driven by real data: therefore, systematic effects on background normalization have no relevance here, and only effects on efficiency ratios will be considered.

6.5.1 Systematic effects from theory

When the mono-jet channel was analyzed in the 14 TeV regime, a detailed account for many effects was provided. The most important uncertainties from theory on the signal have found to be:

- cross section sensitivity to the renormalization and factorization scale $Q = \sqrt{\hat{s}}$. The study was repeated in the new environment and the scale has been varied from $Q/2$ to $2Q$ in the SHERPA generation step for different PDF choices. Results indicate that a $^{+7.5\%}_{-6.7\%}$ uncertainty has to be considered for the ADD signal;

- uncertainties on the parton density function (PDF). They have been evaluated using the reweighting technique and Master Equation on the CTEQ61M model set [84]. We found a $^{+11.5\%}_{-9.5\%}$ cross section fluctuation for the ADD with $M_D = 2\,\text{TeV}$ and $n = 2$.

uncertainty	$n=2$		$n=4$	
	$M_D = 2\,\text{TeV}$	$M_D = 3\,\text{TeV}$	$M_D = 2\,\text{TeV}$	$M_D = 3\,\text{TeV}$
jet p_T and ϕ	3.0	3.4	5.5	3.4
JES	$^{+13.1}_{-16.2}$	$^{+13.2}_{-13.3}$	$^{+11.8}_{-14.9}$	$^{+9.5}_{-11.5}$
Tot from exp.	$^{+13.4}_{-16.4}$	$^{+13.6}_{-13.7}$	$^{+13.0}_{-15.9}$	$^{+10.1}_{-12.0}$

Table 6.6: Relative effect (in percentage) on total number of signal events of systematic uncertainties in jet energy, direction, and calibration, for 4 benchmarks points.

6.5.2 Experimental systematic effects

Most of the selections described above rely on objects obtained from jets: as a consequence, uncertainties associated to energy, direction, and calibration of jets will turn to be the most relevant systematic effects. In this analysis, three mismeasurements of jets have been reproduced:

- jet energy resolution uncertainties. Transverse momenta of all jets have been smeared by a Gaussian function accounting for a 10% resolution;
- jet direction uncertainties. Transverse angle ϕ of all jets have been smeared by a Gaussian centered on zero accounting for a (conservative) 0.1 rad resolution;
- jet energy scale, emulated by shifting the jet 4-vector with a common $(1\pm\alpha)$ factor. For this early LHC stage, $\alpha = 10\%$ can be assumed irrespectively on the jet energy.

These uncertainties have been applied to both to the HLT and off-line (corrected) jets. Their relative shift from the value with no systematic effect, for few benchmark points, is summarized in Tab. 6.6.

When the energy scale of the objects is overestimated, p_T of jets increases faster than the total energy, MHT cut becomes less effective and the signal is enhanced. The effect is smaller when increasing M_D (that boosts the total energy already above 250 GeV) and is partially compensated by the jet veto, that founds more multi-jet events falling in the kinematic requirements. On the other hand, introducing an uncertainty in energy and jet direction results in a loss of few percent of event with respect to no smearing. Since uncorrelation of the two effects is a reasonable assumption, total error from instrumental sources has been evaluated as the quadratic sum.

The value of instantaneous luminosity can be assumed to have a $\pm 10\%$ uncertainty. It has been added in quadrature to the values in Tab. 6.6.

6.6 Data-driven background estimation at 10 TeV

In the following, techniques to estimate the irreducible background of $Z(\nu\nu)$+jets ("invisible Z" background) and $W(e/\mu/\tau\nu)+$ jets are described.

6.6.1 $Z(\nu\nu)$+jets background estimation

As in the case of 14 TeV, the selection defining the control region maintains the same cuts as for the MET+1 jet signal, except for the muon requirement. The muon requirement has been improved and adapted to the 10 TeV case. The different steps of selection are the following:

- the event is triggered by the single-jet stream adopted for the signal;
- a sample with muons is selected, with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.4$;
- the muon is required to fulfill the tracker isolation requirements. The muon is considered isolated when $\mu_{Iso} < 3\,\text{GeV}$;
- the muon must fulfill the calorimeter isolation requirements. The muon isolation variable is defined as in the tracker case (with p_T(trk) replaced by E_T deposit in the cone). In order to minimize scores to be confused with a QCD jet, a muon is considered isolated when the sum is less than 1 GeV;
- only events with one isolated muon are retained;
- jets that are closer than a $\Delta R = 0.5$ distance from the isolated muon direction are ignored;
- the resulting sample undergoes the same set of selections as the signal region, but excluding from TIV the track which is associated to the single muon.

This procedure allows to reproduce the same kinematic region that was designed for the signal, but having one hard isolated muon, for which the hypothesis of coming from W is highly probable. The only systematic effects are the ones associated to the muon selection. Once the contamination from other SM processes has been controlled, the number of Z background events can be obtained from the control sample events after correcting them for the cross section ratio and muon reconstruction and isolation efficiency. Figures 6.5 show muon the reconstruction and isolation efficiency on the W+jets sample.

As in the 14 TeV case, the most significant backgrounds which contaminate the control region are found to be $W(\tau\nu)+$ jets, top pair, and single top events.

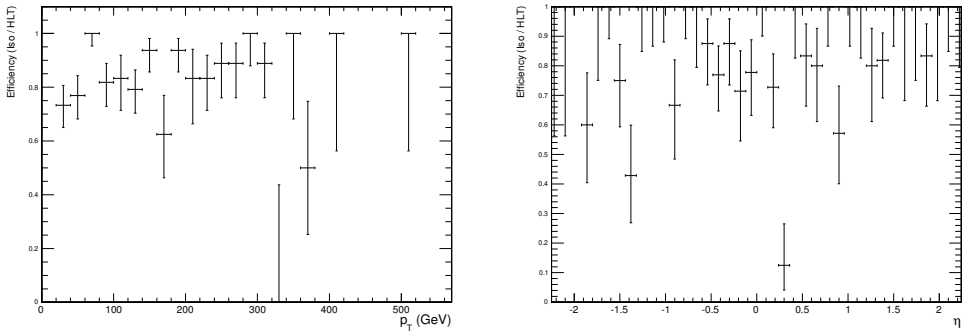


Figure 6.5: Muon selection efficiency as a function of its transverse momentum (left) and pseudorapidity (right). Number of selected muons have been calculated for the $W(\mu\nu)$+jets sample and the other cuts for the control region have been applied. Error bars are statistical only.

The composition of the control region (with number of events (efficiency)) is reported in Tab. 6.7 for the $W(\mu\nu)$+jets channel and the other background sources at different stages (all values quoted for $200\,\mathrm{pb}^{-1}$).

Selection	$W(\mu\nu)$+jets	$W(\tau\nu)$+jets	$t\bar{t}$	single-t
Trigger	19,580	20,290	28,790	4923
	(0.80)	(0.81)	(45.4)	(24.0)
$p_T(\mu) > 20\,\mathrm{GeV}$	14,830	9845	10,820	2195
	(75.7)	(48.5)	(37.6)	(44.6)
Single isolated μ	12,310	8443	5749	1484
	(83.0)	(85.8)	(53.1)	(30.1)
Preselection	310.4	105.6	159.2	28.0
	(2.5)	(1.3)	(2.8)	(1.9)
Indirect Lepton Veto	259.4	68.6	103.5	20.8
	(83.6)	(65.0)	(65.0)	(74.3)
Kinematic cuts	193.6	51.0	56.8	13.2
	(74.8)	(74.3)	(54.9)	(63.5)
Jet multiplicity	177.0	44.5	21.9	10.0
	(91.4)	(87.3)	(38.6)	(75.4)
Angular cuts	163.1	39.8	18.3	8.7
	(92.1)	(89.4)	(83.6)	(87.0)

Table 6.7: Number of events selected selections for the $W(\mu\nu)+$jets control region, along with efficiencies in percentage. All normalizations refer to $200\,\mathrm{pb}^{-1}$.

Due to the hard cuts on calorimeter isolation, high MHT, and angular cuts, less than 1 generated multi-jet event is expected to enter this region. In order to verify that QCD multi-jet contamination can be neglected also in this case, an independent cross-check has been performed by using a factorization method. The selection on control region has been split in single muon selection (with kinematic and isolation requirements) and full selection without muon cuts. Since only 0.5 muons from multi-jet have been estimated in such a control region, it can be assumed that muon faking jets are negligible and the two sets of cuts are not correlated: therefore, the rejection for multi-jet can be obtained by multiplying the two rejection factors separately. The efficiency for the stand-alone single muon selection is $(5.35 \pm 0.01) \cdot 10^{-3}$, while the full set of cuts but those on the muon retains a fraction equal to $(2.5 \pm 0.4) \cdot 10^{-9}$. The total efficiency amounts to[1] $(1.3 \pm 0.2(\mathrm{MC})) \cdot 10^{-11}$: after $200\,\mathrm{pb}^{-1}$, this corresponds to less than 0.01 surviving events.

On the other hand, contaminations of muons from top quark and τ decay is not avoidable and deserve a separate treatment.
Since this analysis is specifically aimed to a very early data-taking, no b-tagging techniques have been introduced to veto the b-jets from top. In this way, systematic effect due to tagging inefficiencies can be avoided and the analysis is feasible even with a not fully commissioned Pixel detector and not optimal alignment among detector parts. As long as it is not possible to identify a region where top pair and single top events are decoupled from W, their contribution should be evaluated from Monte Carlo. On the other hand, recent studies in the 10 TeV regime demonstrate that both processes can be observed in a very early stage with reasonable precision. For the single top, 35% statistic and 11% systematic errors on the cross section is foreseen at $200\,\mathrm{pb}^{-1}$ [80], while for $t\bar{t}$ the result in Ref. [81] can be scaled to 3.35%(10%) statistic(systematic) error. Therefore, it can be assumed that the expected cross sections will be normalized in data with those uncertainties: $N(t\bar{t})^{Contr} = 18.3 \pm 0.6\,(\mathrm{stat}) \pm 1.8\,(\mathrm{syst})$, $N(\mathrm{single-t})^{Contr} = 8.7 \pm 3.0\,(\mathrm{stat}) \pm 1.0\,(\mathrm{syst})$.

A preliminary analysis looking for b-jet using a 'combined secondary vertex' technique has been attempted. It has found that removing jets with a discriminator value for this tagger higher than 0.6, top pair contribution is reduced by $(27 \pm 2)\%$ and single top by $(38 \pm 3)\%$, while the effect on ADD signals and other background is not worse than 90%. Since many of multi-jets events survived to selections can contain b quark, this veto would be also beneficial in reducing hardest multi-jet component down to 30% when $\hat{p}_T > 470\,\mathrm{GeV}$.

Concerning the $W(\tau\nu)+$jet contamination, two important observations hold:

[1] Notation MC refers to the counting error due to the limited number of generated Monte Carlo sample (or the error induced by this source through error propagation).

- simulations showed that the shape of MHT distribution is similar between $W(\tau\nu)$+jet and $W(\mu\nu)$+jet within few percent. Therefore, the ratio calculated by Monte Carlo could be used to retrieve the number $W(\tau\nu)$+jet from $W(\mu\nu)$+jet, with little systematic uncertainty;

- as verified by probing the Monte Carlo truth, all events from $W(\tau\nu)$ events entering the region have a muon from τ decay.

The method described produces the number of events in the control region $N(W(\mu\nu)+\text{jets})^{Contr} = 172.9 \pm 13.1\,(\text{stat}) \pm 13.4\,(\text{syst})$, which is consistent with the Monte Carlo result. Subtracting this value from the control region leads to a $N(W(\tau\nu)+\text{jets})^{Contr} = 30.0 \pm 2.3\,(\text{stat}) \pm 2.3\,(\text{syst})$, to be compared with the result in Tab. 6.7. The small uncertainty on Branching Ratio has been neglected.

To reproduce the number of $Z(\nu\nu)$+jets invisible background, the amount of selected $W(\mu\nu)$+jets has to be rescaled for the following factors[2]:

- ratio between $W(\mu\nu)$+jets and $Z(\nu\nu)$+jets production cross sections. High statistic samples from the two processes have been produced, to cover a broad energy spectrum. The two spectra and their ratio are shown in Fig. 6.6 and Fig. 6.7. They demonstrate that the ratio is constant up to high energies in the hard interaction, and equal to $1.314 \pm 0.008\,(\text{MC}) \pm 0.013\,(\text{PDF})$;

- muon reconstruction and isolation efficiency. In this control region, it was determined to be $0.760 \pm 0.079\,(\text{MC}) \pm 0.001\,(\text{syst})$, where the systematic error is estimated from a fit on the muon efficiency. These values can be checked with data using a standard 'Tag and Probe' method on the muons from the Z decay.

Applying all the correction factors with their uncertainty, the number of invisible Z events in the signal region is found to be $N(Z(\nu\nu)+\text{jets})^{Sign} = 163 \pm 22\,(\text{stat}) \pm 13(\text{syst}) \pm 17(\text{MC})$, to be compared with $N(Z(\nu\nu)+\text{jets})^{MC} = 182 \pm 13\,(\text{stat})$.

In Fig. 6.8(a) the Z background MHT distributions taken directly from Monte Carlo and from our data-driven method are displayed, along with their ratio in Fig. 6.8(b). The two shapes are consistent and confirm that even in an early scenario, $Z(\nu\nu)$+jets background can be estimated from data, with a $\sim 10\%$ uncertainty and relying on robust reconstructed objects.

[2]Here it is implicitly assumed that, when extrapolating a data-set to infer the other, a correction for the acceptance of the two different triggers is included.

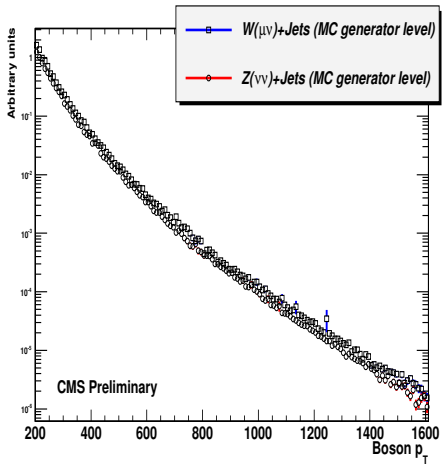


Figure 6.6: $(Z(\nu\nu), W(\mu\nu))$+1 jet spectra produced with PYTHIA, as a function of momentum of the hard interaction.

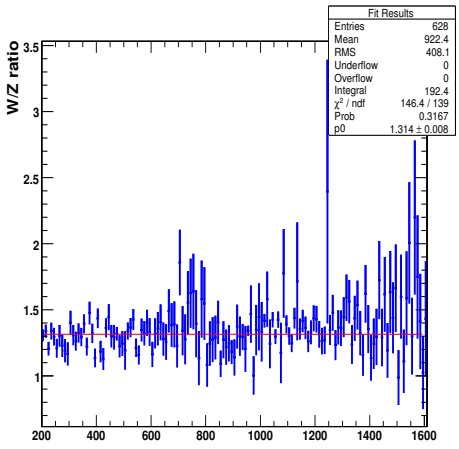


Figure 6.7: Ratio of $(Z(\nu\nu), W(\mu\nu))$+1 jet spectra produced with PYTHIA. This ratio has been assumed constant all over the control and signal region considered.

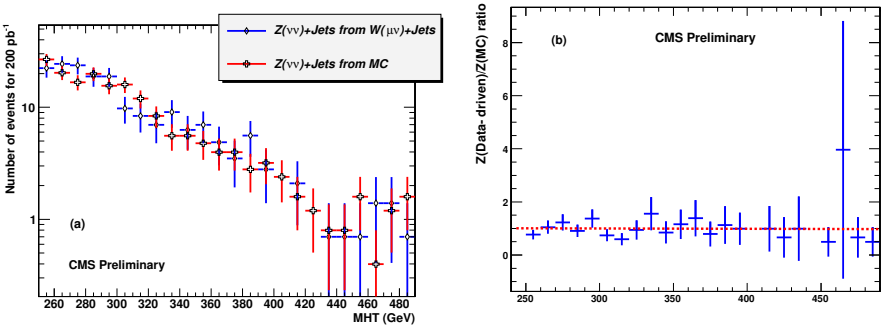


Figure 6.8: Selected $Z(\nu\nu)$+ jets events from Monte Carlo and estimated from $W(\mu\nu)$+ jets procedure (a) and their ratio (b). Within error bars from Monte Carlo error, the ratio is consistent with 1 (red dashed line).

6.6.2 Other background sources

As done in the 14 TeV case, rescaling the $N(W(\tau\nu) + \text{jets})^{Contr}$ by muon reconstruction and isolation efficiencies we obtain: $N(W(\tau\nu) + \text{jets})^{Sign} = 52.4 \pm 7.1\,(\text{stat}) \pm 4.0\,(\text{syst}) \pm 5.4\,(\text{MC})$ for $200\,\text{pb}^{-1}$.
Taking into account the ratio between the different W decay channel, predicted by Monte Carlo in the signal region we obtain: $N(W(\mu\nu) + \text{jets})^{Sign} = 22.1 \pm 3.0\,(\text{stat}) \pm 1.7\,(\text{syst}) \pm 5.9\,(\text{MC})$ and $N(W(e\nu) + \text{jets})^{Sign} = 19.3 \pm 2.6\,(\text{stat}) \pm 1.5\,(\text{syst}) \pm 5.3\,(\text{MC})$.
All these estimates are consistent with the direct Monte Carlo results reported in Tab. 6.4.

Finally, some boosted top can originate large missing energy tail from W and have a not-rejected lepton and hard jets, falling in the signal region. These contribution are treated here in the same way as the control region, by taking the prediction from Monte Carlo as the central value, but using the precision expected to be reached for the actual observation of top, after $200\,\text{pb}^{-1}$ [81]. The assumption leads to $N(t\bar{t})^{Sign} = 6.4 \pm 0.2\,(\text{stat}) \pm 0.6\,(\text{syst})$, $N(\text{single} - \text{t})^{Sign} = 2.3 \pm 0.8\,(\text{stat}) \pm 0.2\,(\text{syst})$.

6.7 Discovery potential and exclusion limits at 10 TeV

The mono-jet discovery reach can be established by considering all the relevant background sources, the ADD signal efficiency, and the impact of systematic effects. Combining together the results from the previous section, the total background N_B is estimated as:

$$N_B = 266 \pm 23\,(\text{stat}) \pm 13\,(\text{syst})\,\text{events} \tag{6.1}$$

after $200\,\text{pb}^{-1}$of integrated luminosity. The much larger number of signal events found compared to the case of 14 TeV must be ascribed not only to the integrated luminosity considered, which is the double of before, but also and especially to the softer cut on the missing energy variable, in this case MHT, which was used in order to gain statistics on the tails of the simulated samples. All the correlations between uncertainties on different background sources have been accounted for. This number has to be compared with the signal number of events obtained from the selection (Tab. 6.5) together with instrumental and theoretical systematic errors.

In the absence of an excess over the dominant invisible background, an upper limit on the M_D parameter can be calculated in the $200\,\text{pb}^{-1}$ data-set. The Profile Likelihood approach [88] is chosen, where the likelihood function is derived by a Poisson distribution for the total number of observed events, multiplied by a Gaussian with N_B as mean and the total background error as sigma. Nuisance

	$n=2$			$n=4$	
	$M_D=1$	$M_D=2$	$M_D=3$	$M_D=1$	$M_D=2$
95% C.L.	0.2	1.1	11	0.5	5.0
5 σ	2	14	163	6.0	69

Table 6.8: Minimum integrated luminosity (pb^{-1}) needed for a 95% C.L. exclusion or a 5 σ discovery, for different ADD points. Fundamental mass values around 1 TeV are already excluded by Tevatron (Tab. 5.1) and can be cross-checked after few integrated pb^{-1}.

parameters due to systematic error on signal are introduced, by assuming a Gaussian prior and convolving with the likelihood function. The 95% C.L. limit was found by scanning the parameter space to minimize the Negative Log Likelihood. Scanning has been repeated for 5 different benchmark points for ADD and the minimum integrated luminosity needed is quoted in Tab. 6.8. When different M_D, n are interpolated, the exclusion plot in Fig. 6.9 is derived. It can be used to indicate the fundamental mass exclusion limits as a function of integrated luminosity.

Since the sensitivity of the search extends significantly the previously established limits on the fundamental scale, it is possible that we will see an evidence for a signal already in the very first LHC run. Therefore, we need to also consider signal discovery and estimate the integrated luminosity necessary to establish the signal with a certain confidence level. A significance estimator based again on Profile Likelihood has been exploited to measure the amount of data needed for a 5 σ discovery and results are reported in Tab. 6.8. When results from some benchmark signals considered are interpolated, a plot like Fig. 6.10 is obtained, that represents the sensitivity for a discovery after $200\,\mathrm{pb}^{-1}$.

These results indicate that the current exclusion limits from Tevatron experiments (Tab. 5.1) can be obtained at LHC with the first physics run. Increasing the amount of data will allow to further constrain both n and M_D: after $200\,\mathrm{pb}^{-1}$, M_D can be excluded better than 3.8(3.2) TeV for $n=2(4)$. Also early discoveries for $n=2(4)$ scenarios are possible, if M_D is below 3.1(2.3) TeV, respectively.

6.8 Conclusions for 10 TeV case

A simulation study of the ADD model in the G+jet channel with a Monojet plus MET signature has been performed with the CMS detector. The previous analysis for a center of mass energy of 14 TeV has been updated to the machine condition expected for 2009/2010 LHC run ($\sqrt{s}$=10 TeV) and extended to $200\,\mathrm{pb}^{-1}$ of data.

Unlike the former analysis, based on a standard missing energy measurement

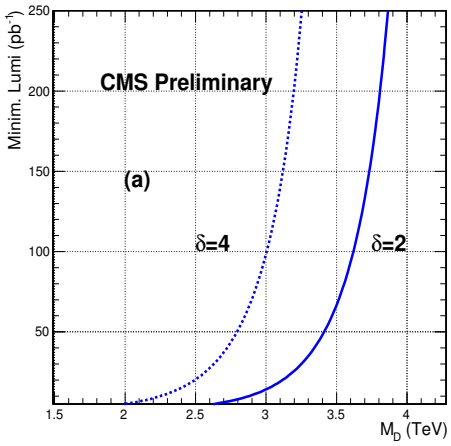


Figure 6.9: Exclusion plot at 95% C.L., showing the minimum luminosity necessary to exclude a given value of M_D. Sensitivity is plotted for two different extra dimension scenarios. In the plot δ is the number of extra dimension, referred as n in the text.

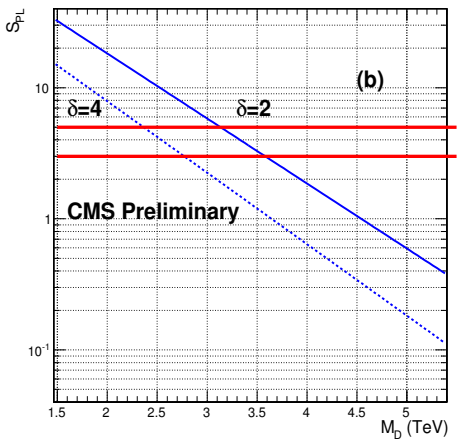


Figure 6.10: Discovery potential of the analysis as a function of M_D and n. Significance estimator S_{PL} [88] is adopted and the assumed integrated luminosity is $200\,\mathrm{pb}^{-1}$. The horizontal thick lines correspond to 3 σ and 5 σ significance level. In the plot δ is the number of extra dimension, referred as n in the text.

from the calorimeters of the detector, here the vectorial sum of corrected jets is exploited. This kinematic variable was found to be significantly more robust for first data stage and considered the best choice for the very early collisions. Other analysis criteria, as jet selection, event topology, and lepton veto have been chosen to maximize rejection for the SM reducible backgrounds (top production, multi-jet, and $W(l\nu)$+ jets). For the irreducible $Z(\nu\nu)$+ jets, the usage of a control region with $W(\mu\nu)$+ jets events was demonstrated to work.

An ADD parameter scan has been performed in order to calculate the CMS sensitivity to the studied model. A 5 σ discovery for a MET+1 jet signal can be obtained for values of the fundamental scale M_D lower than 3.1(2.3) TeV for n =2(4), while 95% C.L. exclusion limits for $M_D = 3\,\text{TeV}$, $n = 2$, $M_D = 2\,\text{TeV}$, $n = 4$ can be reached after only $11\,\text{pb}^{-1}$ and $5.0\,\text{pb}^{-1}$, respectively.

Chapter 7

Conclusions

The Standard Model (SM) is a spectacular scientific achievement, tested at a high level of precision. Despite the beautifulness of its principles and its experimental reliability, the SM is an effective theory valid up to a cut-off energy, affected by a problem known as hierarchy problem. Many are the proposed model capable to solve the hierarchy problem and to answer the several questions arising from the SM: Technicolor, Grand Unified Theory, Superstrings, Compositeness, Extra Dimensions, Extra Color, and Contact Interactions (CIs), just to mention a few. The European Organization for Nuclear Research (CERN) has now concluded the construction of the Large Hadron Collider (LHC) and ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) detectors, and will enable the testing of these models at unprecedentedly explored collision energies.

Awaiting the start-up of the LHC machines, CMS Collaborators are preparing for the data taking and analysis. In this thesis we have presented a study of Exotic physics and Extra Dimensions, in 3 different analyses and 2 different channels.

CMS Collaboration will use 2 different techniques, among others, for dijet searches: the inclusive jet rate versus jet p_T in the search of CIs and the dijet rate versus dijet mass in the search of resonances. The achievement obtained during the work of this thesis are several: We have repeated the basic plots in the dijet resonance analysis of Physics TDR II, and we have compared some basic distributions between ORCA and CMSSW. We find reasonable levels of agreement. Figure 4.6 shows the differential cross section versus dijet mass, where both leading jets have $|\eta| < 1$, and the mass bins have a width roughly equal to the dijet mass resolution. Considering first the QCD processes, the cross section for corrected jets agrees with the QCD prediction from generated jets. To determine the background shape either the Monte Carlo prediction or a parameterized fit to the data can be used. Fig. 4.10 shows a simulation of narrow dijet resonances with a q^* production cross section. This is compared to the statistical uncertainties in the QCD prediction, including trigger prescaling. This comparison shows that

with an integrated luminosity of 100 pb^{-1} a q^* dijet resonance with a mass of 2 TeV would produce a convincing signal above the statistical uncertainties from the QCD background. For comparison, a Tevatron search has excluded q^* dijet resonances with mass, M, below 0.87 TeV. The heaviest dijet resonances that CMS can discover (at five standard deviations) with 100 pb^{-1} of integrated luminosity, using this search technique and including the expected systematic uncertainties, are: 2.5 TeV for q^*, 2.2 TeV for Axigluons or Colorons, 2.0 TeV for E_6 Diquarks, and 1.5 TeV for Color Octet Technirhos. Studies of the jet η cut have concluded that the optimal sensitivity to new physics is achieved with $|\eta| < 1.3$ for a 2 TeV spin 1 dijet resonance decaying to $q\bar{q}$. In conclusion, CMS plans to use measurements of rate as a function of jet p_T and dijet mass to search for new physics in the data sample collected during early LHC running. With integrated luminosity samples in the range 10-100 pb^{-1}, CMS will be sensitive to CI's and dijet resonances beyond those currently excluded by the Tevatron.

For what concern the Monojet channel, a simulation study on the ADD model in the G+jet channel with a Monojet plus missing transverse energy (MET) signature has been performed with the CMS detector, focusing on the conditions expected with 100 pb^{-1} of data. Simple selection cuts based on the signal events topology, MET and jet energy have been used to reduce the SM backgrounds (mainly QCD, $t\bar{t}$ and $W(l\nu)$+ jets), along with a lepton veto algorithm against the $W(\mu/\tau\nu)$+ jets rejection. All cuts thresholds have been tuned in order to obtain the better signal over reducible background ratio. An HT and MHT based trigger stream has been adopted. Strategies to infer the irreducible $Z(\nu\nu)$+ jets, $W(l\nu)$+ jets and QCD background processes have been deployed. They give estimates consistent with Monte Carlo and are effective even at the early data-taking stage. An ADD parameter scan has been performed in order to calculate the CMS sensitivity to the studied model. A first evidence for a MET+1 jet signal can be obtained for values of the fundamental scale M_D lower than 3.30(2.37) TeV for $n = 2(4)$, while 95% C.L. exclusion limits are expected to be 4.18(3.04) TeV for $n = 2(4)$. Current limits from Tevatron are exceeded by almost a factor 3, leaving large room to explore gravity scale with the early LHC.

The analysis for a center of mass energy of 14 TeV has been updated then to the machine condition expected for 2009/2010 LHC run ($\sqrt{s}$=10 TeV) and extended to 200 pb^{-1} of data. Unlike the former analysis, based on a standard missing energy measurement from the calorimeters of the detector, the vectorial sum of corrected jets were exploited. This kinematic variable was found to be significantly more robust for first data stage and considered the best choice for the very early collisions. Other analysis criteria, as jet selection, event topology, and lepton veto have been chosen to maximize rejection for the SM reducible backgrounds (top production, multi-jet, and $W(l\nu)$+ jets). For the irreducible $Z(\nu\nu)$+ jets, the usage of a control region with $W(\mu\nu)$+ jets events was demonstrated to work. An

ADD parameter scan has been performed in order to calculate the CMS sensitivity to the studied model. A 5σ discovery for a MET+1 jet signal can be obtained for values of the fundamental scale M_D lower than 3.1(2.3) TeV for n =2(4), while 95% C.L. exclusion limits for $M_D = 3\,\text{TeV}$, $n = 2$, $M_D = 2\,\text{TeV}$, $n = 4$ can be reached after only $11\,\text{pb}^{-1}$ and $5.0\,\text{pb}^{-1}$, respectively.

The work related to this dissertation has been object of several articles from the same author of the thesis [89].

Bibliography

[1] N. Arkani-Hamed, S. Dimopoulos, G. Dval, Phys. Lett. **B 429**, 263.

[2] V. A. Rubakov, "Large and infinite extra dimensions: An introduction," Phys. Usp. **44**, 871 (2001) [Usp. Fiz. Nauk **171**, 913 (2001)] hep-ph/0104152.

[3] R. Rattazzi, "Cargese lectures on extra dimensions," hep-ph/0607055.

[4] D.J. Kapner *et al*, Phys. Rev. Lett **98**, 021101 (2007).

[5] See, e.g., V. Rubakov and M. Shaposhnikov, Phys. Lett. **125B** (1983) 136; and G. Dvali and M. Shifman, Nucl. Phys. **B504** (1997) 127.

[6] C. Csaki, J. Hubisz and P. Meade, "Electroweak symmetry breaking from extra dimensions," hep-ph/0510275.

[7] Agashe *et al*, Phys. Rev. D 77, 015003 (2008).

[8]

[9]

[10] G. F. Giudice, R. Rattazzi, J. D. Wells, Nucl. Phys. **B544**, 3-38,1999, "Quantum gravity and extra dimensions at high-energy colliders".

[11] CMS Collaboration, The Tracker Project, Technical Design Report, CERN/LHCC 98-6, CMS TDR 5 (15. April 1998).

[12] CMS Collaboration, CMS ECAL Technical Design Report, CERN/LHCC 97-33, CMS TDR 4, 15 December 1997.

[13] CMS Collaboration, The HCAL Technical Design Report, CERN/LHCC 97-31, CMS TDR 2, 20 June 1997.

[14] S. Abdullin et al. Eur. Phys. J. **C55** (2008) 159-171.

[15] S. Abdullin et al. Eur. Phys. J. **C53** (2008) 139-166.

[16] CMS Collaboration, CMS MUON Technical Design Report, CERN/LHCC 97-32, CMS TDR 3, 15 December 1997.

[17] CMS Collaboration, CMS Tridas Project, Technical Design Report, The trigger system, CERN/LHCC 2000 - 38, CMS TDR 6.1, December 15, 2000.

[18] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideFrameWork.

[19] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideFastSimulation .

[20] K. Lane and S. Mrenna, "The collider phenomenology of technihadrons in the technicolor Straw Man Model," Phys. Rev. D **67**, 115011 (2003) hep-ph/0210299.

[21] E. Eichten, I. Hinchliffe, K. D. Lane and C. Quigg, "Super Collider Physics," Rev. Mod. Phys. **56**, 579 (1984) [Addendum-ibid. **58**, 1065 (1986)].

[22] J. L. Hewett and T. G. Rizzo, "Low-Energy Phenomenology of Superstring Inspired E(6) Models," Phys. Rept. **183**, 193 (1989).

[23] U. Baur, I. Hinchliffe and D. Zeppenfeld, "Excited Quark Production At Hadron Colliders," Int. J. Mod. Phys. A **2**, 1285 (1987).

[24] L. Randall and R. Sundrum, "A large mass hierarchy from a small extra dimension," Phys. Rev. Lett. **83**, 3370 (1999) hep-ph/9905221.

[25] J. Bagger, C. Schmidt and S. King, "AXIGLUON PRODUCTION IN HADRONIC COLLISIONS," Phys. Rev. D **37**, 1188 (1988).

[26] E. H. Simmons, "Coloron phenomenology," Phys. Rev. D **55**, 1678 (1997) hep-ph/9608269.

[27] E. Eichten, K. D. Lane and M. E. Peskin, "New Tests For Quark And Lepton Substructure," Phys. Rev. Lett. **50**, 811 (1983).

[28] K. D. Lane, "Electroweak and flavor dynamics at hadron colliders," hep-ph/9605257.

[29] E. Eichten, K. Lane and M. E. Peskin, Phys. Rev. Lett. **50**, 811 (1983).

[30] U. Baur, I. Hinchliffe and D. Zeppenfeld, Int. J. Mod. Phys. **A2**, 1285 (1987).

[31] T. Sjostrand, L. Lonnblad, S. Mrenna, hep-ph/0108264.

[32] [CMS Collaboration], CMS PAS JME-07-003, "Performance of Jet Algorithms in CMS".

[33] CMS Collaboration, CMS PAS JME-08-002, "Jet Corrections to Parent Parton Energy".

[34] Acosta, D. and others, CMS, CMS Physics Technical Design Report, V.1 : Detector performance and software, CERN, Geneva, CERN-LHCC-2006-001, 2006.

[35] K. Gumus, N. Akchurin, S. Esen and R. Harris, CMS Note 2006/070, "CMS Sensitivity to Dijet Resonances", May 2006.

[36] CMS Collaboration, CERN-LHCC-2006-021; CMS-TDR-008-2, "CMS Physics Technical Design Report Volume 2 : Physics Performance", August 2006.

[37] S. Esen and R. Harris, CMS Note 2006/069, "Jet Triggers and Dijet Mass", May 2006.

[38] http://cmsdoc.cern.ch/orca .

[39] S. Agostinelli *et al.* (GEANT4 Collaboration), Nucl.Instrum. Meth. A**506**, 250 (2003).

[40] CMS Collaboration, Physics TDR Volume I, CERN-LHCC-2006-001.

[41] CMS Collaboration 2008 *CMS Physics Analysis Summary* **JME-07-003** at http://cms-physics.web.cern.ch/cms-physics/public/JME-07-003-pas.pdf .

[42] S. Esen and R. Harris, Jet Triggers and Dijet Mass, CMS Note 2006/069.

[43] See for instance: F. Abe *et al.* [CDF Collaboration], "Search for new particles decaying to dijets at CDF," Phys. Rev. D **55**, 5263 (1997) hep-ex/9702004.

[44] See for instance: V. M. Abazov *et al.* [D0 Collaboration], "Search for new particles in the two-jet decay channel with the DØ detector," Phys. Rev. D **69**, 111101 (2004) hep-ex/0308033.

[45] CDF Collaboration, F. Abe et al., Search for New Particles Decaying to Dijets at CDF, Phys. Rev. D55: 5263-5268 (1997), hep-ex/9702004.

[46] B. Abbott *et al.* [D0 Collaboration], "The dijet mass spectrum and a search for quark compositeness in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV," Phys. Rev. Lett. **82**, 2457 (1999) [arXiv:hep-ex/9807014].

[47] J. Pumplin *et al.* (CTEQ Collaboration), JHEP **07**, 012 (2002).

[48] B. Abbott *et al.* (D0 Collaboration), Phys. Rev. Lett. **82**, 2457 (1999).

[49] Hatakeyama K *et al.* (CDF Collaboration) 2008 recent public results at http://www-cdf.fnal.gov/physics/exotic/r2a/20080214.mjj_resonance_1b/ .

[50] Bayatian, G. L. and others, CMS, CMS technical design report, volume II: Physics performance, J. Phys., G34, 2007, 995-1579.

[51] J. Bagger, C. Schmidt and S. King, Phys. Rev. D**37**, 1188 (1988).

[52] E. Simmons, Phys. Rev. D**55**, 1678 (1997).

[53] J. Hewett and T. Rizzo, Phys. Rep. **183**, 193 (1989) and references therein.

[54] K. Lane and S. Mrenna, Phys. Rev. D**67**, 115011 (2003).

[55] CMS Collaboration, 2007 CMS PAS SBM-07-001, at http://cms-physics.web.cern.ch/cms-physics/public/SBM-07-001-pas.pdf .

[56] C. D. Hoyle *et al.*, Phys. Rev. Lett. **86**, 1418, "Submillimeter Test of the Gravitational Inverse-Square Law: A Search for Large Extra Dimensions".

[57] V. H. Satheeshkumar and P. K. Suresh, astro-ph/0805.3429v1, "Bounds on large extra dimensions from photon fusion process in SN1987A".

[58] [CDF Collaboration], hep-ex/0807.3132v1, "Search for large extra dimensions in final states containing one photon or jet and large missing transverse energy produced in p-pbar collisions at sqrt(s) = 1.96 TeV", FERMILAB-PUB-08-247-E, Submitted to Phys. Rev. Lett.

[59] DØ Collaboration, Phys. Rev. Lett. **90**, 251802 (2003), "Search for Large Extra Dimensions in the Monojet + Missing ET Channel at D0".

[60] [D0 Collaboration], Phys. Rev. Lett. **101**, 011601 (2008), "Search for Large Extra Dimensions via Single Photon plus Missing Energy Final States at sqrt(s) = 1.96 TeV".

[61] [DELPHI Collaboration], Eur. Phys. J. **C 38**, 395 (2005), "Photon Events with Missing Energy in e+e- Collisions at sqrts = 130 to 209 GeV";[L3 Collaboration], Phys. Lett. **B 587**, 16 (2004), "Single- and Multi-Photon Events with Missing Energy in e+e- Collisions at LEP".

[62] V. Krutelyov, hep-ex/0807.0645v1, "Searches for Large Extra Dimensions at the Tevatron".

[63] L. Vacavant, Eur. Phys. J. **C33**, S924 (2004), "Search for Extra Dimensions at LHC".

[64] T. Gleisberg, S. Hoeche, F. Krauss *et al.* JHEP 0402, 056 (2004), "SHERPA 1.alpha, a proof-of-concept version", http://www.sherpa-mc.de.

[65] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP 0307, 001 (2003), "ALPGEN, a generator for hard multiparton processes in hadronic collisions".

[66] T. Sjöstrand *et al.* Comput. Phys. Commun. 135 **238** (2001), "High-Energy-Physics Event Generation with PYTHIA 6.1".

[67] https://twiki.cern.ch/twiki/bin/view/CMS/CSA07.

[68] [CMS Collaboration], CMS PAS SUS-08-05.

[69] https://twiki.cern.ch/twiki/bin/view/CMS/CSA07GeneratorInformation.

[70] TS. Frixione, B. R. Webber, JHEP 0206, 029 (2002), "Matching NLO QCD computations and parton shower simulations".

[71] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuidePAT.

[72] [CMS Collaboration], CERN/LHCC/2006-001, CMS TDR 8.1, 2 February 2006; CERN/LHCC/2007-009, CMS TDR 8.2-Add1, 5 March 2007.

[73] M. Spiropulu, T. Yetkin, CMS AN-2006/089, "Inclusive Missing Transverse Energy + multijet SUSY search at CMS".

[74] [CMS Collaboration], CERN-LHCC 2006/021 (2006), "CMS Physics Technical Design Report Volume II: Physics Performance".

[75] [CMS Collaboration], CMS PAS SUS-08-002, "Data-Driven Estimation of the Invisible Z Background to the Early SUSY MET Plus Jets Search".

[76] W.-M. Yao *et al.* [Particle Data Group], Journ. of Phys. G **33**, 1 (2006), http://pdg.lbl.gov.

[77] [CMS Collaboration], CERN CMS PAS-EWK-07-002, "Towards a measurement of the inclusive $W(\mu\nu)$ and $Z(\mu\mu)$ cross section in pp collisions at $\sqrt{s} = 14\,\mathrm{TeV}$".

[78] https://twiki.cern.ch/twiki/bin/view/CMS/HamburgWikiAnalysisSusy and http://indico.cern.ch/conferenceDisplay.py?confId=37172.

[79] [CDF Collaboration], hep-ex/0605101v1, "Search for Large Extra Dimensions in the Production of Jets and Missing Transverse Energy in $p\bar{p}$ Collisions at $\sqrt{s}$ =1.96 TeV".

[80] M. Beneke *et al.*, hep-ph/0003033, "Top Quark Physics".

[81] [CMS Collaboration], CERN CMS PAS-TOP-08-002, "Di-lepton ttbar cross section with 100 pb‘$^{-}$1".

[82] S. I. Bityukov and N. V. Krasnikov, hep-ph/0204326, "Uncertainties and discovery potential in planned experiments".

[83] http://indico.cern.ch/conferenceDisplay.py?confId=30773#3 and http://indico.cern.ch/conferenceDisplay.py?confId=37479#4.

[84] J. Pumplin *et al.*, JHEP **07** 012, (2002) hep-ph/0201195 "New generation of parton distributions with uncertainties from global QCDanalysis".

[85] L. Benucci, M. Cardaci, A. De Roeck, U. Emrah Surat, L. Sala, P. Van Mechelen, CMS AN-2008/073, "Search for Mono-Jet Final States".

[86] https://twiki.cern.ch/twiki/bin/view/CMS/ProductionSummer2008.

[87] CMS Collaboration, CMS PAS EXO-08-011, "Search of extra dimension with monojets".

[88] R. D. Cousins, J. T. Linnemann, J. Tucker, NIM **A595** 480, (2008).

[89] M. Cardaci *et al.*, "CMS search plans and sensitivity to new physics with dijets", 2009 *J. Phys. G: Nucl. Part. Phys.* **36** 015004 (6pp).

M. Cardaci *et al.*, "Search for Mono-Jet Final States from ADD Extra-Dimensions at 10 TeV", CMS PAS EXO-09-013.

M. Cardaci *et al.*, "Search for Mono-Jet Final States from ADD Extra Dimensions", CMS PAS EXO-08-011.

M. Cardaci *et al.*, "CMS Search Plans and Sensitivity to New Physics using Dijets", CMS PAS SBM-07-001.

M. Cardaci for the CMS Collaboration, "Extra Dimensions in Photon or Jet plus Missing Transverse Energy" (SUSY 09 talk report), CMS CR-2009/305

M. Cardaci for the CMS Collaboration, "Extra Dimensions in Photon or Jet plus Missing Transverse Energy" (EPS09 poster report), CMS CR-2009/244.

M. Cardaci *et al*, "Search for Mono-Jet from ADD Extra-Dimensions $\sqrt{s}$ at = 10 TeV", CMS AN-2009/109.

M. Cardaci *et al.*, "Search for Mono-Jet from ADD Extra-Dimensions in different LHC energy scenarios", CMS AN-2009/003.

M. Cardaci *et al.*, "Search for Mono-Jet Final States from ADD Extra-Dimensions", CMS AN-2008/073.

M. Cardaci for the CMS Collaboration, "SEARCHES FOR NEW PHYSICS USING DIJET EVENTS AT THE LHC", CMS CR-2008/050.

M. Cardaci *et al.*, "CMS Search Plans and Sensitivity to New Physics using Dijets", CMS NOTE-2008/019.

M. Cardaci et al., "CMS Search Plans and Sensitivity to New Physics using Dijets", CMS AN-2007/039.

M. Cardaci, B. Bollen and R. Harris, "Dijet Resonance Analysis with CMSSW 1.2.0", CMS AN-2007/016.