



Université catholique de Louvain
Faculté des Sciences
Département de Physique
Unité de physique nucléaire
Centre for Particle Physics and Phenomenology

Observability of an unconventional two-Higgs-doublet model at the LHC

Doctoral dissertation presented by
Simon de Visscher
for the degree of Doctor in Sciences

| | |
|--------------------------------|------------------------|
| Pr. Vincent Lemaitre (Advisor) | UCL |
| Pr. Denis Favart (Chairman) | UCL |
| Pr. Giacomo Bruno | UCL |
| Pr. Roberto Chierici | IN2P3 Lyon |
| Pr. Günther Dissertori | ETH Zurich |
| Pr. Fabio Maltoni | UCL |
| Pr. Nick van Remortel | Universiteit Antwerpen |

October 2009

Abstract

The Brout-Englert-Higgs mechanism proposed in the 60's predicts the existence of the Higgs boson, the last undiscovered elementary particle of the Standard Model of fundamental interactions. It is based on the presence of one doublet field governing the scalar potential of the theory. Alternative models predict more scalar doublet fields and a corresponding increase of the number of Higgs particles. The observation of these Higgs bosons is one of the primary goals of the Large Hadron Collider, whose first collisions are expected in 2010. In this context, particular attention is given in this work on the quality of hadron collision modeling. This includes multi-jet final states which are particularly dangerous backgrounds to many searches at the LHC as well as heavy colored particle production which could be the sign of new physics.

In this thesis an unconventional realization of the two-Higgs-doublet model is presented, which is mainly characterized by an inverted mass hierarchy of the resulting Higgs particles. It is shown that this model cannot be excluded by existing theoretical and experimental constraints and that it could be observed at the LHC via a large variety of experimental signatures in CMS and ATLAS detectors.

In particular, it is shown that the pseudo scalar Higgs particle, the A boson, could be discovered in CMS only after a few inverse femptobarns of integrated luminosity via the production of another Higgs particle decaying into Z and A bosons which subsequently decay into charged leptons.

à la mémoire d'Olivier Garcet, mon ami.

CONTENTS

| | |
|---|-----------|
| Introduction | ix |
| 1 The scalar sectors in the Standard Model and two-Higgs-doublet model | 1 |
| 1.1 The scalar sector of the Standard Model | 3 |
| 1.1.1 Introduction | 3 |
| 1.1.2 Constraints on the SM Higgs mass | 6 |
| 1.1.3 Higgs searches at the LHC | 9 |
| 1.2 The two-Higgs-doublet model and the inverted mass spectrum | 12 |
| 1.2.1 Introduction | 12 |
| 1.2.2 Theoretical and indirect constraints | 15 |
| 1.2.3 Decays of Higgs bosons | 23 |
| 1.2.4 Direct constraints | 27 |
| 1.3 Conclusion | 29 |
| 2 The LHC and the modeling of collision | 31 |
| 2.1 The LHC | 32 |
| 2.1.1 The installation | 32 |
| 2.1.2 p-p interactions | 34 |
| 2.2 Monte Carlo simulation of collisions | 36 |

| | | |
|----------|---|-----------|
| 2.2.1 | Factorization of hard and soft processes | 36 |
| 2.2.2 | Matching/merging of initial state radiation | 38 |
| 2.2.3 | Control of the matching parameters | 41 |
| 2.2.4 | Multi-jet event production in the SM | 46 |
| 2.3 | Multi-jet event production with heavy colored particles | 49 |
| 2.3.1 | Anatomy of E_T^{miss} + multi-jet final states | 52 |
| 2.3.2 | Observability of gluinos and squarks at the LHC | 56 |
| 2.4 | Monte Carlo systematics for b -jet detection | 59 |
| 2.5 | Conclusion | 64 |
| 3 | Studies in the iM2HDM at the LHC, the Higgs hunting | 65 |
| 3.1 | Introduction | 65 |
| 3.2 | Event simulation and reconstruction | 67 |
| 3.3 | Discovery potential of Type I i M2HDM | 70 |
| 3.3.1 | A single process for all neutral Higgs bosons observation | 70 |
| 3.3.2 | Possible processes for charged Higgs bosons observation | 73 |
| 3.4 | Discovery potential of Type II i M2HDM | 78 |
| 3.4.1 | From SM-like h observation | 78 |
| 3.4.2 | A process for H^0 and A bosons observation | 79 |
| 3.4.3 | A process for H^\pm observation | 84 |
| 3.5 | Conclusion | 89 |
| 4 | Study of the $pp \rightarrow H^0 \rightarrow ZA$ in CMS | 91 |
| 4.1 | Event Simulation | 91 |
| 4.2 | Event detection and reconstruction | 93 |
| 4.2.1 | The CMS detector | 95 |
| 4.2.2 | CMSSW | 102 |
| 4.3 | Selection of leptons from Z and A decays | 105 |

| | | |
|-------|--|------------|
| 4.3.1 | Final states and Trigger condition | 105 |
| 4.3.2 | Lepton identification | 108 |
| 4.4 | Cut based analysis | 114 |
| 4.4.1 | Basic cuts | 114 |
| 4.4.2 | Topological cuts | 116 |
| 4.4.3 | Probing a possible deviation from the Standard Model | 125 |
| 4.5 | Data-driven estimation of the reducible backgrounds | 126 |
| 4.5.1 | Normalization | 128 |
| 4.5.2 | Control plots of key variables | 129 |
| 4.6 | Analysis of the benchmark point BP2 | 131 |
| 4.6.1 | Neural net based selection | 131 |
| 4.6.2 | Systematics | 134 |
| 4.6.3 | An estimation of the significance and mass resolution. | 137 |
| 4.7 | Conclusion | 138 |
| | Conclusion | 141 |
| | A Delphes configuration card | 143 |
| | B MatchChecker | 145 |
| B.1 | Introduction | 145 |
| B.2 | Running MatchChecker | 147 |

Introduction

IN the next few years, the Large Hadron Collider (LHC) is expected to lead us towards a better comprehension of high-energy physics up to the TeV scale. Despite the success of Standard Model in the 80's and 90's (*e.g.* confirmation of the existence of the weak currents and direct observation of top quark) some questions of the first importance are left open like the origin of the mass of particles, the mass hierarchy of fermions, the origin of the matter-antimatter asymmetry. The commonly considered solution to explain how particles acquire a mass is the Brout-Englert-Higgs mechanism. It consists in a spontaneous symmetry breaking of a scalar potential, predicting the existence of the so-called Higgs boson(s), still unobserved.

While the Standard Model predicts the presence of one of these scalars, its extensions contain more complex Higgs sectors. The supersymmetric models are such commonly considered solutions, as they have the interesting properties of solving the hierarchy problem and tends to give a good behavior to the coupling constants convergence at a larger scale. However, beside this apparent elegance, the number of degrees of freedom explodes literally with the presence of supersymmetric particles, whose masses are unknown. In this work we prefer to concentrate on a two-Higgs doublet model, not supersymmetric, with a minimal set of free parameters, and characterized by an inverted scalar mass-spectrum compared to the minimal supersymmetric model. The first interest of this model is to provide unusual experimental signatures at the LHC and being still viable by theoretical and experimental constrains.

In the first chapter the Brout-Englert-Higgs mechanism is first presented in the context of the Standard Model. The theoretical, indirect and direct constraints are briefly reviewed. The discussion is then extended to a CP -conserving two-Higgs-doublet model with a twisted custodial symmetry present in its potential. The interest of such a symmetry is to allow a particular scalar sector, with a heavy SM -like Higgs boson h , a degenerated triplet formed by H^\pm and the CP -even H^0 boson and finally a light

pseudoscalar A . As for the Standard Model the result of theoretical, indirect and direct constraints are given.

Before reviewing in details how the scalar sector can be studied at the LHC, in the second chapter we discuss the importance of the events simulation, especially from the point of view of the QCD activity from initial state radiation. For about 10 years, many improvements have been done to get more precise distribution shapes and cross-sections predictions. This is done by merging the matrix-element and parton shower descriptions and by calculating processes production up to next-to-leading orders. In this work we investigate what are the impacts of jet matching techniques in the SM and beyond.

While the first chapter describes the two-Higgs-doublet model from a theoretical point of view, the third chapter reviews the reactions which could be studied for the Higgs boson(s) discoveries at the LHC. The discovery potential of three of these reactions is discussed in more details, using simple analyses and a fast detector simulation. These processes are $pp \rightarrow H^0 \rightarrow ZA$, $pp \rightarrow tH^- \rightarrow tW^- A$ and $pp \rightarrow h \rightarrow ZAZA$.

The reaction $pp \rightarrow H^0 \rightarrow ZA$ followed by the leptonic decay of the Z and A bosons appears to be particularly promising and is the subject of the fourth chapter. The discovery potential of this reaction is evaluated by using the official detector simulation and reconstruction tool of CMS. The analysis of this reaction is realized in whole allowed mass range for the Higgs bosons masses and allows to evaluate the required integrated luminosity to reject the SM hypothesis. The data-driven control of the reducible background processes is also discussed and the study of a particular benchmark point is achieved by means of a more sophisticated analyses including neural network techniques and effects of systematic errors.

The work presented in this thesis is essentially based on the following publications:

- “Unconventional phenomenology of a minimal two-Higgs-doublet model”. S. de Visscher, J-M. Gerard, M. Herquet, V. Lemaitre, F. Maltoni, *JHEP* **08** (2009) 042, arXiv:0904.0705 [hep-ph]
- “QCD radiation in the production of heavy colored particles at the LHC” J. Alwall, S. de Visscher and F. Maltoni *JHEP* **02** (2009) 017, arXiv:0810.5350 [hep-ph]

The scalar sectors in the Standard Model and two-Higgs-doublet model

THE Standard Model (SM) is the successful realization of several decades of research in high-energy physics, both at the theoretical and experimental level. It largely describes the fundamental interactions between particles [1] content as well as their interactions.

The fermions, often referenced as matter fields, have a $1/2$ - spin and are divided themselves into two types. The quarks which carry the electromagnetic, weak and strong charges and the leptons which only carry the electromagnetic and weak charges¹. The fermion electric charges and masses in the SM are summarized in Tab 1.1.

In the Standard Model, interactions result from local invariance of the lagrangian under specific gauge groups. Interaction being mediated by virtual particle exchanges, for each local gauge symmetry, a number of mediators (bosons) is predicted according to the dimension of the corresponding group. These bosons carry an integer spin. There are eight massless gluons for the strong interaction, the massive Z ($91.1786 \text{ GeV}/c^2$) and W^\pm ($80.403 \text{ GeV}/c^2$) bosons (Ref. [2]) for the weak interaction, and finally the massless photon for the electromagnetic interaction.

¹Both types are also sensitive to gravity but the smallness of its strength is such that it is neglected in this work

| Quarks | | | Leptons | | |
|--------|--------|----------------------------------|------------|--------|--------------------|
| name | charge | mass | name | charge | mass |
| u | 2/3 | $(1.5-3.3) \times 10^{-3}$ | ν_e | 0 | < 2 eV |
| d | -1/3 | $(3.5-6) \times 10^{-3}$ | e | -1 | 0.511 keV |
| c | 2/3 | $1.27^{+0.07}_{-0.11}$ | ν_μ | 0 | < 0.19 MeV |
| s | -1/3 | $105^{+25}_{-35} \times 10^{-3}$ | μ | -1 | 0.105 |
| t | 2/3 | 171.3 ± 2.3 | ν_τ | 0 | < 18.2 MeV |
| b | -1/3 | $4.2^{+0.17}_{-0.07}$ | τ | -1 | 1776.99 ± 0.17 |

Table 1.1: List of fermions in the SM, with their respective charge and masses [2]. The mass are expressed in GeV/c^2 unless stated explicitly.

In the 60's Salam, Glashow and Weinberg [1] have suggested that the electromagnetic and weak interactions could be two manifestations of a single *electroweak* interaction, Resulting from the gauge symmetry group $SU(2)_L \times U(1)_Y$ where L recalls the maximal parity violation for the leptons (no right-handed neutrinos exist) and Y is the hypercharge. It implies the presence of four fields: B_μ for the group $U(1)_Y$, and three fields W_μ^i ($i=1,2,3$) for $SU(2)_L$. The corresponding lagrangien (we restrict the fermionic part to the e and ν) can be written as

$$\mathcal{L} = \bar{L}_L \gamma^\mu (i\partial_\mu - g \frac{\tau^i}{2} W_\mu^i + g' \frac{B_\mu}{2}) L_L + \bar{e}_R \gamma_\mu (i\partial_\mu + g' B_\mu) e_R - \frac{1}{4} W^{\mu\nu} W_{\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} \quad (1.1)$$

where the terms in parenthesis are the covariant derivatives required for the local gauge invariance, L_L is the isospin doublet containing the left-handed neutrino and electron, e_R is the right-handed electron, τ^i are the Pauli matrices, g and g' are the coupling constants for the weak isospin and hypercharge currents. Finally $W_{\mu\nu}^i$ and $B_{\mu\nu}$ are the kinetic energy terms for both categories of fields, with the stress tensors defined as

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g \epsilon^{ijk} W_\mu^j W_\nu^k \quad (1.2)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.3)$$

At this stage, no mass term $\propto m^2 W^{i\mu} W_\mu^i$ is present and even allowed. Indeed this would break the gauge invariance of \mathcal{L} . For the fermions, a mass term has the form $m \bar{\psi} \psi$ where $\psi = \psi_L + \psi_R$ with both left and right-handed components. Such couplings break the gauge invariance as well, so this version of the theory forbids massive fermions.

However, experiments tells us that the fermions as well as the W and Z gauge bosons have a mass. A solution for both problems may come from the Brout-Englert-Higgs

mechanism ([3, 4, 5]) that breaks spontaneously the symmetry $SU(2)_L \times U(1)_Y$ into $U(1)_{em}$, allowing to gauge bosons and fermions to acquire a mass.

We first review the main features of this mechanism in the context of the Standard Model as well as the theoretical and experimental constraints on the associated Higgs boson. We then focus on a particular realization of the two-Higgs-doublet-model characterized by the presence of a ‘‘twisted’’ custodial symmetry. The theoretical and experimental constraints are reviewed, defining the parameter space that should be considered for experimental analyses.

1.1 The scalar sector of the Standard Model

1.1.1 Introduction

The Brout-Englert-Higgs mechanism is realized in the SM with the contribution of a new doublet ϕ of complex scalar fields

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta_1 + i\eta_2 \\ \eta_3 + i\eta_4 \end{pmatrix}. \quad (1.4)$$

The lower component of the doublet is a neutral scalar field that ensures that the subgroup $U(1)_{em}$ remains unbroken.

The SM lagrangien (1.1) is extended with two new terms: \mathcal{L}_{Higgs} and \mathcal{L}_{Yukawa} containing ϕ . The first term reads:

$$\mathcal{L}_{Higgs} = |D_\mu \phi|^2 - \mu^2 \phi^2 - \lambda \phi^4. \quad (1.5)$$

where D is the covariant derivative

$$D = i\partial_\mu - g\frac{\tau_i}{2}W_\mu^i - \frac{g'}{2}B_\mu. \quad (1.6)$$

A potential of the form $\mu^2 x^2 + \lambda x^4$ has only two parameterization leading to the presence of a minimum if $\lambda > 0$. The position of the minimum is zero if $\mu^2 > 0$ while different from zero if $\mu^2 < 0$. In the second case, the potential has a continuum of minima in $\partial V/\partial \phi = 0$, hence in $\phi = \mu^2/2\lambda$. This is illustrated in a 1-dimensional case in Fig. 1.1.

To allows the conservation of the electric charge, the position of the minimum can be chosen as $\eta_1 = \eta_2 = \eta_4 = 0$ and $\eta_3 = v$ where v is the vacuum expectation value (v.e.v.). The radial excitation around the v.e.v. can be quantified with a new field h .

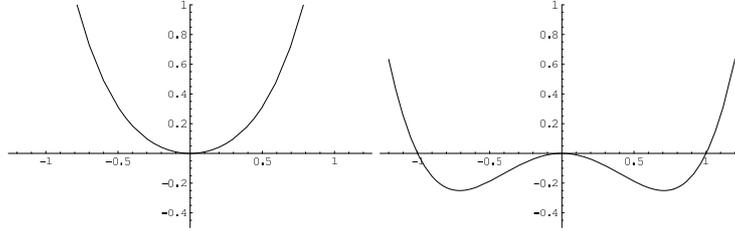


Figure 1.1: Shape of a potential of the form $\mu^2 x^2 + \lambda x^4$ with $\mu > 0$ (left) and $\mu < 0$ (right).

The variation of the potential in the η_3 direction therefore correspond to the generation of a massive Higgs boson. On the contrary, the orthogonal directions generate the massless Goldstone bosons.

The potential part of Eq. (1.5) can be rewritten using the minimum as

$$V = \frac{1}{2}\mu^2(v+h)^2 + \frac{1}{4}\lambda(v+h)^4. \quad (1.7)$$

As we have $\mu^2 = -\lambda v^2$, the potential reads

$$V = -\frac{1}{4}\lambda v^4 + \lambda v^2 h^2 + \lambda v h^3 + \frac{1}{4}\lambda h^4. \quad (1.8)$$

The mass related to the Higgs field is identified with $m_h^2 = 2v^2\lambda$, whereas the terms in h^3 and h^4 correspond to the couplings of the Higgs boson with itself.

From (1.5), and by introducing the fields $W_\mu^\pm = \frac{1}{2}(W_\mu^1 \mp W_\mu^2)$, one gets then

$$\begin{aligned} \left| \left(-g\frac{\tau_i}{2}W_\mu^i - \frac{g'}{2}B_\mu \right) \phi \right|^2 &= \frac{1}{8} \left| \begin{pmatrix} gW_\mu^3 + g'B_\mu & gW_\mu^1 - igW_\mu^2 \\ gW_\mu^1 - igW_\mu^2 & -gW_\mu^3 + g'B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v+h \end{pmatrix} \right|^2 \\ &= \frac{1}{2}(v+h)^2 gW^{+\mu}W_\mu^- + \\ &\quad \frac{(v+h)^2}{8} \left[g^2 W^{3\mu}W_\mu^3 - 2gg'B_\mu W^{3\mu} + g'^2 B_\mu B^\mu \right]. \end{aligned}$$

Two combination of the fields B and W correspond to the observed fields A (photon) and Z (Z boson):

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \quad (1.9)$$

$$Z_\mu = -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W. \quad (1.10)$$

with θ_W the Weinberg mixing angle. The identification of the mass terms of the gauge fields reads

$$m_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2}(m_Z^2 Z_\mu Z^\mu + m_A^2 A^\mu A_\mu). \quad (1.11)$$

It is then straightforward to show that

$$\tan \theta_W = \frac{g'}{g}, \quad (1.12)$$

and hence

$$m_W = \frac{1}{2}vg, \quad m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}, \quad m_A = 0. \quad (1.13)$$

From (1.13) the masses of the Z and W bosons are linked via the relation

$$\rho = \frac{m_W^2}{m_Z^2 \cos^2 \theta_W}, \quad (1.14)$$

that is exactly equal to 1 at the tree-level. This relation is very important as it is an observable measured experimentally very close to the unity.

The value of v can be found via the Fermi constant

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} = \frac{g^2}{8(\frac{1}{4}v^2g^2)} = \frac{1}{2v} \Rightarrow v \sim 246 \text{ GeV}/c^2, \quad (1.15)$$

with $G_F = 1.166353(9) \times 10^{-5} \text{ GeV}^{-2}$ [6]. The value of v gives the scale at which the $SU(2)_L \times U(1)_Y$ symmetry is broken into $U(1)_{em}$. However, since the value of the parameter λ is unknown, the Higgs boson mass is therefore a free parameter of the theory.

Note that the relation (1.5) defines also the possible couplings of the Higgs field to the gauge bosons. One then has the trilinear couplings hWW , hZZ and also four-legs vertices $hhZZ$ and $hhWW$.

Using the same scalar field ϕ , the fermions can also acquire a non-vanishing mass thanks to the Yukawa interaction. The lagrangian (invariant under $SU(2)_L$) can be written

$$\mathcal{L}_{Yukawa} = -\lambda_e \bar{L} \phi e_R - \lambda_d \bar{Q} \phi d_R - \lambda_u \bar{Q} \tilde{\phi} u_R + h.c. \quad (1.16)$$

with the isodoublet $\tilde{\phi} = iT_2 \phi^*$ needed to generate the up-quark masses. For instance, the Yukawa interactions with the τ and ν_τ is written, after spontaneous symmetry breaking as

$$\mathcal{L} = \frac{\lambda_\tau}{\sqrt{2}} \left[(\bar{\nu}_\tau, \bar{\tau})_L \begin{pmatrix} 0 \\ v+h \end{pmatrix} \tau_R + \bar{\tau}_R (0, v+h) \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \right]. \quad (1.17)$$

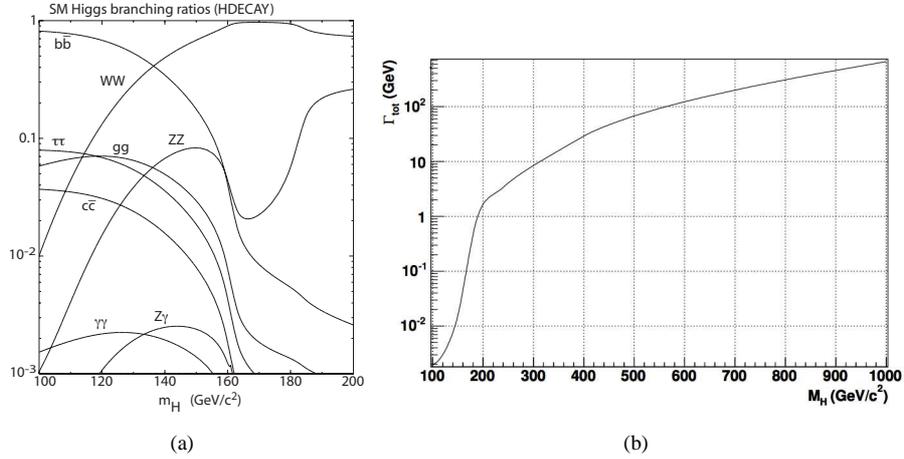


Figure 1.2: (a) Branching ratios of the Standard Model Higgs.(b) Evolution of the Higgs boson width with respect to its mass (from Ref. [7]).

The mass of the τ lepton is identified with $m_\tau = \frac{\lambda_\tau v}{\sqrt{2}}$ and the mass of the neutrino is zero.

Depending on its mass, the SM Higgs boson can decay in fermionic and/or bosonic modes. The evolution of the branching ratios with respect to m_h is shown in Fig. 1.2 shows. In the low mass regime, the Higgs boson decays preferentially into a pair of b quarks, whereas above $130 \text{ GeV}/c^2$ the decay into two W vector bosons becomes dominant, followed at the higher mass regions by the decay into two Z bosons.

As said previously the mass of the Higgs boson is an unknown parameter of the theory. However a range of possibility for m_h can be obtained thanks to a set of theoretical and experimental constraints.

1.1.2 Constraints on the SM Higgs mass

Triviality, vacuum stability, unitarity and perturbativity

Two first constraints on the Higgs boson are illustrated in Fig. 1.3(a) with respect to an energy cutoff Λ . This cutoff is defined as the energy before which no new physics effect can appear. The upper bound (red curve) comes from the *triviality* condition while the lower bound (green curve) is due to the *vacuum stability* condition.

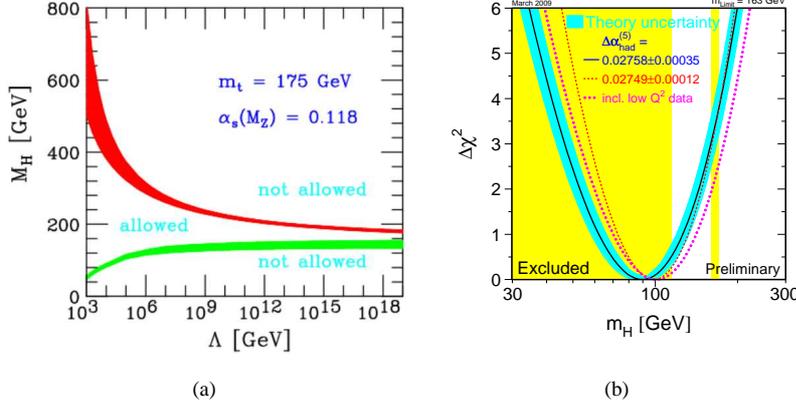


Figure 1.3: (a) Triviality (red) and stability (green) bounds [10]. (b) The result of the electroweak global fit expressed in $\Delta\chi^2 = \chi^2 - \chi^2_{min}$, with respect to the mass of the Higgs boson. [11].

The triviality² condition ensures that the radiative corrections of the λ parameter remain finite below the scale Λ . Close to the TeV scale, the triviality bound is $m_h \lesssim 700$ GeV/c² [8].

The existence of the vacuum expectation value (and its stability) has a meaning if the potential possesses a minimum. The vacuum stability condition is guaranteed if λ is positive at all energies and all all orders. The sign of λ depends on the cutoff Λ and the energy involved. It is shown that if $\Lambda \sim 1$ TeV, the Higgs mass range is bounded from below at around 70 GeV/c² [9].

An additional theoretical constraint is related to the violation of unitarity at high energies in electroweak processes like $W^+W^- \rightarrow W^+W^-$. Indeed the corresponding cross-section diverges as $\sim s/M_W^2$ where s is the squared energy involved in the process[12]. This can be solved by introducing the contribution of the Higgs particle in the diagrams. It has been calculated that this contribution is sufficient if the Higgs boson is lighter than 840 GeV/c²[13].

Finally if $\Gamma_h \sim m_h$, the perturbative expansion involving the Higgs boson does not hold any longer. Moreover, the particle appellation in such a context can also be questioned. From Fig. 1.2 (b), Γ_h becomes large compared to m_h at around 1 TeV. An approximative upper limit of $m_h \lesssim 700$ GeV/c² is in general considered here [14].

²The appellation comes from the possibility to consider $\lambda = 0$ at all orders and energies. This is equivalent to have a non-interacting model and therefore is not satisfying as no spontaneous symmetry breaking is possible.

Indirect and direct constraints

Since the 70's, the parameters of the Standard Model have been evaluated experimentally with a high accuracy and confronted to the theoretical predictions, notably with the calculation of radiative corrections to these parameters.

The level of precision reached by measurements allows to be sensitive to the presence of the Higgs boson in radiative corrections. The comparison between data and the theoretical prediction is englobed in the electroweak fit of the SM parameters. The fit quality expressed by means of $\Delta\chi^2 = \chi^2 - \chi_{min}^2$ is shown in Fig. 1.3(b) with respect to m_h . The position of the minimum is $\chi_{min}^2 \sim 100 \text{ GeV}/c^2$ and represents the best fit to the electroweak precision measurements. The Higgs boson mass should be lighter than around $160 \text{ GeV}/c^2$ at 95 % C.L..

The absence of direct observation of the SM Higgs boson signatures puts also limits on its mass. The main production mechanism of the SM higgs boson at LEP during the second run ($E_{c.m.s.} \sim 210 \text{ GeV}$) is the Higgstrahlung process $e^+e^- \rightarrow Zh$, which has a significant cross section for Higgs lighter than roughly $120 \text{ GeV}/c^2$. Below this mass, the Higgs boson decays essentially in a $b\bar{b}$ or $\tau^+\tau^-$ pair. The main final states considered were therefore two b -jets from the Higgs boson, acoplanar with either two leptons, light jets, b -jets or missing transverse energy \cancel{E}_T , depending on the Z boson decay mode. The channel with $h \rightarrow \tau\tau$ and $Z \rightarrow q\bar{q}$ was also studied. The combined analysis from the four different LEP experiments has given a 95 % C.L. lower bound on the Higgs boson mass at $114.4 \text{ GeV}/c^2$ (Ref. [15]). Note that the corresponding limit for the light Higgs boson h in the context of MSSM fluctuates significantly from the SM value, depending on the scenario (Ref.[16]).

The Tevatron experiments are currently pursuing the mission of LEP, imposing new direct constraints on the SM Higgs mass. The Higgs boson production modes at $\sqrt{s}=1.96 \text{ TeV}$ are dominated by the gluon fusion, quark annihilation $q\bar{q} \rightarrow Wh, Zh$ and vector boson fusion $q\bar{q} \rightarrow q''\bar{q}'h$ processes as shown in Fig. 1.4 (up).

In March 2009, the combined results of CDF and DØ, using these channels and the decays of the Higgs boson $h \rightarrow b\bar{b}, WW, \tau\tau$ and $h \rightarrow \gamma\gamma$, allowed a new exclusion zone for m_h (see Ref. [17]). The result, obtained with a total integrated luminosity of 4.2 fb^{-1} , is shown in Fig. 1.4 (down). This important conclusion will be certainly updated with the expected 11 fb^{-1} of integrated luminosity that should be reached before the Tevatron shuts down, moving the bound across a larger mass range if no discovery is made.

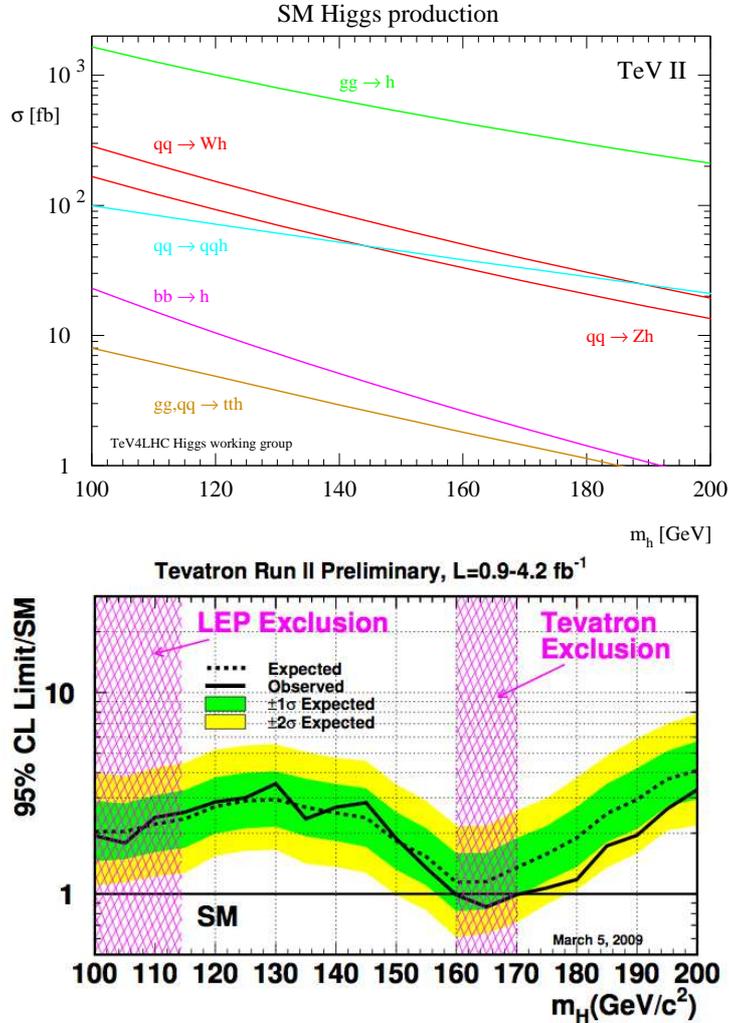


Figure 1.4: (Up) Cross-section for SM Higgs production at Tevatron (from [18]). (Down) The limits on Standard Model Higgs observation at Tevatron, dated from March 2009.

1.1.3 Higgs searches at the LHC

As briefly discussed in Sec. 1.1.2, the mass of the Standard Model Higgs is expected to be much smaller than the TeV/c^2 , most probably smaller than $150 \text{ GeV}/c^2$. However, with a c.m.s. energy of 14 TeV, the Large Hadron Collider will allow to produce the Higgs boson abundantly up to $1 \text{ TeV}/c^2$.

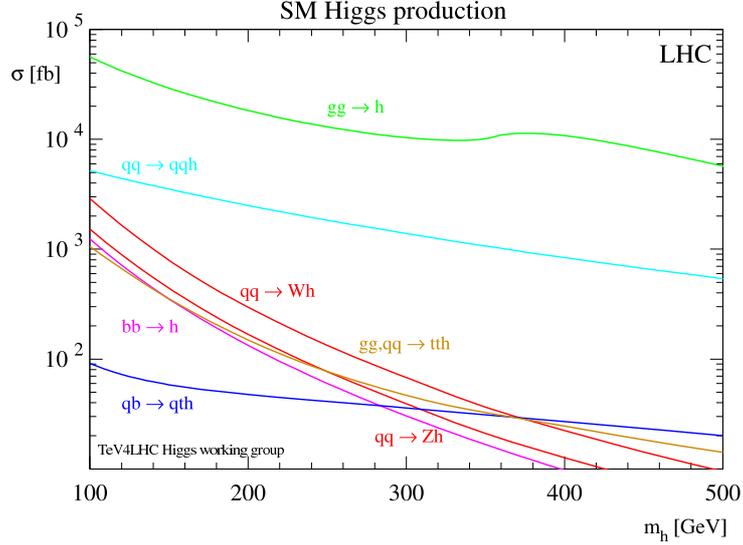


Figure 1.5: NLO cross-sections of the most relevant processes for Higgs production at the LHC. (see Ref. [18])

There are different ways to produce the Higgs boson at the LHC. For instance the gluon fusion has the largest cross-section on the whole Higgs boson mass range, followed by the vector boson fusion and Higgstrahlung process. The NLO production cross-sections evolution with respect to m_h are given in Fig. 1.5.

The observability of the Higgs boson depends strongly on the cross-section (and therefore on its mass), but even more on the decay modes, as seen in Fig. 1.2 (left). It has been shown in Ref. [19] that in the m_h range higher than $180 \text{ GeV}/c^2$, the $5\text{-}\sigma$ evidence could be reached with 10 fb^{-1} with the process $pp \rightarrow h \rightarrow ZZ \rightarrow \mu\mu\mu\mu$ (Fig. 1.6 (left)). With four electrons in final state the required integrated luminosity is 20 fb^{-1} (Fig. 1.6 (right)).

In the lower mass region, *i.e.* between 114 and $150 \text{ GeV}/c^2$, the discovery of a Higgs boson decaying into photons could be achieved with an integrated luminosity larger than roughly 20 fb^{-1} . Up to $180 \text{ GeV}/c^2$, there are other channels of interest. For instance the vector boson fusion process with $h \rightarrow \tau\tau$ should allow to reach the discovery level after 60 fb^{-1} for masses below $145 \text{ GeV}/c^2$. The Higgstrahlung process Wh is also foreseen to be interesting, but would require a larger luminosity ($\sim 100 \text{ fb}^{-1}$) to reach a $5\text{-}\sigma$ evidence on the $[140\text{-}180] \text{ GeV}/c^2$ mass range. See. [19] for more detailed discussions.

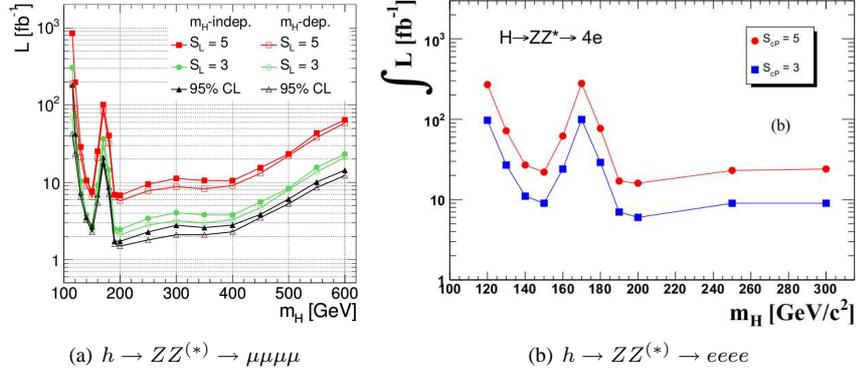


Figure 1.6: (a) The integrated luminosity required reach the 5- σ evidence for SM Higgs boson detection through the process $gg \rightarrow h \rightarrow ZZ \rightarrow \mu\mu\mu\mu$ in CMS . (b) Same as (a) but with four electrons in the final state. Note that the statistical estimators are slightly different in both figures. S_{cp} is related to the probability to observe $\mu_S + \mu_B$ events if the event distribution follows a Poisson law with mean μ_B . This probability is converted in a number of equivalent standard deviation of a gaussian distribution. S_L is $\sqrt{2\ln Q}$ where Q is the likelihood ratio used to evaluate the compatibility between the signal+background and the background hypotheses. [19]

These analyses will be crucial, not only for a possible discovery of the SM Higgs boson, but also to detect possible signs of physics beyond the SM (BSM). Indeed, many of the BSM models contain a scalar sector, and often Higgs boson(s) presenting similar characteristics as the SM one. The structure of this scalar sector could however lead to a modification of the “SM-like” Higgs branching ratios with respect to those in an unaltered SM case. A simple example would be the decay of this Higgs boson into additional scalars. This situation is already possible in one of the simplest extension of the SM scalar sector, the two-Higgs-doublet model. The next section focuses precisely on a particular realization of this model.

1.2 The two-Higgs-doublet model and the inverted mass spectrum

1.2.1 Introduction

The most common extension of the scalar sector of the Standard Model contains a second $SU(2)$ doublet, like for example in the Minimal Super Symmetric Model (MSSM). A large number of studies have been achieved in this context. The MSSM contains a scalar sector with five Higgs bosons: a light h , two charged H^\pm , a neutral H^0 and a pseudo-scalar A . In this work we base our discussion on a complementary alternative with a light pseudo-scalar A boson. This is a particular minimal realization of the non-supersymmetric two-Higgs-doublet-Model (M2HDM) discussed in details in Ref. [20] and Ref. [21]. The M2HDM potential is defined as

$$V^{\text{M2HDM}}(\phi_1, \phi_2) = -m_1^2 \phi_1^\dagger \phi_1 - m_2^2 \phi_2^\dagger \phi_2 + \frac{\lambda_S}{2} (\phi_1^\dagger \phi_1 + \phi_2^\dagger \phi_2)^2 + \frac{\lambda_{AS}}{2} (\phi_1^\dagger \phi_2 - \phi_2^\dagger \phi_1)^2, \quad (1.18)$$

with $m_{1,2}$ and $\lambda_{S,AS}$ a set of free real parameters. Assuming the charge conservation, the mean values of the two doublets $\phi_{1,2}$ reads for instance

$$\langle \phi_1 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix} \quad \text{and} \quad \langle \phi_2 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 e^{i\theta} \end{pmatrix}, \quad (1.19)$$

where v_1 and v_2 are the vacuum expectation values of ϕ_1 and ϕ_2 respectively.

The ‘‘Minimal’’ appellation of this model refers to its low number of free parameters, 4 compared to 14 in a general 2HDM. This choice is motivated by the following points. First the lagrangian is invariant under CP . This is an important statement, it allows to have clearly defined scalar states like in MSSM, notably a CP -odd and a CP -even Higgs boson.

Second, the potential (1.18) is invariant under a custodial symmetry, *i.e.* a symmetry that protects the value of the ρ parameter linking the weak boson masses as

$$m_{W^\pm}^2 = m_Z^2 \cos^2 \theta \quad (1.20)$$

In the SM, the scalar potential, which involves only ϕ_1 , displays an $SO(4)$ global symmetry spontaneously broken into a custodial $SO(3)$. This global symmetry is apparent at the level of the massless Goldstone bosons

$$m_{G^\pm}^2 = m_{G^0}^2 = 0 \quad (\text{with } G^0 \text{ CP - odd}) \quad (1.21)$$

“eaten” by the weak gauge fields whose masses have the same $SO(3)$ degeneracy

$$m_{W^\pm}^2 = m_{W^3}^2. \quad (1.22)$$

In the case of the M2HDM, the radiative corrections to the mass of the W and Z bosons take into account the contributions of the whole scalar sector. The relevant contributions of the scalars are illustrated in Fig. 1.7 [22].

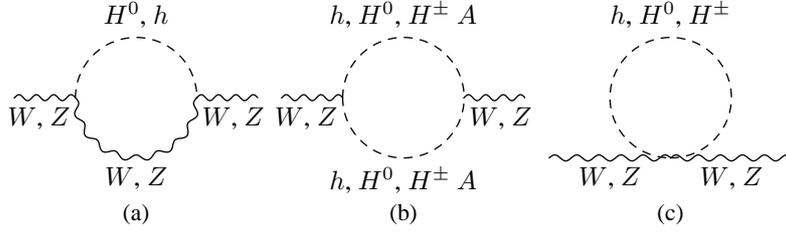


Figure 1.7: Illustration of the scalar contribution for W and Z mass calculation in (a) scalar-vector loop, (b) and (c) scalar-scalar loops.

The stability of the ρ parameter is therefore naturally ensured if both corrections are the same and this is obviously dependent of the scalar mass spectrum.

To understand this from the point of view of the potential, we choose to place ourselves in a *Higgs basis* Ref. [23]. This basis is obtained by redefining through a $U(2)$ transformation the fields $\phi_1 \rightarrow \Phi_1$ and $\phi_2 \rightarrow \Phi_2$ so that one of them has a vanishing vacuum expectation value, for instance Φ_2 . The first doublet plays therefore the same role as the unique doublet in the SM.

The “Standard Model” custodial symmetry acting on Goldstone bosons related to Φ_1 can be extended to the fields contained in Φ_2 if

$$m_{H^\pm}^2 = m_A^2 \quad (\text{with } A \text{ CP - odd}). \quad (1.23)$$

This is the degeneracy foreseen by the MSSM for example.

However, it has been demonstrated recently in Ref. [20] that the particular form of the potential (1.18) holds another $SO(3)$ mass degeneracy, namely

$$m_{H^\pm}^2 = m_{H^0}^2 \quad (\text{with } H^0 \text{ CP - even}). \quad (1.24)$$

This particular degeneracy corresponds to the “twisted” custodial symmetry.

Finally this potential exhibits an additional \mathbb{Z}_2 symmetry that forbids the presence of Flavor-Changing-Neutral-Currents (FCNC). A detailed discussion about these points can be found also in Ref. [8] and Ref. [21]

Scalars masses and couplings

The twisted custodial symmetry forces to have a mass degeneracy between the charged Higgs bosons H^\pm and the CP -even scalar H^0 . The values of the masses with respect to the parameters of the lagrangian can be found by rewriting the potential in terms of the eight real fields η_i , and by finding the second derivatives with respect to these real fields. In the Higgs basis the Higgs bosons masses are the diagonal elements of the 8×8 matrix $\partial^2 V^{M2HDM} / \partial \eta_i \partial \eta_j$ ($i, j=1..8$).

The first derivative of V^{M2HDM} gives the minimization condition $m_1 = -v\sqrt{\lambda_S/2}$ and the calculation shows that the charged and H^0 boson masses are

$$m_{H^\pm}^2 = m_{H^0}^2 = \frac{v^2 \lambda_S}{2} - m_2^2 = m_1^2 - m_2^2, \quad (1.25)$$

The boson H^0 is built from the fields of the second doublet, where the v.e.v. vanishes. As a consequence neither $H^0 WW$ nor $H^0 ZZ$ couplings is allowed.

This potential is also characterized by the presence of neutral SM-like Higgs boson h . This can be understood in the Higgs basis as only the doublet Φ_1 possess a v.e.v. This, associated to the fact that no mixing is present between the two doublets - and this is the choice we adopt here-, means that Φ_1 contains a SM-like Higgs boson h with the Yukawa couplings of the SM. Its mass is

$$m_h^2 = \lambda_S v^2, \quad (1.26)$$

Finally the pseudoscalar state A has the mass

$$m_A^2 = m_2^2 + \frac{1}{2}v^2(\lambda_S - 2\lambda_{AS}) = m_{H^\pm}^2 - \lambda_{AS}v^2. \quad (1.27)$$

In the potential (1.18), four free parameters are present (against 14 in a general 2HDM with charge conservation): m_1 , m_2 , λ_S , λ_{AS} . It turns out that the twisted custodial symmetry suppresses one degree of freedom, since λ_{AS} , m_1^2 and m_2^2 lead to the determination of only two masses: m_A and m_T .

Yukawa couplings

It is shown in Ref. [8] that the presence of the \mathbb{Z}_2 symmetry on the potential (1.18) allows to define two types of models for couplings between the scalars and the fermions.

The Yukawa couplings of the bosons A , H^0 and H^\pm are dependent of the ratio of the v.e.v. $\tan \beta = v_2/v_1$, which can be considered as an extra free parameter of the theory.

- In Type I all fermions couple to one doublet in a generic basis. The Yukawa couplings of H^0 , H^\pm and A are rescaled by $\tan \beta$
- In the Type II the up-fermions couple to one doublet and the down-fermions to the other doublet in a generic basis. This correspond to the situation of MSSM, the couplings with up (down)-fermions are scaled by $\cot \beta$ ($\tan \beta$).

The inverted mass spectrum

The twisted custodial symmetry does not constraint the A boson to be degenerate in mass with the charged Higgs bosons like in the usual custodial case. This leaves us the possibility to choose its mass, and in order to focus on the region of the M2HDM parameter space not covered by the MSSM, we impose the A boson to be the lightest Higgs ($m_A < 90 \text{ GeV}/c^2$). We also assume that the heaviest boson is h , followed by the triplet $T \equiv (H^\pm, H^0)$. This configuration has the virtue to increase the number of possibilities of scalar-to-scalar decays. In addition, if they are kinematically allowed, the unusual decay modes $H^\pm \rightarrow W^\pm A$ and $H^0 \rightarrow ZA$ lead to interesting final states particles.

In the following we therefore restrict our discussion to the "inverted" mass spectrum $m_A < m_T < m_h$ such that the decay of H^\pm and H^0 into a scalar and a gauge boson is kinematically allowed. We will refer to this particular configuration by using the word "*i*M2HDM".

1.2.2 Theoretical and indirect constraints

As explained in the previous section, the *i*M2HDM is described with four parameters m_A , $m_T = m_{H^0} = m_{H^\pm}$, m_h and $\tan \beta$. Although free *a priori*, these parameters have domains restricted by theoretical arguments and must be compatible with precise experimental data. Note that most of the theoretical and indirect constraints are not only limited to this model but also to less minimal version of the 2HDM. A more detailed discussion as well as the calculations can be found in Ref. [21] and Ref. [8].

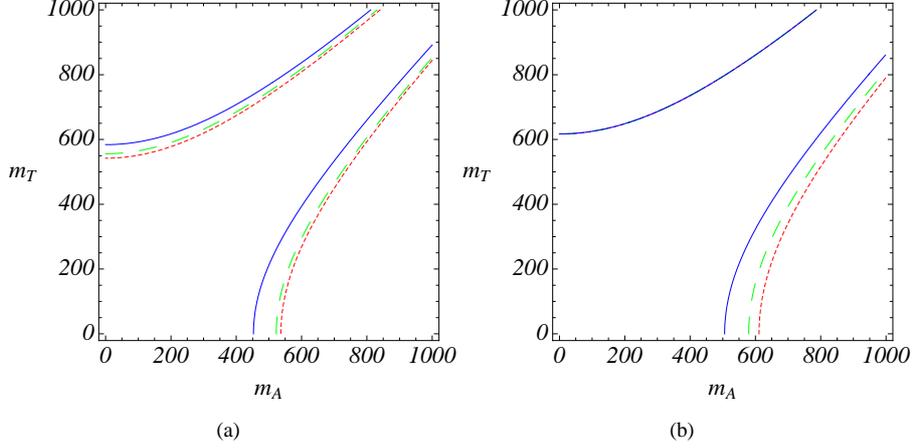


Figure 1.8: Unitarity (a) and perturbativity (b) constraints in the (m_A, m_T) plane for the M2HDM scenario. Dotted red lines are limits for $m_h = 120 \text{ GeV}/c^2$, dashed green lines for $m_h = 300 \text{ GeV}/c^2$ and plain blue lines for $m_h = 500 \text{ GeV}/c^2$. The allowed regions lie between these lines.

Theoretical constraints

The vacuum stability requires that the potential is lower bounded at large values in the (ϕ_1, ϕ_2) plane. This condition is fulfilled if

$$m_{h,A,T}^2 > 0 \quad (1.28)$$

$$m_h^2 > m_T^2 - m_A^2 \quad (1.29)$$

As in the SM, the evaluation of the unitarity constraint has to be estimated as new contributions enters in the scattering amplitudes for vector boson productions. If all scalars masses are non negligible, the unitarity requirement may help to restrict the allowed region in the (m_A, m_T) plane for different values of m_h , see Fig. 1.8(a). Essentially the restriction in the case of the i M2HDM is $m_T \lesssim 500 \text{ GeV}/c^2$

As in the SM the perturbativity condition constrains the masses of the Higgs bosons so that they can be described as resonances. This region in the (m_A, m_T) plane is shown in Fig. 1.8(b) for different values of the h boson mass. This bound is slightly looser than the unitarity condition, as $m_T \lesssim 600 \text{ GeV}/c^2$.

In the following we review the indirect constraints, first related the electroweak sector, then to the measurement of $b \rightarrow s\gamma$ transition rate and $B^0 - \overline{B}^0$ mixing. The effect

of the measurements of $Z \rightarrow b\bar{b}$, the muon anomalous magnetic moment and the $B \rightarrow \tau\nu$ are also briefly discussed.

Electroweak constraints

The electroweak S , T and U parameters [2] allows to quantify all the one-loop BSM electroweak effects.

The T parameter is proportional to the radiative corrections of the ρ parameter defined in (1.14) and is therefore sensitive to the measurement of W and Z masses and widths. More precisely the deviation from the unity is quantified as [2]:

$$\alpha_{EW}T = \Delta\rho = 0.0002^{+0.0007}_{-0.0004} \quad (1.30)$$

As mentioned in Sec. 1.2.1, the h boson contributes in the radiative corrections of W and Z mass. If h boson is heavy, large logarithms are present in ΔT (the difference between the measurement and the theoretical expectation), and therefore affect ρ . This deviation can be compensated by a loss of degeneracy between H^0 and H^\pm . It is shown in [8] that the deviation from the Standard Model ΔT can be maintained at zero even for large masses of h boson is a mass splitting between H^0 and H^\pm is present. This splitting is shown in Fig. 1.9.

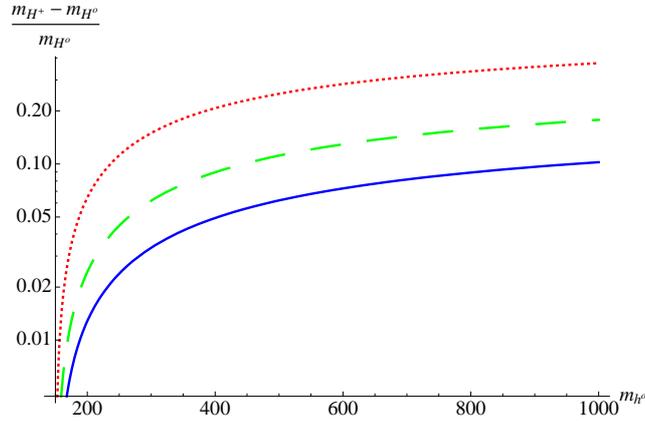


Figure 1.9: Relative mass difference $(m_{H^\pm} - m_{H^0})/m_{H^0}$ required to achieve $\Delta T = 0$ with respect to SM-like Higgs boson mass m_{h^0} in the $iM2HDM$ scenario. The dotted red, dashed green and plain blue lines correspond to $m_{H^0}^0=200, 300$ and 400 GeV/c^2 respectively. The A mass is fixed at 100 GeV/c^2 but this does not affect sizably the result.

The S and U parameters are the other two variables needed to describe the one-loop BSM electroweak corrections. From the electroweak fit, the values of S and U are -0.13 ± 0.1 and 0.2 ± 0.12 respectively [2].

It turns out that the $iM2HDM$ is not strongly constrained by these parameters, as S is lowered by roughly 15 % (0.02) at maximum while U is raised by roughly 2.5 % at maximum. However it is shown in [8] that both S and U seems in favor of an $iM2HDM$.

Constraints from $b \rightarrow s\gamma$ measurement

In the Standard Model, the radiative process $b \rightarrow s\gamma$ involves a flavor change via the presence of a W bosons loop. In the case of an extended scalar sector, the contribution of a charged Higgs boson is also present (Fig. 1.10) and its mass can be constrained.

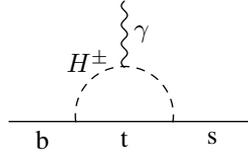


Figure 1.10: Feynman diagram of the transition $b \rightarrow s\gamma$ with a loop involving a charged Higgs H^\pm .

The average value of the transition has been measured by Belle, Babar and CLEO as $\text{BR}(b \rightarrow s\gamma) = (3.53 \pm 0.23(\text{stat}) \pm 0.09(\text{syst})) \times 10^{-4}$ for $E_\gamma > 1.6$ GeV [24]. The Standard Model NNLO prediction is $\text{BR}(b \rightarrow s\gamma) = (3.15 \pm 0.23) \times 10^{-4}$ [25].

The $iM2HDM$ prediction has been normalized to the SM expectation by setting the mass of the charged Higgs boson to infinity, this removes the contribution of the charged Higgs in the loop. In the $iM2HDM$ the leading order constraint is shown in the $(\tan\beta - m_{H^\pm})$ plane in Type I (Fig.1.11 (a)) and Type II (Fig.1.11 (b)).

The scaling of Yukawa coupling in the Type I allows clearly only the small $\tan\beta$ region. A very tight unconstrained region around $\tan\beta=2$ is also present because of an interplay between the contributions associated to the W and H^\pm in Inami-Lim

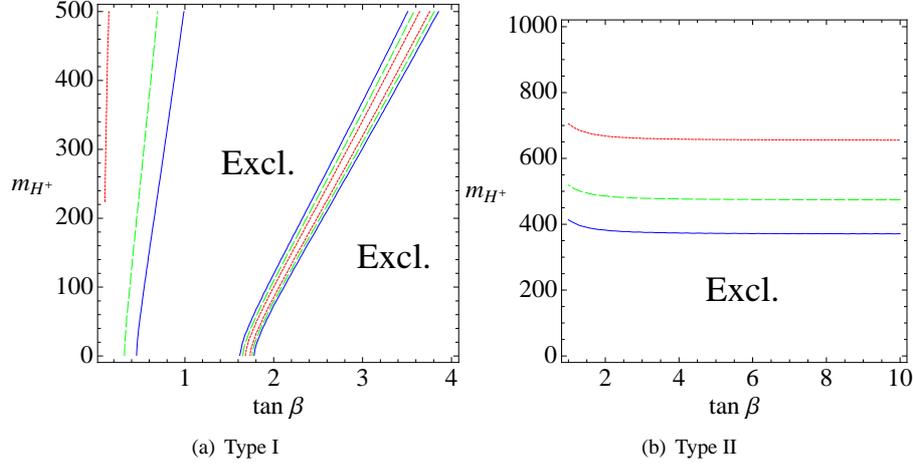


Figure 1.11: Leading order $b \rightarrow s\gamma$ bounds on the charged Higgs mass (in GeV/c^2) with respect to $\tan\beta$ in (a) Type I and (b) Type II scenarios, at one (dotted red), two (dashed green) and three (plain blue) standard deviations.

terms [26], *i.e.* the loop contributions with one vector and one charged Higgs boson. We do not consider this region as a case of interest.

In Type II, the excluded region corresponds essentially to very large values of m_{H^\pm} ($> 500 \text{ GeV}/c^2$), on the whole $\tan\beta$ range. However, as in Ref. [25], the NLO calculation decreases this limit to roughly $m_{H^\pm} > 300 \text{ GeV}/c^2$ ($\sim 95\%$ C.L.).

$B_0 - \bar{B}_0$ mixing

The presence of the charged Higgs bosons might also affect the $B_d - \bar{B}_d$ oscillations [27] as shown in the Feynman diagram in Fig. 1.12

As for the $b \rightarrow s\gamma$ process, the measurement of the oscillation amplitude gives a bound on the charged Higgs mass, depending on $\tan\beta$. The observable $\Delta m_B \equiv |M_{B_L} - M_{B_s}|$ can be calculated (see Ref. [27] for more details) in the context of the *iM2HDM*. Here again, the calculation with the *iM2HDM* is normalized to the SM prediction [28] by setting the mass of the charged Higgs to infinity.

As shown in Fig. 1.13 the resulting constraints in Type I and Type II are typically fairly symmetric around $\tan\beta=1$ and rather insensitive to the charged Higgs mass.

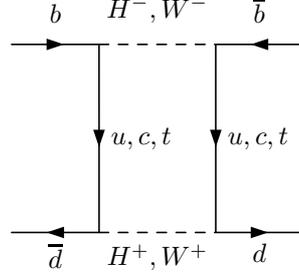


Figure 1.12: Feynman diagram of the $B_0 - \bar{B}_0$ oscillation, including the contributions of the charged Higgs H^\pm bosons.

R_b and A_b constraint

The process $Z \rightarrow b\bar{b}$ yields two observables sensitive to the presence of new scalars. The R_b quantity is the hadronic branching ratio of Z to b quarks

$$R_b = \Gamma_{Z \rightarrow b\bar{b}} / \Gamma_{Z \rightarrow \text{hadr}}. \quad (1.31)$$

and A_b quantifies the asymmetric rate of b-jets produced in the forward or backward direction in a e^+e^- collider. This asymmetry is expressed as

$$A_b = \frac{\sigma(e_L^- \rightarrow b_F) - \sigma(e_L^- \rightarrow b_B) + \sigma(e_R^- \rightarrow b_F) - \sigma(e_R^- \rightarrow b_B)}{\sigma(e_L^- \rightarrow b_F) + \sigma(e_L^- \rightarrow b_B) + \sigma(e_R^- \rightarrow b_F) + \sigma(e_R^- \rightarrow b_B)}, \quad (1.32)$$

where $e_{L,R}^-$ are left and right handed initial state electrons and $b_{F,B}$ the b-quarks in forward and backward direction, defined as the direction and anti-direction of initial electrons movement respectively.

The constraint on a Type I model is relevant for the charged Higgs bosons and turns out to be an upper bound of $\tan \beta < 1$ (Ref. [8]). In the Type II model, the neutral bosons contribute as well and the result gives an upper bound in $|m_T - m_A|$, getting stronger as $\tan \beta$ increases. For instance if $\tan \beta \sim 50$ the Type II is excluded at $2\text{-}\sigma$ for a light pseudoscalar (below $\sim 50 \text{ GeV}/c^2$) since the mass of the triplet should be under $300 \text{ GeV}/c^2$, which is disfavored by the $b \rightarrow s\gamma$ transition. At $\tan \beta \sim 30$, $m_A \gtrsim 30 \text{ GeV}/c^2$ for the same reasons.

a_μ constraint

The measurement of the anomalous magnetic moment of the muon is defined as $a_\mu = \frac{g-2}{2}$ where the g -factor links the spin of the muon to its anomalous magnetic moment.

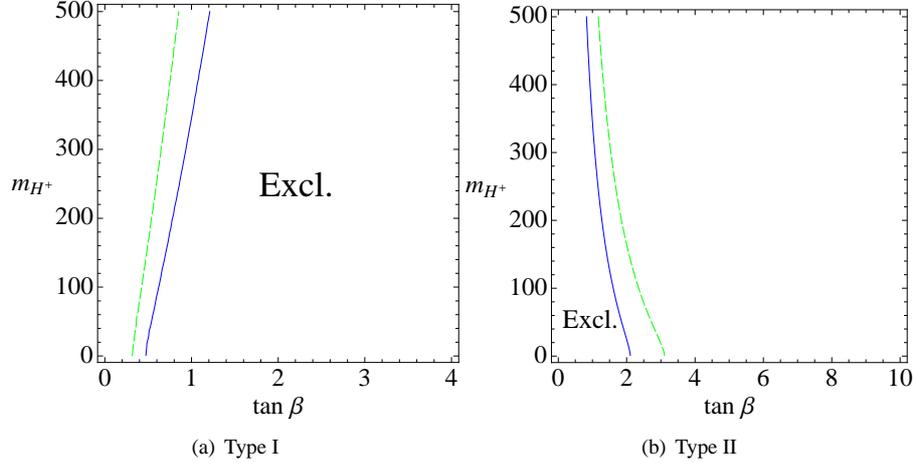


Figure 1.13: $2\text{-}\sigma$ (dashed green) and $3\text{-}\sigma$ (solid blue) bounds on the charged Higgs mass (in GeV/c^2) from the $B_0 - \overline{B}_0$ mixing measurement, with respect to $\tan \beta$ in (a) Type I and (b) Type II scenarios, at two (dashed green) and three (plain blue) standard deviations. There are no 1σ limits on these figures due to the slight discrepancy between the SM prediction and the current experimental measurement.

This quantity is sensitive to the physics beyond the Standard Model through the one and two loops corrections of $\mu\mu\gamma$ vertex shown in Fig.1.14. In our case, This concerns only the Type II since the Higgs bosons must couple to the muons.

For the one loop case, the calculation shows that the contributions of neutral Higgs boson dominate for masses above 0.2 GeV . It is in favor of a light ($\lesssim 10 \text{ GeV}/c^2$) H^0 boson, which is experimentally excluded by the measurement of $\Upsilon \rightarrow A\gamma \rightarrow \tau^+\tau^-\gamma$ done by the CLEO collaboration.

The two loop correction is more interesting since the contribution of the pseudoscalar can compensate the discrepancy observed between the SM prediction [29] and experimental measurements [30]:

$$a_\mu^{SM} = (11659180.4 \pm 5.1) \times 10^{-10} \quad (1.33)$$

$$a_\mu^{exp} = (11659208 \pm 6.3) \times 10^{-10} \quad (1.34)$$

This is the case where the mass of the pseudo-scalar is close to $20 \text{ GeV}/c^2$ and $\tan \beta \sim 30$. We will consider the limit $m_A < 100 \text{ GeV}/c^2$ adopted in Ref. [8].

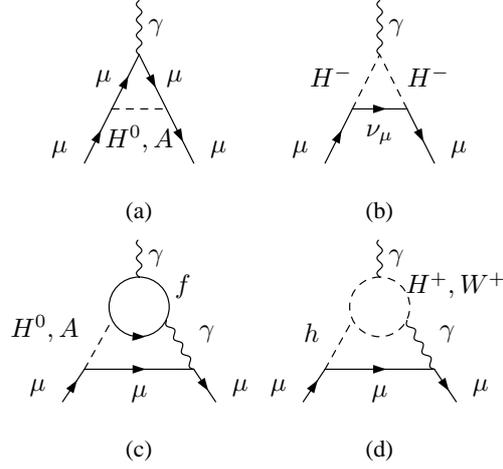


Figure 1.14: One-loop contribution to a_μ due to (a) neutral (pseudo)scalars and (b) charged Higgs boson exchange. Two-loop contributions to a_μ from (c) a light (pseudo)scalar with a fermionic loop or (d) from a light scalar with a charged boson loop.

$B \rightarrow \tau\nu$

In a Type II Model, *i.e.* at large $\tan\beta$, the charged Higgs may be produced by the fusion of a b -quark and a u or c quark, and decay into $\tau\nu$. This should therefore induce a deviation from the SM expectation for the $B \rightarrow \tau\nu$ rate.

The recent result from Babar collaboration has shown that the deviation between the experimental result and the SM expectation is [31]

$$\frac{BR(B \rightarrow \tau\nu)}{BR(B \rightarrow \tau\nu)_{SM}} = 1.13 \pm 0.44. \quad (1.35)$$

This leads, in the context of the $iM2HDM$ Type II (see Ref. [8]) to the constraint

$$\tan\beta \lesssim 0.13 \text{ GeV}^{-1} \times m_{H^\pm}.$$

The $b \rightarrow s\gamma$ limit for the mass of the charged Higgs being close to $300 \text{ GeV}/c^2$, the corresponding limit would be close to $\tan\beta \sim 40$.

For the sake of clarity, the theoretical and indirect constraints are summarized in Tab. 1.2

| Type | Constraint | Origin |
|--------|--|--|
| I & II | $m_h < 500 \text{ GeV}/c^2$ | Unitarity |
| | $\sqrt{m_T^2 - m_A^2} < 400 \text{ GeV}/c^2$ | Perturbativity |
| | $m_{H^0} \sim m_{H^\pm}$ | ΔT if m_h is large |
| | $m_A \ll m_T$ | $\Delta S, \Delta U$ if m_h is large |
| I | $\tan \beta < 0.5$ | B -physics |
| II | $m_T > 300 \text{ GeV}$ | $b \rightarrow s\gamma$ |
| | $m_A > 30 \text{ GeV}/c^2$ | R_b if $\tan \beta \sim 30$ |
| | $m_A < 100 \text{ GeV}/c^2$ | a_μ |
| | $\tan \beta > 2$ | $B_0 - \bar{B}_0$ mixing |
| | $\tan \beta < 40$ | $B \rightarrow \tau\nu$ |

Table 1.2: Summary of the theoretical and indirect constraints in the $iM2HDM$.

1.2.3 Decays of Higgs bosons

In this section we review the decays³ of the Higgs bosons in Type I and Type II models and then consider the direct constraints on the scalar mass spectrum.

Neutral Higgs bosons

As shown in Fig. 1.2 (left), the decay modes of the Standard Model Higgs boson h are, essentially dominated by $h \rightarrow W^+W^-$ and $h \rightarrow ZZ$ above $130 \text{ GeV}/c^2$, while below most of the decays are in fermions pairs.

In the context of $iM2HDM$, the situation can be dramatically different, especially in the low mass regions, due to the presence of additional scalars. Indeed the decay widths of h into a pair of generic scalars ϕ is given by

$$\Gamma_{h \rightarrow \phi\phi} = k \frac{G_F}{64\sqrt{2}\pi m_h} |g_{h\phi\phi}|^2 \left(1 - 4 \frac{m_\phi^2}{m_h^2}\right) \quad (1.36)$$

where k is 1 or 2 depending whether the final state particles are distinguishable or not, $g_{h\phi\phi}$ is the coupling values between h and a scalar ϕ and G_F is the Fermi constant. From the Feynman rules related to the potential (1.18), and depending on the nature of ϕ , the value $g_{h\phi\phi}$ is given by

$$\begin{aligned} g_{h^0 AA} &= m_{h^0}^2 + 2m_A^2 - 2m_T^2 \\ g_{h^0 H^0 H^0} &= m_{h^0}^2 \\ g_{h^0 H^+ H^-} &= m_{h^0}^2. \end{aligned}$$

³Only $1 \rightarrow 2$ decays are considered.

The decay width of h into fermions reads

$$\Gamma_{h^0 \rightarrow f\bar{f}} = \frac{G_F}{\sqrt{2}} \frac{m_h m_f^2}{4\pi} \left(1 - 4 \frac{m_f^2}{m_h^2}\right)^{3/2}. \quad (1.37)$$

As a consequence, the branching ratios of the h boson in the low mass region are totally dominated by the scalar decays, as shown in Fig. 1.15. There is however a small exception: the coupling g_{hAA} depends from both m_A and m_T , and vanishes if $m_h^2 = 2m_T^2 - 2m_A^2$. As a consequence, the non-observation of a Higgs decaying into A boson does not necessarily mean that $2m_A > m_h$.

A typical situation, with $m_A=20 \text{ GeV}/c^2$, and $m_T=150 \text{ GeV}/c^2$ is illustrated in Fig. 1.15 where the $\text{BR}(h \rightarrow AA)$ falls rapidly when m_h is close to $210 \text{ GeV}/c^2$. The remaining decays are $h \rightarrow WW$ and $h \rightarrow ZZ$, very close to the SM values.

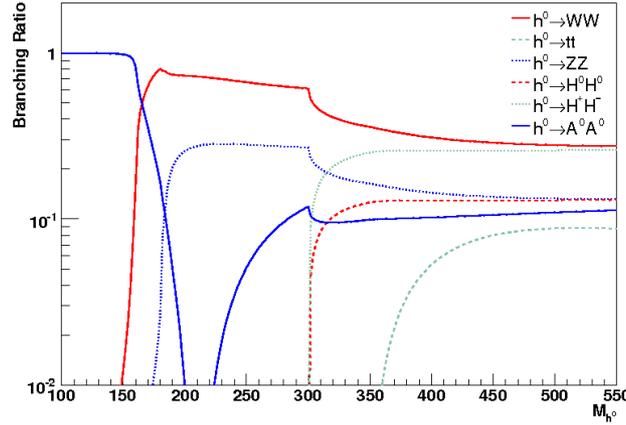


Figure 1.15: Branching ratios of h for $m_A=20 \text{ GeV}/c^2$ and $m_T=150 \text{ GeV}/c^2$.

From $150 \text{ GeV}/c^2$ towards heavier masses, the h boson always decays into vector bosons and, if allowed kinematically, also into the scalars H^0 and H^\pm . All decay modes, *i.e.* $h \rightarrow WW, ZZ, H^+H^-$ and H^0H^0 could then be considered in experimental analyses.

The situation is less complicated for the decay of H^0 as no coupling between H^0 and a pair of gauge bosons is allowed in $iM2HDM$. The only available modes are thus $H^0 \rightarrow f\bar{f}'$ and $H^0 \rightarrow ZA$. The decay width of H^0 for the decay into ZA is given by

$$\Gamma_{H^0 \rightarrow ZA} = \frac{G_F m_W^2 m_Z^2}{8\sqrt{2}\pi m_{H^0}} \left(1 - 2 \frac{m_Z^2 + m_A^2}{m_{H^0}^2} + \frac{(m_Z^2 - m_A^2)^2}{m_{H^0}^4}\right)^{3/2} \quad (1.38)$$

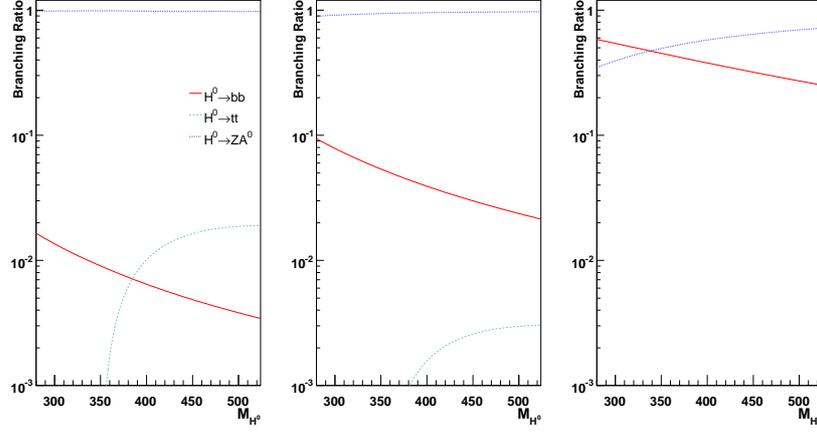


Figure 1.16: Branching ratios of the H^0 Higgs in the Type II for $m_A=40$ GeV/ c^2 and the $\tan\beta$ parameter for the values 4 (left), 10 (center) and 40 (right).

and into fermions ($b\bar{b}$) by

$$\Gamma_{H^0 \rightarrow b\bar{b}} = \frac{3G_F}{\sqrt{2}} \frac{m_{H^0} m_b^2 \tan^2 \beta}{4\pi} \left(1 - 4 \frac{m_b^2}{m_{H^0}^2}\right)^{3/2}. \quad (1.39)$$

In the Type I model, $\tan\beta$ is restricted to be small ($\lesssim 0.3$). It turns out that the branching ratio $\text{BR}(H^0 \rightarrow b\bar{b})$ is suppressed if the decay $H \rightarrow ZA$ is allowed. If $m_A + m_Z > m_H^0$ the H^0 boson decays essentially into $b\bar{b}$ and $\tau\tau$.

In the Type II $iM2HDM$, the mass of the triplet T is bounded from below by the $b \rightarrow s\gamma$ indirect constraint at around 300 GeV/ c^2 , and $\tan\beta > 4$. The main decays are then $H^0 \rightarrow ZA$ and $H^0 \rightarrow b\bar{b}$. However, as shown in Fig. 1.16, the $H^0 \rightarrow ZA$ is seriously affected once $\tan\beta$ is larger than 40.

The decay modes of the pseudoscalar A are quite simple in both types. The main decays are $A \rightarrow b\bar{b}$ (0.85-0.9) and $\tau^+\tau^-$ (0.05-1) if $m_A \gtrsim 10$ GeV/ c^2 .

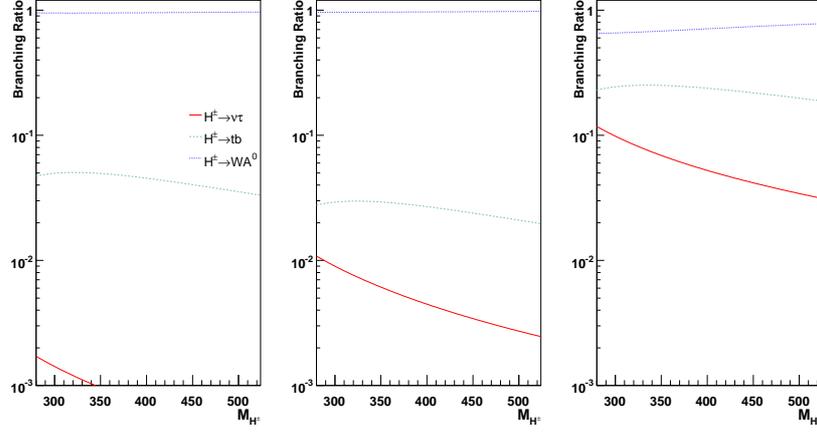


Figure 1.17: Branching ratios of the H^\pm Higgs in the Type II for $m_A=40 \text{ GeV}/c^2$ and the $\tan \beta$ parameter for the values 4 (left), 10 (center) and 40 (right).

Charged Higgs bosons

In the $iM2\text{HDM}$, the possible decays of the charged Higgs are $H^\pm \rightarrow W^\pm A$ and $H^\pm \rightarrow f\bar{f}'$. The decay width is given by

$$\Gamma_{H^\pm \rightarrow W^\pm A} = \frac{G_F m_W^4}{8\sqrt{2}\pi m_{H^\pm}} \left(1 - 2\frac{m_W^2 + m_A^2}{m_{H^\pm}^2} + \frac{(m_W^2 - m_A^2)^2}{m_{H^\pm}^4} \right)^{3/2} \quad (1.40)$$

As for H^0 , the decay into fermions depends on $\tan \beta$. In particular, the experimentally more interesting leptonic partial decay width reads:

$$\Gamma_{H^\pm \rightarrow l^\pm \nu_l} = \frac{G_F m_{H^\pm}}{4\sqrt{2}\pi} m_l^2 \tan^2 \beta \left(1 - \frac{m_l^2}{m_{H^\pm}^2} \right)^3. \quad (1.41)$$

If the decay into WA is kinematically allowed, $\text{BR}(H^\pm \rightarrow W^\pm A) \sim 1$ except in the Type II for a large $\tan \beta$. The possible dominance of this decay is important as it may lead to the suppression of the $H^\pm \rightarrow \tau\nu$ decay often considered in many analyses for the detection of a charged Higgs boson. In the case of a $\tan \beta$ around 40, the branching ratio with m_{H^\pm} is at worst around 0.6, followed by the decay into tb whose branching ratio is close to 0.2. This shows the importance of the decay into $W^\pm A$ in both types.

As for the H^0 case, the evolution of the main branching ratios of the charged Higgs in the Type II is illustrated in Fig. 1.17

1.2.4 Direct constraints

This section present briefly the constraints on the i M2HDM that could be established from the direct searches at past and present colliders. We also list some processes of interest to establish a first set of constraints (or a discovery) on i M2HDM at Tevatron.

Constraints on neutral Higgs bosons h , H^0 and A

As seen in the section 1.1.2, the Standard Model Higgs has been constrained by the LEP experiment to be heavier than $114.4 \text{ GeV}/c^2$ and by the analyses in CDF and $D\phi$ to be below 160 of above $170 \text{ GeV}/c^2$. This holds of course only if the decay modes are exactly those foreseen in the Standard Model, i.e. if they follow the branching ratios given in Fig. 1.2. This statement is not true in i M2HDM since the decay $h \rightarrow AA$ can automatically be considered and $h \rightarrow H^0 H^0$ and $h \rightarrow H^+ H^-$ are possible in Type I.

At LEP, the analyses of topologies $Z+2b$ -jets and $Z+4b$ -jets showed an excess of events. This has been interpreted in the context of NMSSM [32], which allows the presence of a light A boson decaying into a pair of b -quarks. However, the fact that these excess stands for slightly different masses of h led to the conclusion that no significant deviations from the SM expectation was observed. We therefore keep the m_h lower bound at $114 \text{ GeV}/c^2$.

At the Tevatron, to our knowledge, no official search of the decay $h \rightarrow AA$, with h produced by gluon fusion or by associate production with a vector boson has been attempted up to now. This channel has been recently discussed in Ref. [33] and shown to be very difficult. This is due either to the large QCD backgrounds if the pseudoscalars decay into b quarks, or to the too low signal visible cross-section if one or two A decay(s) into τ 's.

From the most recent public results of Tevatron (in fact CDF) [34], a constraint on the A boson mass is made from the analysis of $b\bar{b} \rightarrow A \rightarrow \tau\tau$ channel in Type II. Unfortunately the A mass considered in the analysis is at minimum $70 \text{ GeV}/c^2$ since the research is essentially oriented towards MSSM. The CDF result is shown in Fig. 1.18, we adopt the conservative limit $m_A > 70 \text{ GeV}/c^2$ if $\tan\beta > 35$.

However, an excluded region in the $m_A - \tan\beta$ plane has ben set at LEP in the context of analyses of $e^+e^- \rightarrow b\bar{b}A$ channel with $A \rightarrow b\bar{b}, \tau^+\tau^-$. The tables provided in Ref. [36] gives an approximative limit in $\tan\beta$ versus m_A : if $\tan\beta > 15$, $m_A > 10 \text{ GeV}/c^2$.

A simultaneous constraints on both H^0 and A masses could be established with the study of the reaction $e^+e^- \rightarrow Z^* \rightarrow H^0 A$ done at LEP. In the i M2HDM the decay

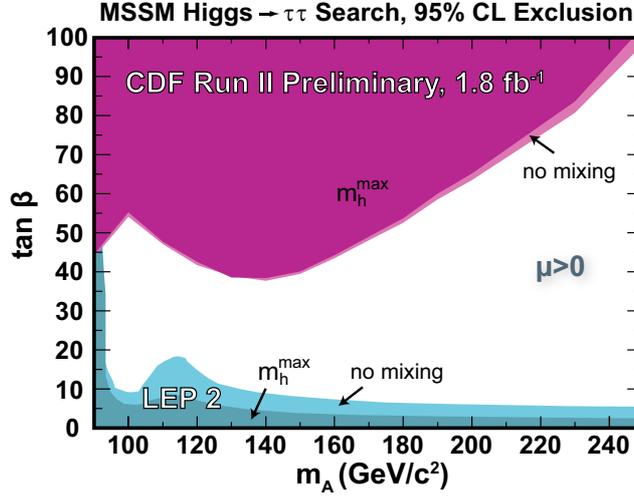


Figure 1.18: The exclusion limits for the Run 2 at Tevatron. The meaning of the benchmark choices can be found in Ref. [35]

$H^0 \rightarrow ZA$ is automatically considered, leading to the final state $Zb\bar{b}b\bar{b}$. As said before, no significant excess in this final state has been observed, we suppose therefore that the sum of the Higgs boson masses is out of reach for the center-of mass energy at LEP2. The constraint adopted is $m_A + m_H \gtrsim 200 \text{ GeV}/c^2$.

At Tevatron, a constraint on m_{H^0} could be achieved via $b\bar{b} \rightarrow H^0 \rightarrow ZA$ in the Type II model. Assuming the B physics restrictions and $\tan\beta=30$, the NLO cross-section is close to 50 fb. With $A \rightarrow \tau^+\tau^-$ branching ratio of a few percents, only the main $A \rightarrow b\bar{b}$ decay could have been observed, leading to the final state $Zb\bar{b}b\bar{b}$. However it seems that such an analysis has not been realized. Therefore up to now no constraint on H^0 mass can be set from this channel.

Constraints on the charged Higgs bosons

At LEP, the main production mechanism of charged Higgs bosons is $e^+e^- \rightarrow Z^* \rightarrow H^+H^-$. Both the $W^{+(*)}W^{-(*)}AA$ and $W^{\pm(*)}A\tau\nu$ final states have been studied for $m_{H^\pm} < 90 \text{ GeV}/c^2$ [37]. The result is a lower bound of $76.7 \text{ GeV}/c^2$ for m_{H^\pm} in Type I scenarios. A limit can also be obtained in Type II, but is not considered here since the mass of the scalar triplet has to remain above $300 \text{ GeV}/c^2$ due to B -physics constraints.

At the Tevatron there are two main ways to produce the charged Higgs: via the top decay $t \rightarrow H^\pm b$ or the single top associated production $p\bar{p} \rightarrow tH^-$. This latter case does not offer a viable alternative to observe the Higgs boson due to its too low cross-section (< 1 fb even in Type I). However the first option provides a reasonable final state associated with a sizable cross-section. The test of Standard Model $t\bar{t}$ production has been done in both CDF and DØ (with different assumptions for BSM decay of the charged Higgs boson) and it turns out that the $1\text{-}\sigma$ limit to branching ratio $\text{BR}(t \rightarrow H^\pm b)$ is lower than 30% on the whole mass range of m_{H^\pm} . This does not constraint the i M2HDM where this branching ratio is at maximum around 10 % (see Ref. [8] for more details).

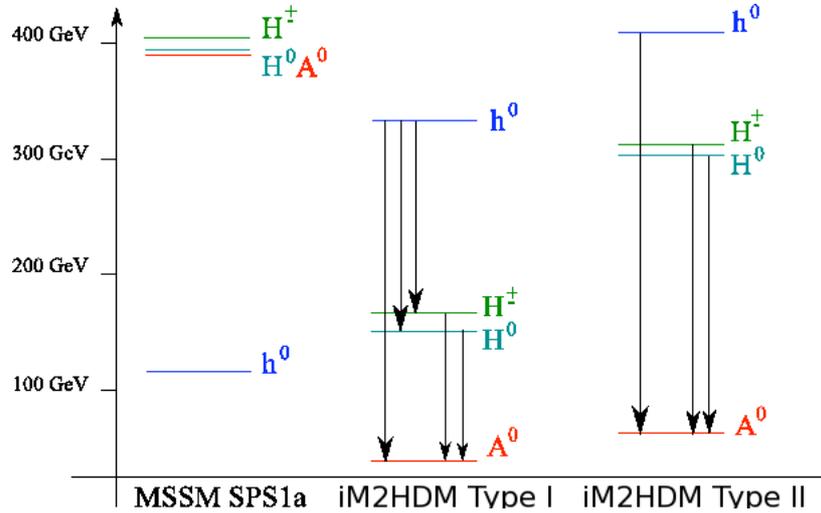
1.3 Conclusion

In this chapter we have briefly reviewed the Brout-Englert-Higgs mechanism and the properties of the associated SM Higgs boson. We have presented a possible extension of the SM case by introducing a particular realization of the two-Higgs-doublet-Model, CP conserving and holding a twisted custodial symmetry imposing a degeneracy between the charged and a CP -even Higgs boson H^0 . This model leaves the possibility to consider an inverted mass spectrum where the lightest Higgs boson is the pseudo-scalar A and the heaviest is the SM-like h

The impact of a set of theoretical and experimental constraints have been discussed and summarized in Tab. 1.3. Both Type I and Type II models leave an unconstrained parameter space where the scalar mass arrangement of i M2HDM is possible. The comparison between scalar masses in the SPS1a benchmark point of MSSM and within both types of the i 2HDM is represented in Fig. 1.19.

The study of such a scalar sector, and more generally, of the BSM phenomenology will require both a large luminosity and high energy. At the LHC it will strongly rely on the comparison with Monte Carlo simulation for the signals but also for the backgrounds. One of the major problem at the LHC is the presence of a potentially hard QCD activity from initial states radiation, potentially penalizing for the establishment of event shapes estimation. This consideration could be therefore crucial to hopefully identify the small BSM event excesses from the Standard Model background. The next chapter describes in more details this problematic and the solutions proposed.

| Type | Constraint | Origin |
|--------|---|-------------------------------|
| I & II | $m_h < 500 \text{ GeV}/c^2$ | Unitarity |
| | $\sqrt{m_T^2 - m_A^2} < 400 \text{ GeV}/c^2$ | Perturbativity |
| | $m_{H^0} \sim m_{H^\pm}$ | ΔT if m_h is large |
| | $m_A \ll m_T$ | ΔS if m_h is large |
| | $m_h > 114 \text{ GeV}/c^2$ | LEP direct constraint |
| I | $m_T + m_A > 200 \text{ GeV}/c^2$ | LEP $Zb\bar{b}b$ |
| | $\tan \beta < 0.5$ | B -physics |
| II | $m_T > 300 \text{ GeV}$ | $b \rightarrow s\gamma$ |
| | $m_A > 30 \text{ GeV}/c^2$ | R_b at $\tan \beta \sim 30$ |
| | $m_A < 100 \text{ GeV}/c^2$ | a_μ |
| | $m_A < 70 \text{ GeV}/c^2$ if $\tan \beta > 35$ | Tevatron $b\bar{b}A$ |
| | $\tan \beta > 2$ | $B_0 - \bar{B}_0$ mixing |
| | $\tan \beta < 40$ | $B \rightarrow \tau\nu$ |

Table 1.3: Summary of the theoretical and experimental constraints in the $iM2HDM$.Figure 1.19: Typical mass spectra in MSSM (SPS1a), and in both Type I and Type II configurations of the $iM2HDM$. The arrows indicate typical scalar decays.

The LHC and the modeling of collision

BEFORE the end of year 2009, the Large Hadron Collider will start to strike protons together, slowly increasing its center-of-mass energy up to 14 TeV. At such an energy the analyses done from the data collected by the detectors will be sensitive to the existence of new particles predicted by different models, for instance the Higgs boson(s), the supersymmetric particles and other heavy resonances present for instance in models with extra-dimension or in technicolor models. Their detection is one of the main challenge of the LHC experiments and certainly a major step towards the understanding of physics up to the TeV scale.

In this chapter the generalities about the Large Hadron Collider are first reviewed. The discussion focuses then on the modeling of hadron-hadron collisions, especially the multi-jet processes generation where additional jets are produced by initial state radiation (ISR). This is particularly important as any event produced at the LHC will be affected by this QCD activity. A difficulty resides in the simulation of this radiation in the large range of energy, from TeV scale down to hadronization scale. A solution to this issue is presented (jet matching/merging), and the consequences of using such technique are described, mostly for the production of heavy colored particles.

2.1 The LHC

2.1.1 The installation

The LHC is installed near Geneva, at the Franco-Swiss border and exactly on the former LEP tunnel, the electron-positron collider that ran from 1989 to 2000.

To reach 14 TeV of colliding energy, different steps of acceleration are needed as it can be seen in Fig. 2.1.

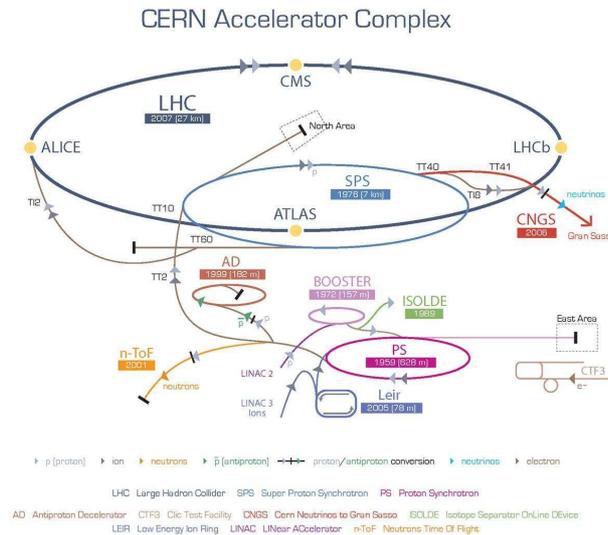


Figure 2.1: Illustration of the acceleration chain for the LHC.

The protons are first produced by hydrogen ionization in the Duoplasmatron Proton Ion Source (see for instance Ref. [38] for more details). After a geometric arrangement into bunches the protons are injected in the Linac-2 accelerator. Coupled to the Proton Synchrotron Booster, the protons accelerate to get a kinetic energy of 1.4 GeV. Then with the Proton Synchrotron the energy reaches around 26 GeV, and the time separation of 25 ns between each bunch is established. Afterwards, the injection is done in the SPS, which accelerates protons up to 450 GeV, and finally the LHC ring itself finalizes the acceleration process with radio-frequency cavities to obtain 7 TeV per proton. Two beams in opposite directions are needed. In order to bend the path of

the proton, a magnetic field of 8.33 Tesla is applied thanks to 1282 dipoles placed all along the curve. The field is generated by NbTi superconductors cooled at 2 K with liquid helium.

The protons are grouped by bunches, each of them will contain at best $1.15 \cdot 10^{11}$ protons. In the whole ring there will be at maximum around 2800 bunches, each of an approximate length of 8 cm. The diameter of a bunch varies in function of the position and should be close to $6 \mu\text{m}$ at the interaction points.

Four main experiments are installed on the LHC ring: CMS and ATLAS are called "multi-purpose" detectors by opposition to the two others, ALICE and LHCb which are oriented towards more specific fields of interest. The main motivation of multi-purpose detectors is to detect the signature of new particles while the direction of research in ALICE is oriented toward the studies of QCD plasma physics and LHCb towards CP -violation in B-physics. While CMS¹, ATLAS and LHCb searches are essentially based on proton-proton collisions, heavy ions (Pb, Au) collisions are mostly looked at by ALICE.

At the interaction points the rate of production associated to a particular type of process is related to its cross-section, but also to the luminosity. The instantaneous luminosity can be defined as

$$\mathcal{L} = fk \frac{N_1 N_2}{\pi \sigma^2} \quad (2.1)$$

where f is the frequency of crossing, $N_{1,2}$ are the number of proton per bunch and σ is the estimated RMS of the bunch distribution in the transverse plane. With the time-spacing of 25 ns induced by the PS, the frequency reaches 40 MHz. The factor k represents the beam occupancy level: the 2800 bunches represents around 77% of the overall number of available spaces. There are a few different steps scheduled to increase the luminosity. For instance within the first month it is expected to reach $1.2 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$, with an accumulated luminosity of 200 nb^{-1} . This is expected to be realized at 5 TeV of energy per beam. The low luminosity regime ($\sim 2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$) will then probably require 5 month of commissioning before being reached. The luminosity should then slowly increase to hopefully reach its nominal regime ($10^{34} \text{ cm}^{-2} \text{ s}^{-1}$). This will normally correspond to 100 fb^{-1} of integrated luminosity per year.

The total inelastic cross-section of proton-proton collision is roughly 70 mb. Assuming that the mean bunch crossing rate is $40 \text{ MHz} \times 0.77 = 31 \text{ MHz}$, and the high instantaneous luminosity is $10^{34} \text{ cm}^{-2} \text{ s}^{-1} = 10^{10} \text{ b}^{-1} \text{ s}^{-1}$, up to 20 simultaneous (piled-up) interactions are expected. At low luminosity, this number follows a Poisson distribu-

¹In fact heavy ion collision should also take place in CMS

tion with an expected mean around 4.5. Pile-up events make many analyses more difficult to do, especially those where the soft jets are important.

2.1.2 p-p interactions

The choice of using protons as colliding particles is related to the desired high luminosity. In $p\bar{p}$ colliders (SPS, Tevatron), this is indeed a major limitation. However in both cases this choice was reasonable: at SPS the discovery of the weak vector boson need quarks as initial states and at Tevatron the observation of top quark pairs events was one of the main motivation. With 1.96 TeV of c.m.s the cross-section is dominated by quark-antiquark interactions, which is not longer the case at 14 TeV, where gg interactions are predominant.

This is clearly visible from the parton density distributions (pdf) at 14 TeV, and more precisely from the computation of inclusive cross-sections for each combinations of initial partons. The cross-section of a process with two hadrons in initial states, each carrying a momentum P_i , can be written as in (2.2).

$$\sigma(P_1, P_2) = \sum_{a,b} \int_0^1 \int_0^1 \hat{\sigma}_{a,b}(p_1, p_2, Q^2/\mu^2) f_a(x_1, \mu^2) f_b(x_2, \mu^2) dx_1 dx_2 \quad (2.2)$$

The terms a and b run over the parton type (*e.g.* u, d, s, c, g, b and their antiparticles), $x_{1,2}$ are the fraction of the momentum carried by the partons, $\hat{\sigma}$ represents the partonic cross-section with $p_i = x_i P_i$, Q^2 is the scale of the process (with Q typically close to the masses involved in the process and/or still the P_T of produced massless particles), μ is the factorization scale and f_i the parton distribution function. The factorization scale is an arbitrary parameter that defines the separation between the hard and soft perturbative regimes at NLO and higher orders. Essentially, the emission of soft partons from the initial states leads to logarithmic divergencies that can be reabsorbed in the pdf at the condition that the factorization scale is close to the scale of the hard scattering process.

We define the quantity $\tau = x_1 x_2 = \hat{s}/s$ where $s = P_1 + P_2$ and $\hat{s} = p_1 + p_2$. The variables x_1 and x_2 can be rewritten as functions of τ and the rapidity of the system y

$$x_1 = \sqrt{\tau} e^y \quad (2.3)$$

$$x_2 = \sqrt{\tau} e^{-y} \quad (2.4)$$

In the τ, y basis the expression (2.2) can be rewritten as

$$\sigma = \sum_{a,b} \int \frac{d\hat{s}}{\hat{s}} dy \frac{dL_{a,b}}{d\hat{s}dy} \hat{\sigma}_{a,b} \quad (2.5)$$

where

$$\frac{dL_{i,j}}{dyd\hat{s}} = \frac{1}{s} \frac{1}{1 + \delta_{i,j}} \left[f_i(x_a) f_j(x_b) + (a \leftrightarrow b) \right] \quad (2.6)$$

is a differential luminosity term that helps to quantify the partons flux in terms of the total momentum and the rapidity. If the integration is performed over y the differential luminosities $dL/d\hat{s}$ corresponding to CTEQ6L1 pdf's and $\sqrt{s}=14$ TeV are shown in Fig. 2.2 (see also [39]). The luminosity corresponding to gg initial state is indicated in green, $\sum gq_i + g\bar{q}_i + q_i g + \bar{q}_i g$ in blue, and $\sum q_i \bar{q}_i + \bar{q}_i q_i$ in red.

We see that at low $\sqrt{\hat{s}}$ (low- x) the processes with at least one gluon dominates whereas this does not hold when \hat{s} tends to s . This implies for instance that a process with colored particles like $t\bar{t}$, corresponding to $\tau \sim 0.25$ will be mainly produced with at least one gluon.

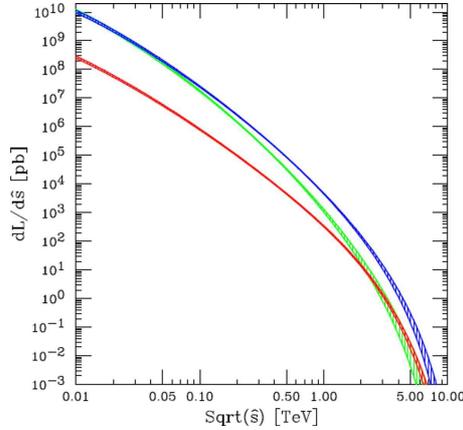


Figure 2.2: Differential luminosities for different combinations of gluons and quarks at the LHC for $\sqrt{s}=14$ TeV: gg in green, $\sum gq_i + g\bar{q}_i + q_i g + \bar{q}_i g$ in blue, and $\sum q_i \bar{q}_i + \bar{q}_i q_i$ in red. .

Both the gluons and the quarks are susceptible to emit resolvable radiation. The Monte Carlo simulation events with hard initial state radiation is a long-term work still in progression. However it is now clear that the control of the different techniques available will be crucial both at the level of the distribution normalization and in point of view of the shapes. This could indeed affect the discovery of new particles with more or less

large impact depending of the signal and background signatures. In the next section we treat the problem of the multi-jet event simulations.

2.2 Monte Carlo simulation of collisions

2.2.1 Factorization of hard and soft processes

Two common approaches exist for the modeling of a collision. The first is the matrix-element technique that allows to calculate exactly the Feynman amplitudes associated to a process, taking into account for instance the possible interferences between diagrams, the spin correlations, etc... The integration over the phase space gives an estimation of the cross-section and allows the generation of events. The second is the parton shower simulation that allows to simulate the successive parton splittings between a given, possibly hard, scattering scale and the typical scale considered for hadronization.

There are several generators that can be used for the matrix-element calculation: Sherpa [40], CalcHep [41], MADGRAPH/MADEVENT[42], CompHep [43], Alpgen [44], HELAC [45] or Whizard [46].

Besides flexibility, the particle content can be adapted to in principle any model in particle physics, this approach is appropriate only in certain regions of the phase-space. The calculation of a process cross-section with emission of massless partons will be correct only if they are sufficiently separated in the phase-space as fixed-order calculations suffer from collinear and infrared divergencies [39, 47].

In the soft and collinear limit of a gluon emission, the cross-section σ_n associated to a process with n partons, is modified by the splitting of one parton as

$$d\sigma_{n+1} = d\sigma_n \frac{\alpha_s}{2\pi} \frac{dt}{t} P(z) dz \quad (2.7)$$

where t is the virtuality of the incoming parton in the splitting, z the ratio between the energy of this parton and the energy of one of the outgoing partons in the splitting and finally P is the Altarelli-Parisi splitting function [39] which depends of the nature of the partons involved in the branching.

The expression shows that in the collinear/soft limit, the problem is factorizable into two pieces: on one hand the Born cross-section ($d\sigma_n$) is evaluated one time for all, while the term containing the divergencies can be considered separately.

From the expression (2.7), it can be shown that the case of a successive number N of gluon emissions leads (we here concentrate on the collinear divergence) to the following expression:

$$\sigma = \sigma_0 \alpha^N \int_{t_0}^{Q^2} \frac{dt_1}{t_1} \int_{t_0}^{t_1} \frac{dt_2}{t_2} \dots \int_{t_0}^{t_{N-1}} \frac{dt_N}{t_N} \quad (2.8)$$

$$\sim \sigma_0 \frac{\alpha^N}{N!} \left[\ln \left(\frac{Q^2}{t_0} \right) \right]^N, \quad (2.9)$$

where Q is the starting scale of the emission and t_0 the threshold under which no split is allowed, typically the hadronization scale.

In the practical context of Monte Carlo generations, the σ_0 term could correspond to the partonic cross-section calculated with a matrix-element method while the logarithmically divergent term is given by the parton shower simulation.

The parton shower approach helps to bridge the simulation of hard-scattering interactions and the detector level in a realistic way, leading to the creation of stable particles. For both initial and final state radiation, the algorithm creates a tree-structure of branching, controlled by the DGLAP evolution equation [48, 49, 50]. An initial state radiation sees its virtuality (or an other ordering variable) increased with the successive parton emissions while in a final state shower it decreases towards the hadronization scale. The shower structure is therefore related to the probability of having emission above a given threshold.

Assuming that a parton is generated at the virtuality t_a , the probability that no resolvable emission, *i.e.* above a given cutoff ϵ , happens before the parton reaches the virtuality t_b is given by the Sudakov form factor $\Delta(t_a, t_b)$:

$$\Delta(t_a, t_b) = \exp \left[- \int_{t_b}^{t_a} \frac{dt'}{t'} \int_{\epsilon}^{1-\epsilon} dz \frac{\alpha_s(t)}{2\pi} P(z) \right] \quad (2.10)$$

Practically, the shower algorithm controls the branchings via the equation $\Delta(t_a, t_b) = \mathcal{R}$ where \mathcal{R} is random number out from a uniform distribution on the interval $[0,1]$. Second, the value of the momentum fraction hold by each parton produced is also calculated from a random variable, according to a probability proportional to the splitting function $P(z)$.

The branchings can happen while the virtuality stays bigger than the hadronization scale in the case of final state radiation or lower than the hard scattering scale in the case of initial state radiation.

Different redefinition of evolution variable to perform showering are available in the parton shower algorithms. In PYTHIA [51], there are the virtuality-ordering and the recently developed P_T ordering. They are defined by using respectively $m^2 = E^2 - \mathbf{p}^2$ and $P_T^2 = z(1-z)m^2$. Note that the forthcoming discussion is restricted to these two cases. The Herwig [52] method is based on the emission angle $m^2/(z(1-z))$ and finally ARIADNE [53] uses $z(1-z)m^2$.

2.2.2 Matching/merging of initial state radiation

The discussion focuses on the problems to simulate properly the multi-jet event production with Monte Carlo techniques, where the additional jets are produced by initial state radiation. In principle this discussion should be extended to final state radiation, *i.e.* jets from decays, but in such a case the available phase-space is related to the mass of decaying particle whereas the ISR can benefit from whole energy available at the collision.

The first option is to simulate processes with parton shower generators only. This is possible thanks to a library of processes integrated to the software. The corresponding matrix-element are mostly limited to $2 \rightarrow 2$ mode. In the case of processes like $t\bar{t}$, $\tilde{g}\tilde{g}$, ZZ , the ISR production is therefore not controlled by the matrix-element but by the shower algorithm. This statement hold partially if one light final state parton is present in the matrix-element of the process; for instance in $Z/W/h^0 + 1$ light parton.

The first problem in this approach is that the shower model is intended to describe the parton splittings (and therefore the radiation) in the soft and collinear limit. It turns then out that the parton shower description can easily break down above a given scale (where the matrix-element calculation should be considered instead). As a consequence, if the scale of the process is hard enough, the kinematics of ISR generated by the shower algorithm only might be underestimated.

The second issue is that the shower can be easily parameterized, for instance in terms of evolution variable or shower starting scale. This therefore affects also the extra-jets² with no initial parton present at the matrix-element level.

A solution to get rid of these two questions is to maximize the number of ISR partons at the matrix-element level and merge together the simulations done by matrix-element and parton shower generators.

If both matrix-element (ME) and parton shower (PS) approaches are considered at the same time without control, a double counting between samples of different multiplicities appears. For instance two collinear partons at the matrix-element level give the

²We refer to this denomination for the jet produced by ISR.

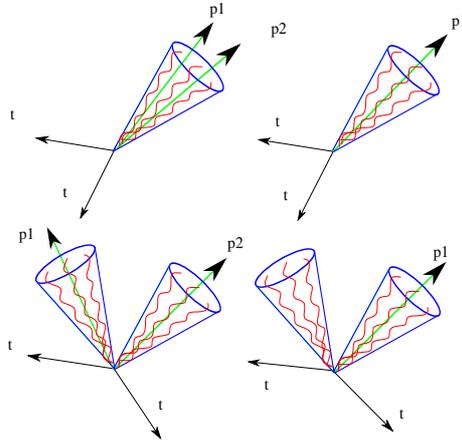


Figure 2.3: Illustration of the issue of double counting between jet topologies: on the upper left part a $t\bar{t} + 2$ collinear ME partons is shown, giving the same topology as a $t\bar{t}+1$ ME parton (upper right) with no additional resolvable PS radiation. The lower left side case if a $t\bar{t}+2$ well separated partons, no additional resolvable PS radiation and it looks the same as a $t\bar{t}+1$ ME parton and one additional resolvable PS radiation. The arrows denoted p_1 and p_2 indicate the ME partons, the blue cone refers to a given jet size definition and the red sine curves stand for PS radiation.

same topology as a single shower confined in a small region of the phase-space. In the same way, a resolvable radiation emitted while the showering is performed may exhibit an additional jet with respect to the initial number of partons. This is illustrated in Fig. 2.3. The problem of double-counting plus the soft/collinear divergencies for fixed order calculation can strongly affect the estimation of the cross-section as well as the distributions. This can be solved by using a jet matching/merging method. The principle of the jet matching is to divide the phase-space into two independent regions characterized by the hardness of QCD emissions. The region containing the soft/collinear emission is naturally ruled by the shower generator while the other is ruled by the matrix-element generator. The separation between the two regions of the phase-space is named Q_{match} and can be expressed either in terms of ΔR and P_T (Cone algorithm) or still with the k_{\perp} distance.

In the following we restrict our discussion to MadGraph and PYTHIA. In this software chain, two commonly used matching methods have been tested and used, respectively

k_{\perp} -MLM [54] and Shower- k_{\perp} [55]. Both use the k_{\perp} measure to define Q_{match} and have their working principle based on the original MLM method[56].

The k_{\perp} -MLM can be used with both virtuality and P_T ordered showers in PYTHIA, it works as follows:

- The matrix-element events are generated in MADGRAPH/MADEVENT with additional light ISR partons (for instance $W^{\pm}+0,1,2,3$ partons). The distance between these partons is defined as $Q_{\text{cut}}^{\text{ME}}$. Since the matrix-element generator is required to only populate the region above the matching scale Q_{match} , there is in principle no reason to choose $Q_{\text{cut}}^{\text{ME}}$ different from Q_{match} . However, as the k_{\perp} -MLM matching method cuts on the combined radiation from the showers, a smearing across the matching scale may happen. Therefore if $Q_{\text{cut}}^{\text{ME}} = Q_{\text{match}}$, events with a jet configuration close to Q_{match} in one of the differential jet rate transition (see Sec. 2.2.3) can be lost while the event rejection takes place. In this case $Q_{\text{cut}}^{\text{ME}}$ must be smaller than Q_{match} .
- A reweighting by $\alpha_s(k_{\perp})$ at each vertex is performed in order to mimic what is done in the shower algorithms. This helps to get a smooth transition between both regions of the phase-space.
- The showering is performed with PYTHIA.
- The showered partons are clustered into jets using the k_{\perp} algorithm. The scale definition of the jets is Q_{match} .
- If the event does not have the highest (partonic) multiplicity (e.g. $W^{\pm}+2$ partons in a $W^{\pm}+0,1,2,3$ parton production), it has to be treated exclusively. This means that each jet has to be matched with a parton and vice-versa within a distance defined by Q_{match} . If the event has the highest multiplicity the additional jets not produced by the matrix-element partons should be kept. Therefore the only requirement is to have each matrix-element parton matched with a jet. If the matching condition is not fulfilled, the event is simply rejected.

The shower- k_{\perp} scheme works exclusively with P_T -ordered showers. The principle of the rejection is the same as for k_{\perp} -MLM, but the criteria are slightly different. The algorithm is the following:

- The events are generated by MADGRAPH/MADEVENT as for k_{\perp} -MLM, including the reweighting of α_s .
- Each event is then passed to PYTHIA and showered using the P_T -ordered showers. PYTHIA reports the scale of the hardest emission, $Q_{\text{hardest}}^{\text{PS}}$ in the shower.

- The events from lower-multiplicity samples are rejected if $Q_{\text{hardest}}^{\text{PS}}$ is above the matching scale Q_{match} , while events from the highest multiplicity sample are rejected if $Q_{\text{hardest}}^{\text{PS}} > Q_{\text{softest}}^{\text{ME}}$, the scale of the softest matrix element parton in the event.

This matching scheme, although simple, effectively mimics the behavior of the k_{\perp} -jet MLM scheme. However, it allows for the matching scale Q_{match} to be set equal to the matrix element cutoff scale $Q_{\text{cut}}^{\text{ME}}$.

As shown in Sec. 2.2.3, the physical outputs of both methods are, above Q_{match} , quite similar.

Note that while the principle of MLM methods is based on event rejection, the complementary point of view is possible and used in the CKKW matching method[57]. In this method, the showers are controlled so that each parton in the events of the lower-multiplicity samples are matched with one jet and vice versa. There is therefore no rejection but a reweighting by the probability to not have additional radiation above the matching scale Q_{match} has to be applied. This is called the Sudakov reweighting since the definition of the Sudakov form factor is precisely linked to this probability.

2.2.3 Control of the matching parameters

As depicted in the previous sections, the matching procedures rely on a small set of parameters, basically $Q_{\text{cut}}^{\text{ME}}$ and Q_{match} . It is said in the description of the k_{\perp} -MLM method that the value of $Q_{\text{cut}}^{\text{ME}}$ has to be chosen below Q_{match} , but except from this precise case, $Q_{\text{cut}}^{\text{ME}}$ can be chosen equal to Q_{match} ; this means that the parameterization of the matching is essentially related to Q_{match} .

The matching scale is not a physical observable, therefore the choice of the value of Q_{match} is arbitrary. However, the transition between parton shower and matrix-element regimes has to be smooth and the cross-section must be stable with respect to the variation of Q_{match} . These requirements can be influenced by the choice of too low or too high values of the matching scale. For instance, if the value is set too high, dead zones may appear, *i.e.* regions of the phase-space that the parton-shower is unable to fill correctly. On the contrary, if the matching scale is too low, divergencies may appear at the level of the matrix-element generation. This is why the cross-section stability with respect to small variations of Q_{match} is a good indicator.

This last argument is illustrated in Tab. 2.1 for the case of W+0,1,2 ME parton, treated with the k_{\perp} -MLM method.

| $(Q_{\text{cut}}^{\text{ME}})Q_{\text{match}}$ | $\sigma(\text{nb})$ | | | $\sum \sigma(\text{nb})$ | $\epsilon_M(\%)$ |
|--|---------------------|---------|----------|--------------------------|------------------|
| | $N_j=0$ | $N_j=1$ | $N_j=2+$ | | |
| (5)7 | 1.9 | 1.56 | 1.82 | 5.27 | 30 |
| (7)10 | 2.4 | 1.37 | 1.19 | 5.0 | 39 |
| (10)15 | 3.0 | 1.2 | 0.7 | 5.0 | 50 |
| (20)30 | 3.4 | 0.9 | 0.4 | 4.8 | 63 |
| (30)40 | 3.8 | 0.6 | 0.2 | 4.7 | 72 |

Table 2.1: Cross-sections at the LHC (14 TeV) by multiplicity, total cross-section and k_{\perp} -MLM matching efficiencies as a function of the $Q_{\text{cut}}^{\text{ME}}$ and Q_{match} parameters, for W^- +jets, with up to 2 ME partons. The "+" prescription indicates the inclusiveness of the treatment of the highest multiplicity.

Since both matching implementations employed here rely (totally or partially) on the Durham k_{\perp} measure to achieve the separation of the phase space, the most revealing distributions to study their features are the differential jet rates defined according to the same measure.

The differential jet rate variable is defined by the k_{\perp} scale at which an event pass from a $N+1$ jet(s) configuration to a N jet(s) configuration while the k_{\perp} algorithm clusters the showered partons. After the matching procedure, all showered partons belonging to a same shower are grouped together within a k_{\perp} distance smaller than the cutoff. On the contrary, the jets issued each from a different ME parton lay at a k_{\perp} distance larger than this cutoff. Therefore in a transition denoted " $N+1 \rightarrow N$ ", the events with a multiplicity smaller than $N+1$ will be situated below the cutoff, while the events with a multiplicity higher or equal to $N+1$ are situated above this cutoff.

This statement is graphically explained in Fig. 2.4 in the case of $t\bar{t}+0,1,2,3$ jets.

In the k_{\perp} -jet MLM scheme, there is at parton level a sharp division in the jet rates between the shower and matrix element regions, making it very easy to see to which extent the transitions are smooth. For the shower k_{\perp} scheme, as well as the cone jet MLM scheme implemented in ALPGEN and the CKKW scheme implemented in SHERPA, the separation is less sharp, but the differential jet rates still tend to be the best variables to study the transition between parton showers and matrix elements.

To illustrate the behavior of the differential jet rates, we consider the process $t\bar{t}+0,1,2$ partons. As it turns out, the P_T -ordered PYTHIA showers allows significantly higher matching scales than the virtuality-ordered showers. The reason for this is that they give significantly harder emissions than the virtuality-ordered showers, and therefore

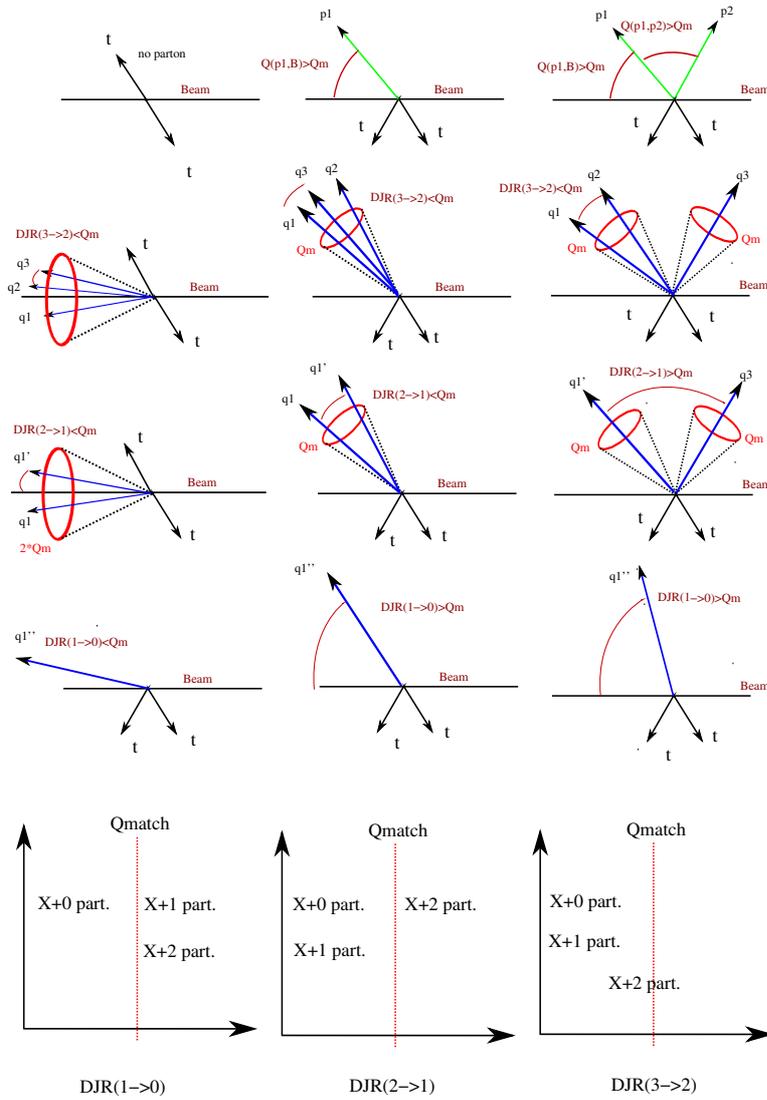


Figure 2.4: Illustration of differential jet rates construction ($3 \rightarrow 2, 2 \rightarrow 1, 1 \rightarrow 0$) for ISR control (here with $t\bar{t} + 0, 1$ and 2 jets). The notation p refers to an original parton at ME level. The notation q, q' and q'' refers to the showered partons, and the clustered partons for the successive clustering iterations.

give distributions more similar to the matrix element distributions, at least up to a given scale. For virtuality-ordered showers, $Q_{\text{cut}}^{\text{ME}} = 20$ GeV and $Q_{\text{match}} = 30$ GeV, for P_T -ordered showers $Q_{\text{cut}}^{\text{ME}}$ and $Q_{\text{match}} = 100$ GeV.

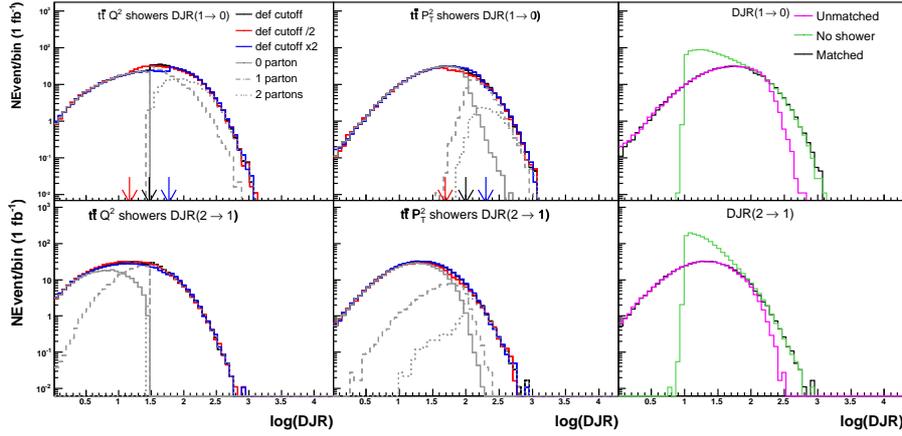


Figure 2.5: Differential jet rates for $1 \rightarrow 0$ and $2 \rightarrow 1$ jets from QCD radiation for $t\bar{t}$ production at the LHC. The first two columns show the distributions for the two types of PYTHIA showers, virtuality-ordered and P_T -ordered, for three different choices for the matching scale. For the default choice, also the contributions from the separate multiplicity samples are shown. The colored arrows show the value of the three Q_{match} used for each kind of shower. The third column shows how the matched curve interpolates between the pure parton shower curve and the pure matrix element curve without parton showering.

In Fig. 2.5 we show the $2 \rightarrow 1$ and $1 \rightarrow 0$ differential jet rates. There is no need to look at higher orders because the maximal parton multiplicity considered here is 2, which means that the transition $3 \rightarrow 2$ is given entirely by the parton shower. The first and second columns show the result for the virtuality and P_T -ordered showers respectively, with the first and second rows indicating the $1 \rightarrow 0$ and $2 \rightarrow 1$ transitions, with a variation by a factor 2 of the default Q_{match} .

The difference between the two matching schemes is visible in the plots, in the different behavior of the parton multiplicity sample contributions. The left-hand column for each particle type shows the k_{\perp} -MLM matching scheme, with the contributions from the different parton multiplicity samples in grey. The matching scale cutoff is, in this scheme, done in the same variable that is plotted, the differential jet rate, and there is therefore a sharp cutoff between the 0- and 1-parton samples in $\text{DJR}(0 \rightarrow 1)$, and between the 1- and 2-parton samples in $\text{DJR}(1 \rightarrow 2)$, so that below the cutoff

only the lower-multiplicity samples contribute and above the cutoff only the higher-multiplicity samples.

In the middle column, the Shower- k_{\perp} method is used (with the P_T -ordered PYTHIA showers). This method cuts on the first emission of the parton shower rather than on the combined radiation of the whole shower, giving some smearing across the matching scale. This scheme allows to use the same cut at matrix element and parton shower levels. The distributions for the P_T -ordered showers have been double-checked using the k_{\perp} -jet MLM matching method, with excellent agreement.

In both left and middle columns, the arrows indicate the different values of Q_{match} considered. The corresponding global DJR curves are shown, showing a good stability with respect to the variation of Q_{match} .

In the right hand-side column, the curve for unmatched PYTHIA showers with default parameters are shown (red), together with the pure matrix element prediction without any parton showering or matching (green) and the matched curve (black). We see that the matched curve smoothly interpolates between the unmatched PYTHIA curve below the matching scale, and the matrix element prediction for large scales. This last statement is very important because at large scale the behavior has to follow the pure matrix element calculation.

It is also interesting to notice the differences in curve shapes depending on the choice of shower type. Below the matching scale, the shape of the curve is given completely by the shower, in particular for the $0 \rightarrow 1$ jet rate. Above the matching scale, however, the shape is mainly given by the matrix element. It is easy to see the reason for the different choices of matching scales for the different showers – the P_T -ordered shower gives significantly harder distributions than the virtuality-ordered shower, and is more similar to the matrix element curve, hence allowing a higher matching scale.

In conclusion the differential jet rates give a lot of information on the matching result both in the behavior of each multiplicity but also at a global (physical) point of view. Using such variable constitutes an important sanity check to be sure that a production with matching has been correctly performed. In the appendix B, we propose a easy solution to obtain these distributions.

In the next sections we review the main effects of jet matching in the case of Standard Model. The production of heavy colored particles in the context of the SM and SUSY is then discussed in details.

2.2.4 Multi-jet event production in the SM

The correct simulation of multi-jet event within the Standard Model is important as for instance the QCD events and productions of a vector boson with jets have, at the LHC, a very large cross-section compared to most production mechanism beyond the SM. The LO cross-section of some of the main SM processes are given for two center of mass energy (10 TeV and 14 TeV) in Tab. 2.2, using PYTHIA and the k_{\perp} -MLM jet matching for virtuality-ordered showers. The cutoffs $Q_{\text{cut}}^{\text{ME}}$ and Q_{match} used both at the ME level and at the PS level are indicated as well as, if needed, the minimal H_T value defined as the sum of the P_T of the partons ($u, d, s, c, (b)$). This cut is useful to get reasonable statistics far in the tails without having to populate the bulks with millions of events. The basic cuts required are $|\eta_j| < 5$ and $P_T = Q_{\text{cut}}^{\text{ME}}$. In addition, for Z^0/γ , the minimal invariant mass of the two leptons is required to be larger than 50 GeV.

| Process | Particular cut | $Q_{\text{cut}}^{\text{ME}}$ | Q_{match} | $\sigma(\text{pb})$ (10TeV) | $\sigma(\text{pb})$ (14TeV) |
|----------------|------------------------------|------------------------------|--------------------|-----------------------------|-----------------------------|
| QCD | $H_T \in [100, 250]$ | 20 | 30 | 1.5×10^7 | 2.4×10^7 |
| QCD | $H_T \in [250, 500]$ | 20 | 30 | 4×10^5 | 7.7×10^5 |
| QCD | $H_T \in [500, 1000]$ | 40 | 60 | 1.4×10^4 | 3.6×10^4 |
| QCD | $H_T \in [1000, \text{inf}]$ | 40 | 60 | 3.7×10^2 | 1×10^3 |
| B-enr. | $H_T \in [100, 250]$ | 20 | 30 | 4.5×10^5 | 9×10^5 |
| B-enr. | $H_T \in [250, 500]$ | 20 | 30 | 1.5×10^4 | 5×10^4 |
| B-enr. | $H_T \in [500, 1000]$ | 20 | 30 | 7×10^2 | 4×10^3 |
| B-enr. | $H_T \in [1000, \text{inf}]$ | 40 | 60 | 13 | 1.5×10^2 |
| $W^{\pm}+j$ | / | 10 | 15 | 4×10^4 | 6×10^4 |
| $Z^0/\gamma+j$ | / | 10 | 15 | 3.7×10^3 | 7×10^3 |
| $t\bar{t}+j$ | / | 20 | 30 | 3.2×10^2 | 7.5×10^2 |

Table 2.2: Cross-sections of some of the main Standard Model Processes. The first column indicates the type of process: the ‘‘QCD’’ appellation concerns the inclusive production of 2,3 and 4 ME partons (j=u,d,s,c,b); B-enr. goes for the b quark enriched samples, defined by $b\bar{b}+0,1,2,3$ ME partons + $b(\bar{b})+1,2,3$ ME partons (jet=u,d,s,c). W^{\pm} and Z^0/γ decays into e, μ, τ and are accompanied by 0,1,2,3,4 ME partons of the five flavors, whereas the $t\bar{t}$ has only up to 3 extra-partons.

Due to their large cross-sections, any unknown systematic effect originating from Monte Carlo generators could easily lead to a misinterpretation of the data.

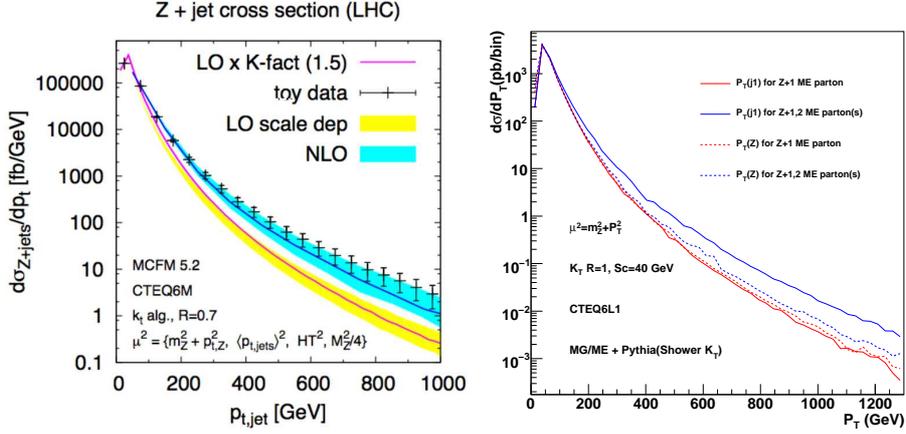


Figure 2.6: Left: P_T of the leading jet with a LO simulation (red line and yellow band) and with a NLO simulation done with the generator MCFM (blue line and blue band). Right: P_T of the leading jet (in blue) and Z^0 with (plain lines) and without (dashed line) the matrix-element calculation of the kinematics of a second emitted parton.

An example of this kind of issue is illustrated by the simulation of the Z +jet process with a LO and a NLO generator (see Ref. [58]). The LO generation is made in a $2 \rightarrow 2$ more precisely the matrix element is $Z+1$ parton (Fig. 2.7(a)). The NLO mode takes into account one additional jet compared to the LO generation.

The P_T of the leading jet simulated with a LO simulation (red line and yellow band) and a NLO simulation from the generator MCFM (blue line and blue band) is shown in Fig. 2.6(left). The 1.5 K-factor helps to normalize the LO cross-section to the NLO estimation. This normalization is largely weighted by the bulk region (very low P_T) and it appears that besides this problem of normalization a discrepancy between the shapes is also visible at large P_T , this difference reaching almost one order of magnitude at the TeV scale.

The second jet ruled by the matching in the process $Z+1,2$ parton can play the role of real corrections to the LO generation, mimicking at least a part of the NLO behavior. To check this and understand the origin of the discrepancy in the large P_T region we have reproduce the result of Fig. 2.6(left) with the Shower- k_\perp matching method.

In the Fig. 2.6(right) we show the P_T spectrum of the Z boson and leading jet in P_T with and without the Shower- k_\perp matching method. The dashed lines indicate

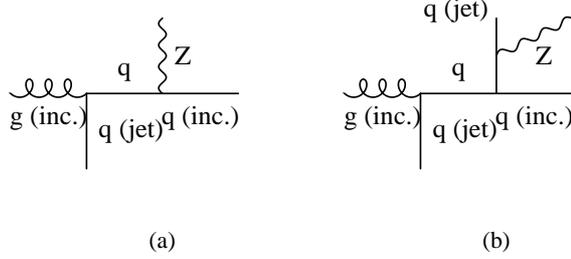


Figure 2.7: Representative processes for $Z+1$ ME parton (left) and $Z+2$ ME partons (right) in which there is a possible collinearity between a jet and the Z .

the distribution in P_T of the Z boson while the solid lines show the distributions of the leading jet in P_T . The blue and red colors denote respectively the matched and unmatched productions.

It is clear that the shape discrepancy observed in Fig. 2.6(left) is also present when passing from $Z+1$ parton to $Z+1,2$ partons. This shows that the matching is able to reproduce trustfully some NLO effects.

Besides this we can see that the P_T distribution of the Z is not affected when passing from $Z+1$ jets to $Z+1,2$ jets. This tends to show that the excessive recoil taken by the leading jet is the consequence of the presence of a second jet relatively close to the Z . This situation cannot be represented in a $2 \rightarrow 2$ process done with a PS simulation standalone as it should mean to have a Z "radiated" by a jet. This is illustrated in Fig. 2.7(b). The distance (here $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) between the Z and the closest jet is shown in Fig. 2.8. This confirms the expectation about their collinearity.

This result shows the importance of considering the jet matching in cases supposed to be rather simple or well known. This suggests also that up to a given level of approximation the simulation using a fast and adaptative tree-level generator can be useful even as a stand-in for NLO calculations.

In the next section we generalize the study of the impact of jet matching. To do so we emphasize on the case of heavy colored particle production, for which the level of QCD radiation is expected to be the largest and therefore the most problematic for analyses.

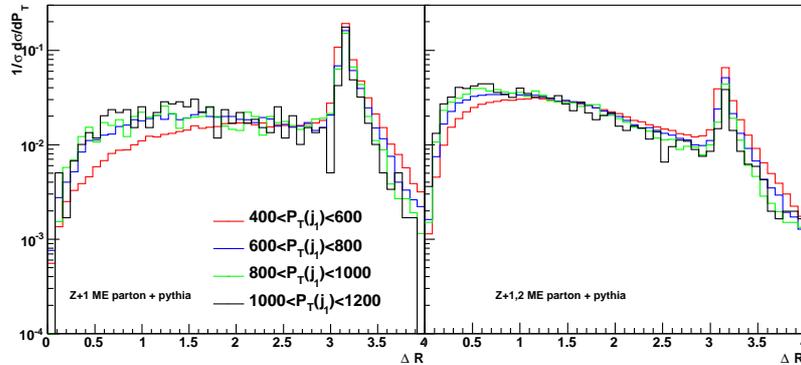


Figure 2.8: The distribution show the distance ΔR between the Z boson and the closest jet. The left hand-side figure is done using $Z + 1$ partons while the distributions of the right hand-side plots takes also into account the second parton at matrix-element level. The different colors give the slice cutting (in P_T) of the phase-space, showing that the small angle region is more populated in the second case and since the P_T range is higher.

2.3 Multi-jet event production with heavy colored particles

In this section we first study the QCD radiation accompanying the production of heavy colored states obtained from either a standalone parton shower (PYTHIA) and from a matrix element plus parton shower (MadGraph + PYTHIA). In the rest of this chapter we consider the production of SUSY particles as a toy-model because it contains the commonly considered heavy colored gluinos and squarks.

Parameter dependence in matched and unmatched generation

As described above, the parton shower Monte Carlo generators have different built in parameters which modify the behavior of the shower, and hence allows tuning to data. This means that the parton shower has a lack of predictability: it is not in general clear that a tune done for one type of initial state will be applicable to other initial states, or that a tune done for a particular mass of a pair-produced particle will be applicable for other masses.

The discussion is here focused on the showers emitted by initial state QCD radiation, so no jets from decays are taken into account due to the fact that in principle their

kinematics are well defined and the result shouldn't be too much dependent of the shower scheme/parameterization chosen. In order to study and compare the systematic uncertainties involved in parton showering, we use the two different showering implementations of PYTHIA described before: the virtuality-ordered and P_T -ordered showers, with a range of shower parameters similar to that in [59]³. We also use two different matching schemes, the k_\perp -jet MLM scheme for virtuality-ordered showers and the shower k_\perp scheme for P_T -ordered showers. The description of these techniques has been done in Sec. 2.2.2.

For most of the studies here below, the other shower parameter that is varied is the starting scale of the shower. This is the most important parameter determining the hardness of radiation allowed in parton shower emissions, and the default value has varied over the years and is still dependent of the processes stored in the software bank.

For both types of shower, we follow the nomenclature of Ref. [59] and call the appellation "wimpy" and "power" for the lowest and uppermost limits of hardness of the showers. The wimpy shower use the factorization scale as the maximal starting scale while for the power shower the whole available phase space may be used.

In order to see how this lack of predictability is reduced by using the kinematics calculated at the matrix-element level, we first compare the results obtained with $g\tilde{g}+0, 1, 2$ ME and $g\tilde{g}+0$ ME partons (equivalent to the $2 \rightarrow 2$ mode of PYTHIA standalone). The P_T distributions of the two leading extra-jets is shown in Fig. 2.9. We include also the variation of both factorization and renormalization scales by a factor 1/2 and 2 for one of the showers parameterization, in the matrix element as well as the parton showers⁴. The mass of the gluinos is fixed at 607 GeV/c² (benchmark SPS1a). The different curves correspond to the different shower parameters settings: the virtuality-ordered shower and the P_T -ordered implementation, each with two different choices for the starting scale of the shower, wimpy and power showers. Several interesting features can be noted from Fig. 2.9:

- The spread in predictions for the parton shower is very large and strongly affects the shapes of the distributions. This uncertainty due to shower parameters is almost completely removed when matching is applied.
- The region where the shower predictions start to diverge, and the rate of this divergence, is strongly correlated with the mass of the produced particles. This

³Note that the virtuality and P_T ordering of the showers are called in PYTHIA by `MSTP(81)=0` and `20`, or `1` and `21` if the multiple parton interaction switch is activated, which is not the case in this study

⁴The factorization and renormalization scales can be changed respectively with the parameters `scalefact` and `alpsfact` in MADGRAPH/MADEVENT and `PARP(64)` and `PARP(72)` in PYTHIA

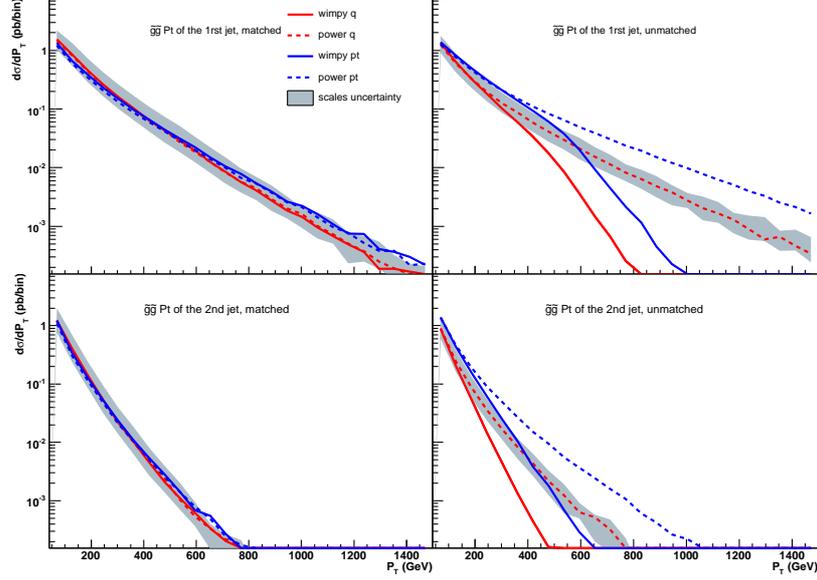


Figure 2.9: High- P_T spectrum for the first and second hardest radiated jet in the production of $\tilde{g}\tilde{g}$ pairs. The upper row shows the spread of PYTHIA predictions with different choices of the shower evolution variable (virtuality- and p_T^2 -ordered) and starting scale for the evolution (labeled as “wimpy” and “power” showers respectively). The lower row presents the results obtained after matching ME with PS in the same four shower scenarios. The grey band shows the systematic uncertainty associated with a variation of a factor of two of the renormalization and factorization scales with respect to their central values. The different curves have a relative normalization corresponding to their relative cross-section.

correlation is due to the choice of starting scale for the wimpy showers since the factorization scale, which is close to the mass of the produced particle.

- The power shower curves consistently overshoot the matched curves, and hence give too hard predictions, while the wimpy showers give too soft distributions.
- The uncertainty due to scale variations is considerable, but mainly affects the normalization and only to a small degree the shape of the curves.

Furthermore, it is observed that the discrepancy between the unmatched and matched productions of ISR is also dependent of the type of the particles produced (gluinos or squarks) as well as their masses. The full discussion can be found in Ref. [55].

2.3.1 Anatomy of E_T^{miss} + multi-jet final states

While the previous section dealt with jets from QCD radiation only, in real life we can expect new heavy colored particles produced at the LHC to decay into jets and missing transverse energy. It is therefore rather interesting to study the impact of jet matching in the context where we also have jets from decays.

In general, one could expect that jets from QCD radiation should be relatively unimportant compared to the hard decay jets from heavy particles. For example a pair of squarks with a mass around 600 or 700 GeV/c^2 will easily give jets with a $\langle P_T \rangle \sim 200 \text{ GeV}/c^2$ while the initial state radiation are peaked to zero.

In order to clarify the discussion, and keep our results conservative, in this section we will use a set of simplified supersymmetric benchmark scenarios, summarized in Table 2.3. In all the scenarios we assume all light-flavour squarks to have the same masses and that they all decay directly to the LSP, i.e. we ignore the existence of intermediate weak states. Introduction of cascade decays will have as main effect that jets from decays get softer, while the jets from QCD radiation are not affected, and will hence mainly further accentuate our results.

| Scenarios | | I | II | III |
|-----------------------------|-------------|-----|-------|-------|
| | \tilde{g} | 607 | 607 | 607 |
| Masses (GeV/c^2) | \tilde{q} | 560 | heavy | heavy |
| | χ_1^0 | 100 | 100 | 500 |

Table 2.3: Benchmark scenarios employed in this work as modifications of SPS1a [60]. We always assume the squarks decaying 100% into quark+lightest neutralino.

For ease of comparison between the scenarios, we have chosen to use the same masses, around 600 GeV/c^2 for the active heavy QCD states in all scenarios. Scenario I has a SUSY QCD spectrum similar to the SPS point 1a [60], with a gluino at 607 GeV/c^2 which decays to squarks at roughly 560 GeV , while the LSP is at 100 GeV . In scenario 2, the gluino has a mass of 607 GeV/c^2 but all squarks are heavy, so that the gluino decays through offshell squarks to two quarks and the LSP. A representative feynman diagram is shown in Fig. 2.10.

Finally scenario III has a gluinos at 607 GeV/c^2 decaying through offshell heavy squarks, but the LSP mass is 500 GeV , only 100 GeV/c^2 lighter than the gluino. These scenarios will act as case studies to illustrate different effects of QCD radiation in the production of new heavy QCD states.

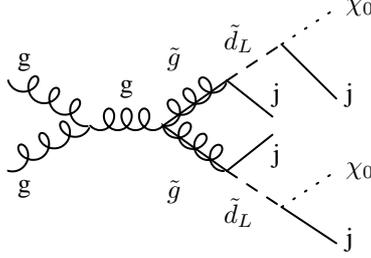


Figure 2.10: Representative Feynman diagram for the productions in the scenario I and II.

All plots and results in this section are generated using matched samples with the default PYTHIA parameter choices for virtuality-ordered showers, unless otherwise stated.

H_T^{jet} variables in gluinos production

In order to study squark pair production one should select 2-jet observables, for associated gluino-squark production 3-jet observables and for gluino pair production 4-jet observables. While it is true that gluinos decay to two quarks and χ_1^0 , it should be kept in mind that the visibility of these jets depends strongly on the mass hierarchy of QCD states. In Fig. 2.11 we show the $H_T^{\text{jet}}(n)$ distributions, defined as

$$H_T^{\text{jet}}(n) \equiv \sum_{i=1}^n |P_{T_i}^{\text{jet}}| \quad (2.11)$$

for $n = 2, 3, 4$, for the scenarios I and II. The sum in eq. (2.11) is taken over jets defined using the SISCone [61] algorithm with a radius of 0.5 and $P_T > 40$ GeV. This choice is made to get close to standard jet algorithm tunes used for jet reconstruction in experimental analyses. We also show the composition of $H_T^{\text{jet}}(n)$ in terms of jets from the gluino decay and ISR's, as well as show the average fraction of the H_T coming from the decay. For scenario II, where the gluinos decay through off-shell squarks to two quarks of similar energies and an LSP, the majority of events in the peak of $H_T^{\text{jet}}(4)$ include only jets from the decay, while the tail of the distribution is dominated by 2-3 jets from the decay and one jet from radiation. For scenario I however, where

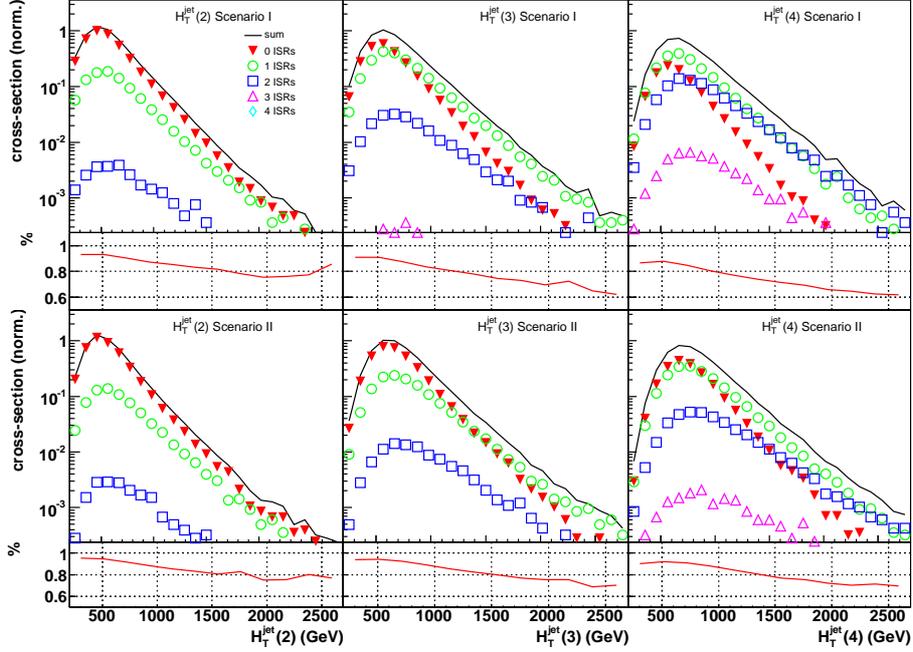


Figure 2.11: $H_T^{\text{jet}}(n)$ for $n = 3, 4, 5$ in $\tilde{g}\tilde{g}$ production. The different point markers show how many of the jets entering in the definition of $H_T^{\text{jet}}(n)$ come from QCD radiation. Upper row: Scenario I, lower row: Scenario II. Below each main plot, the red curve indicates the percentage of H_T coming from the decays. Jets are defined using the SIScone algorithm with a P_T^{min} of 40 GeV and a radius of 0.5.

the 607 GeV/ c^2 gluino decays into a fairly soft jet (with an energy around 50 GeV) and a squark which in turn decays to a hard jet and an LSP, the distributions are quite different. Here, $H_T^{\text{jet}}(4)$ is dominated by events where at least one of the jets come from radiation, and we need to go down to $H_T^{\text{jet}}(2)$ to be dominated by events with only decay jets across the whole H_T range. The energy fraction of $H_T^{\text{jet}}(4)$ coming from decay is still fairly high, even in the high- H_T tail, since most of the transverse energy comes from the squark decay jets.

The immediate interpretation of this result is that in a scenario with a small mass splitting between gluinos and squarks, gluino production might be difficult to distinguish from squark production with additional QCD radiation.

Jet multiplicities for different scenarios

A question that the proper matching of jets is particularly apt to answer, is to specify the number of jets typically present in production of different particles. We here present a table with the jet multiplicities, for matched and unmatched (Pythia virtuality-ordered default) production, for the scenarios studied. In order to make the table as useful as possible, we have used jet cuts close to what is used in many preparatory analyses for squark and gluino searches: $P_T^{\text{jet}_1} > 180 \text{ GeV}/c, P_T^{\text{jet}_2} > 110 \text{ GeV}, P_T^{\text{jet}_j} > 50 \text{ GeV}/c$ for $j > 2$ and $|\eta^{\text{jet}}| < 3$ for all jets. We require all events to have at least two jets. The jet multiplicities are exclusive, and so add up to 100% of the events passing the 2-jet cut. The jet algorithm used is SIScone with a radius of 0.5 and $P_T > 40 \text{ GeV}$.

The difference between matched and unmatched generation is twofold, as emphasized in Ref. [55]: the ISR jets get harder (compared to default Pythia starting scale), what helps them to get above the threshold of $50 \text{ GeV}/c$. The matching tends therefore to increase the transverse boost of the produced pair, and hence the P_T of the softest jets from the decays. The matched production therefore in general populates higher jet number bins than the unmatched ones.

| Process | $N = 2$ | | $N = 3$ | | $N = 4$ | | $N \geq 5$ | | Signal eff. | |
|-----------------------------|---------|------|---------|------|---------|------|------------|------|-------------|------|
| | M | U | M | U | M | U | M | U | M | U |
| $\tilde{g}\tilde{g}$ sc.I | 15.7 | 27.1 | 30.0 | 33.9 | 24.4 | 21.0 | 29.7 | 17.8 | 43.7 | 40.4 |
| $\tilde{g}\tilde{q}$ sc.I | 35.2 | 39.4 | 32.5 | 33.8 | 17.5 | 16.0 | 14.6 | 10.4 | 31.9 | 28.3 |
| $\tilde{q}\tilde{q}$ sc.I | 40.2 | 48.1 | 33.4 | 32.0 | 15.9 | 12.9 | 10.3 | 7.2 | 16.9 | 16.0 |
| $\tilde{g}\tilde{g}$ sc.II | 4.4 | 4.7 | 19.5 | 22.1 | 27.1 | 29.2 | 49.1 | 43.7 | 43.8 | 40.1 |
| $\tilde{g}\tilde{g}$ sc.III | 21.5 | 28.4 | 32.6 | 37.0 | 23.9 | 21.0 | 21.8 | 13.6 | 4.7 | 3.0 |

Table 2.4: Contribution of events with N jets for matched and unmatched processes: $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$ and $\tilde{q}\tilde{q}$ in Scenario I (for squark production this is very similar to scenario III), and $\tilde{g}\tilde{g}$ in Scenario II and III. All numbers are in percent. “Signal efficiency” shows the percentage of events that pass the 2-jet cut. The jet cuts are described in the text. The label “M” means “matched” while “U” means “unmatched”.

The first three lines in Table 2.4 represent Scenario I, which is very similar to the benchmark point SPS1a. In this scenario, the gluinos decay as $\tilde{g} \rightarrow \tilde{q}q \rightarrow \tilde{q}q\chi^0$ (with $m_{\chi^0} \sim 100 \text{ GeV}$), so there are typically two hard and two soft jets from the gluino decay. Here the effect of the matching is large, due to the increase in P_T for the soft jets as well as increased hardness of the ISR jets. For scenario II on the other hand, where gluinos decay as an equivalent three body decay with two jets and a light

neutralino, the sensitivity to the matching is much lower, since there, typically at least four reasonably hard jets are present from the decays. It is only in this and similar cases that the statement that gluino pair production generally corresponds to four hard jets in the event is true. In the last row in the table, the produced gluinos decay to a near-degenerate LSP, meaning that most hard jets are due to QCD radiation. Here, a large recoil against initial state jets is needed in order to even pass the 2-jet cut, hence the very low signal efficiency.

For squark-squark and gluino-squark production for Scenario I (row 2 and 3 in Table 2.4, there are only two hard partons from the decays (and one additional soft parton from the gluino decay) and additional QCD radiation. Again the main consequence of the matching is to increase the mean number of jets. It is interesting to note that in this particular case, the addition of matching to the generation of squark pair production gives very similar numbers to the unmatched $\tilde{g}\tilde{q}$ associated production (although the selection efficiency is different), indicating that the matching has a similar impact to the addition of one extra jet with $P_T \sim 50$ GeV.

The discussion about the matching impact on inclusive production is relatively broad and several pathologic case can be identified. These cases are discussed in details in Ref. [55].

2.3.2 Observability of gluinos and squarks at the LHC

We consider the effects of an accurate simulation of QCD radiation in the typical observables employed in the BSM inclusive searches characterized by high P_T jets and high missing transverse energy. In order to see the effect of the matching on the sensitivity to shower after a smearing of the signal, and in order to be as complete as possible, we consider the production of $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$, $\tilde{q}\tilde{q}^{(*)}$ and $\tilde{t}_{1,2}\tilde{t}_{1,2}^*$, with \tilde{q} defined as $\tilde{u}_{L,R}$, $\tilde{d}_{L,R}$, $\tilde{s}_{L,R}$, $\tilde{c}_{L,R}$, $\tilde{b}_{1,2}$. The signal is produced in both the matched (2 \rightarrow 2, 3, 4) and unmatched (2 \rightarrow 2) modes in the the SPS1a benchmark scenario [60]. For the background, we consider the most important processes leading to four hard jets and potentially large missing transverse energy: $W^\pm \rightarrow l^\pm\nu + 4$ jets, $Z^0 \rightarrow \nu\nu + 4$ jets, $W^\pm \rightarrow \tau_{jet}\nu + 3$ jets and finally the inclusive $t\bar{t}+0,1,2,3$ jets. We do not include QCD multi-jets production, since we have no means of realistically performing simulations of the missing energy distribution, which is due to decays of heavy quarks to neutrinos and jet mismeasurement in the detector. We instead base our analysis on cuts similar to those used in Refs. [19] and [62], and keep this contribution in mind. All background simulations are done using jet matching. Many comparisons have been done between matched and unmatched background simulations, which are well known to differ by up to several orders of magnitude for this type of multi-jet observables [63, 64, 65]. We therefore here look at the effects of including matching only in the signal simulation.

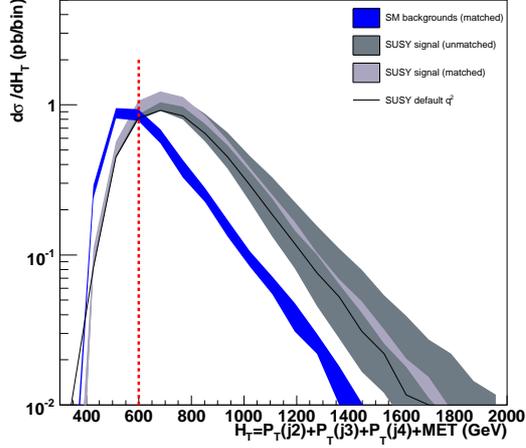


Figure 2.12: H_T for SUSY production in the SPS1a benchmark scenario, compared to (matched) Standard Model backgrounds. The grey band shows the spread of unmatched PYTHIA predictions with varying shower parameters, while the green band shows the corresponding matched predictions. H_T definition and cuts are described in the text.

In order to get a more realistic perception of the jets, we choose to simulate a detector response using PGS 4 [66]. The MidPoint cone algorithm is considered and the jet definition based on a with a minimum P_T of 40 GeV/c and a radius of 0.5. To be conservative, we use only kinematic variables associated to the jets and the missing transverse energy. The cuts used are

- $N_{\text{jet}} \geq 4$
- $|\eta_1| < 1.7, |\eta_{2,3,4\dots}| < 3$
- $P_T^{\text{jet}_1} > 180 \text{ GeV}/c, P_T^{\text{jet}_2} > 110 \text{ GeV}/c, P_T^{\text{jet}_{>2}} > 50 \text{ GeV}/c$
- $\cancel{E}_T > 150 \text{ GeV}/c$
- $\Delta\phi(\cancel{E}_T, \text{jet}_1) > 0.5$ and $\Delta\phi(\cancel{E}_T, \text{jet}_2) > 1$
- $\sum_{i=2}^4 P_T^{\text{jet}_i} + \cancel{E}_T > 600 \text{ GeV}/c^2$.

The SPS1a scenario is affected by several of the difficulties described in Sec. 2.3.1. The gluino has a mass higher than, but close to, the squarks, and hence decays to a squark and a soft jet, making the jet counting complicated. Since gluinos and squarks

are of similar mass, the QCD SUSY production includes associated squark-gluino production, gluino pair production and squark pair production (both $\tilde{q}\tilde{q}$ and $\tilde{q}\tilde{q}^*$ which are of similar cross-section, and to a negligible degree $\tilde{q}^*\tilde{q}^*$), in order of cross-section. This means that the PYTHIA shower cannot simultaneously describe all production modes, as demonstrated in Ref. [55]. Furthermore the separation of the different production modes based on jet counts or jet kinematics will be non-trivial. The production cross-section is however in this scenario dominated by associated gluino-squark and gluino pair production, where the default PYTHIA description is reasonably good, and only undershoots the matched description by about 10-30% for the first couple of QCD radiation jets, so we expect the inclusive PYTHIA description to be reasonably close to the matched curve.

The result is illustrated in Fig. 2.12, which shows the

$$H_T = \sum_{i=2}^4 P_T^{\text{jet}_i} + E_{\perp} \quad (2.12)$$

for inclusive supersymmetric production of gluinos and squarks (including \tilde{t} and \tilde{b}).

The effect of the matching is as expected – a significant reduction in the sensitivity to parton shower parameters, and a shift of the prediction as compared to the default virtuality-ordered PYTHIA shower (black line) by about 10-30%. Even with the smearing due to the detector simulation and more complex decays than the simplified scenarios used in Sec. 2.3.1, the power P_T -ordered shower continues to overshoot the result obtained with the matching whereas the wimpy virtuality-ordered showers undershoot the matched curve.

Since we require four hard jets, the strongest impact is on the squark-(anti)squark pair production, with an efficiency increase close to 40% when passing from the unmatched default virtuality-ordered PYTHIA shower to the matched production. This happens for two reasons; as described in Sec. 2.3.1, only two hard jets are produced by the decay of the squarks, which means that two jets from QCD radiation are needed. Second the difference between the unmatched and matched radiation is particularly large for squark production. On the other hand, the effect is much smaller for gluino pair production (around 5%). This has to do first with the large fraction of events with at least one top quark in the decay, second the large presence of τ (and therefore τ -jets) from chargino and neutralino decays and finally the production of lighter right-handed squarks providing harder partons q in the decay $\tilde{g} \rightarrow q\tilde{q}_R$. This gives rise to multiple hard jets and hence a small sensitivity to the matching. More explicitly, any decay to stop (around 10%) gives at minimum one top (and therefore one b-jet) and other jets from stop decay via production of neutralinos and charginos decaying into tau-jets or b-jets. Also the decay of left-handed squarks (15%) $\tilde{u}_L, \tilde{d}_L, \tilde{s}_L, \tilde{c}_L$ provide easily neutralinos and charginos decaying onto tau's and then possibly tau-jets. Finally the

sbottom decay gives a b quark and charginos as well as neutralinos with as consequence the production of tau-jets. Note that in this case the b jets produced benefits from a larger phase space than in the case of the jet from squark decay in the case of $\tilde{u}_L, \tilde{d}_L, \tilde{s}_L, \tilde{c}_L$.

The effect on the associated gluino-squark production lies between these two extremes, with an efficiency increase close to 10%. \tilde{t} and \tilde{b} production is dominated by $\tilde{t}_1\tilde{t}_1^*$ pairs (due to a low \tilde{t}_1 mass, around 390 GeV), which decay to top quarks and the LSP. For those, the effect of the matching is negligible.

2.4 Monte Carlo systematics for b -jet detection

The simulation of the QCD radiation, beside the aspect of hardness or shower evolution variable, is also related to the type of initial parton especially the b -quarks. There is a long-standing problem related to this, whose representative case is the t channel of single top production at hadron colliders. This problem concerns the NLO predictions of the rates but also the differential distributions and is related to the choice between viewing b quarks as massless constituents of the proton or massive emerging particles.

The first option is to calculate the NLO corrections in the 5-flavor scheme. This means that the b -quark is considered also as initial state and therefore for the single-top case the calculations are based on a $2 \rightarrow 2$ process. While this method benefits from easy calculations, the modeling of the accompanying (massive) b -quark, needed at the experimental level is not accurately performed at LO.

The second possibility is to consider the four-flavor scheme, *i.e.* the massive b -quarks are not considered as initial state, and are generated by gluon splitting only. A main difficulty is due to the addition of a third particle in the final state, but the NLO prediction of the kinematics of the b -quark is expected to be better than in the five-flavor scheme.

As discussed in details in Ref. [67] and [68], the NLO calculations for single-top in the four-flavor scheme behaves very well and is substantial improvement with respect to the five-flavor scheme.

The problem is essentially the same with the presence of two “spectator” b -quarks. A typical example in the SM is the production of Z +jets. In the Higgs sector, this situation is also possible if the Yukawa coupling $b\bar{b}H$ is enhanced, for instance by $\tan\beta$ like in the two-Higgs-doublet-model presented in the first chapter. We do not propose theoretical solutions here, instead the present discussion can be seen as strong remark for analyses where the presence of these accompanying b -quarks is important.

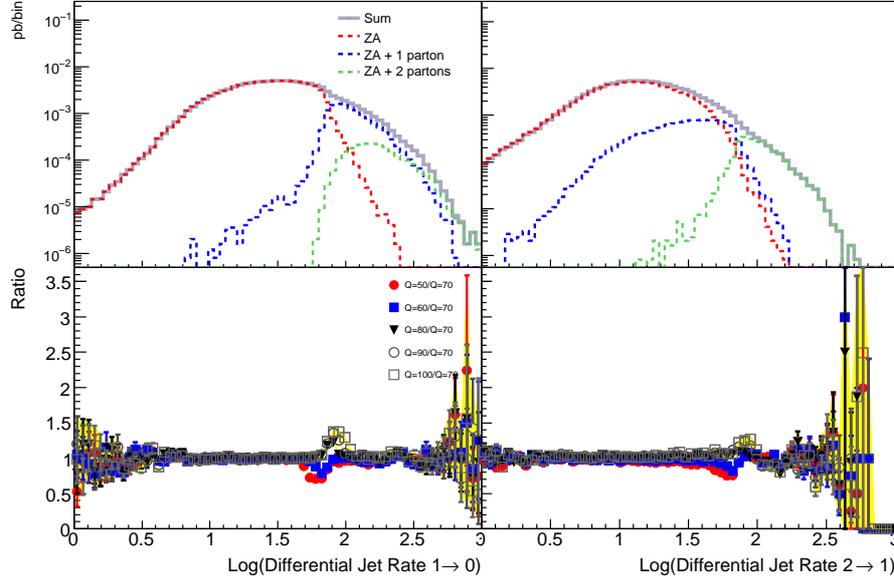


Figure 2.13: $1 \rightarrow 0$ and $2 \rightarrow 1$ differential jet rates for ZA production with 0,1 and 2 additional partons at the matrix element level. $Q_{\text{match}}=70$ GeV for the upper figures. The distributions are normalized to the cross-section (equivalent to a luminosity of 1 pb^{-1}). The lowest plots indicate the ratio of distributions for $Q_{\text{match}} = 50, 60, 80, 90, 100$ GeV with respect to the case where $Q_{\text{match}} = 70$ GeV.

Practically we consider the production of a heavy scalar h with $m_h=350$ GeV.

In the context of Monte Carlo simulations, the main point is to establish which simulation choice reflects correctly the calculations; there are essentially two options. The first possibility is the generation of an inclusive sample of the form $2 \rightarrow h+0,1,2$ partons, using a matching technique as it has been discussed in this chapter. The jet-matching is an elegant solution for this type of problem as a decreasing cross-section is associated to each multiplicity. The other possibility is to consider the $gg \rightarrow h\bar{b}\bar{b}$ process, with or without matching while the showering is performed.

To do the comparison, the matched production is done with shower- k_{\perp} , and the choice of Q_{match} is shown to be the best from around 50 GeV up to 100 GeV. This is illustrated with the case in Fig. 2.13 for the $1 \rightarrow 0$ and $2 \rightarrow 1$ jet rates. The evolution of the global (physical) shape is very small if the matching scale is changed. The only variation happens in a small region around Q_{match} and the smoothest transition is obtained for $Q_{\text{match}}=70$ GeV.

At the moment of writing this text, the implementation of the CKKW matching method in MadGraph+PYTHIA has been just done. The jet rates are shown to be in perfect agreement to what is obtained with the shower- k_{\perp} scheme. We also consider this possibility for the comparison.

A study of the P_T and rapidity distributions gives a good indication of the differences between the matched (both CKKW and shower- k_{\perp} methods) and the unmatched case $gg \rightarrow b\bar{b}h$. For the comparison we also display the results for the unmatched $2 \rightarrow 1$ channel $b\bar{b} \rightarrow h$. Figure. 2.14 (up) shows these variables for the two hardest b -quark at the shower level. No cut in rapidity or P_T is applied at this level and the curves are normalized to 1.

While the tails of the P_T distributions tend to be parallel for the matched productions and $gg \rightarrow b\bar{b}h$, the matched case shows an overshooting effect in the bulk region, mostly visible for the leading b -quark in P_T . Correlated to this, the b -quarks are produced more centrally in the $gg \rightarrow b\bar{b}h$ case than in the matched cases. This is due to the fact that in the matched case a substantial fraction of the leading b -quarks in P_T are produced by the shower algorithm and not via a matrix-element. Besides, the $2 \rightarrow 1$ solution description tends to be close to the matched description at low P_T scales while large discrepancies appear in the tails. This is of course the same observation as for the production of heavy colored particles.

As no cut is applied, these distributions are biased compared to what could be observed in a detector. The b -tagging of jets is indeed based on minimal P_T cut as well as on the acceptance of the detector. To fix the ideas, we assume the acceptance cut $P_T(\text{jet}) > 40$ GeV/c and $|y| < 2.5$. The corresponding P_T and rapidity distributions are shown in Fig. 2.14 (low). The matched and $gg \rightarrow b\bar{b}h$ shapes tends to be very similar even in the bulk region. This sanity check not only confirms what is expected, that the matched and four-flavor scheme behave similarly at high P_T , but also shows that the remaining between Monte Carlo systematic bias can be translated as a normalization factor between choices of simulation. This conclusion is however not true for $b\bar{b}h$ production as it still displays an undershoot in the tails.

This factor is estimated as the ratio

$$R = \frac{\epsilon_{\text{matched}}}{\epsilon_{\text{unmatched}}} \quad (2.13)$$

where ϵ is defined by the ratio between the number of events for which a given number of b -quark fulfill the P_T and rapidity cut and the total number of events. This ratio is shown in Fig. 2.15 (left) for events with at least one and two b -quarks passing the cuts.

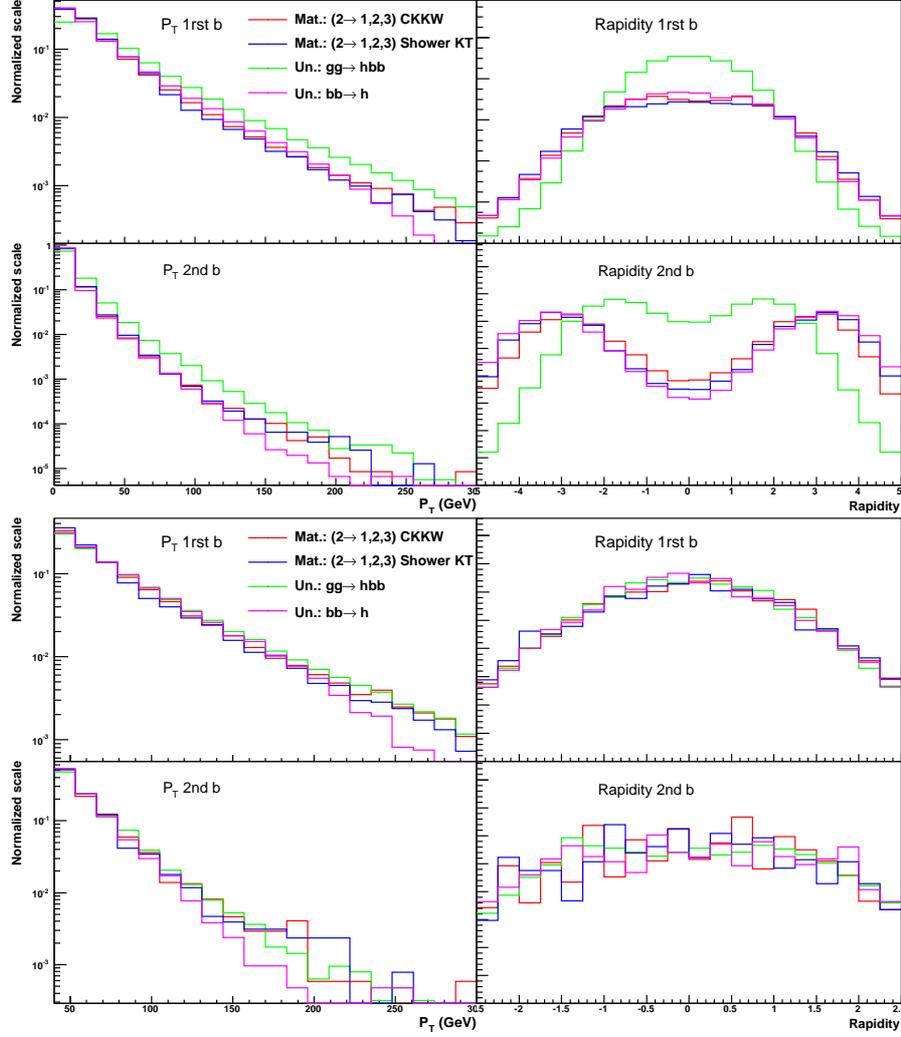


Figure 2.14: P_T and rapidity distributions for the two leading b -quark in P_T for (up) no cut in P_T nor rapidity and (down) for $P_T > 40$ GeV/c and the $|y| < 2.5$. The blue and red lines are the distributions for CKKW and Shower- k_\perp samples respectively, the green and purple lines correspond to the unmatched $gg \rightarrow hb\bar{b}$ and $b\bar{b}$ samples respectively.

In a first approximation we can conclude that the b -tagging efficiency of at least one ISR with a P_T larger than 30 GeV/c would suffer from a variation of at least 30 %

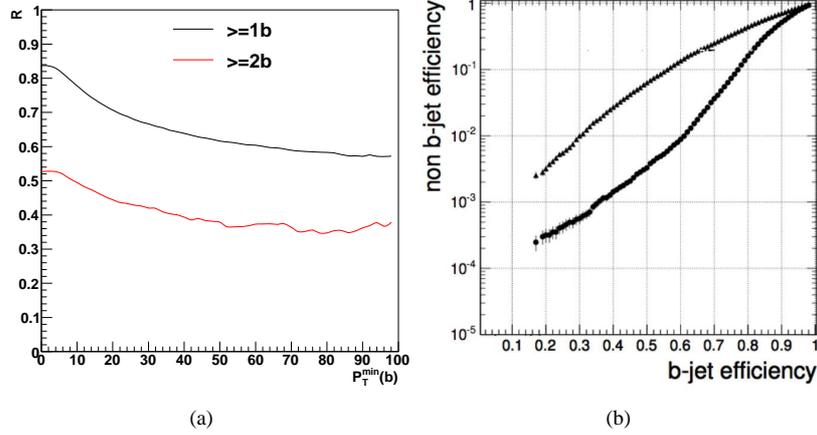


Figure 2.15: Left: The ratio of efficiencies between matched and unmatched sample, for the detection of at least one or two b quark with a minimal $P_T > P_T^{min}$ and with the pseudo-rapidity restricted to 2.5. Right: The b-tagging versus the mistagging rate for c (triangle) and uds (round) jets in CMS. $P_T > 30$ GeV/c and $|\eta| < 1.4$.

with respect to the Monte Carlo method used as reference. If two ISR b -jet are tagged, this variation is raised to more than 50 %.

However the estimation of this effect in the context of an experimental analysis is slightly more complicated since the contribution of mis-tagged light jets and the proportion between light, c and b-jets when the acceptance cuts are applied have to be taken into account. Note that this proportion is also different with respect to the Z +jets case.

It is possible to estimate the contamination of the sample by light jets as the following. We define B_f the probability to b -tag a jet initiated by a parton of flavor f and $N_{j,f}^{obs}$ the number of event produced with j parton(s) of flavor f , all passing the acceptance cut. Therefore

$$N^{tagged} = \sum_{j=1}^{\infty} [1 - (1 - B_f)^j] N_{j,f}^{obs} \quad (2.14)$$

represents the actual number of events with at least one jet b-tagged. If we sum over the flavors and divide by the actual number of events produced, the expression can be rewritten as

$$\tilde{\epsilon} = \sum_f \sum_{j=1}^{\infty} [1 - (1 - B_f)^j] \epsilon_{j,f} \quad (2.15)$$

where $\tilde{\epsilon}$ is the proportion of events with at least one jet is b -tagged and $\epsilon_{j,f}$ is the proportion of events providing exactly j jets of flavor f passing the acceptance cuts. The parameter $\epsilon_{j,f}$ is submitted to the Monte Carlo systematic uncertainty illustrated in Fig. 2.15 (left), fluctuating between 20 and 60 %.

The parameter B_f is experimental. There are different methods [69] to identify the b -jets: secondary vertex detection, track counting and soft lepton identification. The efficiencies of b -jet tagging versus the mis-tagging rates is shown on Fig. 2.15 (right).

A 40% efficiency for b -jets shows that a bit more than the percent of c -jets and the per mil of light (udsg) jets are b -tagged. This estimation shows that weight of light jets in the tagging estimation shouldn't be too large and therefore the systematic error shown in Fig. 2.15(left) can be propagated from $\epsilon_{j,b}$ to $\tilde{\epsilon}$. This ratio could then be taken into account directly in the significance prediction, using the normalization to the NNLO cross-section.

To conclude, it is important to keep an eye on this point as it might influences the estimation of the significance if such b -tagged jet are required to be visible. A mismatch between the data and the simulation could lead to wrong interpretations at the level of the physical production of the ISR b -jet.

2.5 Conclusion

This chapter focuses on the importance of simulating correctly the QCD radiation for the study of the Standard Model and beyond. The jet matching methods implemented in MadGraph are discussed and their impact on the simulation of initial state radiation is studied.

It is shown that a matched matrix element plus parton shower approach for heavy particle production is globally much more accurate than a parton shower alone. Indeed the P_T spectra of the extra-jets in the parton shower approach is extremely sensitive to the shower starting scale and the shower algorithm, while the matched simulations are not.

The discussion is also extended to a more inclusive context, where the extra-jets are produced in addition to hard jets coming from heavy particles decays. It has been demonstrated that even in this situation the initial state radiation cannot be neglected and is susceptible to influence the jet counting and the energy flux. This is a very strong argument in favor of using jet matching as a standard simulation method since the discovery of many heavy colored particles (in MSSM for instance) is expected to be achieved with final states composed by jet and missing transverse energy.

Studies in the i M2HDM at the LHC, the Higgs hunting

IN this chapter we study how the Type I and Type II i M2HDM could be discovered at the LHC. The aim is to propose analyses covering most of the parameter space left free from the theoretical and experimental constraints.

This chapter is organized as follows. First a brief reminder about the i M2HDM and its constraints (discussed in the Sec. 1.2.2 and 1.2.4) is done. We then define two benchmark points for the analyses. In order to evaluate the discovery potential at the LHC, we want the detector effects to be taken into account. To do so, a fast simulation software is presented. Several channels of interest are then discussed, and for three of them an analysis using the fast detector simulation is done.

3.1 Introduction

Let's first recall the main features of the i M2HDM scalar sector. Its Higgs boson mass spectrum is defined by the relation $m_h > m_{H^0, H^\pm} > m_A$ with $m_A < 90 \text{ GeV}/c^2$, and so that the decays $H^\pm \rightarrow W^\pm A$ and $H^0 \rightarrow Z A$ are always allowed and dominant.

Two Types of Yukawa couplings mode are possible, in the Type I the couplings of A , H^\pm and H^0 to fermions are proportional to $\tan \beta$ and in the Type II the couplings to up (down) fermions are rescaled by $\cot \beta$ ($\tan \beta$). On the contrary, the Yukawa couplings of the SM-like Higgs boson h are insensitive to $\tan \beta$.

In the Type I, the decays $h \rightarrow AA$, $H^0 H^0$, $H^+ H^-$ are kinematically allowed while only $h \rightarrow AA$ is possible in Type II due to the theoretical and indirect constraints. Indeed the unitarity condition restricts the scalar masses to be smaller than roughly $500 \text{ GeV}/c^2$ and the B -physics measurements constraints the triplet mass to be larger than $300 \text{ GeV}/c^2$. Finally the twisted custodial symmetry forbids the H^0 boson to couple to two gauge bosons.

The unconstrained region of the parameter space in Type I and Type II scenarios defines the choices of benchmark points in the mass spectrum. Figure. 3.1 (a) and (b) summarize the set of constraints presented in the first chapter. We do not cover the whole available parameter space. Instead, we restrict ourselves to two representative cases (the purple stars) one for each Type. These benchmark points will be used for the more detailed analyses of this chapter (3 processes) on the contrary of the smaller discussions (3 processes).

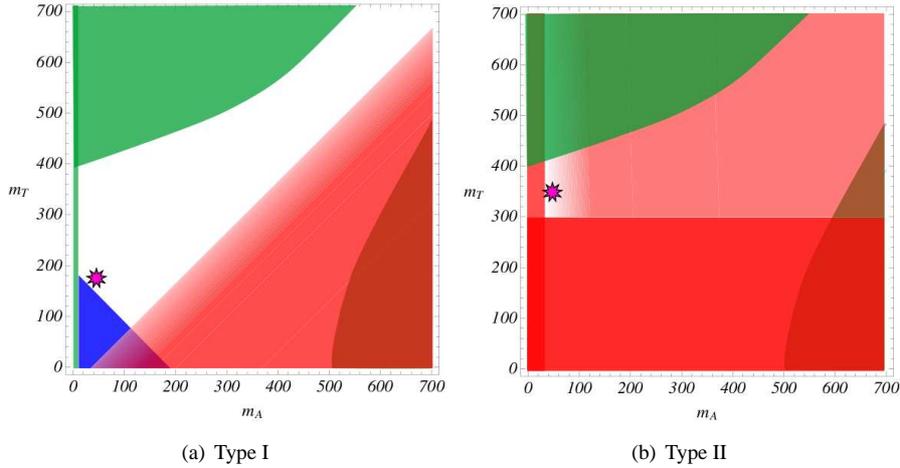


Figure 3.1: Summary of all relevant theoretical (green), indirect (red) and direct (blue) constraints in the Type I and Type II twisted two-Higgs-doublet model, in the plane $[m_A, m_T]$. Gradient bounds indicate milder indirect constraints, like constraints associated to the S and a_μ parameters, which should not be consider too strictly. The purple stars indicate the two choices of benchmark points detailed in the text.

The parameterization of both benchmark points is given in Tab.3.1, labelled “BP1” for the Type I and “BP2” for the Type II.

| Parameter | BP1 | BP2 |
|------------------------------|------------------------|------------------------|
| m_h | 400 GeV/c ² | 400 GeV/c ² |
| m_T | 180 GeV/c ² | 350 GeV/c ² |
| m_A | 30 GeV/c ² | 40 GeV/c ² |
| $\tan \beta$ | 0.2 | 30 |
| Branching ratio (%) | BP1 | BP2 |
| $A \rightarrow b\bar{b}$ | 86 | 90 |
| $A \rightarrow \tau^+\tau^-$ | 10 | 10 |
| $H^0 \rightarrow ZA$ | ~100 | 63 |
| $H^\pm \rightarrow W^\pm A$ | ~100 | 79 |
| $h \rightarrow H^+H^-$ | 20 | – |
| $h \rightarrow H^0H^0$ | 10 | – |

Table 3.1: Parameters values for the i M2HDM that define the BP1 and BP2 benchmark points. The branching ratios relevant for the analyses presented in the following are also given.

3.2 Event simulation and reconstruction

The signal and background events have been simulated using the implementation of a generic 2HDM in the tree-level matrix-element based event generator MadGraph v4.4. The parameters of the model have been calculated using the TwoHiggsCalc calculator [70]. The PDF set used is CTEQ6L [71] and the factorization (μ_F) and renormalisation (μ_R) scales are evaluated on an event-by-event basis using the relation

$$\mu_F^2 = \mu_R^2 = (M_{max}^2 + \sum_j P_T^2) \quad (3.1)$$

where M_{max} is the larger mass among the final state particles and j runs over the visible particles.

The showering/hadronization phase, as well as the decay of unstable SM particles, are simulated using Pythia 6.4 with virtuality ordered showers (MSTP(81)=0) and default starting scale (PARP(67)=4) [72]. No underlying event nor piled-up events are simulated.

In order to take into account the efficiency of event selection under realistic experimental conditions, the fast detector simulator Delphes [73] is used. Characteristics of the simulated detector, *i.e.*, its geometry, granularity and resolution, are close to those associated with the ATLAS and CMS detectors (see Fig. 3.2).

As a reminder, the pseudo-rapidity η is defined as

$$\eta = -\ln \left[\tan \frac{\theta}{2} \right] \quad (3.2)$$

where θ is the polar angle calculated from the longitudinal direction (beam axis). The second often considered coordinate is the azimuthal angle ϕ , running on the transverse plane of the detector and calculated from an arbitrary transversal x axis.

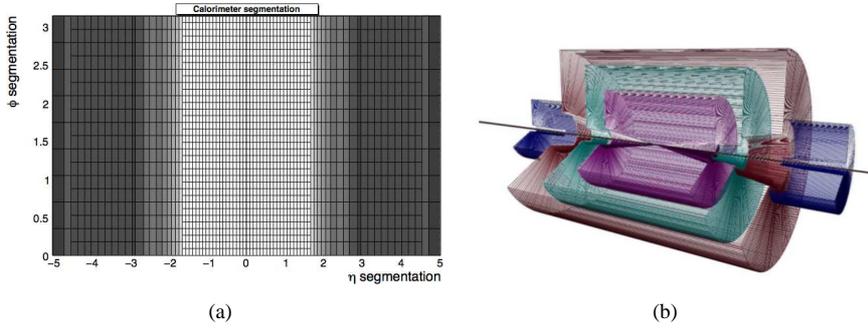


Figure 3.2: (a) The η and ϕ segmentation of the calorimetric towers in DELPHES. (b) Simulated view of the simulated detector in DELPHES. The central purple part in the barrel is the tracker, it is surrounded by the calorimeter system (green) and the muon chamber system (red). Two endcaps for the calorimeters are also present (blue).

The tracker is assumed to reconstruct tracks within $|\eta| < 2.5$ with a 100% efficiency and the calorimeters cover a pseudo-rapidity region up to $|\eta| < 3$ with an electromagnetic and hadronic tower segmentation of $\Delta\eta \sim 0.1$ and $\Delta\phi \sim 0.1$. The energy of each quasi stable particle is summed up in the corresponding calorimeter tower. The resulting energy is then smeared according to resolution functions assigned to the electromagnetic calorimeter (EC) and the hadronic calorimeter (HC) parameterized by:

$$\frac{\sigma_{EC}}{E} = 0.005 + \frac{0.25}{E} + \frac{0.05}{\sqrt{E}} \quad (3.3)$$

$$\frac{\sigma_{HC}}{E} = 0.05 + \frac{1.5}{\sqrt{E}}. \quad (3.4)$$

where the energy E is expressed in GeV. Note that the configuration card is given in the appendix A.

The acceptance criteria are summarized in the Tab. 3.2. For the lepton, we demand a tight isolation criterion by requiring that the number of additional tracks with $P_T > 1$

| Final state | $ \eta^{max} $ | P_T^{min} (GeV) |
|-------------|----------------|-------------------|
| e, μ | 2.4 | 5 |
| jets | 3 | 40 |
| b -jets | 2.5 | 40 |

Table 3.2: Acceptance of the different final states in the simulated detector.

GeV/c (denoted N_{tracks}^{cone}) present in a cone $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ centered on the lepton track be either 0 or 1. In so doing, we accommodate cases where collinear hard leptons are produced. The jets are reconstructed using only calorimeter towers and making use of the SIScone algorithm, as defined in the FastJet package [74] and implemented in Delphes. Unless stated explicitly, a cone size radius of 0.7 is applied for the jet algorithm. The b -tagging efficiency is assumed to be 40% for all b -jets, independently of their transverse momentum, with a fake rate of 1% (10%) for light (charm) jets. Finally, the total missing transverse energy \cancel{E}_T is reconstructed using information from the calorimetric towers and muon candidates only.

In order to avoid repetitions in the forthcoming analyses, we define here a set of cuts related to the acceptance of produced particles $A(l_i^n, j^m)$, the selection of n Z boson C_{nZ} , the b -tagging of a jet C_b and the selection of leptons coming from a light pseudoscalar $C_A(l_1, l_2)$:

- $A(l_i^n, j^m)$: m jets and n leptons (electrons or muons) are required in the acceptance region with the isolation corresponding to $N_{tracks}^{cone} \leq i$. In other words, we keep the freedom to redefine the track isolation by accepting i more tracks inside de isolation cone around the track from the lepton.
- C_{nZ} : n Z boson(s) are reconstructed from lepton kinematics. Lepton candidates fulfilling the acceptance cuts must have the same flavour, opposite charges, and a $P_T > 10$ GeV/c to reduce the amount of leptons from B meson semi-leptonic decays. A Z boson is then reconstructed if the di-lepton invariant mass lays in a 10 GeV/c² mass window around the Z mass.
- C_b : at least one of the jet passing the acceptance cuts is b -tagged.
- $C_A(l_1, l_2)$: The two leptons l_1 and l_2 have different flavours and opposite charges, belong to the same $\Delta R < 1.2$ cone, and have an invariant mass smaller than 25 GeV/c².

In the following, we present simple strategies that can lead to promising Signal-over-Background (S/B) ratios. For some of the processes, detailed information on the

efficiencies and the visible cross sections are given. The possibility for additional, more sophisticated, selection methods is also briefly discussed.

3.3 Discovery potential of Type I $iM2HDM$

In this section we review simple ways to discover all Higgs bosons in the Type I with a minimum of channels.

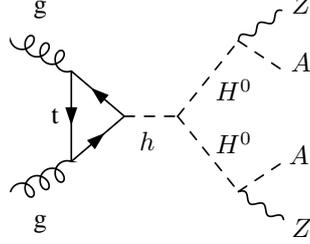
In Sec. 3.3.1 we focus on $pp \rightarrow h \rightarrow ZAZA$ process for the simultaneous discovery of h , A and H^0 bosons. The discussion in Sec. 3.3.2 is focused on the detection of the charged Higgs.

3.3.1 A single process for all neutral Higgs bosons observation

The main NLO cross-sections [18] for the Higgs boson h production at the LHC are shown in Fig. 1.5. For instance the $gg \rightarrow h$ fusion process is dominant on the whole mass range. The other production modes are $qq' \rightarrow qq'h$ with vector boson fusion, the heavy flavored $pp \rightarrow q\bar{q}h$ and the Higgstrahlung channels of the type $pp \rightarrow Vh^0$.

In the context of SM Higgs searches at the LHC, it has been shown that the discovery of h at high mass could be achieved after less than 10 fb^{-1} (see Fig. 1.6 (left)), notably with the observation of $h \rightarrow ZZ$ decays [75, 76]. A deviation from the expected visible cross section σ_{vis} (defined as $\sigma \times BR$ after the application of cut based selections) could reveal the presence of additional decays of the Higgs boson such as those predicted by the $iM2HDM$. In this context, it is interesting to study the process $gg \rightarrow h \rightarrow H^0 H^0 \rightarrow ZAZA$ since beside the large cross-section it benefits from a sizable $h \rightarrow H^0 H^0$ branching ratio (see Fig. 1.15). To study this channel, we assume that the mass-spectrum is defined with the benchmark point BP1.

A first requirement is the decay of both Z boson into leptons e and μ . This provides a very clear signature (see Fig. 3.3), helping for a large background reduction. Under this decay hypothesis, the total rate decreases considerably, such that only the dominant $A \rightarrow b\bar{b}$ decay mode can reasonably be retained. At the parton level the signal final state is then $l^+ l^- l'^+ l'^- b\bar{b} b\bar{b}$, with $l^\pm = e^\pm, \mu^\pm$. In average the b -quarks produced by the pseudoscalar A are reasonably separated ($\Delta R \gtrsim 1$) but have a low average transverse momentum. In order to avoid the experimental complication of tagging these low- P_T b -jets separately, a large cone size radius (≈ 1) is used. This therefore leads to a final state with only two main jets in the final state. It should be

Figure 3.3: Feynman diagram for the $gg \rightarrow h \rightarrow H^0 H^0 \rightarrow ZAZA$ process.

noted, however, that with a heavier A the individual detection of all four jets could be attempted.

The main backgrounds to be taken into account are $t\bar{t}Z$ and $ZZjj$. The process $gg \rightarrow h \rightarrow ZZ$ can be neglected as well as W^+W^-Z because of their relatively low cross section for the second case and low probability to provide a b -tagged jet for both cases. The background process $t\bar{t}(h \rightarrow ZZ \rightarrow l^+l^-l'^+l'^-)$ is also neglected thanks to its low cross-section (~ 5 ab). For $ZZjj$ the jets are produced with a minimal P_T of 20 GeV, a maximal pseudo-rapidity of 5 and a $\Delta R(jj) > 0.3$. The cross sections times branching ratio for the signal and background processes are given in Tab. 3.3

| Process | Decay (MC) | σ (fb) |
|-------------|--|---------------|
| $ZAZA$ | $(Z \rightarrow l^+l^-)(Z \rightarrow l'^+l'^-)b\bar{b}b\bar{b}$ | 3.2 |
| $ZZjj$ | $(Z \rightarrow l^+l^-)(Z \rightarrow l'^+l'^-)jj$ | 16 |
| $t\bar{t}Z$ | $(t \rightarrow \tilde{l}^+b\cancel{E}_T)(\bar{t} \rightarrow \tilde{l}^-b\cancel{E}_T)(Z \rightarrow \tilde{l}^+\tilde{l}^-)$ | 3.5 |

Table 3.3: List of processes considered in the analysis of the $ZAZA$ channel. The notation l means that only electron and muons are considered. If the notation \tilde{l} is used instead, all flavours are included and the taus are decayed in e or μ .

In order to increase the S/B ratio, the acceptance cut $A(l_0^4, j^2)$ and the C_b cut are applied. The efficiency of the C_b cut is assumed to be the same as for single b quark induced jets. We then apply the C_{2Z} cut, where the invariant mass of the two pairs of same-flavour leptons are the closest to the actual Z mass. The relative and total efficiencies for all processes are listed in Tab. 3.4, as well as their visible cross sections. The visible cross section around 0.3 fb and a S/B ratio close to 3 suggest

that, using the simple algorithm depicted here above, the evidence of such a signal could be reached with a total integrated luminosity smaller than 30 fb^{-1} . However, as said in the introduction no pile-up is considered here, this estimation is most probably optimistic.

| | | $ZAZA$ | $ZZjj$ | $t\bar{t}Z$ |
|------------------|------|--------|--------|-------------|
| $A(l_0^A, j^2)$ | (%) | 27 | 11 | 18 |
| C_b | (%) | 50 | 7.9 | 54 |
| C_{2Z} | (%) | 72 | 75 | 4.1 |
| ϵ_{tot} | (%) | 9.6 | 0.63 | 0.4 |
| σ_{vis} | (fb) | 0.32 | 0.1 | 0.014 |

Table 3.4: Relative efficiencies (in percent) for each cut presented in the text. Combined efficiencies and resulting visible cross sections after all cuts are also shown.

The S/B ratio could be further improved by applying a more sophisticated cut, taking advantage of the fact that invariant mass of the h can, in principle, be fully reconstructed. First, the two H^0 masses are reconstructed, each from one Z and one jet (which for the signal, has to be understood as the single jet induced by the A boson decay), such that the difference between the two possible H^0 mass combinations must be minimal and smaller than $100 \text{ GeV}/c^2$. Furthermore, the mass difference between the h candidate directly reconstructed from the sum of all 4-vectors of the four leptons and the two jets, and the mean of H^0 candidate masses, must be smaller than $400 \text{ GeV}/c^2$. The relevance of this cut is illustrated in Fig. 3.4(left) showing the distribution of signal and background events as a function of the difference between m_h and the mean of m_{H^0} , after applying all other cuts of Tab. 3.4. The S/B ratio could be easily increased up to 5, to the price of a lower signal visible cross section (by roughly 30%).

The distribution of the two H^0 reconstructed masses for the events passing all cuts is also shown in Fig. 3.4 (right). It illustrates the possibility to measure the H^0 mass with the simple algorithm described above. The resolution could certainly be improved as well as the signal significance if a proper jet reconstruction with an optimal cone size and the tracker information were taken into account. The determination of the invariant mass of the pseudoscalar A could also be attempted.

Note however that the conclusions of this section has to be interpreted with some caution since b -tagging efficiency and jet kinematics have to be re-evaluated when

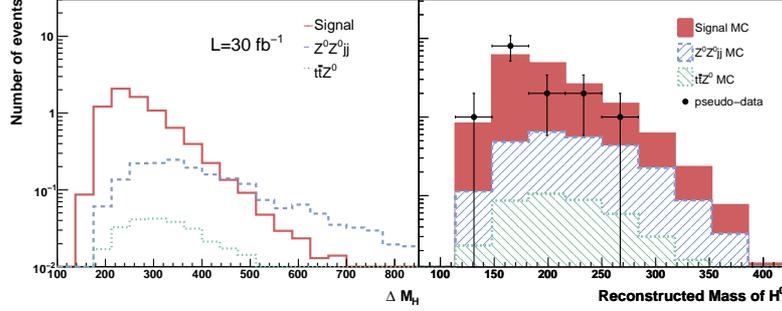


Figure 3.4: Left: the variable ΔM_H is calculated for each event as the difference between m_h and the mean of m_{H^0} after $A(l_0^A, j^2)$, C_b and C_{2Z} cuts. Right: invariant masses of the pseudoscalar H^0 bosons (two entries per event). The markers show a hypothetical event excess for an integrated luminosity of 30 fb^{-1} .

two soft and/or collinear b -induced jets are merged into one jet. This question can only be precisely addressed with a proper full simulation of detector effects.

3.3.2 Possible processes for charged Higgs bosons observation

The possibilities for discovering a charged Higgs bosons is particularly important to reveal the Higgs doublet structure of the proposed model. In the Type I $iM2HDM$ there are essentially three processes of interest named according to the presence of heavy particles:

- ($WAZA$): $pp \rightarrow W^{*\pm} \rightarrow H^\pm H^0 \rightarrow W^\pm AZA$
- ($WAWA$): $pp \rightarrow h \rightarrow H^+ H^- \rightarrow W^+ AW^- A$
- ($TWAb$): $pp \rightarrow t\bar{t} \rightarrow tH^- \bar{b} \rightarrow tW^- A\bar{b}$.

This section does not consider the benchmark point BP1, instead we focus on the regions of the parameter space where the visibility of the processes is expected to be the highest, *i.e.* with lighter masses to enhance the cross-sections.

$WAZA$, a process for H^\pm and H^0 discovery

In the Type I, the discovery of the charged Higgs bosons could be attempted via the channel $pp \rightarrow W^* \rightarrow WAZA$ (see Fig. 3.5).

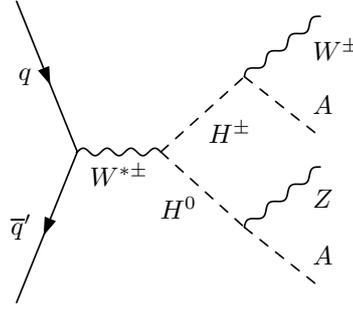


Figure 3.5: Feynman diagram for the $pp \rightarrow W^{*\pm} \rightarrow H^\pm H^0 \rightarrow W^\pm AZA$ process.

The production cross-section is close to 120 fb at best, *i.e.* for $m_{H^0} = m_{H^\pm} \sim 140$ GeV/ c^2 and $m_A \sim 50$ GeV/ c^2 . As for the $ZAZA$ channel, the interesting point of this topology is to provide a Z boson and therefore a clear signature. Assuming that both the Z and W bosons decay leptonically and the pseudo-scalars into b quarks, the final state is then $l^+l^-l'^\pm + 4 \text{ jets} + \cancel{E}_T$ and the corresponding cross-section is close to 1.5 fb. With such a mass configuration, the boost of the A boson is sufficiently high to produce the b quarks collinearly ($\langle \Delta R \rangle \lesssim 1$), this means that using a large cone size, the experimental final state is defined by two large jets.

This allows to require at least one b -tagged jet to reduce strongly the largest WZ +jets background with both vector bosons decaying into leptons. Indeed the main backgrounds are WZ with one mistagged jet, $WZb\bar{b}$ and $t\bar{t}Z$ with one top decaying leptonically and the other top hadronically. If we assume that the b -tagging requires at least 30 GeV/ c in P_T , a b -jet tagging rate of 40% and a mistagging rate around the percent, the visible cross-section of WZ is close to 2 fb, $WZb\bar{b}$ to 0.5 fb and $t\bar{t}Z$ to 5 fb. The visible cross-section of the signal is estimated to be around 0.5 fb. The improvement of S/B ratio can be then improved by using specific selection cuts, notably by the reconstruction of the Higgs bosons masses, which is possible as there is only one main source of missing transverse energy. Indeed the missing information to fully constraint the event is the longitudinal momentum of the neutrino, which can be

found using the relation

$$q_z = \frac{P_{z,\text{lep}}U \pm P_{\text{lep}}\sqrt{U^2 - P_{T,\text{lep}}^2 \cancel{E}_T^2}}{P_{T,\text{lep}}^2} \quad (3.5)$$

where

$$U = \frac{m_W^2}{2} + \vec{E}_T \cdot \vec{P}_{T,\text{lep}} \quad (3.6)$$

Depending on the kinematic arrangement of the neutrino and the lepton, this system gives zero, one or two solution. In the latter case, simulation shows that the smallest root is the best choice. It then leads to the determination of both H^0 and H^\pm masses simultaneously. Assuming a good reconstruction of the jets, the A boson mass can also be estimated.

$WAWA$, a process for H^\pm and h discovery

At the LHC, the production cross-section of $gg \rightarrow h \rightarrow H^+H^- \rightarrow W^+AW^-A$ process reaches at best 2 pb, with a H^\pm mass close to 140 GeV/c² as $m_A + m_W < m_T$ and $m_T + m_A \gtrsim 200$ GeV/c².

The topology resulting from the decay of A into a pair of b -quarks as to be treated in the same way as in the $ZAZA$ boson analysis. However in this case the main background is $t\bar{t}$ +jets for a di-leptonic final state. The S/B ratio is close to 10^{-3} , this case can be then considered as hopeless.

The other possibility is to have one of the pseudo-scalar decaying into τ 's, themselves providing $e^\pm\mu^\mp + \cancel{E}_T$. The cross-section reaches at best around 1 fb for the final state $l^\pm l'^\pm e^\pm\mu^\mp b\bar{b} + \cancel{E}_T$ where the pair $e^\pm\mu^\mp$ is collinear (l and l' refer to the leptons from the W bosons). The reason why the leptonic decay of the two τ lepton from A boson is restricted to one electron and one muon is motivated by the rejection of backgrounds with $Z/\gamma^* \rightarrow ll'$. Indeed the probability for these backgrounds to produce $e^\pm\mu^\mp$ compared to any lepton pair containing e or μ is

$$\frac{BR(Z/\gamma^* \rightarrow \tau\tau) \times 2 \times BR(\tau \rightarrow e\nu\nu) \times BR(\tau \rightarrow \mu\nu\nu)}{BR(Z/\gamma^* \rightarrow ee \vee \mu\mu) + BR(Z/\gamma^* \rightarrow \tau\tau) \times BR^2(\tau \rightarrow l\nu\nu)} \sim 3\%, \quad (3.7)$$

while this choice affects by a factor 1/2 the signal.

The main backgrounds are

- ZZ -like: $(Z/\gamma^* \rightarrow \tau\tau \rightarrow \text{lept.})(Z/\gamma^* \rightarrow \tau\tau \rightarrow e^\pm\mu^\mp + \cancel{E}_T)$

- WWZ : $WW(Z/\gamma^* \rightarrow \tau\tau \rightarrow e^\pm \mu^\mp + \cancel{E}_T)$
- $t\bar{t}Z$: $t\bar{t}(Z/\gamma^* \rightarrow \tau\tau \rightarrow e^\pm \mu^\mp + \cancel{E}_T)$
- tWZ : $tW(Z/\gamma^* \rightarrow \tau\tau \rightarrow e^\pm \mu^\mp + \cancel{E}_T)$

if a Z veto is required. The contribution of the two first channels can be strongly reduced with a b -tagging, leaving $t\bar{t}\tau\tau$ and $tW\tau\tau$, having a respective cross-section times branching ratio of 13 fb and 0.1 fb.

At this stage the S/B ratio is close to 0.1. The collinearity of the electron and the muon should help to increase this ratio as no large boost of the two τ 's in $t\bar{t}Z$ is expected. This channel seems to have a good potential for the charged Higgs discovery and the measurement of its mass. Therefore it clearly deserves a more detailed analysis.

TWAb

In the Type I the production of charged Higgs produced by top quark decay could be interesting since the process $pp \rightarrow t\bar{t} \rightarrow tH^\pm b \rightarrow tW^\pm Ab$ benefits from the large cross-section of $t\bar{t}$. However, this concerns only a small region of the parameter space, first because of the kinematical constraints (for instance BP1 cannot be considered), and second due to the low branching ratio $BR(t \rightarrow H^\pm b)$ as the mass of the Higgs boson approaches the top quark mass.

Let's consider first the decay of the A boson into a pair of b -quark, the final state at parton level is then $W^\pm W^\pm b\bar{b}b\bar{b}$. The observation of such final state seems particularly challenging. Indeed for a high m_{H^\pm} the available phase space left for the b -quark is small while if m_{H^\pm} is low, the b quarks produced by the A boson are softer. The P_T of the four b -quarks in $pp \rightarrow (t \rightarrow W^+ b)(\bar{t} \rightarrow (H^- \rightarrow W^-(A \rightarrow b\bar{b}))\bar{b})$ are shown in Fig. 3.6 for two masses of the charged Higgs, 120 and 160 GeV/c².

Assuming a minimum P_T of 30 or 40 GeV/c for jet reconstruction, all jets from decay cannot be seen. A maximum of two jets can be required and assuming a di-leptonic final state, the background level is very high, composed notably by $t\bar{t}$, WZ +jets with $Z \rightarrow b\bar{b}$. Even with a cross-section for $WWb\bar{b}b\bar{b}$ reaching at best 30 pb, this choice of channel seems not optimal and shouldn't probably be considered in priority.

The alternative solution is to use the same argument as in the $WAWA$ case, with the decay of the A boson into a pair of tau leptons, themselves providing leptons of different flavors.

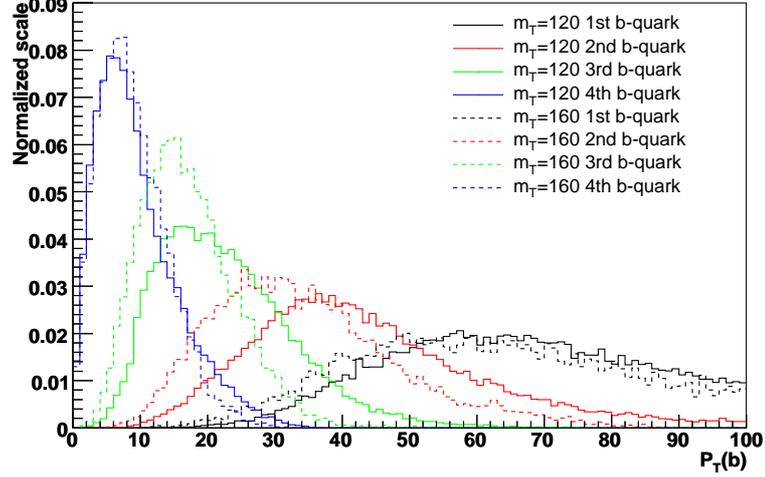


Figure 3.6: The P_T of the four b -quarks produced by $t\bar{t} \rightarrow tH^\pm\bar{b}$. The solid and dotted lines indicate respectively the distributions for $m_{H^\pm}=120$ GeV/ c^2 and $m_{H^\pm}=160$ GeV/ c^2 .

If the two W bosons provide one lepton and two jets, the background level falls strongly due the particular final states (3 leptons, two jets and missing transverse energy). The cross-section of production corresponding to the final state $l^\pm e^\pm \mu^\mp jj\bar{b} + \cancel{E}_T$ with $e^\pm \mu^\mp$ from the A boson is close to 15 fb in the favored regions of the parameter space ($m_{H^\pm}=140$ GeV/ c^2 and $\tan\beta = 0.2$). A strategy for background reduction is to require at least a b -tagged jet with $P_T > 30$ GeV/ c and a veto on Z presence from lepton pairing. The main SM processes delivering such topology are

- $t\bar{t}W$
- $t\bar{t}Z$ with $Z \rightarrow \tau\tau$
- WZ with the Z decaying into τ 's.

The corresponding LO cross-sections are 9 fb, 0.5 fb and < 1 fb respectively. The S/B ratio is close to 1 and could certainly be increased in a more detailed analysis, notably by exploiting the collinearity of the electron-muon pair.

Besides the $t\bar{t}$ channel, the other way to produce the charged Higgs is via $pp \rightarrow tH^\pm$. The cross-section is lower but does not suffer from the decay of the top into a charged

Higgs and also is much less limited by the kinematics. The visibility of this process is evaluated in BP2 (Type II) in Sec. 3.4.3

3.4 Discovery potential of Type II $iM2HDM$

In this section we review rather simple ways to discover the Higgs bosons at the LHC in the Type II.

We start by a short introduction, and then in Sec. 3.4.2 we focus on $pp \rightarrow H^0 \rightarrow ZA$ process for the simultaneous discovery of A and H^0 bosons. The main method to detect the presence of the charged Higgs via the process $pp \rightarrow tH^\pm \rightarrow tW^\pm A$ is reviewed in Sec. 3.4.3 reviews .

3.4.1 From SM-like h observation

In Type II the indirect constraints imply a minimal m_T value of $300 \text{ GeV}/c^2$. A decay of the h into the triplet cannot be considered because of the unitarity constraint essentially. We consider therefore that the h Higgs boson decays are limited to the modes of Standard Model (essentially into top quarks and vector bosons) and $h \rightarrow AA$. The simultaneous evaluation of the h mass as well as the cross-section of $h \rightarrow WW, ZZ$ (and the comparison with the theoretical predictions within the SM) constraints the relation $2m_T^2 - 2M_A$ as the $h \rightarrow AA$ suffers from the "vanishing coupling" effect discussed in the first chapter (see Fig. 3.7).

The estimation of the branching ratios of the Standard Model Higgs boson is then important, but the investigation of channels like $tH^\pm \rightarrow tW^\pm A$ and above all $pp \rightarrow H^0 \rightarrow ZA$ are needed to ensure that the degeneracy between the heavy Higgs boson is present and that the mass relation between the A boson and the triplet does not escape the model prediction with respect to the h boson mass. In the next sections we focus exclusively on the discovery potential of those two channels: the associated production of top and Higgs $pp \rightarrow tH^-$ is probably the key¹ to confirm the existence of the charged Higgs bosons whereas the ZA channel allows a very fast discovery of two neutral Higgs bosons.

¹This notation used for the tH channel denotes both signs if nothing is precised.

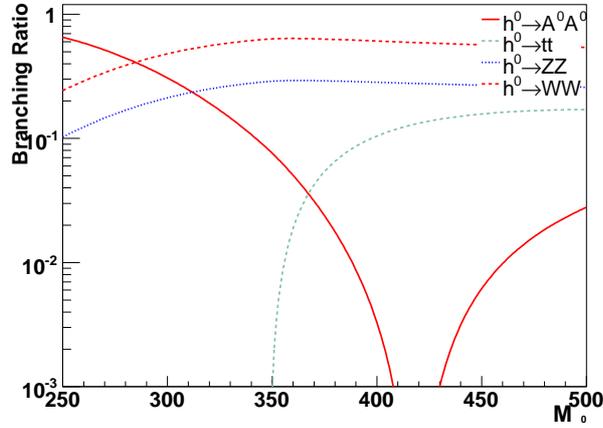


Figure 3.7: Branching ratio of the Standard Model h for a triplet mass of $300 \text{ GeV}/c^2$ and a pseudo-scalar mass of $50 \text{ GeV}/c^2$.

3.4.2 A process for H^0 and A bosons observation

In a Type II 2HDM (e.g., the MSSM scalar sector), the cross section of $b\bar{b} \rightarrow H^0$ is enhanced as $\tan\beta$ increases. This process has been shown to offer a promising discovery channel at the LHC when the Higgs boson decays into a $\tau^+\tau^-$ pair (e.g., see Ref. [77] and references therein). In the BP2 benchmark point of $iM2HDM$ the cross-section of $b\bar{b} \rightarrow H^0$ is about 15 pb , after normalization to the NNLO value (with a theoretical uncertainty of 5 %) [78]. For the mass spectrum defined by the benchmark point BP2 of the $iM2HDM$, a particularly interesting decay mode is $H^0 \rightarrow ZA$ (see Fig. 3.8).

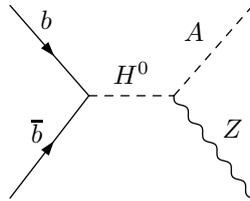


Figure 3.8: Feynman diagram for the $b\bar{b} \rightarrow H^0 \rightarrow ZA$ process.

In order to reduce the background level, the Z boson is assumed to decay leptonically. The decay where the A boson decays into a pair of τ , which in turn decay into leptons is also considered. Following the argumentation of Sec. 3.3.2 (WAWA), we restrict the A boson decay to two leptons of different flavors, one e and one μ . The selected final state is thus $l^+l^-e^\pm\mu^\mp + \cancel{E}_T$ among which l^+ and l^- reconstruct a Z boson mass and $e^\pm\mu^\mp$ are close to each other in ΔR due to the important A boson boost.

Such a multi-lepton final state is extremely clean and does not suffer from jet reconstruction uncertainties. The relevant backgrounds are $Z(Z/\gamma^* \rightarrow \tau\tau)$, $t\bar{t}Z$ and W^+W^-Z with the decay of W and Z bosons into e, μ, τ and $\tau \rightarrow e, \mu + \cancel{E}_T$. The cross sections for the signal and background processes after all decays are given in Tab. 3.5. The selection proceeds as follows. The acceptance cut $A(l_1^4)$ is applied,

| Process | Decay (MC) | $\sigma \times BR(\text{fb})$ |
|---------------|--|-------------------------------|
| $Z A$ | $(Z \rightarrow l^+l^-)(A \rightarrow \tau\tau \rightarrow e^\pm\mu^\mp \cancel{E}_T)$ | 4.2 |
| $Z(Z/\gamma)$ | $(Z \rightarrow \tilde{l}^+\tilde{l}^-)(Z/\gamma^* \rightarrow \tau\tau \rightarrow l^+l^- (\cancel{E}_T))$ | 10 |
| $t\bar{t}Z$ | $(t \rightarrow \tilde{l}^+b \cancel{E}_T)(\bar{t} \rightarrow \tilde{l}^-\bar{b} \cancel{E}_T)(Z \rightarrow \tilde{l}^+\tilde{l}^-)$ | 3.5 |
| W^+W^-Z | $(W^+ \rightarrow \tilde{l}^+ \cancel{E}_T)(W^- \rightarrow \tilde{l}^- \cancel{E}_T)(Z \rightarrow \tilde{l}^+\tilde{l}^-)$ | 0.4 |

Table 3.5: Cross sections of signal $Z A$ and background processes taking into account the leptonic final state considered in the analysis. The notation l includes only e and μ , whereas \tilde{l} also contains τ decaying into e or μ . All cross sections correspond to the final states in the second column.

followed by the C_{1Z} cut. The same flavour opposite sign leptons paired whose mass is the closest to the Z mass is retained. The two remaining leptons $l_A(1)$ and $l_A(2)$ are then assumed to come from the light and boosted pseudoscalar, and forced to satisfy $C_A(l_A(1), l_A(2))$. The relative and total efficiencies, as well as the visible cross sections for all processes listed in Tab. 3.5 are reported in Tab. 3.6.

As can be clearly seen, this channel is very promising: the S/B is high enough so that an excess over the SM could be identified after a few inverse femtobarns of integrated luminosity. Note that the $(Z \rightarrow l^+l^-) + \text{jets}$ background has also been considered due to its very large cross section ($\mathcal{O}(\text{nb})$), the possibility for jets to produce fake electrons, and the possible presence of leptons from heavy meson decays. An inclusive sample of 10^6 events was generated using the matching procedure [55] and no event has passed the isolation cuts. This background is therefore neglected. However a more detailed study should be performed with a more realistic detector simulation and event reconstruction.

| | | ZA | $Z(Z/\gamma^*)$ | $t\bar{t}Z$ |
|------------------|------|------|-----------------|-------------|
| $A(l_1^A)$ | (%) | 51 | 18 | 42 |
| C_Z | (%) | 74 | 63 | 60 |
| C_A | (%) | 85 | 3.6 | 3.3 |
| ϵ_{tot} | (%) | 32 | 0.39 | 0.84 |
| σ_{vis} | (fb) | 1.4 | 0.039 | 0.029 |

Table 3.6: Relative efficiencies of the considered cuts, together with the total efficiencies after all cuts and corresponding visible cross sections for signal and background processes. The WWZ process is omitted since its visible cross section is four orders of magnitude smaller than that of the signal.

Besides a pure counting experiment a more exclusive study can also be attempted. The mass of the two neutral resonances could be measured with an accuracy depending mostly on the \cancel{E}_T reconstruction quality. In the signal, the main source of missing transverse energy originates from the τ 's. If the direction of the \cancel{E}_T is required to lay between the transverse position of the two leptons $l_A(1)$ and $l_A(2)$ and the condition $\cancel{E}_T > 50$ GeV/c imposed, then a proper reconstruction of the invariant mass m_A can be achieved (see Fig. 3.9): assuming that each lepton is emitted collinearly to two neutrinos summed as ν_i , we can define the fraction x_i as

$$x_i = \frac{P^{l_i}}{P^{l_i} + P^{\nu_i}} \quad (3.8)$$

where P^X is the momentum of either the visible leptons or the neutrinos. If we define $m_{e\mu}$ as the invariant mass from the two leptons $l_A(1, 2)$, We can therefore rewrite the invariant mass of the A boson as

$$m_A = m_{e\mu} / \sqrt{x_1 x_2} \quad (3.9)$$

The value of x_i can be found from the information contained in the transverse plan: the projection of the missing transverse energy on the direction of each lepton and the momentum of the leptons are sufficient to solve the problem. This calculation holds only if the two τ 's are back-to-back and if the missing transverse energy stands between the transverse direction of both leptons. For the latter case we require then that

$$\frac{\Delta\phi(l_A(1), l_A(2))}{\sum_i \Delta\phi(l_A(i), \cancel{E}_T)} = 1.$$

A substantial improvement of the S/B ratio is also gained. Finally, the H^0 mass can be estimated from the A and Z boson 4-vectors.

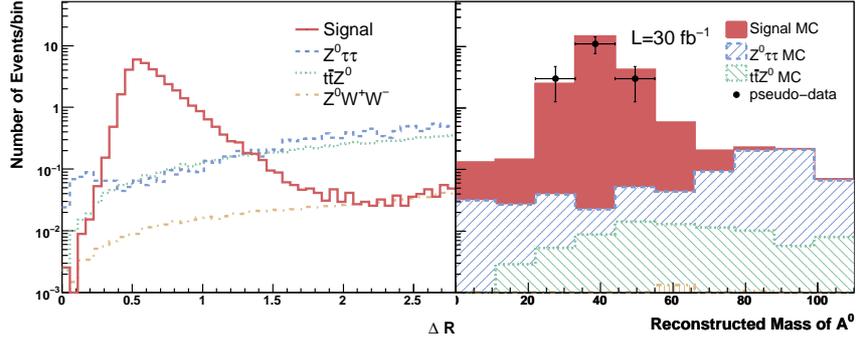


Figure 3.9: Left: the number of events in function of the distance ΔR between the two leptons not assigned to the Z , after applying the $A(l_1^3)$ and C_Z cuts. Right: the number of events in function of the reconstructed mass of A using the leptons 4-vectors and the missing E_T after applying the additional C_A cut (which require $\Delta R < 1.2$). Both figures correspond to an integrated luminosity of 30 fb^{-1} . The markers show a hypothetical event excess for this luminosity.

Beside the fully-leptonic final state, the case where the A boson decays into a $b\bar{b}$ pair could be also considered. It provides a larger cross-section, $\sigma_{\text{NNLO}} \in [250, 60] \text{ fb}$ for H^0 masses between 300 and 400 GeV/c^2 [78]. It is obvious that a certain number of b -tagged jets have to be requested in order to escape to too large backgrounds.

In case of a mass spectrum with a very light A boson, the two b -quarks are very collinear. Using a large cone size, this lead to a topology with one jet, where this jet should be b -tagged. Even in this case the background remains very large: Z + light jets with one mis-tag, $Zb\bar{b}$, ZZ with τ -jets, light jets and b jets from Z decay, and the same for WZ . The case where only one jet is visible seems therefore extremely challenging. And the probability to disentangle the two jet is very low due to their collinearity. This is shown in Fig. 3.10 where the events are generated at the matrix-element level only, requiring a minimal P_T cut of 30 GeV/c for each b -quark, $|\eta|_b < 2.5$ for being in tracker acceptance and $\Delta R_{bb} > 0.5$.

For higher A boson masses ($m_A > 40 \text{ GeV}/c^2$), the efficiency to detect the two b -jets is better, reaching a approximative plateau around 40-50 %. In such condition, the requirement of two b -tagged jets is reasonable. The dominant background processes are

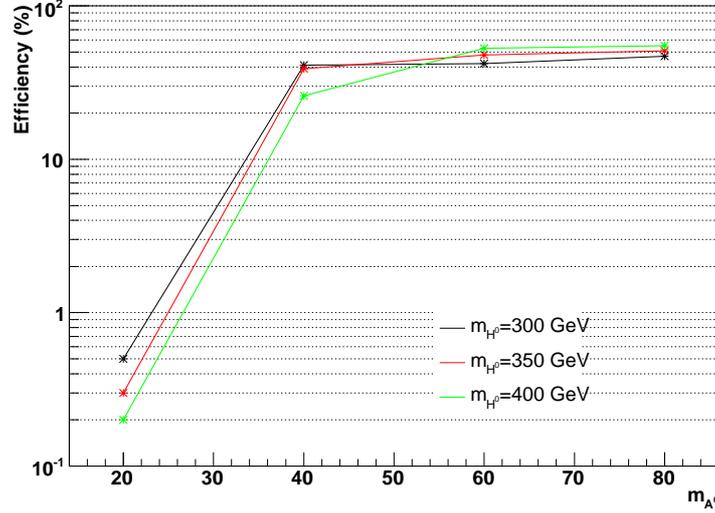


Figure 3.10: The efficiency for having two jets with $P_T > 30$ GeV, $|\eta| < 2.5$ and $\Delta R(jj) > 0.5$, for the $Z A$ signal with $m_A = 20, 40, 60, 80$ GeV/c² and $m_{H^0} = 300, 350$ and 400 GeV/c².

- Z + light jets with two mis-tagged jets,
- Z + b-jets
- $Z Z$ with a Z decaying into b-quarks

Assuming a b mis-tagging rate close to the percent, a minimal P_T of 30 GeV/c for each jet, the LO cross-section of ($Z \rightarrow ee$ or $\mu\mu$) plus at least two light mis-tagged jets falls around 10 fb whereas for ($Z \rightarrow ee$ or $\mu\mu$)+ bb , both tagged, the LO cross-section is close to 300 fb. The $Z Z$ background is the most problematic for benchmarks where $|m_Z - m_A|$ is small since the paired jets invariant mass are similar. The main discriminant variables are the boost of the Z (s), the total invariant mass (m_{H^0}) and as explained in chapter 4, the $\Delta\phi$ (coplanarity) between the 4-vectors $l_1 + l_2$ and $j_1 + j_2$. This option seems interesting but requires a very good understanding of the topologies with jets, and definitely deserves a more detailed study.

3.4.3 A process for H^\pm observation

$$g(b/\bar{b}) \rightarrow (t/\bar{t})H^\pm \rightarrow W^-(b/\bar{b})W^+A$$

In order to fully determine the structure of an extended scalar sector, it is crucial to observe a charged Higgs boson. In the $iM2HDM$, we expect it to be nearly degenerate in mass with H^0 as a consequence of the twisted custodial symmetry.

The associated production of a charged Higgs with a top quark, $g(b/\bar{b}) \rightarrow (t/\bar{t})H^\pm$ (see Fig. 3.11 (a)), is in general considered as a challenging channel at the LHC. The discovery potential strongly depends on $\tan\beta$, the mass of the charged Higgs boson and the considered decay mode. However, as distinct from models such as the MSSM, the $iM2HDM$ offers the possibility for the H^\pm to decay into $W^\pm A$. Its observation would therefore be a very strong evidence that the scalar sector originates from the $iM2HDM$. We consider this possibility in the benchmark point BP2. To normalize the expected signal, we use the NLO prediction for the charged Higgs production cross section from Ref. [79], *i.e.*, 465 fb.

As in the previous analysis, we focus exclusively on the decay $A \rightarrow \tau^+\tau^-$ where the $\tau^+\tau^-$ pair decays into $e^\pm\mu^\mp$. Despite the fact that the total signal cross section is reduced by almost two orders of magnitude compared to the $A \rightarrow b\bar{b}$ case, a strong reduction of the background is foreseen if one of the W bosons decays leptonically. The considered final state is therefore $l^\pm jjbe^\pm\mu^\mp + \cancel{E}_T$. If the light quark pair comes from the W boson produced in the charged Higgs decay, the resulting jets tend to be collinear due to the large boost. As a consequence, they might not be resolved but merged into a single “large” jet (noted J) by the reconstruction algorithms. We include both possibilities.

The relevant backgrounds are $t\bar{t}(Z/\gamma^*)$, $W(Z/\gamma^*)jj$, $Z(Z/\gamma^*)jj$, $tW(Z/\gamma^*)$, and $W^+W^-W^\pm$, with j standing for all light and b quarks (see Fig. 3.11 ((b)–(e))). The cross sections for the signal and the considered background processes, as well as the corresponding final states, are summarized in Tab. 3.7. For $W^\pm/Z + \tau\tau + jj$ and $W^\pm W^+W^-jj$, the jets are initially produced with a minimal P_T of 10 GeV, a maximal pseudo-rapidity of 5, and an angular separation of $\Delta R(jj) > 0.1$ for the firsts and $\Delta R(jj) > 0.2$ for the latter. The details of the decay modes and the corresponding rates are shown in Tab. 3.7. In order to increase the S/B ratio, the acceptance cut $A(l_1^3, j^2)$ and the C_b cut are applied. These are followed by a veto on the presence of a Z boson \overline{C}_Z . The two closest leptons (l_1, l_2) with opposite charges and different flavours are assumed to come from the light and boosted pseudoscalar Higgs boson A , and therefore are required to satisfy the $C_A(l_1, l_2)$ cut. The relevance

| Process | Decay (MC) | $\sigma \times BR$ (fb) |
|----------------------------|---|-------------------------|
| $(\bar{l}/t)H^\pm$ | $((t/\bar{l}) \rightarrow \tilde{l}^\mp b \cancel{E}_T)(H^\pm \rightarrow (W^\pm \rightarrow jj)(A \rightarrow \tau\tau \rightarrow e^\pm \mu^\mp \cancel{E}_T))(b)$ $((t/\bar{l}) \rightarrow jjb)(H^\pm \rightarrow (W^\pm \rightarrow \tilde{l}^\pm \cancel{E}_T)(A \rightarrow \tau\tau \rightarrow e^\pm \mu^\mp \cancel{E}_T))(b)$ | 0.75 |
| $t\bar{t}(Z/\gamma^*)$ | $(t \rightarrow incl.)(\bar{t} \rightarrow incl.)(Z/\gamma^* \rightarrow \tilde{l}^+ \tilde{l}'^- \cancel{E}_T)$ | 4.5 |
| $W(Z/\gamma^*)jj$ | $(W^\pm \rightarrow \tilde{l}^\pm \cancel{E}_T)(Z/\gamma^* \rightarrow \tilde{l}'^+ \tilde{l}''^- \cancel{E}_T)jj$ | 48 |
| $Z(Z/\gamma)jj$ | $(Z \rightarrow \bar{l}^+ \bar{l}^-)(Z/\gamma^* \rightarrow \tilde{l}^+ \tilde{l}'^- \cancel{E}_T)jj$ | 10 |
| $(t/\bar{l})W(Z/\gamma^*)$ | $(t \rightarrow incl.)(l^\pm \rightarrow \tilde{l}'^+ \tilde{l}''^- \cancel{E}_T)$ | 0.6 |
| $W^+W^-W^\pm jj$ | $(W^\pm \rightarrow \tilde{l}^\pm \cancel{E}_T)(W^+ \rightarrow \tilde{l}^+ \cancel{E}_T)(W^- \rightarrow \tilde{l}^- \cancel{E}_T)jj$ | 13 |

Table 3.7: Cross sections of the signal $gb \rightarrow tH^\pm \rightarrow W^-bW^+A$ and the relevant background processes, taking into account leptonic and jet final states considered in the analysis. The notation \tilde{l} means that the three flavour of leptons are taken into account, and the tau leptons decay into e or μ . On the contrary, \bar{l} means that the tau leptons decay inclusively ($Z/(Z/\gamma^*)$ case). All the quoted cross sections correspond to the final states in the second column.

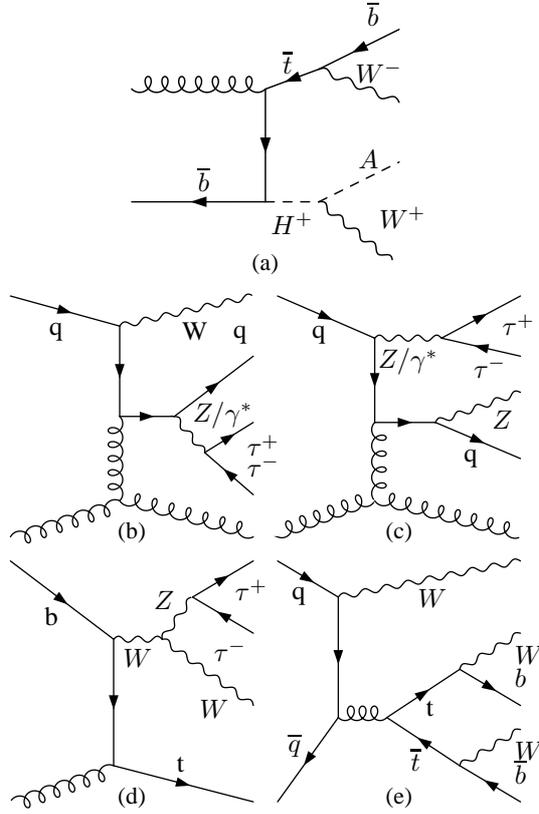


Figure 3.11: Representative Feynman diagrams for (a) the $g\bar{b} \rightarrow \bar{t}H^+ \rightarrow W^+W^- \bar{b}A$ process, and for the background processes (b) $W/Z/\gamma^*jj$, (c) $Z(Z/\gamma^*)jj$, (d) $tW(Z/\gamma^*)$ and (e) $WWWWjj$.

of this last cut is illustrated in Fig. 3.12, where the left-hand side plot shows the dilepton invariant mass after applying all cuts except $C_A(l_1, l_2)$. The relative and total efficiencies, as well as the visible cross sections for all processes listed in Tab. 3.7, are reported in Tab. 3.8.

A rather low visible signal cross section confirms that this channel is also very challenging with the unusual $H^\pm \rightarrow W^\pm A$ decays. However, the S/B ratio of order $\mathcal{O}(1)$ leaves some hope that a charged Higgs could still be discovered after a large integrated luminosity ($\sim 300 \text{ fb}^{-1}$). In any case it should be kept in mind that the benchmark

| | | tH^\pm | $t\bar{t}(Z/\gamma^*)$ | $W(Z/\gamma^*)jj$ | $Z(Z/\gamma^*)jj$ | $tW(Z/\gamma^*)$ | $W^\pm W^+ W^- jj$ |
|------------------|------|----------|------------------------|-------------------|-------------------|------------------|--------------------|
| $A(l_1^3, j^2)$ | (%) | 35 | 16 | 5.5 | 3.6 | 14 | 21 |
| C_b | (%) | 39 | 48 | 6 | 6.3 | 39 | 49 |
| \overline{C}_Z | (%) | 98 | 98 | 91 | 60 | 98 | 95 |
| C_A | (%) | 84 | 15 | 19 | 12 | 11 | 6.5 |
| ϵ_{tot} | (%) | 11 | 1.1 | 0.056 | 0.017 | 0.61 | 0.64 |
| σ_{vis} | (fb) | 0.083 | 0.051 | 0.027 | 0.0017 | 0.0037 | 0.083 |

Table 3.8: Relative efficiencies of the various cuts together with the total efficiencies after all cuts and corresponding visible cross sections for signal and background processes.

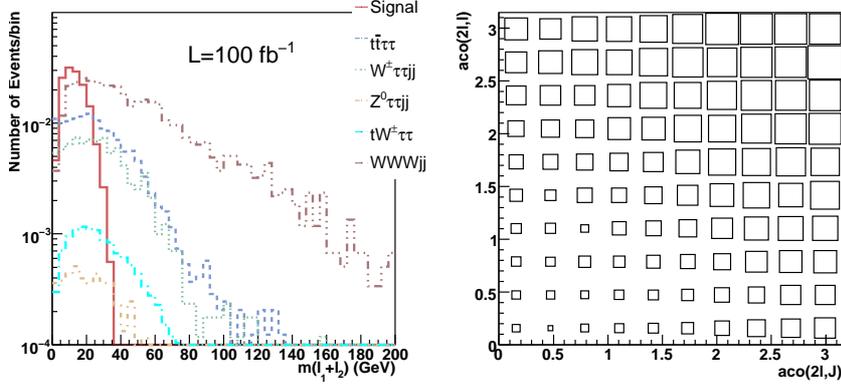


Figure 3.12: Left: invariant mass distribution of the two closest leptons (with different charges and flavours) for both signal and background events. Right: two dimensional distribution of events after $A(l_1^3, j^2)$, C_b , \overline{C}_Z and C_A cuts, as a function of the acoplanarity between the sum of collinear leptons and the hardest non b -tagged jet acoplanarity $\text{aco}(2l, J)$, and the third lepton $\text{aco}(2l, l)$.

point BP2 is not the most optimistic scenario: a lighter H^\pm associated with a larger $\tan\beta$ would sizably increase the production cross section.

In addition, more exclusive discriminant variables could be used to exploit further the characteristics of the typical topology. As an example, let us consider the fact that the heaviest particle in the process is the charged Higgs boson with at least twice the mass of the top quark. As a result, it is typically produced with a small transverse momentum, giving acoplanar W^\pm and A bosons with large boost. This acoplanarity ($\Delta\phi$ between considered final states) can be estimated from the two collinear lepton $e^\pm\mu^\mp$ together with the decay products of the W originating from the charged Higgs. This decay product is either the third lepton, or the “large” jet J if the W boson from the charged Higgs decays hadronically. Since the two topologies are a priori not known, the two acoplanarity definitions (resp. $\text{aco}(2l, l)$ and $\text{aco}(2l, J)$) are built for each event. The distribution of signal events with respect to these two variables is illustrated on the right plot of Fig. 3.12. Since the distribution of background events is much more uniform in this plane, an enhancement of the S/B ratio of around 10% can be achieved if a cut $\text{aco}(2l, l) + \text{aco}(2l, J) > 3$ is applied.

3.5 Conclusion

In this chapter we show that both Types of $iM2HDM$ can be studied with a small set of processes.

In particular, the discovery of the neutral Higgs bosons is possible with a quite small integrated luminosity. For instance, in Type I, all neutral Higgs bosons are expected to be discovered simultaneously via the process $pp \rightarrow h \rightarrow ZAZA$ with an integrated luminosity smaller than 30 fb^{-1} . In the Type II, the $5\text{-}\sigma$ evidence for the existence of the SM-like Higgs could be reached with a few tenths of fb^{-1} , through the analysis of the process $pp \rightarrow h \rightarrow ZZ$. To discover the H^0 Higgs boson, the channel $b\bar{b} \rightarrow H^0 \rightarrow ZA$ is the most promising. With a sizable cross-section and a very low background, less than 10 fb^{-1} of integrated luminosity is required.

The observation of the charged Higgs bosons is also possible in both Type models. In Type I a promising reaction is $pp \rightarrow W^{*\pm} \rightarrow W^\pm AZA$, of which a great advantage is to provide an easy mean to evaluate the charged Higgs boson mass. In Type II such a discovery is based on the single-top associated process $pp \rightarrow tH^- \rightarrow tW^-A$. We show that if one W and A bosons decay leptonically, the required integrated luminosity could be larger than a few hundreds of femtobarns. Despite the more challenging character of this process, its study could be very important as being probably the only mean to discover the charged Higgs in Type II $iM2HDM$.

Study of the $pp \rightarrow H^0 \rightarrow ZA$ in CMS

IT has been suggested in the previous chapter that a discovery of the process $pp \rightarrow H^0 \rightarrow ZA$ modeled by the Type II $iM2HDM$ could be achieved with an integrated luminosity of the order of 10 fb^{-1} . In addition, this process offers the possibility to measure the mass of the two neutral Higgs bosons simultaneously. This chapter aims at confirming these promising estimations with a more realistic detector simulation in the context of the Compact Muon Solenoid (CMS) experiment. In addition, a more exhaustive list of background processes are taken into account and a more sophisticated analysis is performed.

4.1 Event Simulation

In Sec. 3.4.2 the benchmark point BP2 ($m_A=40 \text{ GeV}/c^2$ and $m_{H^0}=350 \text{ GeV}/c^2$) was motivated by the average values of A and H^0 bosons masses in their allowed ranges. However, the extrapolation of this result to situations where other mass choices are considered is not straightforward. We extend our study to nine particular mass configurations, with $m_A= 20, 40, 80 \text{ GeV}/c^2$ and $m_H= 300, 350, 400 \text{ GeV}/c^2$. The decay chain for the signal is the one considered in the analysis made with `Delphes` (see Tab. 3.5): the Z boson decays into a pair of electrons or muons and the pseudo-scalar A boson into a pair of τ 's, themselves decaying into $e^\pm \mu^\mp \cancel{E}_T$.

Another important improvement is the reliable reproduction of interactions between final state particles and the CMS detector. This can only be performed with the official

CMS simulation software suite. It gives a more realistic estimation of the visibility of the signal as it takes into account the exact detector geometry, the (dead) material effects, reconstruction efficiencies and piled up events. In Sec. 3.4.2 we have only considered the presence of irreducible backgrounds, *i.e.* providing the same final states as the signal, $t\bar{t}Z$, WWZ and $Z\tau\tau$. For the sake of completeness the processes tWZ is also taken into account.

As opposed to the irreducible processes, the reducible backgrounds are defined as the processes containing reconstructed leptons that may come either from jet fluctuations or photons, or still non-prompt leptons production. The leptons are classified as non-prompt leptons if they are produced by the semi-leptonic decay of mesons. For instance from B mesons these leptons can be produced directly ($b \rightarrow l^- \nu_l X$, $c \rightarrow l^+ \nu_l X$) or from cascade decay ($b \rightarrow c \rightarrow l^\pm \nu_l X$).

Thus the processes Z +light jets, $Zb\bar{b}$, WZ +jets and $WZ\gamma$, previously neglected, have been included in the present analysis.

As presented in the second chapter, the quality of simulation techniques for the generation of initial states radiation is important. This radiation can, for instance, induce a transverse boost of the final states (in both signal and backgrounds), and influence the modeling of the missing transverse momentum.

For most of the processes considered in this analysis, up to two partons are used at the matrix-element level and therefore a jet matching technique is needed. The simulation of events is done using k_\perp -MLM method, with a matching cutoff Q_{match} that varies from 70 GeV for the signal down to 20 GeV for the backgrounds with smallest scales (masses). .

Other remarks

In order to approach the conditions of the LHC runs, we include the pile-up (PU) for low luminosity ($2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$). Assuming that the inelastic cross-section for a p-p collision is 70 mb and the bunch crossings happen at a rate of 32 MHz, the mean number of PU event is close to 4.4 per bunch crossing. We then choose to simulate, on average, 5.0 PU events.

A possible source of background is the multiple parton interaction. Indeed, the superimposition of a Z events and the product of another partonic interaction (underlying event) could lead to a final state similar to the one of the signal. As an electron and a muon are required beside the leptons from the Z boson the additional process is required to be either $b\bar{b}$ with semi-leptonic decays of B -mesons or $Z/\gamma^* \rightarrow \tau\tau \rightarrow e\mu + \cancel{E}_T$.

The corresponding cross-section can be calculated as the product between the Z cross-section times the probability for two quarks/gluons to provide $b\bar{b}$ and Z/γ^* . We therefore have

$$\sigma_{Z+UE} = \sigma_Z \frac{\sigma_{UE}}{\sigma_{pp}} \quad (4.1)$$

where UE is the underlying event and σ_{pp} is the inelastic cross-section of protons at 14 TeV, close to 70 mb. The inclusive cross-section of $Z \rightarrow ee, \mu\mu$ is close to 4 nb, and the cross-section of $pp \rightarrow b\bar{b}$ with a minimal P_T of 10 GeV/c is roughly 20 μb , leaving $\sigma_{Z+UE(b\bar{b})} \sim 1$ pb. The equivalent cross-section for the production of $Z \rightarrow ee, \mu\mu$ associated to the UE events $Z/\gamma^* \rightarrow \tau\tau \rightarrow e\mu + \cancel{E}_T$ is $\mathcal{O}(100)$ ab.

While the second case seems not problematic, the $b\bar{b}$ case displays a high cross-section compared to the one of the signal. However, we show in Sec.4.4.2 that the probability to get one isolated electron and one isolated muon from the two b -jets is very small, leaving less than one femtobarn of visible cross-section, before applying any topological cut. Furthermore, the production of the two b -quark in the UE event should be essentially acollinear, and therefore very efficiently rejected by the collinearity cut in the analysis algorithm whose simplified version is depicted in Sec. 3.4.2.

This, associated to a bad behavior of UE with jet matching (discontinuity at the matching scale in the differential jet rates) allows us to ignore this source of background in the present work.

The relevant informations relative to the events simulation is summarized in Tab. 4.1. The $\sigma \times BR$ of ZA signal and backgrounds processes, are indicated as well as the number of matrix-element events generated. In case the jet matching is used, Q_{match} is also given. For all processes, the factorization and renormalization scales used are calculated on a event-by-event basis, and no kinematical cut on the vector bosons nor top quarks nor fermions from decay is applied. For WZ +jets, $Zb\bar{b}$ and Zjj , the maximum pseudo-rapidity of partons/jets is 5. The term Zjj encompasses the light jets only: u, d, s, c, g .

The next section describes briefly the CMS detector and its corresponding software CMSSW.

4.2 Event detection and reconstruction

The Compact Muon Solenoid (CMS) detector is installed at the point 5 of the LHC ring, 100 meters under the local mean altitude of the ground. The project was first introduced and the main lines of the design drawn in the early 90's. The main objective

| Process | Q_{match} | Decay (MC) | $\sigma \times BR(\text{fb})$ | # Evts |
|----------------------|-------------|---|-------------------------------|--------|
| ZA(20,300) | 70 | | 6.9 | 73k |
| ZA(20,350) | 70 | | 4.5 | 70k |
| ZA(20,400) | 70 | | 2.9 | 67k |
| ZA(30,350) | 70 | | 4.3 | 70k |
| ZA(40,300) | 70 | | 6.5 | 73k |
| ZA(40,350) | 70 | $(Z \rightarrow l^+ l^-)(A \rightarrow \tau\tau \rightarrow e^\pm \mu^\mp \cancel{E}_T)$ | 4.2 | 350k |
| ZA(40,400) | 70 | | 2.7 | 67k |
| ZA(50,350) | 70 | | 4.15 | 70k |
| ZA(60,350) | 70 | | 4.1 | 70k |
| ZA(80,300) | 70 | | 5.6 | 73k |
| ZA(80,350) | 70 | | 3.9 | 70k |
| ZA(80,400) | 70 | | 2.6 | 67k |
| $W^\pm Z$ | 40 | $(W^\pm \rightarrow \tilde{l}^+ \cancel{E}_T)(Z \rightarrow l^+ l^-)$ | 780 | 1.2M |
| $Zb\bar{b}$ | - | $(Z \rightarrow l^+ l^-)$ | 14000 | 600k |
| Zjj | 20 | $(Z \rightarrow l^+ l^-)$ | 443000 | 2M |
| $Z\tau\tau$ | 30 | $(Z \rightarrow \tilde{l}^+ \tilde{l}^-)(Z/\gamma^* \rightarrow \tau\tau \rightarrow l^+ l'^- (\cancel{E}_T))$ | 10 | 580k |
| $t\bar{t}Z$ | 60 | $(t \rightarrow \tilde{l}^+ b \cancel{E}_T)(\bar{t} \rightarrow \tilde{l}^- \bar{b} \cancel{E}_T)(Z \rightarrow \tilde{l}^+ \tilde{l}^-)$ | 3.5 | 63k |
| $W^\pm Z\gamma$ | 40 | $(W^\pm \rightarrow \tilde{l}^\pm \cancel{E}_T)(Z \rightarrow l^+ l^-)$ | 4 | 280k |
| $W^+ W^- Z$ | 60 | $(W^+ \rightarrow \tilde{l}^+ \cancel{E}_T)(W^- \rightarrow \tilde{l}^- \cancel{E}_T)(Z \rightarrow l^+ l^-)$ | 0.4 | 104k |
| $(t/\bar{t})W^\pm Z$ | - | $(t/\bar{t} \rightarrow \tilde{l}^+ \cancel{E}_T b/\bar{b})(W^- \rightarrow \tilde{l}^- \cancel{E}_T)(Z \rightarrow \tilde{l}^+ \tilde{l}^-)$ | 0.5 | 200k |

Table 4.1: Cross sections of signal ZA and background processes taking into account the leptonic final state considered in the analysis. The notation l includes only e and μ , whereas \tilde{l} also contains τ decaying into e or μ . All cross sections correspond to the final states in the second column. The fourth column indicates the number of Monte Carlo events used for the analysis.

of the CMS detector is to allow the detection of a large variety of new physics signatures. This requires good identification of leptons and jets as well as precise evaluation of their kinematical properties. This is achieved by means of highly segmented sub-detectors and the presence of a string magnetic field (3.8 T). Moreover the hermiticity of CMS allows to measure the transverse missing energy, which is also very sensitive to new physics.

The original design of CMS (1990) was optimized for the detection of muons such that reaction like $pp \rightarrow h \rightarrow ZZ \rightarrow \mu\mu\mu\mu$ could be observed [80]. In 1992, the conceptual design of a more complete, real multi-purpose detector was presented, with central tracking devices, electromagnetic and hadronic calorimeters [81]. The funding and construction of these sub-detectors were contributed by many laboratories such that the CMS collaboration is now composed of more than 180 institutes around the world, and involve more than 2000 physicists and a few hundreds of engineers.

This section describes the main features of these sub-detectors starting from the region close to the interaction point and going outward.

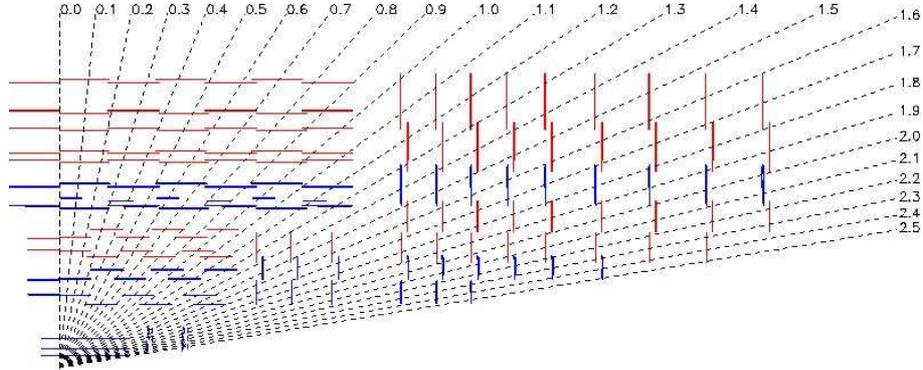


Figure 4.1: View of the active parts of the tracker with the pseudo-rapidity information and de modular structure of layer and ring. The blue color indicates that the layer/ring is equipped with double sided modules while equipped with single modules in red.

We describe the different sub-detectors here, from the center to the outward direction. Note that more details can be found in various sources like Ref. [69].

4.2.1 The CMS detector

The Tracker

The role of the tracker is to detect the passage of charged particles and to allow an accurate calculation of their kinematical properties. The geometry of the tracker is shown in Fig. 4.1. The innermost part is occupied by a pixel detector, whose innermost layer is at roughly 10 cm from the interaction point. It is composed by 3 layers in the barrel and 2 disks in the endcaps. The whole pixel detector contains more than 60 millions of pixels, each with a size of $100 \times 150 \mu m^2$. This allows a localization of the tracks with a resolution better than $20 \mu m$. The occupancy (percentage of pixels crossed by particles) is very low, reaching for instance $\sim 10^{-4}$ at the nominal luminosity of $10^{34} cm^{-2} s^{-1}$. This helps to disentangle the tracks in dense jet environment, facilitates the seeding (first step to build a track) and the identification of primary and secondary vertices.

Surrounding the pixel tracker, the silicon strip tracker is composed, in the barrel, of the Tracker Inner Barrel (TIB) and the Track Outer Barrel (TOB). A similar distinction is present in the endcaps with the Tracker Inner Disk (TID) and the Tracker End Cap (TEC). In the transverse plane, the spatial resolution is a few tenth of micrometer.

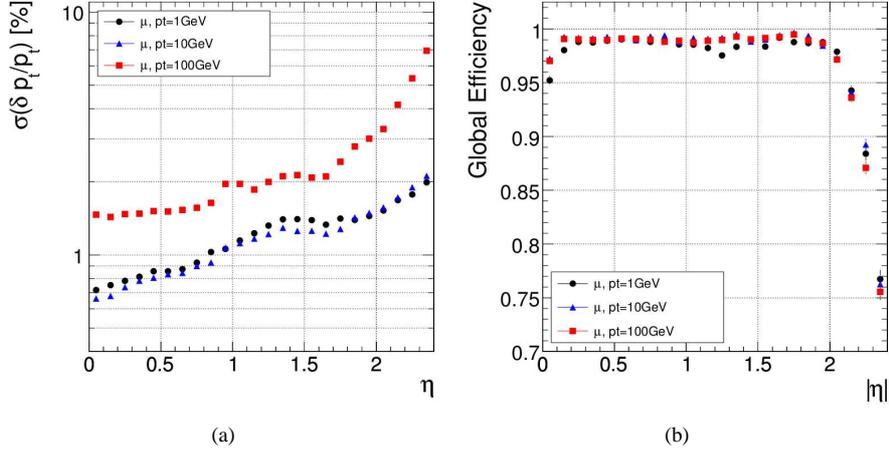


Figure 4.2: (a) Resolution in transverse momentum for muons at 1, 10 and 100 GeV. (b) Efficiency of reconstruction of a track issued from muon. [69]

For the stereo layers (blue in Fig. 4.1), the resolution in the longitudinal direction fluctuates between $230 \mu\text{m}$ in TIB and $530 \mu\text{m}$ in TOB.

The tracker allows an excellent determination of the charged particle momentum, with an efficiency above 95 % on most of the pseudo-rapidity range (see Fig. 4.2 (right) and Ref. [69]). The resolution, expressed as the gaussian fit of $\delta P_T/P_T$ is shown in Fig. 4.2 (left) for muons of different P_T . It is close to the percent on the entire η -range once the transverse momentum is larger than 10 GeV. For such muons, it has been also shown that the resolution on the impact parameter is better than $30 \mu\text{m}$ which is important for the secondary vertices detection (the typical flight distance of B mesons is a few hundred microns).

In CMS, the standard track reconstruction algorithm is based on five steps: hit identification, seeding, trajectory building, ambiguity removal and final fit.

The hit identification consists of finding particle hits in both pixel and silicon strip detectors. A hit is made from the clustering of detection units above a given signal-to-noise ratio. The next step is the track seeding (Ref. [82]), mostly done with the pixel detector. The hits are used to build track seeds parameterized as helices. The seeding is also the starting point of the standard track reconstruction method, the Combinatorial Track Finder (CTF) (Ref. [83]). The trajectory building is made by extrapolating the estimation of the track 3-momentum given by the seed to the outermost layers, taking into account material effects and the magnetic field. For each layer, a new trajectory is created by compatible hit, plus one trajectory in the hypothesis that no hit have

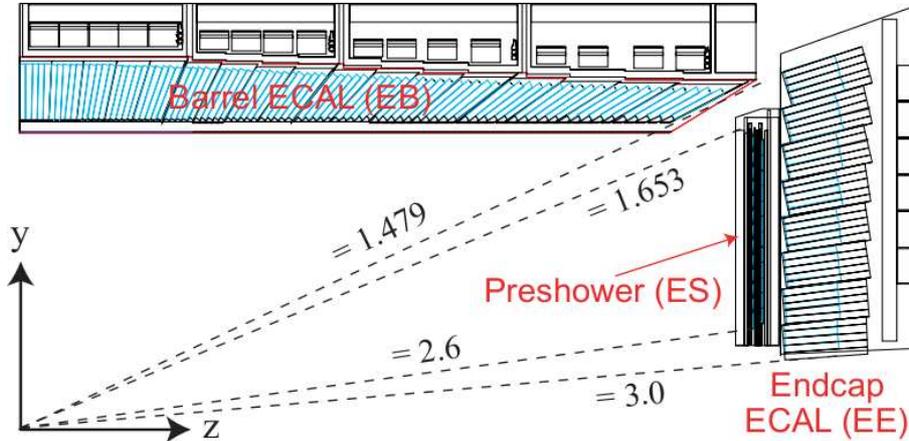


Figure 4.3: Longitudinal view of the ECAL system of CMS.

been left. Each trajectory is passed through the Kalman Filter [84], that combines the informations and outputs an updated track. The track quality is determined at each extrapolation step, using the χ^2 and number of missing hits. If a track does not satisfy a tunable minimum quality conditions, it is rejected from the collection. This avoids an exponential growing of the track number.

If different tracks contain the same seed or if a track is compatible with two different sets of seeds, a double counting problem appears. The comparison of track quality is then achieved, based on the number of hits and χ^2 . In order to reevaluate the track parameters, a final step is performed from outside towards the beam line with the CTF.

An additional step to decrease the number of fake tracks can be performed by running several times the CTF algorithm. At each iteration (3 or 4) the unused hits are removed from the collection, the remaining ones are used for a new track reconstruction.

The electromagnetic calorimeter

In order to allow a good reconstruction of electrons and photons, the tracker is surrounded by an electromagnetic calorimeter (ECAL). The structure of ECAL is shown in Fig. 4.3.

The barrel region ($|\eta| < 1.6$) is composed of more than 60000 crystals of $PbWO_4$, each covering an $\eta - \phi$ area of 0.017^2 (\sim front face surface of 22×22 mm² in $\eta = 0$). Their length is 23 cm, which corresponds to 25.8 radiation lengths (X_0) defined as the mean distance needed for an electron to loose all but $1/e$ of its initial energy

by bremsstrahlung. The crystal width is of the same order of magnitude than the Moliere radius which defines the width of the cylinder where 90 % of the showered energy is contained. This allows to contain the electromagnetic radiation and to have a significant part of the energy deposited in one crystal. Furthermore this material allows a fast response (80% of the shower light emitted in 25 ns), which is crucial to avoid overlap between radiation from different bunch crossings.

The endcaps cover the pseudo-rapidity from 1.479 to 3.0. The crystals are slightly larger with a constant front surface of $28.6 \times 28.6 \text{ mm}^2$ and shorter, with a length of 22 cm.

The energy resolution can be parametrized as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{s}{\sqrt{E}}\right)^2 + \left(\frac{n}{E}\right)^2 + c^2 \quad (4.2)$$

where s is the term sensitive to the shower variations and the photodetector response. The term n is related to the noise from electronics read-out and pile-up (simultaneous energy deposition by uncorrelated particles), and the constant c value is sensitive to mis-calibrations, non-uniformity of the detector, instabilities of temperature¹.

In test beam condition (see Ref. [85]), these parameters were estimated as $s = 0.028 \text{ GeV}^{1/2}$, $n = 0.125 \text{ GeV}$ and $c = 0.03$.

In order to improve the identification of electrons, their energy and position resolution in the electromagnetic calorimeter, two clustering algorithms have been developed[86]. The first method is called hybrid method, is applied only in the barrel and designed to reconstruct high energy electrons. It consists of collecting the energy in a fixed set of adjacent crystals, 3 or 5 in the η axis, while along the ϕ axis the algorithm looks for isolated energy deposit due to bremsstrahlung effect. The fixed number of crystals to be used give a good sensitivity to centroid position. The movement along the ϕ axis makes of the hybrid method a super cluster algorithm, where a super cluster is defined as the association of the cluster from the electron energy with the clusters from the photons radiated.

The second method is the island algorithm and is used in the endcap regions. It starts by identifying the most energetic crystal (seed), which is used as a starting point for the clustering. This is illustrated in Fig. 4.4(left): the search is first performed in both direction in ϕ . The energy of each encountered crystal is summed until a crystal with a larger energy deposit is met. The algorithm redo the search in the η line besides, along the ϕ direction. As for the ϕ direction, the η direction is explored until a crystal with a larger energy is encountered. A crystal cannot belong to two clusters, this guarantees the absence of double counting. A clustering of several island clusters defines a super

¹The temperature regulation is critical since the crystal light yield is highly dependent of it.

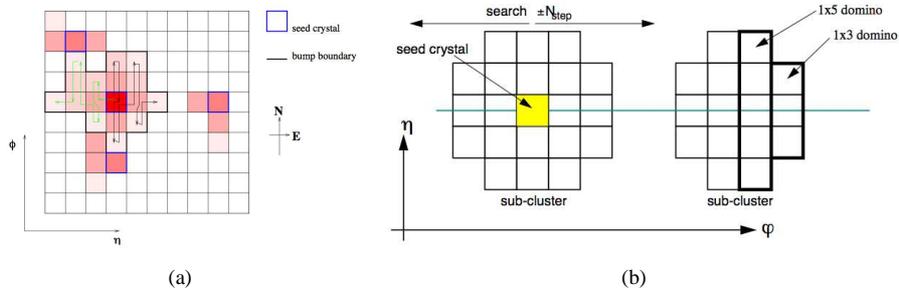


Figure 4.4: Illustration of the Hybrid and Island algorithms on the crystals of CMS.

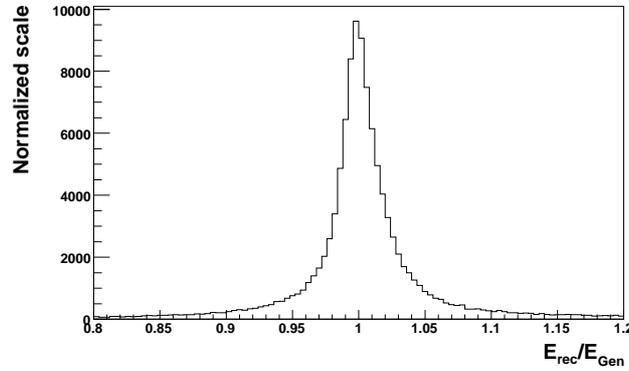


Figure 4.5: The energy resolution for electrons emitted from Z and A bosons in ZA process (BP2). The transverse momentum is larger than 5 and $|\eta| < 2.5$.

cluster.

The offline (reconstructed) electrons are defined as an association between an ECAL supercluster (cluster of clusters to take into account the bremsstrahlung effect) and the tracker information. First the matching is done between the supercluster and the hits in the pixel detector. The energy of the SC allows to deduce the position of the pixel seeds since the radiated photon are assumed to be counted in the SC average position.

The ECAL resolution can be expressed as the ratio between the energy in ECAL and the energy of the generated electron E_{ECAL}/E_{gen} . This is shown in Fig. 4.5 for electrons emitted from the Z and A boson decays in the ZA process (BP2).

The hits in the pixel detector are then used as a seed to find the track corresponding to the electron candidate with the Gaussian Sum Filter algorithm (Ref. [87]), which allows a better reconstruction of the track than the Kalman Filter as it takes into account

the possibility of large energy loss by bremsstrahlung. This method is based on the hypothesis that the energy loss distribution is not modeled by a single gaussian but by a sum of gaussians.

When the measurement of electron energy (E_e) is performed in association with the tracker, the energy resolution is rather good. It varies from 5 % if $E_e=5$ GeV to roughly 1.5 % once the $E_e > 20$ GeV. It is observed that the track information helps to get a better resolution if the electron energy is lower than roughly 15 GeV while above, the calorimetric information is more precise.

The hadronic calorimeter

The hadronic calorimeter (HCAL) is installed around the electromagnetic calorimeter. It allows to estimate the energy and position of hadronic radiation (both the charged and neutral hadrons). The detector is composed of layers of brass and scintillators, reaching 5 radiation lengths and up to 11 radiation lengths with the absorber layer placed just before the magnet so that only muon escape from HCAL and pass the magnet. The barrel covers up to $\eta = 1.4$ the endcap system up to 3.0 and the forward system up to 5.0. The tower dimension is 0.087×0.087 in η and ϕ in the barrel while in the endcaps the $\Delta\eta$ increases up to 0.35 and $\Delta\phi=0.174$ radians.

Test beam using pions has shown that the energy resolution of HCAL can be parameterized as

$$\frac{\Delta E}{E} = \frac{1.22}{\sqrt{E}} + 0.05 \quad (4.3)$$

with E in GeV.

The muon chamber system

Surrounding the magnet coil, the muon system aims at an unambiguous identification of the muons as well as an estimation of their kinematics. This is achieved by combining the information from the muon chambers and the silicon tracker system in a way that depends on the muon transverse momentum.

The muon system consists of an alternative arrangement of iron yokes (for conduction of the magnetic field) and gas detectors. There are three kind of muon detectors: the drift tubes (DT) in the central part of the barrel, the cathode strip chambers (CSC) in the endcaps and the resistive plate chambers (RPC) in both regions of the volume.

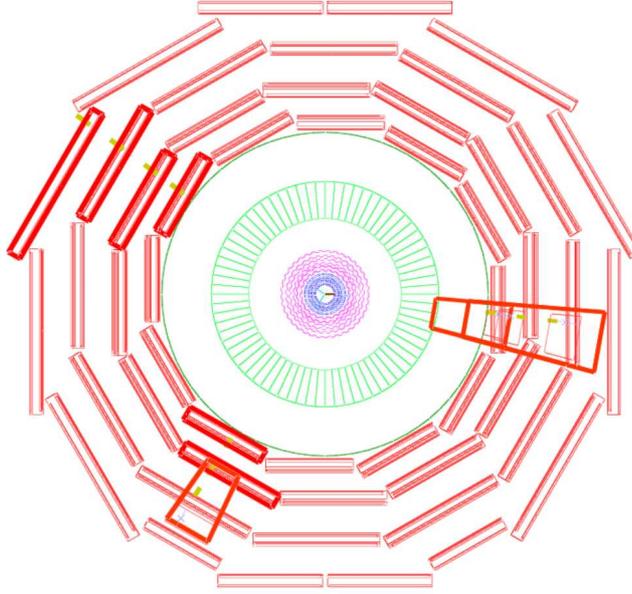


Figure 4.6: Transverse view of the muon system of CMS with 3 muons from a $pp \rightarrow (Z \rightarrow \mu^+\mu^-)(A \rightarrow \tau^+\tau^- \rightarrow \mu^+e^-)$ where one hits the barrel only, another the endcap only, and the third hits both regions.

The muon reconstruction in CMS can be done in different ways. The standalone muon reconstruction is based only on the information from muon detectors, the global muon reconstruction starts with the standalone muons and matches them with tracks in the tracker. Finally the tracker muon reconstruction starts with the tracks in the tracker and matches with segments informations in the muon system. This has the virtue to allow the reconstruction of muons with low transverse momentum that do not leave enough hits in the muon system to be considered as standalone muons.

In this analysis, the muons are reconstructed using the following algorithm.

The muon tracks, *i.e.* the tracks built from muon system only, reconstruction starts with the seed identification. A seed is defined as a segment or a set of segments built from the hits in DT and CSC. At this stage a first estimation of the transverse momentum is realized by evaluating the segment bending with respect to the vertex position. The seed trajectory is propagated down to the innermost muon detector layer. Then a first Kalman Filter is applied to refine the track in the outward direction. When the outmost muon chamber layer is reached, the track reconstruction is done in the reverse direction with again the application of the Kalman Filter. The matching between muon and tracker is then realized as follows. The first step is to select, around

each muon track a region of interest in $\eta - \phi$ and select a subset of tracks in it. Then those tracks are extrapolated up to the muon system tracks and the best pairing in terms of space and momentum is chosen.

4.2.2 CMSSW

The large data flow of each LHC experiment is a challenge in itself. It has necessitate a particular treatment of the data, from the raw level up to reconstructed objects. To manage this, the CMSSW software has been developed and is used in the rest of this work.

The original data flow is formatted as *raw* data. It contains for example the hits informations, the L1 trigger results, and HLT bit pattern.

The raw data are then treated to output reconstructed (RECO) objects: track collections, electrons (for b-tag or not), ECAL cluster collections, jets, missing transverse energy, τ -jets. In order to stay adaptive it is possible to keep both raw and reconstructed collections in the data files.

There are two levels of simulation to get reconstructed objects, the fast and the full simulation. A problem inherent to the full simulation is the large CPU consumption due to the use of GEANT [88] for the simulation of the interaction of particles with the material of CMS. Instead, the fast simulation uses the result of accurate GEANT-based simulation, allowing to by-pass the main time consuming steps. The gain in time can reach two orders of magnitude. In the rest of this work all presented result are obtained with the fast simulation only.

Using as input the events in RECO format, the layers forming the Physics Analysis Tools (PAT) provide a more user-friendly interface and then make easier the access to reconstructed variables for analysis purpose. The path followed by the data from the RECO level up to the user analysis level is shown in Fig. 4.7.

The PAT flow is composed of three layers but here we concentrate on the two first that are actually used in this work. The PAT Layer-0 sequence extracts needed informations from the reconstructed level (generator, isolation, tagging,...), performs a cleaning of the data collections (notably duplicate removal) and the matching with MonteCarlo events. The PAT Layer 1 step groups informations output from the Layer-0, allowing a straightforward access to them and therefore constituting the basis for an analysis. More details about what is used in the present analysis is given below.

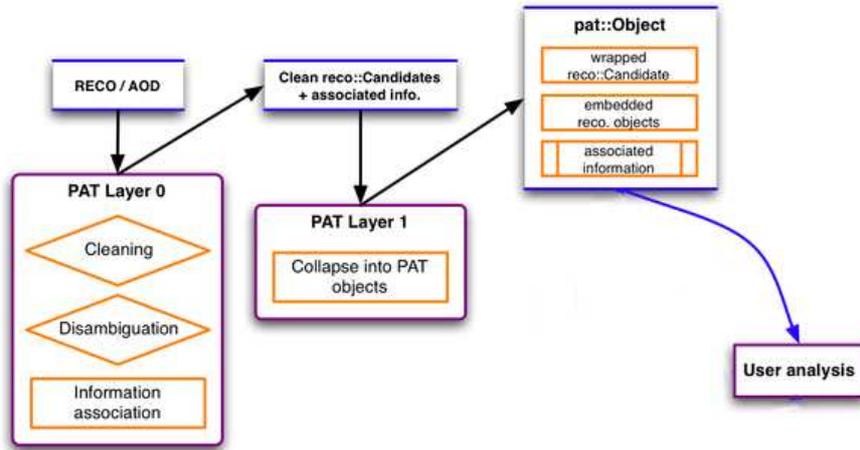


Figure 4.7: Flowchart of the data treatment from reconstructed level (RECO) to User Analysis level.

Electrons

For the electrons the PAT Layer-1 level allows the access to several variables that can be used to define the identification and isolation criteria of the electrons:

- The ratio H/E quantifies the ratio of the energy deposits of the closest HCAL cell to the supercluster energy. For a real, isolated and hard electron, the value of H/E is supposed to tend to zero as most of its energy is absorbed by the electromagnetic calorimeter.
- The variable $\Delta R(\text{track}, \text{SC})$ gives the distance in the $\eta - \phi$ plane between the track extrapolated from the pixel layers up to the supercluster seed. This helps to be convinced that the SC is issued from the same particle that produced the track. For a real, isolated and hard electron, the value of $\Delta R(\text{track}, \text{SC})$ is expected to be close to zero.
- The ratio E_s/P_{out} of the energy of the super cluster seed (E_s) to the momentum of the electron's track calculated at the outermost level of the tracker (P_{out}). This helps to be convinced that the seed is issued from the same particle that built the last segments of the track. For a real isolated and hard electron, the value of E_s/P_{out} is supposed to be close to 1.
- The ratio E_{SC}/P_{in} of the super cluster energy (E_{SC}) to the associated track momentum (P_{in}) at the vertex. This helps to be convinced that the SC is issued

from the same particle that produced the track. For a real isolated and hard electron, the value of E_{SC}/P_{in} is supposed to be close to 1.

- The variables iso_{ECAL} and iso_{HCAL} represent the isolation variables at the level of the electromagnetic and hadronic calorimeters respectively. Both are calculated as the sum of the energy deposits in a 0.4 cone centered on the electron candidate position, excluding the energy of the reconstructed electron for iso_{ECAL} . For a real isolated and hard electron, the value of both isolation is supposed to tend to zero.
- The variable $iso_{track}(elec)$ is defined as the sum of transverse momentum of tracks in a cone of 0.3 or 0.5 around the electron track direction estimated at the vertex. For a real isolated electron, $iso_{track}(elec)$ is therefore supposed to low. A standard requirement for these tracks to be taken into account in the sum is to have a P_T larger than 1 GeV. For a real, isolated and hard electron, the value of track isolation is supposed to tend to zero. In this analysis, the decay of the A boson can produce collinear electrons and muons. A redefinition of the isolation criterion must be done with care.

The output of the PAT layer is kept very loose, *i.e.* no restriction on these variable is considered. This means of course that at the level of the PAT, the muons, jets and photons surrounded by tracks may also be recorded as electrons.

Muons

Reconstructed Global muons, as defined in Sec. 4.2.1 are available at the level of the PAT Layer 1 and are used in the present analysis. The variable $iso_{track}(\mu)$ is defined (as for the electrons) as the sum of transverse momentum of tracks in a cone of 0.3 or 0.5 around the muon track direction estimated at the vertex. As for the electron case, the standard requirement for these tracks to be taken into account in the sum is to have a P_T larger than 1 GeV.

Missing transverse energy

As shown in the Delphes analysis, the missing transverse energy plays a major role in the reconstruction of both A and H^0 bosons. There are different definitions of this variable, and the standard one provided by the reconstruction tools of CMS is defined as the vector sum of calorimetric towers transverse energy deposits and muons transverse momenta. This quantity is accessible from PAT Layer 1. Another possibility that is also briefly discussed in Sec. 4.4.2 is to consider only the four leptons. However this solution is very sensitive to initial state radiation.

4.3 Selection of leptons from Z and A decays

As already mentioned in Sec.3.4.2, the leptons originating from the cascade decays of the A boson are expected to be relatively collinear. This is due to the large mass difference between m_A and m_{H^0} . The isolation criteria must be therefore defined more carefully than the standard choices used on leptons issued from W or Z for instance.

In addition, the expected energy spectrum of the electrons from these cascade decays is relatively soft. We then consider reconstructed electrons down to 5 GeV. However, the danger comes from the higher probability for a soft jet to be seen as an electron. In this section, after a short discussion about the issue of triggering the signal events, we study how the rejection of fake electrons from jets together with non-prompt electrons originating from other sources like heavy flavor meson decays can be performed.

A cut-based analysis specific to each A and H^0 mass points is then proposed. This allows to see how fast a deviation from the SM expectations could be seen. A method is then proposed to control the most dangerous background with data. Finally a more sophisticated analysis is developed which includes an estimation of the systematic errors.

4.3.1 Final states and Trigger condition

At the LHC, the bunch crossing is expected to happen each 25 nanoseconds at the interaction points. In order to reduce the amount of information to be recorded, a strong filtering of the events is needed, reducing the events rate from 40 MHz to about 100 Hz.

The trigger system has two components: the Level-1 and High Level triggers. The Level-1 trigger involves the calorimeters and muons systems, and identifies electrons, muons, photons, \cancel{E}_T and jets using coarse granularity and low resolution data. Schematically, it is divided in 3 parts. The Muon trigger has independent triggers logic for each kind of detector (DT, CSC, RPC), the HCAL trigger uses the tower information. The ECAL trigger is based on the information from 5×5 crystals in the barrel and roughly similar to the layout used for HCAL in the endcaps. The level of output rate is limited to 100 kHz.

The High Level Trigger (HLT) treats the events output by the Level-1 trigger. Many final states can be used at this level, thanks to the lower event rate and hence the access to more sophisticated reconstruction algorithms. The HLT may use for instance the electrons built from the ECAL and the tracker information, the muons with tracker

and muon chambers informations, the photons, the jets, τ , \cancel{E}_T or still b -tagged jets. Considering that one events size around 1 MB, the rate allows the transfer and storage of events.

In this analysis, we only consider the cleanest signal final states, namely $eee\mu$ and $\mu\mu\mu e$, where a pair of electrons or muons comes from the Z decay, and an electron and a muon from by the A boson decay chain. In principle, each of the four leptons is susceptible to contribute to the triggering procedure.

However, the probability to observe both topologies is first subject to the respective reconstruction efficiencies for electrons and muons, which depends on the transverse momentum. This point is illustrated in Fig. 4.8(a): the reconstruction efficiencies for both electrons and muons is shown, with respect to their origin (Z or A) in a signal process (BP2). Note that a generated lepton is considered as reconstructed if a reconstructed lepton (with no identification nor isolation cut) lays at a maximal ΔR of 0.01.

It turns out that in the low transverse momentum region the leptons from A are slightly better reconstructed than the leptons from the Z boson. This can be explained by the difference of the pseudo-rapidity distribution of these leptons, as shown in Fig 4.8(b) for leptons with $P_T < 30$ GeV.

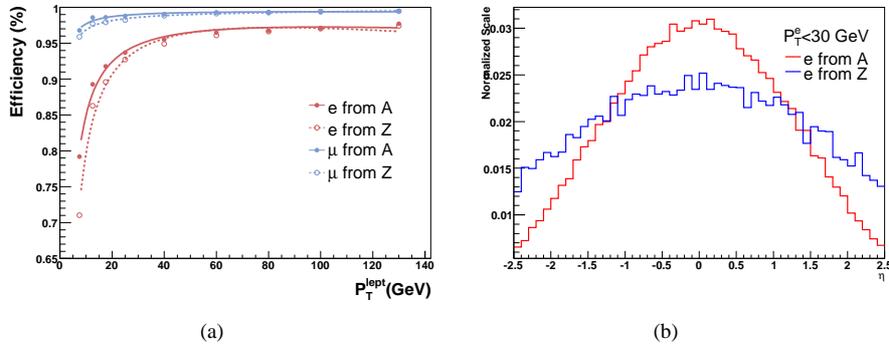


Figure 4.8: (a) Efficiency of reconstruction for electron and muons from Z/A , depending on their transverse momentum. (b) η distribution of electrons from Z/A with $P_T < 30$ GeV.

Depending on the transverse momentum the reconstruction efficiency varies between 70 % and 90% for the electrons and remains above 95% for the muons. The two studied topologies having similar efficiencies and background sources, both final states

| Channel | P_T Threshold (GeV) |
|----------|-----------------------|
| e | 26 |
| ee | 12 |
| μ | 19 |
| $\mu\mu$ | 7 |

Table 4.2: P_T thresholds used to simulate the (HLT) trigger response.

are considered simultaneously in the following analysis and treated together in the triggering procedure.

The configuration of the L1 and HLT triggers should correspond to trigger tables applicable after a time equivalent to an accumulated integrated luminosity of 10 fb^{-1} . However, at the time of writing this text, such possibility has not yet been implemented in the fast simulation of CMSSW. This said, it is possible to get an approximative estimation of the trigger effect with electrons and muons streams.

The choice of trigger is governed by the following requirements.

- I. Keep the signal efficiency on the signal as high as possible.
- II. Choose a pattern trigger as simple as possible to ease the evaluation of its efficiency.
- III. Allow to select events required for possible data driven background estimation.

Given the signal final states, the possible trigger patterns are single electron (e), di-electrons (ee), single muon (μ) and di-muons ($\mu\mu$). The expected transverse momentum threshold corresponding to these trigger patterns can be found in Ref. [19] and are summarized in Tab. 4.2. These threshold are set for isolated leptons.

In order to characterize more precisely (in forecast of the practical redistribution of the data from LHC run) what are the best trigger combinations for our signal, we evaluate the percentage of accepted events, after applying all the selection cuts described in the next sections. The leptons considered are isolated and defined as described in Sec. 4.3.2 and 4.3.2.

The efficiencies possible exclusive trigger patterns based on single and di-leptons patterns are shown in Fig. 4.9. It turns out that the patterns of interest always contain at least two muons or two electrons. The equivalent condition is $ee \vee \mu\mu$. This satisfy

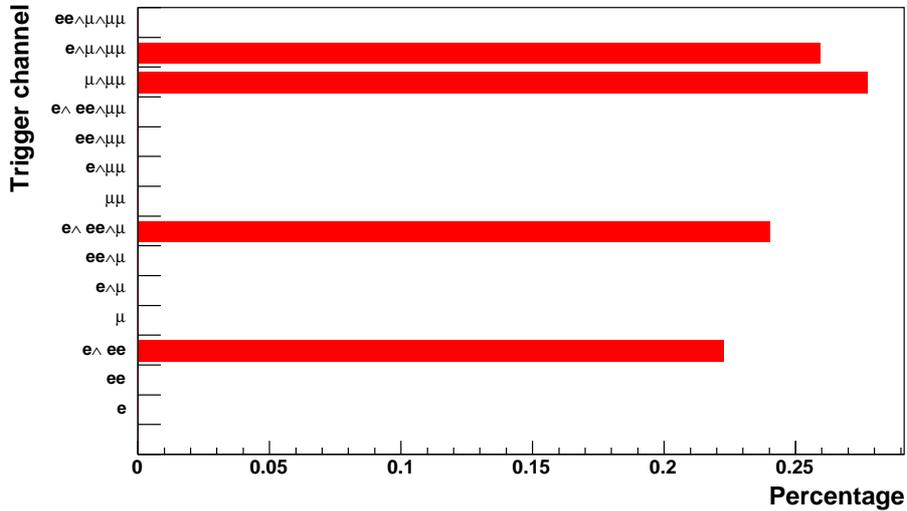


Figure 4.9: List of exclusive (HLT) trigger patterns for the ZA process. These values are obtained after the entire cut based selection. The sum of all trigger pattern efficiency is set to the unity.

the first condition as, on average, 80% of the signal is kept (including the acceptance of the leptons), the pattern is also very simple and the control of the dangerous backgrounds is possible as they contain a Z boson decaying into a pair of electrons or muons.

4.3.2 Lepton identification

The P_T distributions for the four leptons coming from the signal (BP2) are shown in Fig. 4.10. While the leptons issued from the Z boson have P_T most of the time larger than 20 GeV, the situation is dramatically different for the products of the cascade decay of the A boson. The softest lepton has in this case a mean transverse momentum close to 5 GeV/c only.

The minimal P_T requirement is driven by two things. First it has to be low in order to keep as much signal as possible. Second the CMS analyses rarely consider electron P_T below 5 GeV[89] as in lower P_T region the distinction between real electrons and fake electrons from jet fluctuation becomes more difficult. We then simply require the electrons to have a transverse momentum larger than 5 GeV/c .

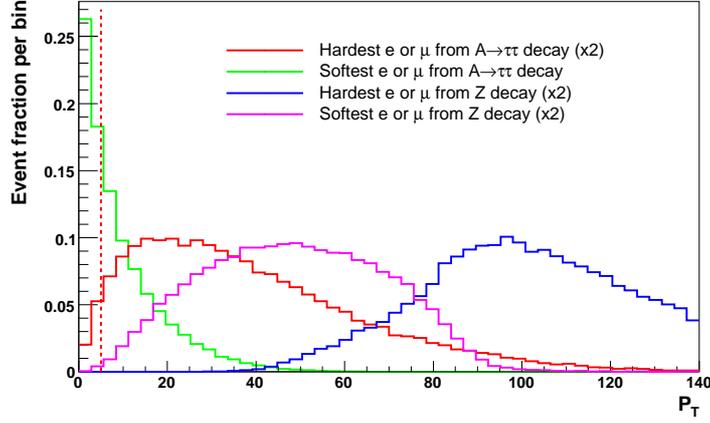


Figure 4.10: Transverse momentum of the leptons originating from $A \rightarrow \tau\tau$ decay chain or Z boson decay, with $|\eta| < 2.5$ in the signal (BP2). The red dashed line shows the 5 GeV/c P_T threshold.

To improve the electron definition, we may want to consider the variables H/E , $\Delta(track, SC)$, E_s/P_{out} and E_{SC}/P_{in} defined in Sec. 4.2.2. For a sake of uniformity, we choose to display a normalized definition $x \rightarrow 1/(1+x)$ so that the infinite values of x correspond to zero while $x=0$ correspond to 1; a normalized distribution of x is denoted x^N

We show in Fig. 4.11 these distribution for four categories of electrons.

- The first categories are the electrons from Z from the signal (BP2).
- The second categories are the electrons from $A \rightarrow \tau\tau$ cascade decay from the signal (BP2).
- The third category is composed by mostly fake electrons reconstructed from misidentified light jets from WZ +jets.
- The fourth category is composed mostly by real electrons from semi-leptonic decay of B mesons in $Zb\bar{b}$ events

The definition of fake and non-prompt leptons is given in Sec. 4.1. This distinction is useful as it allows to see that the distributions are quite sensitive to the amount of non-prompt electrons. In order to reduce the rate of undesired electrons, we apply a cut on each of these variables.

- $(H/E)^N > 0.975$.

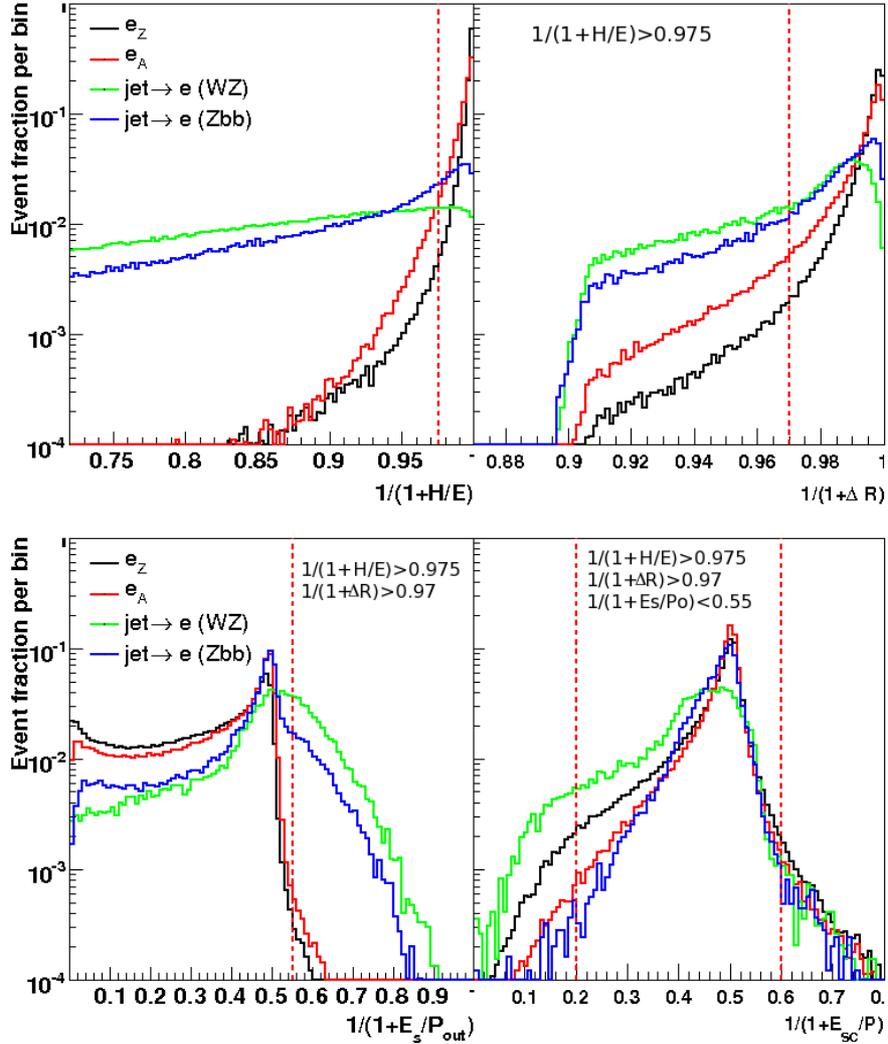


Figure 4.11: The distribution of the variable $(H/E)^N$ (upper left), $(\Delta R(track, SC))^N$, $(\Delta R(track, SC))^N$ (upper right), $(E_s/P_{out})^N$ (lower left) and $(E_{SC}/P_{in})^N$ (lower right) for the electron originated from Z or A boson decay and the other electrons. The distributions from upper left to lower right are shown with the cuts applied to previous figures. Each distribution is normalized to the unity.

- $(\Delta R(\text{track}, \text{SC}))^N > 0.97$.
- $(E_s/P_{out})^N < 0.55$
- $(E_{SC}/P_{in})^N \in [0.2, 0.6]$

The positions of these cuts are also shown in Fig. 4.11 with red dotted lines. With such choices 90% of signal events are kept.

As already discussed in Sec. 4.2.2 the muon must fulfill the global muon requirement, *i.e.* using both tracker and muon chambers information.

As an illustration, the probability for the process WZ and $Zb\bar{b}$ to give a fake or non-prompt electron or muon is given in Tab. 4.3. This table also gives the rejection efficiency of these lepton using the electron or muon identification criteria depicted here above.

| Process | lepton type | $\langle N/\text{event} \rangle$ (%) | ϵ (%) | total% |
|------------------|------------------|--------------------------------------|----------------|--------|
| $WZ+\text{jets}$ | fake e | 39 | 4 | 1.5 |
| | non-prompt e | 1.2 | 10 | 0.12 |
| | fake μ | 2 | 26 | 0.5 |
| | non-prompt μ | 0.5 | 80 | 0.4 |
| $Zb\bar{b}$ | fake e | 20 | 5 | 1 |
| | non-prompt e | 7 | 47 | 3.3 |
| | fake μ | 2 | 23 | 0.46 |
| | non-prompt μ | 9 | 90 | 8.1 |

Table 4.3: For both WZ and $Zb\bar{b}$ processes, the mean number of a given lepton type per event is indicated in the column labelled $\langle N/\text{event} \rangle$. The right hand-side columns give the efficiency when the electron or muon identification criteria are applied, and the mean number of a given lepton type per event after identification (the product of the 3rd and 4th columns).

It turns out that the main danger comes from the non prompt muons and electrons in $Zb\bar{b}$ and from fake electrons in $WZ+\text{jets}$. This motivates the presence of an isolation criterion.

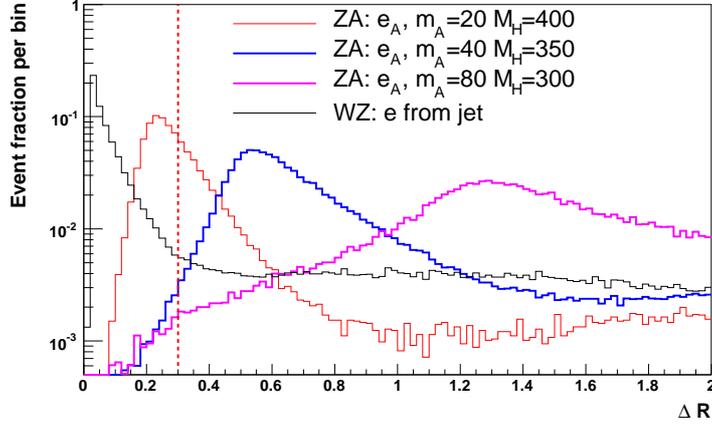


Figure 4.12: The distribution for the distance ΔR between a reconstructed electron and the closest track with $P_T > 5$ GeV. For the signal, the reconstructed electron is matched with a generated electron issued from the A boson whereas for the background it comes from a jet.

Lepton isolation

The rejection of fake and non-prompt leptons can be improved by using $iso_{track}(\text{lepton})$ for both electrons and muons, and iso_{ECAL} and iso_{HCAL} for the electrons. These isolation criteria are defined in Sec. 4.2.2.

However, as emphasized in Sec. 3.4.2, such simple isolation condition in the tracker cannot be considered here as there are configurations of the signal where a high collinearity of the leptons is present, *i.e.* if m_A is small and m_T is large.

The evolution of this collinearity with the Higgs bosons masses is shown in Fig. 4.12 for the ZA signal with $m_A/m_H = 20/400$, $40/350$ and $80/300$ GeV/c^2 . Each histogram displays the distance between a reconstructed electron and the closest track with $P_T > 5$ GeV. In the signal events, the reconstructed electron is matched with generated electron from A boson decay while for WZ +jets, the reconstructed electron comes from a light jet misidentification, *i.e.* it is not matched with any generated electron.

The dashed line refers to a standard cone size of 0.3 used for the track isolation. It is clear that by using the standard definition $iso_{track}(\text{lepton})$, the efficiency on the most boosted case is affected, more than half of the events are indeed rejected. The lepton isolation $iso_{track}(\text{lepton})$ should be therefore redefined at the level of the tracker in order to accept at most one track with $P_T > 5$ GeV/c in a 0.3 cone.

In the signal the closest track to the electron from A boson is supposed to be the track from the muon produced also by the A boson cascade decay. Therefore the associated energy deposit in both ECAL and HCAL around the electron is then small such that iso_{ECAL} and iso_{HCAL} can be included in our new isolation definition.

The electron isolation iso_e is defined as

$$iso_e = \frac{P_T^e}{P_T^e + \sum_{soft} P_T^{tr} + \sum_{2+}^{hard} P_T^{tr} + iso_{ECAL} + iso_{HCAL}} \quad (4.4)$$

and the muon isolation iso_μ is defined as

$$iso_\mu = \frac{P_T^\mu}{P_T^\mu + \sum_{soft} P_T^{tr} + \sum_{2+}^{hard} P_T^{tr}}, \quad (4.5)$$

where P_T^{tr} is the transverse momentum of a track, soft and hard mean respectively with $1 < P_T < 5$ GeV/c and $P_T > 5$ GeV/c respectively. The 2+ notation recalls that the sum runs over the list of tracks with $P_T > 5$ GeV/c in the 0.3 cone from which the track with the highest transverse momentum is removed.

The distributions iso_e (left) and iso_μ (right), defined in the beginning of this section, for the lepton candidates fulfilling the identification criteria are shown in Fig. 4.13. The distinction is done between the leptons from A or Z boson decays (in signal) and the fake and non-prompt leptons from WZ and $Zb\bar{b}$.

The last bin of the iso_μ distribution is surprisingly filled by the muons issued from jets. A short investigation has shown that this comes from generated electrons from W and Z decays, mostly hard and isolated. Different attempts have been done to reject this contribution but they were mostly unsuccessful. However, this bin represents roughly only 10 % of the muons from jets, and the corresponding event will be rejected through the cut-based analysis. We therefore choose to neglect this problem.

The final choice of iso_e and iso_μ cuts is largely conditioned by the rejection of Zjj , $Zb\bar{b}$ and WZ +jets backgrounds. As it will be shown in the following analysis, we choose $iso_e > 0.8$ and $iso_\mu > 0.9$.

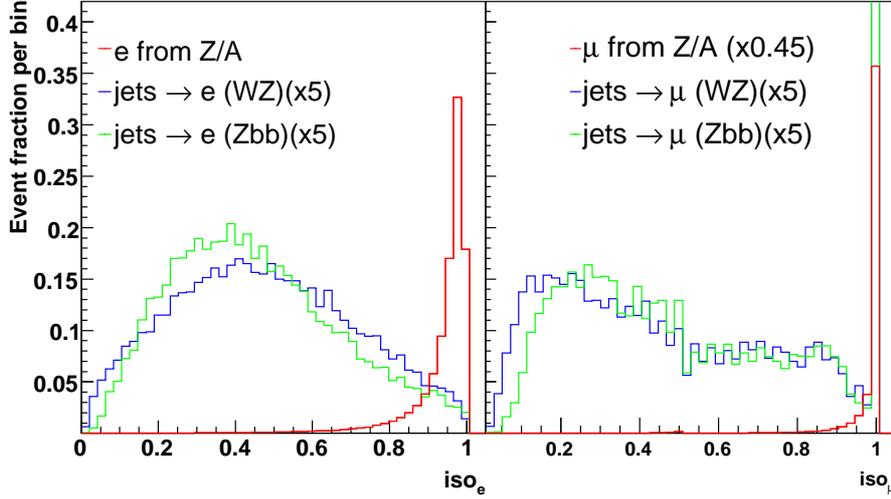


Figure 4.13: Left: The distribution for the iso_e variable. The red line represents the contribution of the reconstructed electrons issued from Z and A in the signal while the blue and green line are for the fake and non-prompt electrons in WZ and $Zb\bar{b}$ events. Right: The same but for the muons, the distribution shown is iso_μ . All distribution are normalized to the the unity times the scaling factor indicated in parenthesis.

4.4 Cut based analysis

4.4.1 Basic cuts

C_Z

After the trigger conditions, events have to pass the acceptance cuts defined by requiring that four leptons must be present with at least one muon and one electron. The first selection cut is C_Z and is defined as in the Sec. 3.4.2: exactly one pair of electrons or muons with opposite charges must reconstruct the Z boson invariant mass in a 10 GeV/c^2 mass window.

Let us use the notation similar to the one introduced in Sec. 3.4.2: $l_Z(1)$ and $l_Z(2)$ are the leptons reconstructing the Z boson and $l_A(1)$ and $l_A(2)$ those from the A boson. We require that the two leptons $l_Z(1,2)$ must have at least 10 GeV/c of transverse momentum, with same flavors and opposite charges. In addition their invariant mass must lay in a mass windows of 10 GeV/c^2 around the Z mass.

After this cut the Z +light jets and $Zb\bar{b}$ as well as the WZ +jets backgrounds are by far the biggest background sources. An illustration of their dominance over the other processes is shown in Fig. 4.14(left).

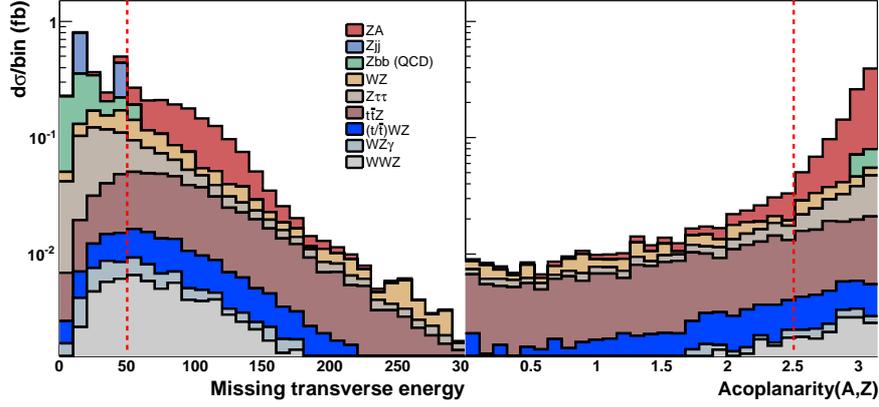


Figure 4.14: (Left) The missing transverse energy of the signal and the backgrounds. (Right) Acoplanarity between $l_Z(1) + l_Z(2)$ and $l_A(1) + l_A(2)$ also after C_Z . The distribution are normalized for a luminosity of 1 fb^{-1} . The red dotted lines indicate the value of the cut considered in the analysis.

The Zjj , $Zb\bar{b}$ and WZ +jets backgrounds can be further reduced by imposing a cut on the missing transverse energy. In order to keep enough signal, we choose to cut at 50 GeV. This allows to reduce strongly the Z +jet background by a factor close to 98%.

We have seen in Sec. 4.2.2 that the missing transverse energy in CMSSW^2 is based on both calorimetry and muons. Instead of this definition, the missing transverse energy can be reconstructed also from the four leptons only. The latter has the advantage to be less sensitive to the calibration of the calorimetric towers. The obvious drawback is that the initial states radiation is not taken into account, leading to a biased estimation of the missing transverse energy direction. This affects the selection efficiency when the mass of the A boson is reconstructed.

We see indeed in Fig. 4.15 that the norm of the missing transverse energy is not strongly affected while the loss of efficiency for the A mass reconstruction is large. This is due to the mass reconstruction strategy described in Sec. 3.4.2, in which \cancel{E} must lay between transverse direction of $l_A(1, 2)$. Therefore it could be a good al-

²In version 2_2_9. This statement is not true for the earlier versions.

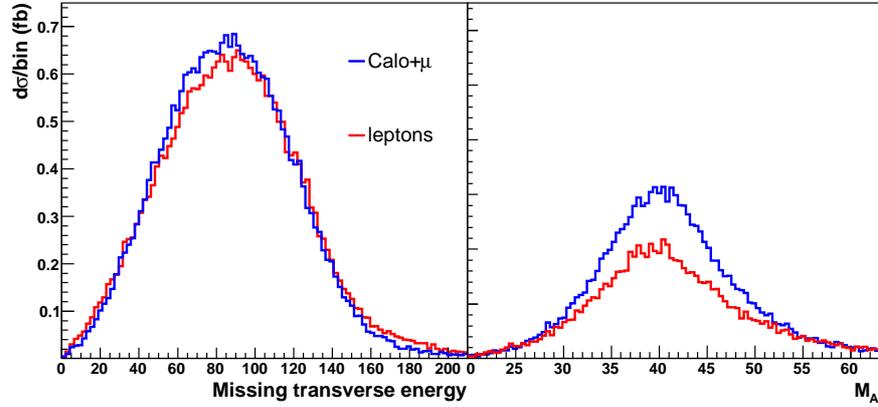


Figure 4.15: Left: The missing transverse energy build from (in blue) the calorimeters+muons and (in red) only the leptons. Right: the same but for the reconstructed mass of the A boson. The events used are from ZA process (BP2).

ternative for the measurement for very early data and provide a cross-check with the value obtained with calo+muon definition.

Flavor and charge

The next condition is that the leptons $l_A(1)$ and $l_A(2)$ have opposite charges and different flavors.

4.4.2 Topological cuts

Acoplanarity

A kinematical characteristic of the ZA process is to produce the Z and the A bosons relatively back-to-back in the transverse plan. The only source of deviation from the perfectly acoplanar situation is due to the presence of initial-state radiation.

This acoplanarity can be measured from the pair issued from the Z and $l_A(1)$ and $l_A(2)$, we show its distribution for the different backgrounds and the signal (BP2) in Fig. 4.14(right). We choose to fix the minimal value of this acoplanarity $\Delta\phi(A, Z)$ to 2.5 in order to keep at least 90% of the signal.

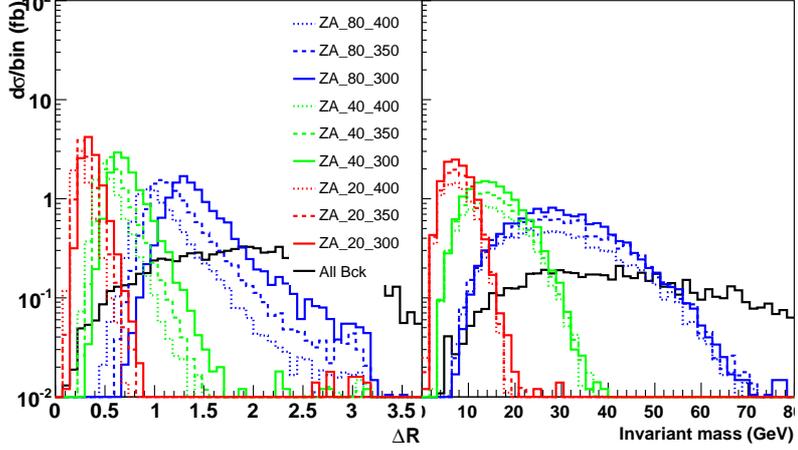


Figure 4.16: The $\Delta R(l_A(1), l_A(2))$ (left) and $\text{Minv}(l_A(1), l_A(2))$ (right) distributions for the signal declined in the nine benchmark points as well as for the sum of backgrounds. These distributions are shown after the application of the trigger, the acceptance, C_Z , the flavor and charge and $\Delta\phi(A, Z)$ conditions.

Lepton collinearity, invariant mass and Z boson energy

The collinearity between $l_A(1)$ and $l_A(2)$ is a strong signature of the signal. Therefore we cut on the distance between these leptons $\Delta R(l_A(1), l_A(2))$, on their invariant mass $\text{Minv}(l_A(1), l_A(2))$. In addition, we also use the energy of the Z boson E_Z as a discriminant variable. Since we want to consider nine mass points in the $(m_A - m_{H^0})$ plane, the three selection cuts should be adapted to each of the kinematical characteristics.

The dependency of $\Delta R(l_A(1), l_A(2))$ and $\text{Minv}(l_A(1), l_A(2))$ is illustrated in Fig. 4.16 where the corresponding distributions are shown for the nine benchmark points, after the application of the trigger, the acceptance, C_Z , the flavor and charge and $\Delta\phi(A, Z)$ conditions. A clear dependence of the distributions with respect to m_A is observed, while the dependence on m_{H^0} is less pronounced. It only shifts the ΔR which can be understood by the variation of the boost of the A boson.

For each of the nine benchmark points different selection windows for $\Delta R(l_A(1), l_A(2))$ and $\text{Minv}(l_A(1), l_A(2))$ would be required. However in first approximation we can restrict our strategy to three choices driven by the value of m_A . For $m_A=20, 40$ and 80 GeV/c^2 , we choose respectively $\Delta R(l_A(1), l_A(2)) \in [0, 0.6], [0, 1.2], [0.6, 2.5]$ and $\text{Minv}(l_A(1), l_A(2))$ to be inside the mass windows $[0, 15], [0, 25]$ and $[15, 50]$ GeV/c^2 .

The energy of the reconstructed Z boson, E_Z , is more sensitive to m_{H^0} . This is shown in Fig. 4.17. Therefore, for each of the three possibilities we set a minimal cut, at 125,

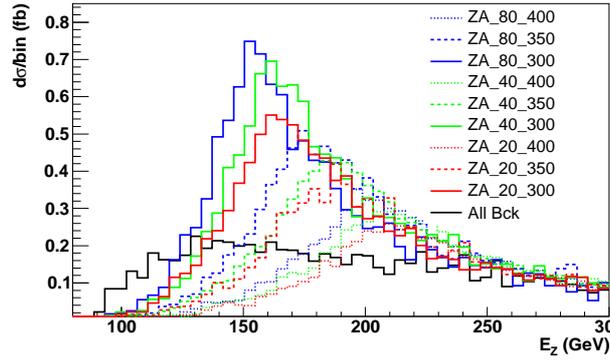


Figure 4.17: The distributions of the energy of the Z boson for the signal (for the nine benchmark points) and for the sum of the background processes. These distributions are shown after the application of the trigger, the acceptance, C_Z , the flavor and charge and $\Delta\phi(A, Z)$ conditions.

150 and 170 GeV/c^2 for $m_{H^0}=300, 350$ and $400 \text{ GeV}/c^2$ respectively.

Reconstruction of m_A and m_{H^0}

As discussed in Sec. 3.4.2, the mass of the two Higgs bosons can be reconstructed. This requires to have the missing transverse energy vector pointing between $l_A(1)$ and $l_A(2)$. This condition is therefore

$$\text{MissIn} \equiv \frac{\Delta\phi(l_A(1), L_A(2))}{\sum \Delta\phi(\vec{E}_T, l_A(i))} = 1. \quad (4.6)$$

For events fulfilling this condition and all the previous cuts, we show in Fig. 4.18 and 4.19 the reconstructed masses for m_A varying from 20 to 80 GeV/c^2 and m_{H^0} from 300 to 400 GeV/c^2 . See Sec.3.4.2 for the details about m_A estimation.

One sees that it is possible to slightly increase the S/B ratio by cutting on both m_A and m_{H^0} . A loose cut on $m_{H^0} < 500 \text{ GeV}/c^2$ is applied while we select three mass windows $m_A^{reco} \in [0, 35] \text{ GeV}$, $m_A^{reco} \in [15, 65] \text{ GeV}/c^2$ and $m_A^{reco} \in [50, 100] \text{ GeV}/c^2$ for $m_A=20, 40$ and $80 \text{ GeV}/c^2$ respectively.

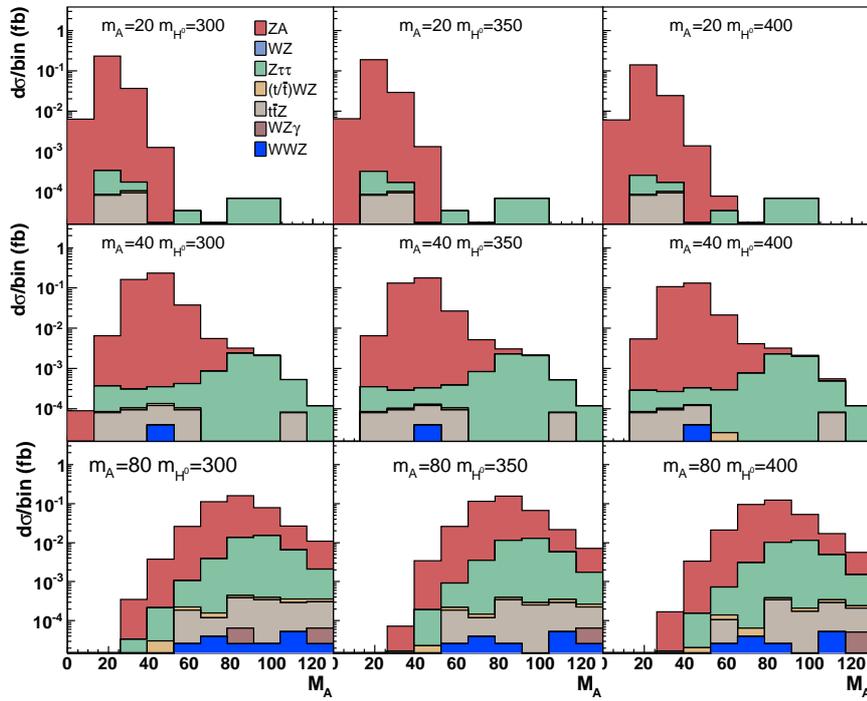


Figure 4.18: The mass of the A boson evaluated at the nine benchmark points, running m_A from 20 to 80 GeV/c^2 and m_{H^0} from 300 to 400 GeV/c^2 . The background Zjj and $Zb\bar{b}$ are not labelled in the legend as none event contribute at the level of Higgs boson mass reconstruction.

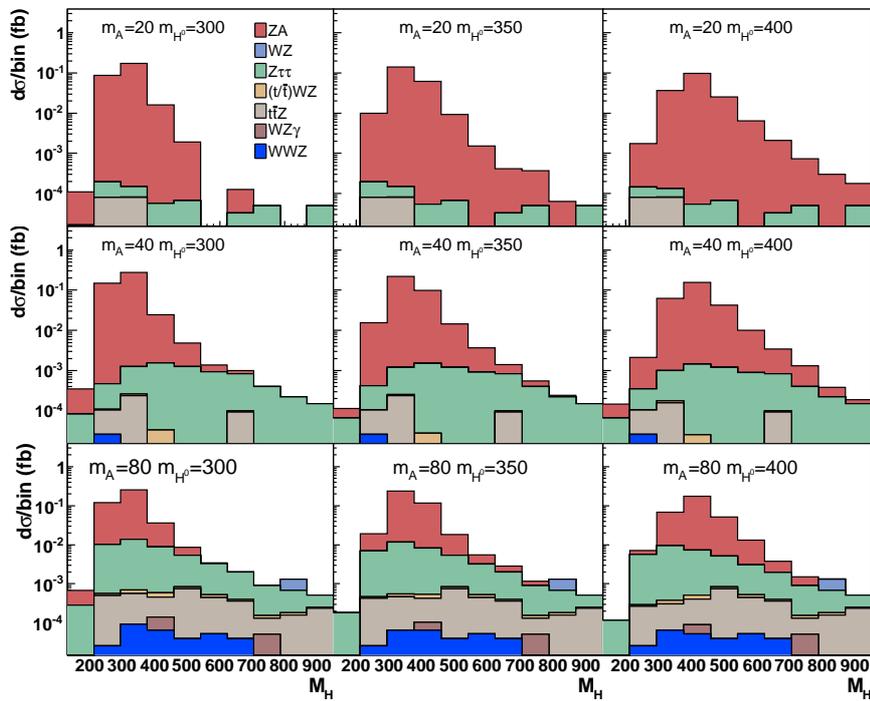


Figure 4.19: The mass of the H^0 boson evaluated at the nine benchmark points, running m_A from 20 to 80 GeV/c^2 and m_{H^0} from 300 to 400 GeV/c^2 . The background Zjj and $Zb\bar{b}$ are not labelled in the legend as none event contribute at the level of Higgs boson mass reconstruction

Summary tables

To summarize what is described here, there are nine different configurations of cuts, as illustrated in Tab. 4.4. For the sake of simplicity we use in the rest the nomenclature xX under quotation marks, where x is l, m, h and X is L, M, H , referring to the low, mean or high cuts for the A boson and the H^0 boson respectively.

| | | m_{H^0} (GeV/ c^2) | | | | | |
|---------------------|----|---|------|--|------|---|------|
| | | 300 | 350 | 400 | | | |
| m_A (GeV/ c^2) | 20 | $\Delta R \in [0, 0.6]$ Minv $\in [0, 15]$ $E_Z > 125$ $m_A^{reco} \in [0, 35]$ | "lL" | $\Delta R \in [0, 0.6]$ Minv $\in [0, 15]$ $E_Z > 150$ $m_A^{reco} \in [0, 35]$ | "lM" | $\Delta R \in [0, 0.6]$ Minv $\in [0, 15]$ $E_Z > 170$ $m_A^{reco} \in [0, 35]$ | "lH" |
| | 40 | $\Delta R \in [0, 1.2]$ Minv $\in [0, 25]$ $E_Z > 125$ $m_A^{reco} \in [15, 65]$ | "mL" | $\Delta R \in [0, 1.2]$ Minv $\in [0, 25]$ $E_Z > 150$ $m_A^{reco} \in [15, 65]$ | "mM" | $\Delta R \in [0, 1.2]$ Minv $\in [0, 25]$ $E_Z > 170$ $m_A^{reco} \in [15, 65]$ | "mH" |
| | 80 | $\Delta R \in [0.6, 2.5]$ Minv $\in [15, 50]$ $E_Z > 125$ $m_A^{reco} \in [50, 100]$ | "hL" | $\Delta R \in [0.6, 2.5]$ Minv $\in [15, 150]$ $E_Z > 150$ $m_A^{reco} \in [50, 100]$ | "hM" | $\Delta R \in [0.6, 2.5]$ Minv $\in [15, 50]$ $E_Z > 170$ $m_A^{reco} \in [50, 100]$ | "hH" |

Table 4.4: Set of cut used to cover the nine benchmark points of the ZA analysis. ΔR is the distance between the collinear leptons, Minv is their invariant mass (GeV) and E_Z is the energy of the Z boson (GeV). The letters under quotation marks is the label of the cut set.

The impact of the different cut-based selections is quantified in Tab. 4.5 for the signal, Tab. 4.6 for the reducible backgrounds and Tab. 4.7 for the irreducible backgrounds. In the first and third case, the relative efficiencies between each step of the analysis, the total efficiency and visible cross-section are displayed. Furthermore the number of remaining simulated events after all cuts is also given in order to quantify the associated statistical error. For the reducible background, the number of simulated events is given for each analysis step, the total visible cross-section is also provided. As a reminder, the total number of simulated events used for each process is given in Tab. 4.1.

Note that in these tables we group the three cuts on $\Delta R(l_A(1), l_A(2))$, Minv($l_A(1), l_A(2)$) and E_Z under the appellation D-M-E (distance-mass-energy). The tables are divided into two parts: the first (from Trigger to $\Delta\phi(A, Z)$) contains the selection conditions independent of the benchmark points while the second is dependent. There are therefore nice visible cross-sections for each process, one per masses choice.

For the signal, we see that in the mass-independent part, most of the efficiencies are similar, except the acceptance which is dependent on the boost of the final states. In

the second part (mass dependent), the cut `MissIn` affects the signal efficiency differently, depending on M_A . This is due to the fact that the probability that \cancel{E}_T lays between $l_A(1)$ and $l_A(2)$ is strongly dependent of the boost of the A boson. For the backgrounds the following observations are worth mentioning:

- There is a sensitive difference between the Trigger efficiencies for the signals and the background. We have seen in Sec. 4.3.1 that the Trigger patterns taken into account are related to the presence of the leptons from the Z boson. In the signal, the presence of H^0 boosts the system and hence increase the probability to fulfill the conditions imposed by the P_T thresholds.
- As expected, the rejection of the reducible backgrounds by the acceptance condition is very good. It varies between the percent for $WZ\gamma$ to 10^{-5} for Zjj . The Zjj process and $Zb\bar{b}$ are also strongly reduced by the cut imposed on the missing transverse energy, as seen in Fig. 4.14 (left).
- It is interesting to notice the effect of the cut on m_A for the $Z\tau\tau$ process. Indeed, for low m_A and high m_A cuts, the efficiency is higher than for the middle case. This is due to the fact that the production of this process is dominated by the $Z\gamma$ and ZZ processes at high and low invariant mass of $\tau\tau$ respectively.

| | $ZA(20, 300)$ | $ZA(20, 350)$ | $ZA(20, 400)$ | $ZA(40, 300)$ | $ZA(40, 350)$ | $ZA(40, 400)$ | $ZA(80, 300)$ | $ZA(80, 350)$ | $ZA(80, 400)$ |
|---------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Trigger | 80 | 82 | 83 | 81 | 83 | 84 | 82 | 84 | 85 |
| Acc. | 21 | 23 | 26 | 24 | 28 | 30 | 27 | 30 | 33 |
| C_Z | 91 | 91 | 90 | 91 | 86 | 91 | 90 | 91 | 91 |
| \cancel{E}_T | 74 | 83 | 89 | 77 | 86 | 90 | 70 | 82 | 89 |
| $\neq flav.$ | 99 | 99 | 100 | 99 | 100 | 100 | 100 | 100 | 100 |
| $\neq ch.$ | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| $\Delta\phi(A,Z)$ | 92 | 93 | 94 | 92 | 93 | 94 | 85 | 90 | 92 |
| | IL | IM | IH | mL | mM | mH | hL | hM | hH |
| D-M-E | 94 | 93 | 92 | 93 | 90 | 89 | 85 | 81 | 78 |
| MissIn | 41 | 40 | 40 | 61 | 59 | 59 | 69 | 69 | 69 |
| $m_{A,H}$ | 99 | 97 | 89 | 98 | 97 | 90 | 87 | 88 | 83 |
| Eff. Tot. | 4 | 4.8 | 5.2 | 6.8 | 8 | 9 | 5.9 | 8.2 | 9.2 |
| σ_{vis} (fb) | 0.276 | 0.216 | 0.15 | 0.44 | 0.34 | 0.24 | 0.33 | 0.32 | 0.24 |
| #MC events | 2.9e3 | 3.4e3 | 3.5e3 | 5e3 | 2.8e4 | 6.1e3 | 4.1e3 | 5.8e3 | 6.2e3 |

Table 4.5: The selection efficiencies (in %) of the signal for the different benchmark points. The corresponding visible cross-sections and number of simulated events (MC events) are also shown.

| | | Zjj | | | Zbb | | | WZ | | | $WZ\gamma$ | | |
|---------|---------------------------|---------|---|---|---------|---|---|---------|---|---|------------|---|---|
| | Trigger | 8.6e+05 | | | 3.4e+05 | | | 7.1e+05 | | | 6.5e+04 | | |
| | Acc. | 5 | | | 37 | | | 7.4e+02 | | | 7.4e+02 | | |
| | C_Z | 3 | | | 29 | | | 5.5e+02 | | | 6.3e+02 | | |
| | \cancel{E}_T | 0 | | | 2 | | | 3e+02 | | | 3.6e+02 | | |
| | $\neq \text{flav.}$ | | | | 2 | | | 2e+02 | | | 2.9e+02 | | |
| | $\neq \text{ch.}$ | | | | 2 | | | 1.3e+02 | | | 1.4e+02 | | |
| | $\Delta\phi(A,Z)$ | | | | 2 | | | 56 | | | 60 | | |
| | | L | M | H | L | M | H | L | M | H | L | M | H |
| low(A) | D-M-E | | | | 0 | 0 | 0 | 2 | 2 | 2 | 1 | 1 | 1 |
| | MissIn | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 |
| | $m_{A,H}$ | | | | | | | | | | | | |
| | $\sigma_{vis}(\text{fb})$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mid(A) | D-M-E | | | | 0 | 0 | 0 | 5 | 4 | 4 | 3 | 2 | 2 |
| | MissIn | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 |
| | $m_{A,H}$ | | | | | | | | | | | | |
| | $\sigma_{vis}(\text{fb})$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| high(A) | D-M-E | | | | 0 | 0 | 0 | 10 | 6 | 5 | 8 | 6 | 6 |
| | MissIn | | | | | | | 1 | 1 | 1 | 3 | 2 | 2 |
| | $m_{A,H}$ | | | | | | | 0 | 0 | 0 | 1 | 0 | 0 |
| | $\sigma_{vis}(\text{fb})$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3.84e-06 | 0 | 0 |

Table 4.6: The number of simulated events remaining after each cut of the analysis, as well as the equivalent visible cross-section (in fb) for the reducible backgrounds. The results are divided into two parts: the first (from Trigger to $\Delta\phi(A, Z)$) contains the selection conditions independent of the benchmark points while the second is dependent.

| | | $Z\tau\tau$ | | | $t\bar{t}Z$ | | | tWZ | | | WWZ | | |
|---------|---------------------|-------------|---------|---------|-------------|---------|---------|---------|----------|---------|----------|----------|----------|
| | Trigger | 53 | | | 76 | | | 80 | | | 72 | | |
| | Acc. | 15 | | | 25 | | | 29 | | | 25 | | |
| | C_Z | 71 | | | 69 | | | 69 | | | 69 | | |
| | \cancel{E}_T | 34 | | | 78 | | | 75 | | | 75 | | |
| | $\neq flav.$ | 67 | | | 67 | | | 67 | | | 67 | | |
| | $\neq ch.$ | 99 | | | 99 | | | 99 | | | 99 | | |
| | $\Delta\phi(A,Z)$ | 70 | | | 36 | | | 34 | | | 38 | | |
| | | L | M | H | L | M | H | L | M | H | L | M | H |
| low(A) | D-M-E | 1.1 | 1.1 | 0.99 | 1.1 | 0.96 | 0.96 | 1.4 | 1.4 | 1.4 | 1.8 | 1.8 | 1.7 |
| | MissIn | 54 | 54 | 52 | 11 | 13 | 13 | 7.9 | 7.9 | 7.9 | 7.7 | 7.7 | 8.3 |
| | $m_{A,H}$ | 50 | 48 | 41 | 1e+02 | 1e+02 | 1e+02 | 67 | 67 | 67 | 0 | 0 | 0 |
| | Eff.tot. | 0.0028 | 0.0026 | 0.0019 | 0.003 | 0.003 | 0.003 | 0.002 | 0.0015 | 0.0015 | 0 | 0 | 0 |
| | σ_{vis} (fb) | 0.00027 | 0.00025 | 0.00018 | 0.0001 | 0.0001 | 0.0001 | 1e-05 | 7.5e-06 | 7.5e-06 | 0 | 0 | 0 |
| | #MC events | 16 | 15 | 11 | 2 | 2 | 2 | 4 | 4 | 4 | 0 | 0 | 0 |
| mid(A) | D-M-E | 11 | 10 | 9.9 | 4.5 | 4.2 | 3.8 | 4.7 | 4.1 | 3.7 | 6.2 | 6.1 | 6 |
| | MissIn | 74 | 74 | 75 | 7 | 7.7 | 6.8 | 12 | 12 | 13 | 11 | 12 | 12 |
| | $m_{A,H}$ | 11 | 10 | 9.3 | 80 | 80 | 75 | 38 | 33 | 31 | 60 | 60 | 60 |
| | Eff.tot. | 0.0081 | 0.0072 | 0.0062 | 0.0063 | 0.0063 | 0.0048 | 0.006 | 0.004 | 0.004 | 0.001 | 0.001 | 0.001 |
| | σ_{vis} (fb) | 0.00077 | 0.00069 | 0.00059 | 0.00022 | 0.00022 | 0.00017 | 3e-05 | 2.25e-05 | 2e-05 | 3.92e-05 | 3.92e-05 | 3.92e-05 |
| | #MC events | 47 | 42 | 36 | 4 | 4 | 3 | 12 | 9 | 8 | 3 | 3 | 3 |
| high(A) | D-M-E | 68 | 59 | 51 | 15 | 14 | 12 | 16 | 14 | 12 | 15 | 14 | 13 |
| | MissIn | 71 | 71 | 72 | 17 | 18 | 19 | 28 | 28 | 29 | 29 | 29 | 30 |
| | $m_{A,H}$ | 56 | 54 | 53 | 19 | 18 | 14 | 25 | 23 | 21 | 16 | 14 | 14 |
| | Eff.tot. | 0.24 | 0.2 | 0.17 | 0.013 | 0.011 | 0.0079 | 0.03 | 0.025 | 0.021 | 0.02 | 0.01 | 0.01 |
| | σ_{vis} (fb) | 0.023 | 0.019 | 0.016 | 0.0004 | 0.0004 | 0.0003 | 0.00015 | 0.00012 | 0.00010 | 6.4e-05 | 5.2e-05 | 5.2e-05 |
| | #MC events | 1.4e+03 | 1.2e+03 | 1e+03 | 8 | 7 | 5 | 60 | 49 | 41 | 5 | 4 | 4 |

Table 4.7: The selection efficiencies (in %) of the irreducible backgrounds for the different benchmark points. The corresponding visible cross-sections and number of simulated events (MC events) are also shown.

4.4.3 Probing a possible deviation from the Standard Model

As shown in the Tab. 4.5, 4.6 and 4.7, the visible cross-section for the background is very low compared to the signals. This clearly means that in the very first fb^{-1} of integrated luminosity, any event appearing in the data could be the sign of the presence of “something else” beyond the Standard Model. It is useful to check how much luminosity has to be considered to reject the Standard Model option. To do so we use the S_{CP} estimator defined as the following:

$$\frac{1}{\sqrt{2\pi}} \int_{S_{CP}}^{\infty} e^{-x^2/2} dx = \sum_{i=\mu_s+\mu_b}^{\infty} \frac{\mu_b^i e^{-\mu_b}}{i!} \quad (4.7)$$

The right hand-side term in the relation (4.7), gives the probability to have at least $\mu_{tot} = \mu_s + \mu_b$ events, assuming that the background provides μ_b on average. This estimation is useful when the total number of events is low, and when the gaussian approximation for statistical fluctuation is not valid. The integral term translates this probability as a number of standard deviation for a normal gaussian distribution.

We see in Fig. 4.20 that the set of rectangular cuts depicted in this section allows to estimate the luminosity required to see a deviation from SM expectation. The 95 % C.L. limits shows that a luminosity between roughly 2 and 8 fb^{-1} would be required, depending on the mass point.

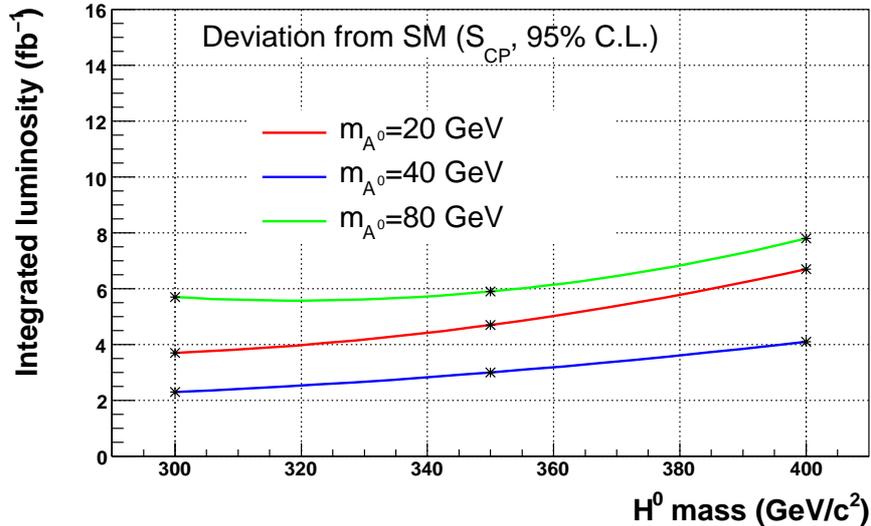


Figure 4.20: The 95 % C.L. limits for a deviation from SM calculated from S_{CP} estimator, for the nine Higgs boson mass points.

Despite the fact that the systematic uncertainties are not taken into account, it confirms the very high discovery potential for the ZA channel as already highlighted in Sec. 3.4.2 and in Ref. [21].

4.5 Data-driven estimation of the reducible backgrounds

As it can be seen in the previous section, the S/B ratio fluctuates with the Higgs mass spectrum but remains excellent for all mass points. Nevertheless, even if the contribution of the WZ +jets, Z +jets and $Zb\bar{b}$ processes seems to be very low, large disagreement between prediction and experimental data in the selected phase-space could be observed. This can be a real issue since their respective cross-sections are very large compared to the one of the signal. The differences that could be observed may come for example from the difficulty to simulate the very few jets identified as electrons, or from a large error on the theoretical prediction of cross-sections.

One way to handle this issue, is to control *in situ* the contribution of dangerous backgrounds directly from data. This is achieved by selecting a phase-space region (not too far from the one defined in the analysis cuts) where the background processes alone contribute. From the total event count in this phase-space region the corresponding background contribution can be extrapolated in the region of interest for the signal selection. In this analysis, the two considered phase-space regions rely on key variables used for the electrons identification and their isolation which must be properly described by the simulation. Another important quantity is the missing transverse energy. The cuts applied to define these two regions are presented in the flowchart of Fig. 4.21.

The “nominal analysis” cuts have been presented in the previous sections, and their impact on signal and background is summarized in Tab. 4.5, 4.6 and 4.7. For the “control region”, we first define “loose acceptance” criteria by requiring no cut on iso_e , $(H/E)^N$, $(E_s/P_{out})^N$, $(E_{SC}/P)^N$ and $(\Delta R(track, SC))^N$. Concerning the muon isolation, we assume that it will be well known before the first 1 fb^{-1} , we therefore keep the cut iso_μ at 0.9 in order to not overpopulate the events by Z +jets background. In addition, the presence of new backgrounds is avoided by requiring C_Z cut with the “tight acceptance” criteria for electrons reconstructing the Z boson.

In order to control the contribution of the reducible backgrounds with 1 fb^{-1} of integrated luminosity, we define the “no cut scheme” by requiring only $\cancel{E}_T > 10 \text{ GeV}$ and different flavors for the two leptons not coming from the Z boson. At this stage,

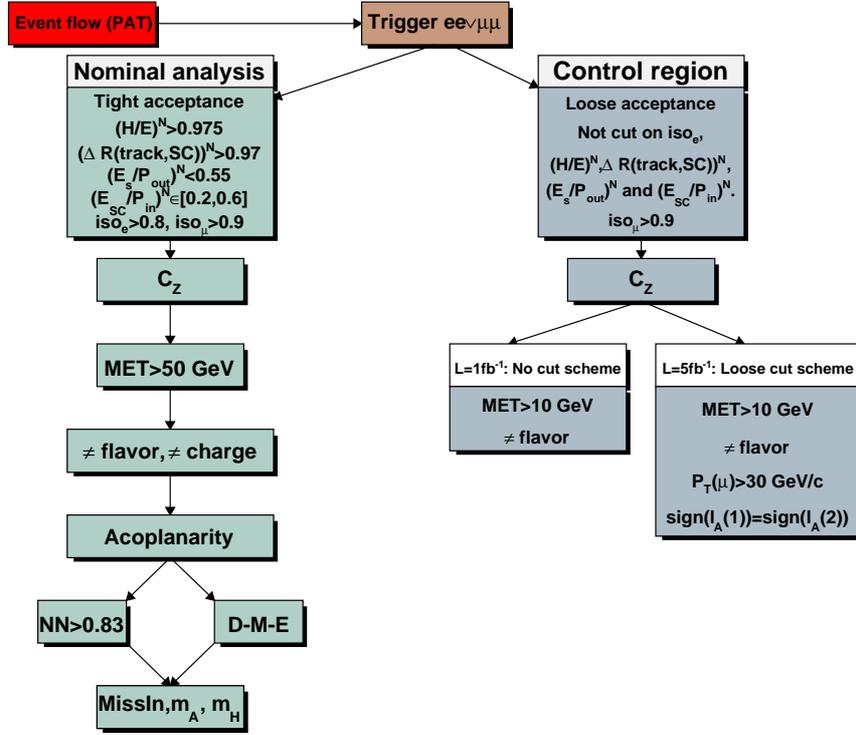


Figure 4.21: Flowchart that illustrates the data flow through either the nominal analysis (left), or through the background normalization region (right). When needed, the boxes contain the details of cut considered.

the P_T spectrum of the muon (not coming from the Z boson) can be used to validate the "No cut scheme" selection. This distribution is shown (after C_Z cut) in Fig. 4.22 (upper left) and is largely dominated by contributions from the three reducible backgrounds. Visible cross-sections of the signal and irreducible background processes are much lower so they do not have any significant impact. The key variable related to electrons identification and isolation could also be obtained. The extrapolation of reducible backgrounds from this control region is discussed in Sec. 4.5.1.

With an integrated luminosity of 5 fb^{-1} , the statistics become sufficient to improve the reconstruction of key variables, notably by suppressing the contribution of $Zb\bar{b}$ background process. To do so we define the "Loose cut scheme" where $P_T(\mu) > 30 \text{ GeV}/c$ is applied on the muon which does not come from the Z boson. Furthermore, this muon and the remaining electron are required to have the same electric charge in order to

remove the contribution of the signal and irreducible backgrounds. The discussion about possible control plots of key variables is presented in Sec. 4.5.2.

4.5.1 Normalization

With 1 fb^{-1} of integrated luminosity, the goal is first to estimate the number $N_{\text{final}}^{\text{analysis}}$ of reducible background events after all cuts in the analysis stream. This number can be written as

$$N_{\text{final}}^{\text{analysis}} = f R_{A/C} N^{\text{control}}. \quad (4.8)$$

The factor N^{control} is the number of reducible background events observed in the control stream using the “no cut scheme” and observed via the muon P_T spectrum of Fig. 4.22 (upper left). It reads:

$$N^{\text{control}} = \mathcal{L}(\sigma_{WZ,vis}^{\text{control}} + \sigma_{Zjj,vis}^{\text{control}} + \sigma_{Zbb,vis}^{\text{control}}), \quad (4.9)$$

where $\sigma_{WZ,vis}^{\text{control}} \sim 33 \text{ fb}$, $\sigma_{Zjj,vis}^{\text{control}} \sim 57 \text{ fb}$ and $\sigma_{Zbb,vis}^{\text{control}} \sim 16 \text{ fb}$ are the corresponding visible cross-sections and \mathcal{L} is the luminosity. In this discussion we consider the sum of all reducible backgrounds instead of their individual contributions. With $\mathcal{L} = 1 \text{ fb}^{-1}$, $N^{\text{control}} \sim 110$, meaning that its relative statistical error $\delta_{N^{\text{control}}}$ is close to 10%.

The ratio $R_{A/C}$ is defined by the following expression:

$$R_{A/C} = \frac{N_{\text{acceptance}}^{\text{analysis}}}{N^{\text{control}}}, \quad (4.10)$$

where $N_{\text{acceptance}}^{\text{analysis}} = \mathcal{L}(\sigma_{WZ,vis}^{\text{acc.}} + \sigma_{Zjj,vis}^{\text{acc.}} + \sigma_{Zbb,vis}^{\text{acc.}})$, with the corresponding visible cross-sections quantified from the “Acceptance” line in Tab. 4.6. This ratio $R_{A/C}$ is found to be close to 1/60. Its associated statistical error is large as the number of simulated events for reducible backgrounds passing the acceptance in the analysis stream is low. However an increase of the statistics is always possible, and is not expected to be a problem when the collection of data will have reached 1 fb^{-1} . The relative error $\delta_{R_{A/C}}$ is therefore largely driven by the systematic uncertainty related to the key variables for electrons identification and isolation. From the discussion in Sec. 4.6.2 it is expected to be smaller than 20 %.

Finally, the factor f is defined as the ratio of the number of simulated events passing all cuts to the number of simulated events passing the acceptance cuts ($N_{\text{acceptance}}^{\text{analysis}}$). This number cannot be estimated accurately with the present MC sample since the number of simulated events of reducible backgrounds is reduced to zero after all cuts. Again, this could be solved with a longer generation time. However at this stage events

passing the acceptance and analysis cuts are signal-like and are therefore expected to be submitted to similar systematic uncertainties related to reconstruction effects. In Sec. 4.6.2, we show that this relative error is close to 10 %.

In conclusion, the relative error associated to $N_{final}^{analysis}$ defined as

$$\delta_{N_{final}^{analysis}} = \sqrt{\delta_f^2 + \delta_{R_{A/C}}^2 + \delta_{N^{control}}^2}, \quad (4.11)$$

is found to be 25 % and is dominated by the systematic uncertainty related to the definition and isolation of electrons.

4.5.2 Control plots of key variables

With a larger integrated luminosity of 5 fb^{-1} and using the ‘‘Loose cut scheme’’, an estimation of the distributions of key variables is expected to be possible by looking to WZ +jets and Zjj events.

The missing transverse energy and the relevant variables related to the electron definition and isolation are shown in Fig. 4.22. The data-driven evaluation of key variables for electrons identification and isolation is obtained with the electron not assigned to the Z boson. At this stage a large deviation in the shapes between the data and the Monte Carlo could probably be seen, and hopefully corrected accordingly.

Based on these distributions, a set of cut can be established to reject the Zjj background completely:

- $\cancel{E}_T > 30 \text{ GeV}$
- $1/(1 + H/E) > 0.975$
- $1/(1 + \Delta R(track, SC)) > 0.97$
- $1/(1 + E_s/P_{out}) < 0.55$
- $1/(1 + E_{SC}/P) \in [0.2, 0.6]$.

The electron isolation iso_e is shown with these cuts, in Fig. 4.23. It turns out that after 5 fb^{-1} , the number of expected reducible background events would be close to 1 while the number of signal events in the phase-space region selected in the main analysis is also close to the unity for the same integrated luminosity.

As we have seen in Sec. 4.4.3, the critical period to find traces of the $iM2HDM$ in Type II and therefore reject the SM hypothesis will require luminosities between 2

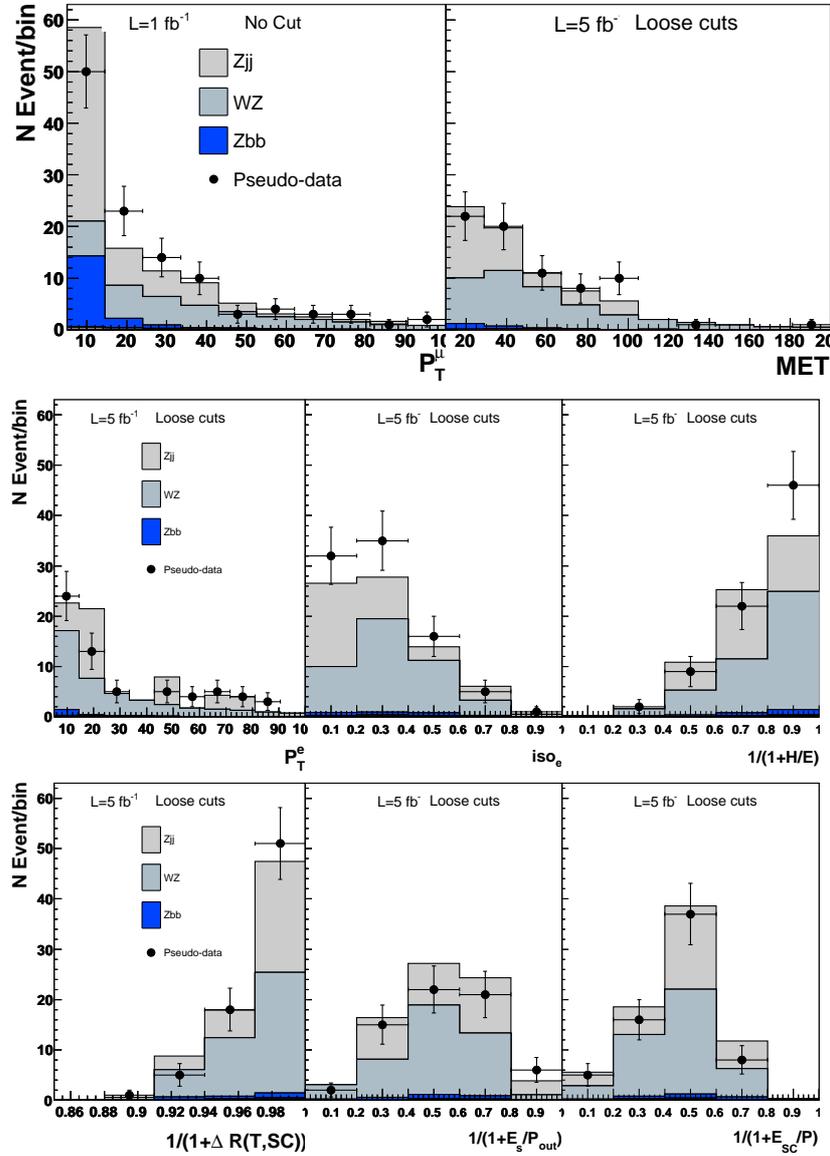


Figure 4.22: Upper left: the P_T distribution of the muon (not from Z) for WZ , $Zb\bar{b}$ and Zjj in the “no cut scheme”. Upper right: the \cancel{E}_T distribution in the “loose cut scheme”. From middle-left to lower-right, the key variables used for the identification and isolation of electrons in the same scheme. They are obtained with the electron not assigned to the Z boson. Each pseudo-data is a random fluctuation of the bin content following a Poisson distribution.

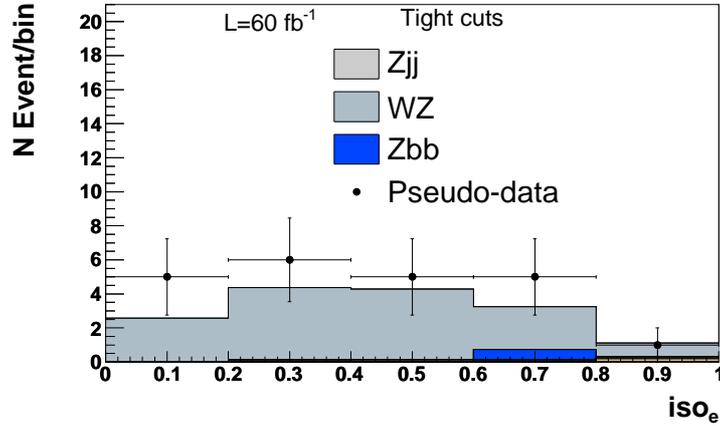


Figure 4.23: The quantity iso_e , assuming the tight cuts on the missing transverse energy and electrons identification variables.

and 8 fb^{-1} . It turns out after these luminosities, the cross-section of the most dangerous background can be normalized and experimentally understood, using sensitive variables used to reduce it. The present method has therefore the virtue to put a limit on the contribution of these background in the ZA analysis.

4.6 Analysis of the benchmark point BP2

In the next sections we restrict our discussion to the BP2 mass point ($m_A=40 \text{ GeV}/c^2$ and $m_{H^0}=350 \text{ GeV}/c^2$) since it corresponds to the average masse of the various benchmark points, and it is close to be the most promising one. Furthermore, we have seen that the predominant background process is $Z\tau\tau$, we therefore concentrate mostly on this process and the signal to test a neural net based selection and discuss the systematic uncertainties.

4.6.1 Neural net based selection

In Sec. 4.4.2, the results are based on a series of rectangular cut on each discriminant variable. This means that the left region is an hypercube in the variable space, which is not necessarily the most optimal selection to reject the background processes. Given the correlation between these variables, a good alternative is to make use of a neural network that rather selects a region which is an hyper-surface in the variable space.

The structure of a multi-layer perceptron neural network contains a set of input neurons, each in connection with a variable, a given number of hidden layers containing neurons and an output neuron providing the variable on which a selection can be achieved. In addition, each neuron of each layer is connected to all neurons of nearby layer(s). These connections are characterized by a weight w .

To fix the idea we consider only one hidden layer. For each event, the input neurons get a value and each neuron of the hidden layer combines these values linearly, according to the weight of each connections. This weighted sum is used as input of a sigmoid function $S(x) = 1/(1 + e^{-x})$. In other words, the output of each hidden neuron is

$$a_j^k = S\left(\sum_{i=0}^N w_{ij} a_i^{k-1}\right) \quad (4.12)$$

where a_q^p is the value taken by the q -th neuron of the p -th layer, S is a sigmoid function and w_{ij} is the weight of the link between the neuron i and j . The output defines whether the event is signal or background like.

The response of the neural net is obviously dependent of the weight. The training phase is dedicated to their adjustment so that the distance between the hyper-surface and each of the event is minimized. Different error minimization exists, we use the simplest one, which is the stochastic minimization [90]. We have seen that for most of mass points considered, the dominant background is $Z\tau\tau$. The training of the neural net, *i.e.* the learning step that helps to determine what is signal-like and what is background-like is therefore based on this process and the signal.

We propose to see how the use of a simple neural net class³ can improve the S/B ratio. The input is composed by the normalized variables $\Delta R^N = 1/(1 + \Delta R(l_A^1, l_A^2))$, the invariant mass term $(\text{Minv})^N = 1/(1 + \text{Minv}(l_A^1, l_A^2))$ and $(E_Z)^N = 1/(1 + E_Z)$. The distributions of these variables are shown in Fig. 4.24 for both the signal and $Z\tau\tau$ processes.

To each event corresponds a value

$$v = F(\Delta R^N, \text{Minv}^N, E_Z^N)$$

with F the output function on which a cut determines if an event is considered either as a "signal" or "background" event.

The training is realized on roughly 30000 events of both signal and background and uses a structure with 3 neurons in input and 4 neurons in the hidden layer. The neural net response on both signal and background is shown in Fig. 4.25. The separation

³MultiLayerPerceptron (see for example [91])

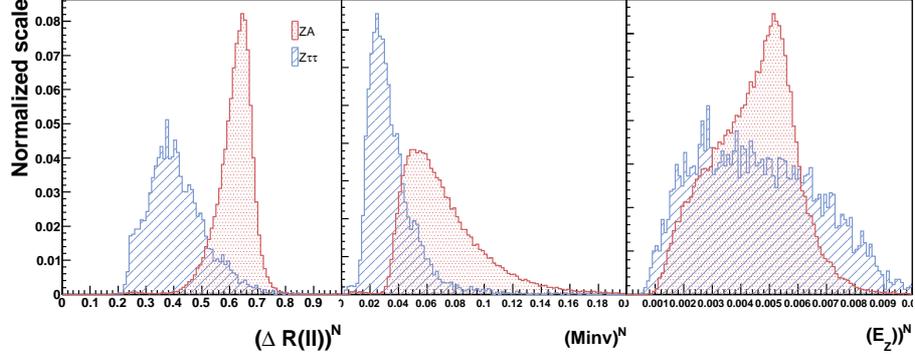


Figure 4.24: The distribution of the normalized variables ΔR^N , $(\text{Minv})^N$ and $(E_Z)^N$ used as input of the neural network. Each distribution is normalized to 1.

between the two processes is very clear. Beside a good discriminant power, the main advantage of such a method is its stability with respect to systematical uncertainties by cutting in a rather flat region.

This distribution can also be used to control the level of (essentially) irreducible backgrounds. To illustrate this point, we show in Fig. 4.26 the neural net response for an integrated luminosity of 5 fb^{-1} , after the acoplanarity cut. The position of this neural net cut is shown (NN) in the analysis stream in Fig. 4.21. The right hand-side of the distribution is almost completely filled by the signal whereas the left hand-side contains only the different irreducible backgrounds and also a small contribution of WZ +jets

Another interesting observation is the relative insensitivity of the neural net response to the light pseudoscalar Higgs mass. As seen in the left hand-side plot of Fig. 4.27, the response of the neural net for events with $m_A \neq 40 \text{ GeV}/c^2$ is similar to the response for events from BP2. For much larger masses, this does not hold, and would require training with other new masses. This said, the reconstruction of the mass peak is very good for $m_A = 20$ to $50 \text{ GeV}/c^2$ and remains acceptable for higher masses.

In the context of our analysis, we require the neural network output to be above 0.83 in order to have the same efficiency on the signal as the one obtained with the rectangular cuts analysis. Compared to the cut-based analysis, the neural net technique is 10 % more efficient. However, translated in the S_{CP} significance, the gain is negligible.

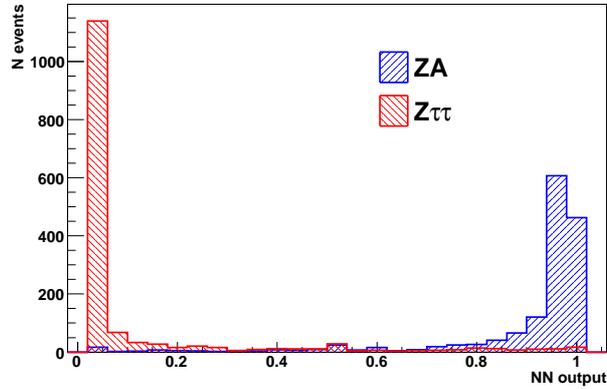


Figure 4.25: Output from the 3:4:1 neural network. The small values indicates that the event is $Z\tau\tau$ -like while the values approaching 1 are signal-like.

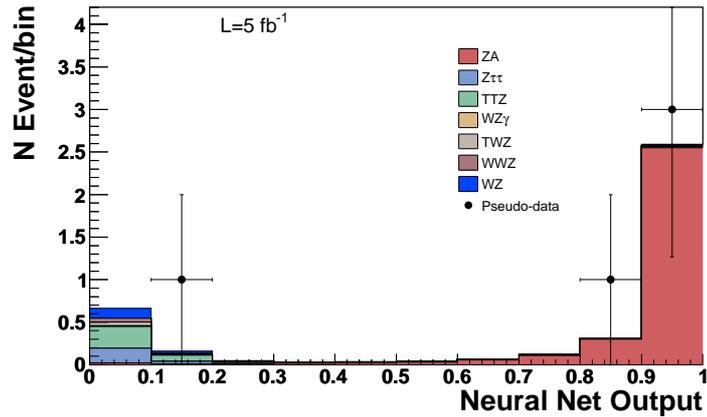


Figure 4.26: The neural network output used as a controller for irreducible backgrounds. The number of events correspond to an integrated luminosity of 5 fb^{-1} .

4.6.2 Systematics

We discuss here the main sources of systematic uncertainties and evaluate their impact on the analysis. As an example we consider the neural-network based algorithm. As the backgrounds are largely dominated by $Z\tau\tau$ process after all cuts, we therefore consider the effects of possible uncertainties on this process and on the signal only.

At the time of writing this text, the number of simulated $Z\tau\tau$ events passing the analysis cuts is rather low (45 whereas for the signal this number approaches 30000) what

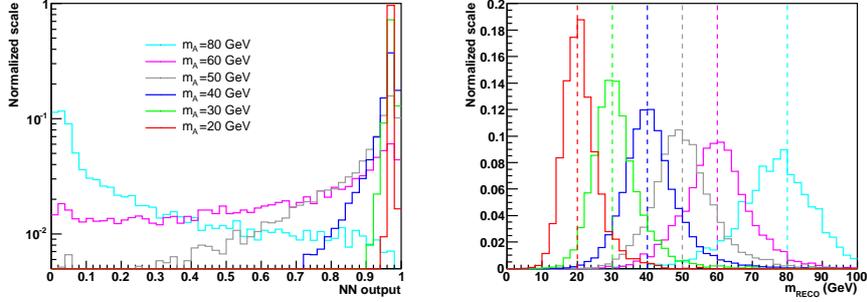


Figure 4.27: Left: neural net output for different signal mass points (always with $m_{H^0}=350$ GeV/ c^2), varying m_A from 20 to 80 GeV/ c^2 . Right: the reconstruction of the theoretical mass when passing through the neural net trained with $m_A=40$ GeV/ c^2 . The vertical dotted lines indicate the respective theoretical masses.

gives a statistical error close to 15 %. The evaluation of the systematic uncertainties related to $Z\tau\tau$ process would require a much larger statistics, which has not been done in the context of the present work. We therefore choose to neglect the systematic errors related to the detector response and final states reconstruction for this process, which is a reasonable statement since the visible cross-section of this background is two orders of magnitude lower than the one of the signal.

ECAL and HCAL resolutions and energy scale

The uncertainty related to the calorimeters comes from both the resolution effect and the energy scale determination.

For ECAL, we choose to adopt the initial miscalibration condition, *i.e.* based on both cosmic (barrel) and test beams for the measurement of crystal light yields and photo-detector gains (endcaps). This is a pessimistic estimation since the calibration should be better at the time of potential discovery, *i.e.* with an integrated luminosity of 5 fb $^{-1}$.

For the barrel, σ_E/E evolves from 1.5% with $|\eta| < 1$ to 2.2% with increasing $|\eta|$ beyond 1. For the endcap, the resolution is fixed to 15%. Concerning HCAL, the σ_E/E is chosen to be 5 % as recommended by the CMS collaboration[92]. It turns out that the resolution affects the determination of the missing transverse energy direction, which lowers the signal efficiency is by 4 %.

The calibration of the calorimeter energy scale is the other main source of systematic error on the measurement of the missing transverse energy. We assume the extreme

situation where the offset is 10 % towards higher and lower values. If the offset is +10 % the efficiency of the algorithm is similar to the case where no offset is present. On the contrary, it decreases by roughly 6 % for the negative offset.

Lepton acceptance and reconstruction

After an integrated luminosity of 5 fb^{-1} the systematic error affecting the transverse momentum evaluation from the tracker is supposed to be very low, close to 2 % in dP_T/P_T [69]. We have not observed a significant effect on the global efficiencies is observed.

We however take into account a 1% systematic uncertainty for the lepton reconstruction efficiency [69]. As we have four reconstructed leptons a total of 4 % uncertainty is considered.

Luminosity and cross-sections

The luminosity measurement can be made from the evaluation of W/Z production rates as these processes have a very large cross section and a clean signature. The associated systematic uncertainty is expected to be 5 % [75].

The cross-section uncertainties for the signal and the background are based on the assumed NLO calculations. For the signal, it is estimated to be roughly 10 % (see Ref. [93]) whereas for the background these uncertainties are rather large. However, if the $Z\tau\tau$ cross-section estimation is derived from the data driven measurement of ZZ cross-section, the uncertainty could be strongly reduced [94]. We take 5 % as a conservative estimation.

The table 4.8 summarizes the different sources of systematic uncertainties as well as their respective values (in percent) for the signal and $Z\tau\tau$ processes.

This shows that with an pessimistic approach for the evaluation of detector systematic uncertainties, the global relative error remains smaller than roughly 15 %. Since the level of background is expected to be very low, this implies the same variation of luminosity to reach the 95 % C.L. to reject the pure SM supposition.

Systematic errors from reducible backgrounds rejection

It is shown in Sec. 4.5 that the normalization of reducible backgrounds visible cross-sections is possible after a few inverse femtobarn. Furthermore, the reconstruction of

| Syst. error | $Z A$ | $Z \tau \tau$ |
|------------------|--|---------------|
| Resol. calo. | -4 | - |
| Offset calo. +10 | 0 | - |
| Offset calo. -10 | -6 | - |
| Lepton acc. | ± 4 | ± 4 |
| Luminosity | ± 5 | ± 4 |
| cross-section | ± 10 | ± 5 |
| quad. sum | $\begin{matrix} +12.5 \\ -13.8 \end{matrix}$ | ± 7.5 |

Table 4.8: Sources of the systematic uncertainties and their respective values (in percent) for the signal and $Z\tau\tau$ processes. The quadratic sum of all systematic errors is also shown.

the sensitive variables \cancel{E}_T , $1/(1+H/E)$, $1/(1+\Delta R(track, SC))$, $1/(1+E_s/P_{out})$, $1/(1+E_{SC}/P)$ and iso_e can also be achieved, even if requires a larger luminosity. These distributions are susceptible to vary with respect to the sub-detectors calibration and therefore affect the rejection of the reducible background processes.

In order to estimate this effect, the variation of rejection efficiency has been evaluated on WZ +jets process by changing the ECAL and HCAL energy scale.

With and offset of +10%, the selection efficiency resulting from the five cuts does not vary with respect to the ideal detector configuration considered in the analysis. On the contrary, a decrease of 20% is observed when the detectors gain is lowered by 10 %.

This estimation shows that the presence of events from reducible backgrounds at a given luminosity is well under control. This is a other argument to claim that any event seen after a few inverse femtobarn after all analysis cut is a sign for BSM.

4.6.3 An estimation of the significance and mass resolution.

We have shown in Sec. 4.4.3 that a deviation from the SM expectation could take place with an integrated luminosity of the order of 5 fb^{-1} . To evaluate the significance, *i.e.* the confidence associated to the signal-plus-background hypothesis with such a luminosity, we follow the recommendation of CMS and consider the estimator S_{L2} . It is defined in Ref. [95] as

$$S_{L2} = \sqrt{2 \ln Q}, \quad \text{with } Q = \left(1 + \frac{S}{B}\right)^D e^{-S} \quad (4.13)$$

where S , B and D are respectively the expected number of signal, the expected number of background and the number of data events. Taking $D = S + B$, the significance reaches 4σ . No systematic error are taken into account in this calculation.

The study of the scalar sector of the $iM2HDM$ aims, among others, at determine the masses m_A and m_{H^0} .

The mass resolution of both A and H^0 bosons is estimated as the following. At a given luminosity, the mass is reconstructed and fills an histogram with a bin size smaller than the expected resolution. The content of each bin is updated according to a Poisson distribution a large number of times. For each iteration, the mean of the obtained histogram is recorded, providing a distributions for which the peak is centered on the true mass value. By fitting a gaussian distribution to the peak region, the standard deviation gives a confidence level on the measured mass. The evolution of this standard deviation with respect to the luminosity is shown in Fig. 4.28.

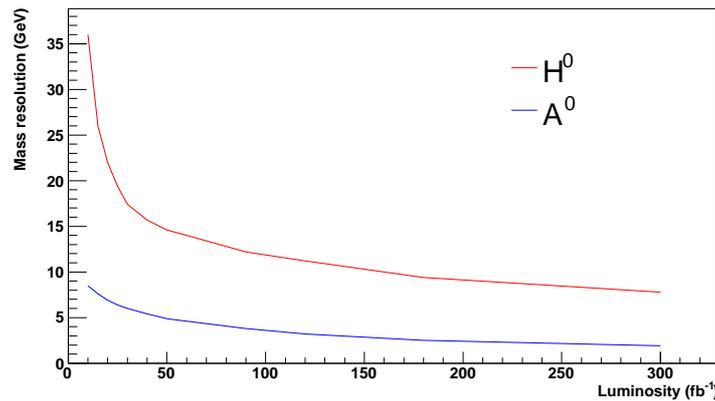


Figure 4.28: The expected evolution of the mass windows width (1σ) with respect to the luminosity. Both A and H^0 graphs are shown.

4.7 Conclusion

The promising discovery potential deduced in the third chapter for the $pp \rightarrow H^0 \rightarrow ZA$ channel deserved a more detailed study. To do so, this reaction is analyzed in the context of the CMS experiment, with low luminosity conditions.

Using several selection cuts, we show that the visibility of a fully leptonic final state can be strongly enhanced for the different benchmark points covering the free param-

eter space in Type II model. This confirms the expectations obtained with the fast simulator `Delphes`. Depending on the Higgs bosons masses, a 95 % C.L. deviation from the Standard Model hypothesis is expected with between roughly 2 and 8 fb⁻¹ thanks, notably, to the very low background level.

We show that, using a data-driven method, the cross-section normalization of the reducible background processes is possible with integrated luminosities as small as 5 fb⁻¹. In addition this method would allow an *in-situ* control of the sensitive variables used for the electron definition and isolation. Since these quantities may be strongly dependent of the calibration of CMS sub-detectors, such an estimation is required if a discovery with a small integrated luminosity is expected.

The consequence of this analysis is important. The observation of an event with the adequate final state and with only a few inverse femtobarns of integrated luminosity, would be a strong sign for a deviation from the Standard Model. Furthermore we show that with a larger luminosity the reconstruction of the Higgs boson masses is possible. This first estimation leads to the conclusion that one year of integrated luminosity would be sufficient to reach a precision between 5 and 10 %.

Conclusion

WITH one or two years of LHC running, the data collected from the experiments will give us the possibility to test the presence of a heavy Standard Model Higgs via its decay into vector bosons. Since the production cross-section and decay rates are relatively well known, a deviation from these expectations could be a sign that the scalar sector is more complex and therefore would motivate the testing of models with an extended scalar sector.

The $iM2HDM$ is an interesting alternative to the SM and also to MSSM. First because in both Type I and Type II Yukawa coupling modes, a large region of the parameter space is left free by the theoretical and experimental constraints. Second this model authorizes unusual decay modes and therefore experimental signatures.

In particular, we show that for each Type, the discovery of the Higgs bosons could be achieved with a minimal set of processes. For the neutral Higgs bosons discovery, two reactions deserve a particular attention, $pp \rightarrow h \rightarrow ZAZA$ in Type I and $pp \rightarrow H^0 \rightarrow ZA$ in Type II. While the first would lead to a discovery of all neutral Higgs bosons with 30 fb^{-1} of integrated luminosity, the second is expected to allow an observation of H^0 and A within a much shorter time.

The very high discovery potential of the reaction $pp \rightarrow H^0 \rightarrow ZA$ has been quantified in a detailed analysis in the context of the CMS experiment. The rejection of the Standard Model is expected to be possible with an integrated luminosity smaller than 10 fb^{-1} , thanks to a very low background level and a high S/B ratio. In order to enhance the credibility of this conclusion, we show that a data-driven control of the main background processes, particularly the reducible ones, is possible after a few inverse femtobarn of integrated luminosity. This is important as it corresponds to period where the analyses will start to constraint the $iM2HDM$ scalar sector. Besides the normalization of the background processes, this method allows also to control

the variables sensitive to the detector calibration, particularly those used to define the reconstructed electrons and their isolation.

This process deserves therefore a particular attention from the CMS community, as its discovery potential is close to those expected for SM process $pp \rightarrow h \rightarrow ZZ$ or light Higgs boson discovery in MSSM. In particular the results obtained in this work should be confirmed by using the full-simulation software of CMS as well as an official triggers table.

Besides the discovery of the neutral Higgs boson, the observation of the charged Higgs boson is also crucial as it would be a strong evidence of a scalar sector structured in Higgs doublets. We have shown that the $iM2HDM$ is characterized by unusual branching ratios for the charged Higgs such that the commonly considered decays into τ lepton might be totally suppressed. In this condition, processes like $pp \rightarrow tH^\pm \rightarrow tWA$ could be the only relevant process to achieved the H^\pm discovery despite the fact that, as we show, it would require a few hundreds of fb^{-1} of integrated luminosity.

The discovery of the Higgs boson(s) and, by extension of the new signatures at the LHC, will pass by the comparison between the data and correct simulations of the backgrounds but also of the signals. In particular we have shown that the simulation of the initial states radiation is important even in presence of hard jets from heavy particles decays and it is now clear that the jet matching technique is expected to give a trustful estimation of such a QCD activity. This method is now more broadly used by the CMS and ATLAS community, helping to refine the results of many analyses made in the past. The predictive power of the Monte Carlo tools is important, but discrepancies between the generators and simulation techniques still exists. An important work is therefore still left to do in order to understand and quantify precisely these differences for a better comparison with the LHC data.

Delphes configuration card

For the sake of reproducibility we give here the configuration card used for the analyses of the chapter 3.

```
# Detector extension, in pseudorapidity units
CEN_max_tracker 2.5 // Maximum tracker coverage
CEN_max_calor_cen 3.0 // central calorimeter coverage
CEN_max_calor_fwd 5.0 // forward calorimeter pseudorapidity coverage
CEN_max_mu 2.4 // muon chambers pseudorapidity coverage

# Energy resolution for electron/photon
# sigma/E = C + N/E + S/sqrt{E}, E in GeV
ELG_Scen 0.05 // S term for central ECAL
ELG_Ncen 0.25 // N term for central ECAL
ELG_Ccen 0.005 // C term for central ECAL
ELG_Cfwd 0.107 // S term for forward ECAL
ELG_Sfwd 2.084 // C term for forward ECAL
ELG_Nfwd 0.0 // N term for central ECAL

# Energy resolution for hadrons in ecal/hcal/hf
# sigma/E = C + N/E + S/sqrt{E}, E in GeV
HAD_Shcal 1.5 // S term for central HCAL // hadronic calorimeter
HAD_Nhcal 0. // N term for central HCAL
HAD_Chcal 0.05 // C term for central HCAL
HAD_Shf 2.7 // S term for HF // forward calorimeter
HAD_Nhf 0. // N term for HF
HAD_Chf 0.13 // C term for HF

# Muon smearing
MU_SmearPt 0.01 // transverse momentum Pt in GeV

# Tracking efficiencies
TRACK_ptmin 0.9 // minimal pt needed to reach the calorimeter in GeV
TRACK_eff 100 // efficiency associated to the tracking (percent)

# Calorimetric towers
TOWER_number 40
TOWER_eta_edges 0. 0.087 0.174 0.261 0.348
0.435 0.522 0.609 0.696 0.783 0.870
0.957 1.044 1.131 1.218 1.305 1.392 1.479 1.566 1.
653 1.740 1.830 1.930 2.043 2.172 2.322 2.500
2.650 2.868 2.950 3.125 3.300 3.475 3.650 3.825
4.000 4.175 4.350 4.525 4.700 5.000
# Thresholds for reconstructed objects, Pt in GeV
PTCUT_elec 3.0
```

```

PTCUT_muon      3.0
PTCUT_jet       10.0
PTCUT_gamma     10.0
PTCUT_taujet    10.0

# General jet variable
JET_coneradius  0.7           // generic jet radius ; not for tau's !!!
JET_jetalgo     3             // 1 for Cone algorithm, 2 for MidPoint algorithm, 3 for SIScone algorithm, 4 for kt algorithm
JET_seed        2             // minimum seed to start jet reconstruction, in GeV

# Tagging definition
BTAG_b          40            // b-tag efficiency (percent)
BTAG_mistag_c   10           // mistagging (percent)
BTAG_mistag_l   1            // mistagging (percent)

# FLAGS
FLAG_bfield     0             //1 to run the bfield propagation else 0
FLAG_vfd        1            //1 to run the very forward detectors else 0
FLAG_trigger    1            //1 to run the trigger selection else 0
FLAG_frog       1            //1 to run the FROG event display

# In case BField propagation allowed
TRACK_radius    129           //radius of the BField coverage, in cm
TRACK_length    300           //length of the BField coverage, in cm
TRACK_bfield_x  0             //X composant of the BField, in T
TRACK_bfield_y  0             //Y composant of the BField, in T
TRACK_bfield_z  3.8           //Z composant of the BField, in T

# Very forward detector extension, in pseudorapidity
# if allowed
VFD_min_calor_vfd 5.2         // very forward calorimeter (if any) like CASTOR
VFD_max_calor_vfd 6.6
VFD_min_zdc       8.3
VFD_s_zdc         140         // distance of the Zero Degree Calorimeter, from the IP, in [m]

RP_220_s         220         // distance of the RP to the IP, in meters
RP_220_x         0.002       // distance of the RP to the beam, in meters
RP_420_s         420         // distance of the RP to the IP, in meters
RP_420_x         0.004       // distance of the RP to the beam, in meters

# In case FROG event display allowed
NEvents_Frog     100

```

MatchChecker

B.1 Introduction

The jet matching procedure requires essentially to tune the matching scale Q_{match} so that the transition between hard and soft scales during jet clustering is made smoothly. The sanity check needed to ensure that this choice of Q_{match} is good can be done via the differential jet rates (DJR) distributions. This is explained in the Sec. 2.2.3 .

The code `MatchChecker` has been written to provide these distributions easily. In addition, the distributions related to the kinematics of massive particle(s) present in diverse productions as well as of the jets (ISR or not) are also provided. This helps for instance to do in-one-go the comparison between productions done with different parton showers generators or still different matching methods.

This software works in a very user-friendly way. The basic use is indeed limited to a card edition and running a command. Furthermore, all the results are provided in a PostScript report.

The package

The `MatchChecker` package requires bash, a recent version of Perl, a C++ compiler, the software ROOT (version 5 or later), latex and dvips to work properly. The input files are only STDHEP files.

It is downloadable from the MadGraph CVS by doing:

```
export CVSROOT=:pserver:anonymous@cp3wks05.fynu.ucl.ac.be:/usr/local/CVS
cvs co MatchChecker
```

It contains a set of codes and directories, namely:

| File or directory | Role |
|-----------------------|--|
| MatchChecker.sh | The main script to run |
| MatchCheckerCard.dat | Parameters Card |
| pdgcard.dat | Contains all possible legends |
| /MatchingPlots | Where plots are stored |
| Cosmetics.dat | Contains the plots specifications |
| /ExRootAnalysis_X.tar | Contains the routines producing needed rootfiles. |
| /TextFiles | Textfiles used by the perl scripts |
| Convertor.pl | A first perl script |
| PlotMaker.pl | A second perl script |
| QparPlot.C | The C++ file producing DJR plots |
| DistribPlots.C | The C++ file producing the kinematic variables plots |

ExRootAnalysis

The version of ExRootAnalysis included in the package is not the same as the one available on the MG/ME download page. The ExRootAnalysis package allows to store events generated by MadGraph and Pythia in a ROOT tree format [91] in order to perform analysis in a ROOT environment.

The ExRootAnalysis package can be subdivided into several subsystems: basic framework of few classes providing event loop, event selection and basic operations with a ROOT tree file; modules selecting events and objects to be analysed at per event and per object level; modules analyzing selected events; converters from different formats (LHEF, STDHEP, LHCO, *etc.*) to ROOT tree format.

For example, a selector module can select and group partons generated by MadGraph into several classes (such as leptons, jets, top quarks, *etc.*) according to their status and particle identification number.

For the present work two selector modules have been developed: matrix element parton selector and shower parton selector. The selection algorithms are based on the following rules:

- For matrix-element parton selection: keep final state partons
- For shower parton selection : keep final state shower partons
- For both selections:
 - skip beam Particles and initial state partons
 - skip intermideate (not final state) partons,
 - skip partons originating (directly or indirectly) from a QED vertex.

The selected shower partons are then used for jet reconstruction. The result of the parton selection and jet reconstruction is registered in form of a ROOT tree and further analysed in QparPlots.C and DistribPlot.C.

Documentation on the content of the ROOT tree is available on the web [96].

B.2 Running MatchChecker

The first manipulation is to untar the ExRootAnalysis package, enter in the directory and run make. This compiles the ExRootAnalysis package.

The next manipulation is the edition of the MatchCheckerCard.dat file that contains all informations relative to files to be analyzed but also the report content.

MatchCheckerCard

The configuration card is presented as follows:

```
### BLOCKNAME
test

### BLOCKFILES: production files

BEGINSAMPLE
tag = tag_for_first_production
comment = "your comment here"
banner = non
files = adress_of_first_prod_sample1
ENDSAMPLE

### BLOCKPDG: Write the PDG code of "X" in a "X" + jets process
pdg_code_of_particles_considered_as_in_the_central_system
###BLOCKSCALE: Scale used for Kt jet definition
40
###BLOCKCUTS: Cuts to apply on jets to calculate the rapidity distribution of jets
20 50 100
###BLOCKNORM: Normalize Jet rates to the cross section or 1 (A or B)
```

```

A
###BLOCKPLOT: Plots to appear in the Report (1=yes, 0=no)
DJR 1
KinCentral 1
KinJet 1

```

- The field `BLOCKNAME` defines the generic name of the files produced. This helps to not delete old results.
- The field `BLOCKFILES` contains the addresses of STDHEP sources. There are as many `BEGINSAMPLE . . . ENDSAMPLE` as productions. For instance if we want to compare the productions made with $t\bar{t}$ and $t\bar{t} + 0, 1, 2, 3$ partons, two blocks are needed.
 - I. The label “files” contains the addresses of all STDHEP files (relative or absolute). A production can be contained in any number of files. Increasing this number may help to decrease statistical fluctuations. These files can be inclusive, *i.e.* with all multiplicities in each file, or on the contrary the production can be composed of several files with each only one multiplicity. This is useful to increase the statistics in higher multiplicities (with lower cross-sections).
The syntax is the following: for cases where all multiplicities are contained in each file, the files names have to be written, separated by a comma. In the one-multiplicity-per-file option, the name of each file has to be followed by a space and the number indicating the multiplicity.
 - II. The label “tag” gives an additional name for the production. This helps to identify different sources of files, and therefore allows the possibility to run on only one production and not on the others.
 - III. The label “comment” lets the possibility for the user to add one comment per production at the beginning of the report.
 - IV. The label “banner” is used to include the banner of the ME and pythia runs. This is of course optional.
- The fields `BLOCKPDG` contains the PDG codes [2] of particles belonging to the hard scattering part of the STDHEP listing for which we want to have the kinematic distributions. For instance if “6 -6” is written (for $t\bar{t}$), the distributions are $P_T(t + \bar{t})$, $\eta(t)$, $P_T(t)$, $\Delta R(t, \bar{t})$, the invariant mass of $t\bar{t}$.
- The field `BLOCKSCALE` defines the k_{\perp} definition of the jets to use.
- The field `BLOCKCUTS` defines the P_T cuts to apply to evaluate the rapidity distributions of the extra-jets

- The field BLOCKNORM defines if the jet rates plot have to be normalized to the unity or to the cross-section

The second part of the card should not be tuned for basic use

```

###BLOCKCONDOR condor use or not? (y or n)
n

###BLOCKJETEXCL Kt specification for extra-jets (jets from additionnal partons)
CollisionType 4
DistanceScheme 3
RecombinationScheme 3
ParameterR 1.0
Exclusive true
ECut 1.0

###BLOCKJETINCL Kt specification for inclusive jets (all jets)
CollisionType 4
DistanceScheme 3
RecombinationScheme 3
ParameterR 1.0
Exclusive false
ECut 1.0

###BLOCKPARTONDEF
partons = {1 2 3 4 5 21}
ExcludeAncestor = {6 24 23 25 35 36 37 1000006 1000021 22 32 1000011
1000012 1000013 1000014 1000015 1000016 1000001 1000002
1000003 1000004 1000005 1000006 2000011 2000012 2000013
2000014 2000015 2000016 2000001 2000002 2000003 2000004
2000005 2000006 1000022 1000023 1000024 1000025 1000035
1000037 1000039}

###BLOCKSHOWEREDPARTONDEF
partons = {1 2 3 4 5 21}
ExcludedAncestorIDs = {6 24 23 25 35 36 37 1000006 1000021 22 32
1000011 1000012 1000013 1000014 1000015 1000016 1000001
1000002 1000003 1000004 1000005 1000006 2000011 2000012
2000013 2000014 2000015 2000016 2000001 2000002 2000003
2000004 2000005 2000006 1000022 1000023 1000024 1000025
1000035 1000037 1000039}

```

- The field BLOCKCONDOR says if a parallel usage of matchchecker is wanted.
- The field BLOCKJETEXCL gives the specification of extra-jets reconstruction using k_{\perp} .
- The field BLOCKJETINCL gives the specification of all jets reconstruction using k_{\perp} .
- the field BLOCKPARTONDEF allows to define what is a initial state radiation. The `parton` array contains the possible PDG codes and `ExcludeAncestor` contains the PDG code of particle that should not ancestors of ISR.

- the field `BLOCKSHOWEREDPARTONDEF` allows to define what are the showered partons from initial state radiation. The `parton` array contains the possible PDG codes and `ExcludeAncestorIDs` contains the PDG code of particle that should not ancestors of these showered partons.

Edit the `Cosmetics.dat` file

This card allows to control the axis ranges of each kind of plot. For instance the DJR plots are controlled by the lines.

```
Qparmin 0 #defines the x min range for DJR plots
Qparmax 3.5 #defines the x max range for DJR plots
QparLog 1 #Logy or not?
QparYmax 5 #ratio between y value of maximum bin and top of graph
QparYmin 1000 #ratio between y value of maximum bin and bottom of graph
```

Running the scripts

Once the card is filled, the production of the report can be done. The user has two possibilities

- `./MatchChecker.sh MatchCheckerCard.dat` for the automatic run
- Run by hand each script. This is useful if only a part of the run has to be redone.

The step-by-step manipulations work as the following:

- `perl Convertor.pl MatchCheckerCard.dat`
- `chmod u+x Convertor.sh`
- `./Convertor.sh`
- `perl PlotMaker.pl MatchCheckerCard.dat`
- `chmod u+x PlotMaker.sh`
- `./PlotMaker.sh`
- `perl ReportMaker.pl MatchCheckerCard.dat`
- `chmod u+x ReportMaker.sh`
- `./ReportMaker.sh`

Running the `Convertor.pl` and `Convertor.sh`

The perl script `Convertor.pl` creates the shell script `Convertor.sh` (+other secondary files) that will convert in the right way all STDHEP files into rootfiles. As the jet calculation (including Ktjet [97]) is done during this step, it can take a quite long time to be completed.

Running the `PlotMaker.pl` and `PlotMaker.sh`

The perl script `PlotMaker.pl` creates, from the reading of the card, an other shell script `PlotMaker.sh` and two others, `LaunchQpar.sh` and `LaunchDistrib.sh` (+other secondary files). The role of the first one, `PlotMaker.sh` is mainly to execute the two other scripts that will produce respectively differential jet rate plots and kinematic variables plots. Beside this, other cards needed for internal purpose are also created.

Running the `ReportMaker.pl` and `ReportMaker.sh`

The perl script `ReportMaker.pl` creates, from the reading of the card, an other shell script `ReportMaker.sh` and a \TeX file. The role of the shell script is to compile the \TeX file and produce the PostScript file `Report.ps`.

The Report

The report contains the following:

- List of banners
- Differential jet rate for each production with the multiplicity details
- Comparison of DJR between productions
- Kinematic plots of X in a $X+$ jets process for each production (with multiplicity details)+comparison of kinematic plots between different productions.
- P_T of the four leading jets in P_T for each production + comparison plots
- η of the four leading jets in P_T for each P_T cut chosen by the user for each production + comparison plots

- $H_T(j)$: Give the scalar sum

$$H_T(j) = \cancel{E}_T + \sum_{i=1}^j P_T^i$$

where \cancel{E}_T is the missing transverse energy, and P_T^i the transverse momentum of the i -th jet. $H_T(1 \rightarrow 4)$ is given for each P_T cut applied on jets. The plots H_T of the four first jets between different production are also given

- Plot of \cancel{E}_T + comparison plots between different productions.

BIBLIOGRAPHY

- [1] S. Weinberg, *What is an elementary particle?*, *SLAC Beam Line* **27N1** (1999) 17–21.
- [2] **Particle Data Group** Collaboration, C. Amsler *et. al.*, *Review of particle physics*, *Phys. Lett.* **B667** (2008) 1.
- [3] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, *Phys. Rev. Lett.* **13** (1964) 321–322.
- [4] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.
- [5] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.
- [6] **FAST** Collaboration, A. Barczyk *et. al.*, *Measurement of the Fermi Constant by FAST*, *Phys. Lett.* **B663** (2008) 172–180, [[arXiv:0707.3904](https://arxiv.org/abs/0707.3904)].
- [7] C. Delaere, *Study of $W W$ decay of a Higgs boson with the ALEPH and CMS detectors*, . CERN-THESIS-2006-041.
- [8] M. Herquet, *A twisted two-Higgs-doublet model: motivation and phenomenology*. PhD thesis, UCL, 2007.
- [9] T. P. Cheng, E. Eichten, and L.-F. Li, *Higgs Phenomena in Asymptotically Free Gauge Theories*, *Phys. Rev.* **D9** (1974) 2259.
- [10] A. Djouadi, *The anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, [hep-ph/0503172](https://arxiv.org/abs/hep-ph/0503172).
- [11] L. electroweak group, <http://lepewwg.web.cern.ch/lepewwg/>, 2009.

- [12] F. Halzen and A. D. Martin, “Quarks and leptons.” Wiley, 1985.
- [13] M. S. Chanowitz, *Strong $W W$ scattering at the end of the 90’s: Theory and experimental prospects*, hep-ph/9812215.
- [14] K. Riesselmann and S. Willenbrock, *Ruling out a strongly-interacting standard Higgs model*, *Phys. Rev.* **D55** (1997) 311–321, [hep-ph/9608280].
- [15] **LEP Working Group for Higgs boson searches** Collaboration, R. Barate *et. al.*, *Search for the standard model Higgs boson at LEP*, *Phys. Lett.* **B565** (2003) 61–75, [hep-ex/0306033].
- [16] **ALEPH** Collaboration, S. Schael *et. al.*, *Search for neutral MSSM Higgs bosons at LEP*, *Eur. Phys. J.* **C47** (2006) 547–587, [hep-ex/0602042].
- [17] **CDF** Collaboration, *Combined CDF and DZero Upper Limits on Standard Model Higgs-Boson Production with up to 4.2 fb-1 of Data*, arXiv:0903.4001.
- [18] T. Hahn, S. Heinemeyer, F. Maltoni, G. Weiglein, and S. Willenbrock, *SM and MSSM Higgs boson production cross sections at the Tevatron and the LHC*, hep-ph/0607308.
- [19] **CMS** Collaboration, G. L. Bayatian *et. al.*, *CMS technical design report, volume II: Physics performance*, *J. Phys.* **G34** (2007) 995–1579.
- [20] J. M. Gerard and M. Herquet, *A twisted custodial symmetry in the two-Higgs-doublet model*, *Phys. Rev. Lett.* **98** (2007) 251802, [hep-ph/0703051].
- [21] S. de Visscher, J.-M. Gerard, M. Herquet, V. Lemaitre, and F. Maltoni, *Unconventional phenomenology of a minimal two-Higgs- doublet model*, arXiv:0904.0705.
- [22] M. Herquet, *Symetries custodiales dans les Modeles a Deux Doublets de Higgs*, DEA. PhD thesis, UCL, 2004.
- [23] S. Davidson and H. E. Haber, *Basis-independent methods for the two-Higgs-doublet model*, *Phys. Rev.* **D72** (2005) 035004, [hep-ph/0504050].
- [24] M. Misiak, *QCD Calculations of Radiative B Decays*, arXiv:0808.3134.
- [25] M. Misiak *et. al.*, *The first estimate of $B(\text{anti-}B \rightarrow X/s \text{ gamma})$ at $O(\alpha(s)**2)$* , *Phys. Rev. Lett.* **98** (2007) 022002, [hep-ph/0609232].

- [26] T. Inami, C. S. Lim, and A. Yamada, *Radiative correction parameter s in beyond the standard models*, *Mod. Phys. Lett.* **A7** (1992) 2789–2798.
- [27] C. Q. Geng and J. N. Ng, *Charged Higgs effect in $B(d)0$ - anti- $B(d)0$ mixing, $K \rightarrow \pi \nu \bar{\nu}$ decay and rare decays of B mesons*, *Phys. Rev.* **D38** (1988) 2857.
- [28] C. Q. Geng and J. N. Ng, *Charged-higgs-boson effect in b_d - \bar{B}_d mixing, $k \rightarrow \pi \nu \bar{\nu}$ decay, and rare decays of b mesons*, *Phys. Rev. D* **38** (Nov, 1988) 2857–2863.
- [29] K. Hagiwara, A. D. Martin, D. Nomura, and T. Teubner, *Improved predictions for $g-2$ of the muon and $\alpha_{QED}(M_Z^2)$* , *Phys. Lett.* **B649** (2007) 173–179, [hep-ph/0611102].
- [30] **Muon G-2** Collaboration, G. W. Bennett *et. al.*, *Final report of the muon E821 anomalous magnetic moment measurement at BNL*, *Phys. Rev.* **D73** (2006) 072003, [hep-ex/0602035].
- [31] K. Ikado *et. al.*, *Evidence of the purely leptonic decay $B^- \rightarrow \tau^- \text{anti-}\nu_\tau$* , *Phys. Rev. Lett.* **97** (2006) 251802, [hep-ex/0604018].
- [32] E. Accomando *et. al.*, *Workshop on CP Studies and Non-Standard Higgs Physics*, hep-ph/0608079.
- [33] T. Stelzer, S. Wiesenfeldt, and S. Willenbrock, *Higgs at the Tevatron in extended supersymmetric models*, *Phys. Rev.* **D75** (2007) 077701, [hep-ph/0611242].
- [34] CDF Collaboration, *Search for Neutral MSSM Higgs Bosons Decaying to Tau Pairs with 1.8 fb^{-1} of Data*, *CDF Note* (2007), no. 9071.
- [35] M. S. Carena, S. Heinemeyer, C. E. M. Wagner, and G. Weiglein, *Suggestions for improved benchmark scenarios for Higgs- boson searches at LEP2*, hep-ph/9912223.
- [36] T. D. Collaboration and J. Abdallah, *Searches for neutral higgs bosons in extended models*, *European Physical Journal C* **38** (2004) 1.
- [37] CDF Collaboration, *A search of charged Higgs in the decay products of pair-produced top quarks.*, *CDF Note* (2006), no. 7712.
- [38] C.E.Hill. <http://linac2.home.cern.ch/linac2/seminar/seminar.pdf>.
- [39] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1–435.

- [40] T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann, and J. Winter, *Sherpa 1.alpha, a proof-of-concept version*, *JHEP* **0402** (2004) 056, [hep-ph/0311263].
- [41] A. Pukhov, *Calcchep 3.2: Mssm, structure functions, event generation, batches, and generation of matrix elements for other packages*, hep-ph/0412191.
- [42] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, *Madgraph/madevent v4: The new web generation*, 2007.
- [43] E. Boos, V. Bunichev, M. Dubinin, L. Dudko, V. Ilyin, A. Kryukov, V. Edneral, V. Savrin, A. Semenov, and A. Sherstnev, *Comphep 4.4 - automatic computations from lagrangians to events*, *NUCL.INSTRUM.METH.A* **534** (2004) 250.
- [44] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *AlpGen, a generator for hard multiparton processes in hadronic collisions*, hep-ph/0206293.
- [45] A. Cafarella, C. G. Papadopoulos, and M. Worek, *Helac-phegas: a generator for all parton level processes*, hep-ph/0710.2427.
- [46] W. Kilian, T. Ohl, and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, 0708.4233.
- [47] B. R. Webber, *Monte Carlo Simulation of Hard Hadronic Processes*, *Ann. Rev. Nucl. Part. Sci.* **36** (1986) 253–286.
- [48] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.
- [49] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B126** (1977) 298.
- [50] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics*, *Sov. Phys. JETP* **46** (1977) 641–653.
- [51] M. Bertini, L. Lonnblad, and T. Sjostrand, “Pythia version 7-0.0 - a proof-of-concept version.”
- [52] G. Corcella et. al., *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010, [hep-ph/0011363].

- [53] L. Lönnblad, *Ariadne version 4: A program for simulation of qcd cascades implementing the color dipole model*, *Comput. Phys. Commun.* **71** (1992) 15–31.
- [54] J. Alwall *et. al.*, *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *Eur. Phys. J.* **C53** (2008) 473–500, [0706.2569].
- [55] J. Alwall, S. de Visscher, and F. Maltoni, *QCD radiation in the production of heavy colored particles at the LHC*, arXiv:0810.5350.
- [56] M. L. Mangano, “Merging multijet matrix elements and shower evolution in hadronic collisions.” Available at <http://cern.ch/~mlm/talks/lund-alpgen.pdf>, 2004.
- [57] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063, [hep-ph/0109231].
- [58] G. Salam and M. Rubin, *Lhc searches: what role for qcd? talk given in ucl*, 2008.
- [59] T. Plehn, D. Rainwater, and P. Skands, *Squark and gluino production with jets*, *Phys. Lett.* **B645** (2007) 217–221, [hep-ph/0510144].
- [60] B. C. Allanach *et. al.*, *The Snowmass points and slopes: Benchmarks for SUSY searches*, hep-ph/0202233.
- [61] G. P. Salam and G. Soyez, *A practical Seedless Infrared-Safe Cone jet algorithm*, *JHEP* **05** (2007) 086, [0704.0292].
- [62] ATLAS Collaboration, *ATLAS detector and physics performance. Technical design report. Vol. 2*, . CERN-LHCC-99-15.
- [63] F. Krauss, A. Schälicke, S. Schumann, and G. Soff, *Simulating $w/z + jets$ production at the tevatron*, *Phys. Rev.* **D70** (2004) 114009, [hep-ph/0409106].
- [64] F. Krauss, A. Schalicke, S. Schumann, and G. Soff, *Simulating $W/Z + jets$ production at the CERN LHC*, *Phys. Rev.* **D72** (2005) 054017, [hep-ph/0503280].
- [65] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *Matching matrix elements and shower evolution for top- quark production in hadronic collisions*, *JHEP* **01** (2007) 013, [hep-ph/0611129].
- [66] J. Conway *et. al.*, *PGS: Pretty Good Simulation of high energy collisions*, . See <http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm>.

- [67] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, *NLO predictions for t-channel production of single top and fourth generation quarks at hadron colliders*, arXiv:0907.3933.
- [68] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, *Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders*, *Phys. Rev. Lett.* **102** (2009) 182003, [arXiv:0903.0005].
- [69] CMS Collaboration, G. L. Bayatian *et. al.*, *CMS physics: Technical design report*, . CERN-LHCC-2006-001.
- [70] J. Alwall *et. al.*, *MadGraph/MadEvent v4: The New Web Generation*, *JHEP* **09** (2007) 028, [0706.2334].
- [71] J. Pumplin *et. al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **07** (2002) 012, [hep-ph/0201195].
- [72] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [hep-ph/0603175].
- [73] S. Ovyn, X. Rouby, and V. Lemaitre, *Delphes, a framework for fast simulation of a generic collider experiment*, arXiv:0903.2225.
- [74] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the $k(t)$ jet-finder*, *Phys. Lett.* **B641** (2006) 57–61, [hep-ph/0512210].
- [75] CMS Collaboration, G. L. Bayatian *et. al.*, *CMS technical design report, volume II: Physics performance*, *J. Phys.* **G34** (2007) 995–1579.
- [76] ATLAS Collaboration, *Atlas: Detector and physics performance technical design report. volume 1*, . CERN-LHCC-99-14.
- [77] A. Djouadi, *The anatomy of electro-weak symmetry breaking. II: The Higgs bosons in the minimal supersymmetric model*, hep-ph/0503173.
- [78] R. V. Harlander and W. B. Kilgore, *Higgs boson production in bottom quark fusion at next-to-next-to-leading order*, *Phys. Rev.* **D68** (2003) 013001, [hep-ph/0304035].
- [79] T. Plehn, *Charged Higgs boson production in bottom gluon fusion*, *Phys. Rev.* **D67** (2003) 014018, [hep-ph/0206121].
- [80] M. Pimia, M. Della Negra, K. Eggert, M. Lanzagorta, and F. Szoncsó, *Compact muon solenoid*, . Prepared for ECFA Large Hadron Collider (LHC) Workshop: Physics and Instrumentation, Aachen, Germany, 4-9 Oct 1990.

- [81] M. Della Negra and J. C. Lottin, *Towards the lhc experimental programme : General meeting on lhc physics and detectors*. oai:[cds.cern.ch:423084](https://cds.cern.ch/record/423084), tech. rep., CERN, Geneva, Mar, 1992. VHS not held by the CERN library.
- [82] S. Cucciarelli, M. Konecki, D. Kotlinski, and T. Todorov, *Track reconstruction, primary vertex finding and seed generation with the pixel detector*, . CERN-CMS-NOTE-2006-026.
- [83] T. Speer *et al.*, *Track reconstruction in the CMS tracker*, *Nucl. Instrum. Meth.* **A559** (2006) 143–147.
- [84] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth.* **A262** (1987) 444–450.
- [85] P. A. *et al*, *Energy resolution performance of the cms electromagnetic calorimeter*, *CMS NOTE* [CMS NOTE-2006/140].
- [86] M. *et al*, *Electron Reconstruction in the CMS Electromagnetic Calorimeter*, . CERN-CMS-NOTE-2001-034.
- [87] R. Frühwirth, *Track fitting with non-gaussian noise*, *Computer Physics Communications* **100** (1997), no. 1-2 1 – 16.
- [88] S. A. *et al.*, *G4—a simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3 250 – 303.
- [89] **CMS** Collaboration, *CMS, tracker technical design report*, . CERN-LHCC-98-06.
- [90] H. Robbins and S. Monro, *A stochastic approximation method*, *The Annals of Mathematical Statistics* **22** (1951), no. 3 400–407.
- [91] L. Moneta, I. Antcheva, and R. Brun, *Recent developments of the ROOT mathematical and statistical software*, *J. Phys. Conf. Ser.* **119** (2008) 042023.
- [92] D.Futyán. CMS Wiki:
<https://twiki.cern.ch/twiki/bin/view/CMS/AICaScenarios#STARTUP>.
- [93] J. M. Campbell *et al.*, *Higgs boson production in association with bottom quarks*, [hep-ph/0405302](https://arxiv.org/abs/hep-ph/0405302).
- [94] S. Abdullin *et al.*, *Study of PDF and QCD scale uncertainties in $p p \rightarrow Z Z \rightarrow 4\mu$ events at the LHC*, . CERN-CMS-NOTE-2006-068.

- [95] V. Bartsch and G. Quast, *Expected signal observability at future experiments*, Tech. Rep. CMS-NOTE-2005-004. CERN-CMS-NOTE-2005-004, CERN, Geneva, Feb, 2005.
- [96] P. Demin. <http://madgraph.phys.ucl.ac.be/...../Tree.html>.
- [97] J. M. Butterworth, J. P. Couchman, B. E. Cox, and B. M. Waugh, *Ktjet: A c++ implementation of the kt clustering algorithm*, *Computer Physics Communications* **153** (2003) 85.

