

UNIVERSITÀ DI SIENA 1240



*dedica a ...*



## Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.



## Ringraziamenti

*Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.*

*Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.*



# Contents

<b>Introduction</b>	<b>xi</b>
<b>1 The Standard Model, Higgs Boson and New Scalar Particles</b>	<b>1</b>
1.1 The Standard Model . . . . .	1
1.2 The Higgs Boson . . . . .	1
1.3 New Scalar Particles . . . . .	1
1.3.1 Study of the interference effects . . . . .	1
<b>2 The CMS experiment at LHC</b>	<b>3</b>
2.1 The Large Hadron Collider . . . . .	3
2.2 The Compact Muon Solenoid experiment . . . . .	3
<b>3 Monte Carlo Generators</b>	<b>5</b>
3.1 Hard process . . . . .	7
3.2 Parton shower . . . . .	9
3.3 Multiple Interaction . . . . .	13
3.4 Hadronization . . . . .	14
3.5 Hadronic Decays and Electromagnetic Radiation. . . . .	16
3.6 Jets Reconstruction . . . . .	16
3.7 Main Monte Carlo generators . . . . .	17
3.8 Monte Carlo sample in High Mass Analysis . . . . .	18
3.8.1 The DY sample . . . . .	21
3.8.2 The WW sample . . . . .	23
<b>4 Event Reconstruction</b>	<b>25</b>
<b>5 High mass resonances searching</b>	<b>27</b>
5.1 Introduction . . . . .	27
5.2 Discriminating variable . . . . .	28
5.3 Signal interpretation . . . . .	28
5.4 Opposite Flavor final state . . . . .	33
5.5 Same Flavor final state . . . . .	46
5.6 2HDM and MSSM interpretations . . . . .	56
5.6.1 Introduction to 2HDM and MSSM . . . . .	56
5.6.2 Statistical inference . . . . .	56
5.7 Systematic uncertainties . . . . .	58

<b>6 Results and Interpretation</b>	<b>63</b>
<b>A Special commands</b>	<b>67</b>

# Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.



## Chapter 1

# The Standard Model, Higgs Boson and New Scalar Particles

1.1 The Standard Model

1.2 The Higgs Boson

1.3 New Scalar Particles

1.3.1 Study of the interference effects



## Chapter 2

# The CMS experiment at LHC

**2.1** The Large Hadron Collider

**2.2** The Compact Muon Solenoid experiment



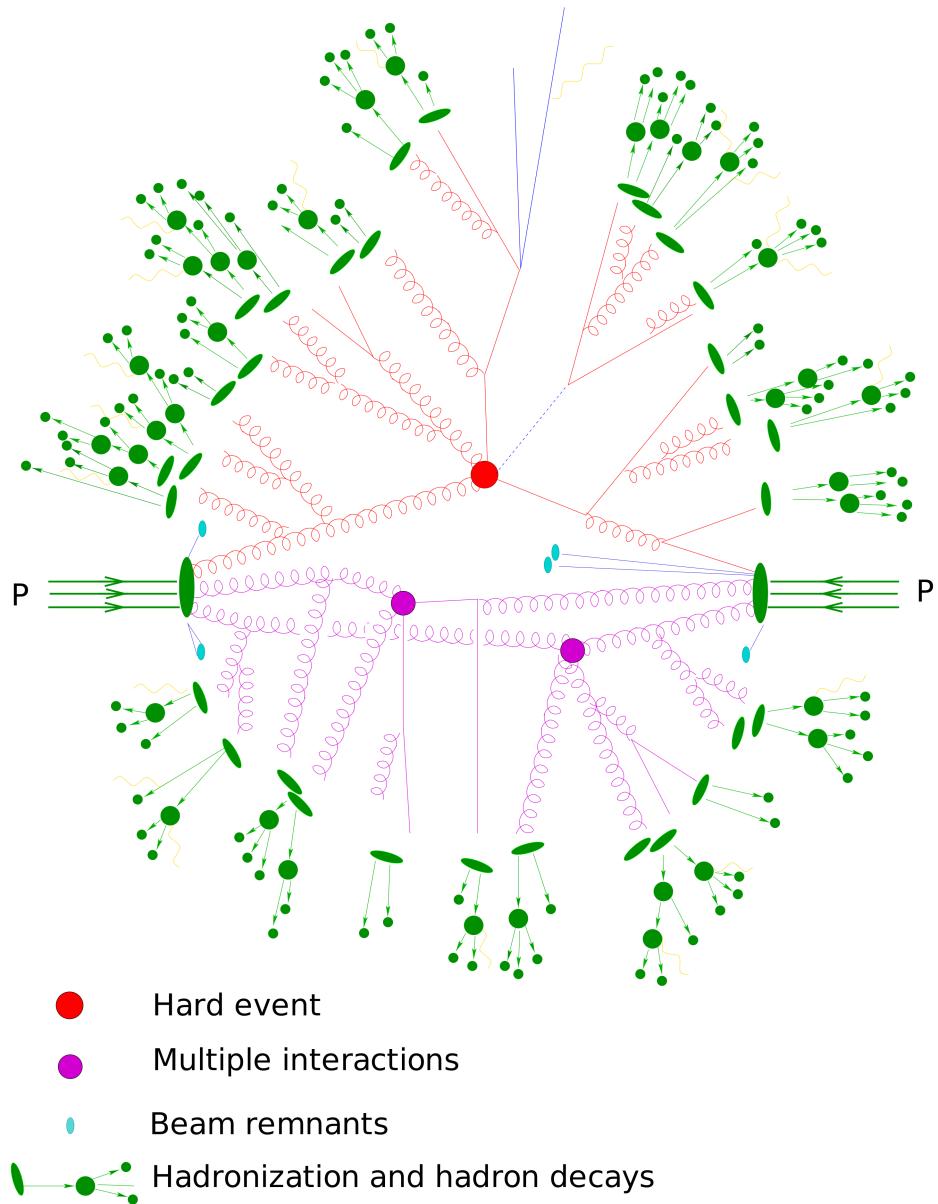
## Chapter 3

# Monte Carlo Generators

As we have seen in the collisions between high energy protons, hundreds of particles are generally present in the final state. Given the complexity of the events it is necessary to use Monte Carlo generators, i.e. programs that allow to simulate the realistic result of the collisions assuming a certain model for the processes involved. The use of Monte Carlo generators is necessary because it is impossible to predict what happens event-by-event: in fact, in quantum mechanics we can only calculate the probability of having a certain result. The simulation of an event is carried out in successive steps [1, 2], as schematized in Fig. 3.1, thus subdividing the problem into several parts of lower complexity.

The various steps are summarized here:

- Hard process: the incident protons are composed by partons (quarks and gluons) and the hard process consists in a collision between two partons, coming from different hadrons. The matrix element of the process is evaluated perturbatively and often only the lowest perturbative order, called leading order (LO), is calculated.
- Parton shower: the incoming or outgoing partons participating in the hard process can emit gluons: in fact, in analogy with the electromagnetic interaction, a particle with an accelerated color charge can radiate for the bremsstrahlung. The gluons in turn, can produce quark-antiquark pairs thus generating the parton showers. The emission of additional partons takes place mainly in the collinear space respect to the initial parton and progressively with less energies. In the final state there will be a set of partons, called jet, located in the collinear respect to the initial parton. This probabilistic process can be simulated as a Markov process and is implemented in the parton shower algorithms we will discuss later.
- Multiple interactions: in a single collision, it may happen to have more pairs of partons interacting. In this case it is said that there are multiple interactions in addition to the hard process.
- Hadronization: in the evolution of the event the partons are gradually generated with ever lower relative momenta. For momentum values of 1 GeV the confinement forces prevail. At these energy scales, the perturbation theory



**Figure 3.1.** Schematic representation of an event generated within an event generator. The partons coming from the protons indicative participate in both the hard process and multiple interactions. Subsequently there is the hadronization.

fails in the description, so we resort to non perturbative models which describe the formation of real hadrons. This hadronization process preserves the jet structure which can therefore be observed experimentally.

- Decaying of unstable particles: many of the particles produced in the primary process are unstable and they decayed unless they interact directly with the detector.

In the Monte Carlo simulation all these steps are considered sequentially: the result of each step is the starting point of the next. At the end, in a single event, there are hundreds of particles each of which has a dozen degrees of freedom (mass, flavor, impulse, average life, spin, vertex production, etc.), so there is a high number of parameters that came into play and must be simulated for each event. The final aim is to provide a realistic description of what happens in high-energy collisions, in order to compare the Monte Carlo model with the experimental data. Schematically, the cross section of the final state is given by,

$$\sigma_{\text{final state}} = \sigma_{\text{hard process}} \mathcal{P}_{\text{tot, hard process} \rightarrow \text{final state}}, \quad (3.1)$$

integrated over the total phase space and summed over all possible final states (e.g. the production of two or more jets). This is the measurable quantity associated with the hard process.

### 3.1 Hard process

In many processes of interest to LHC high momenta come into play, to produce high mass particles or energetic jet. The simulation of these events is the main goals of the Monte Carlo generators. The cross section for a scattering  $ab \rightarrow n$  process is given [2] by,

$$\begin{aligned} \sigma &= \sum_{a,b} \int_0^1 dx_a dx_b \int f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) d\hat{\sigma}_{ab \rightarrow n}(\mu_F, \mu_R) \\ &= \sum_{a,b} \int_0^1 dx_a dx_b \int d\Phi_n f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \\ &\times \frac{1}{2\hat{s}} |\mathcal{M}_{ab \rightarrow n}(\Phi_n, \mu_F, \mu_R)|^2, \end{aligned} \quad (3.2)$$

where

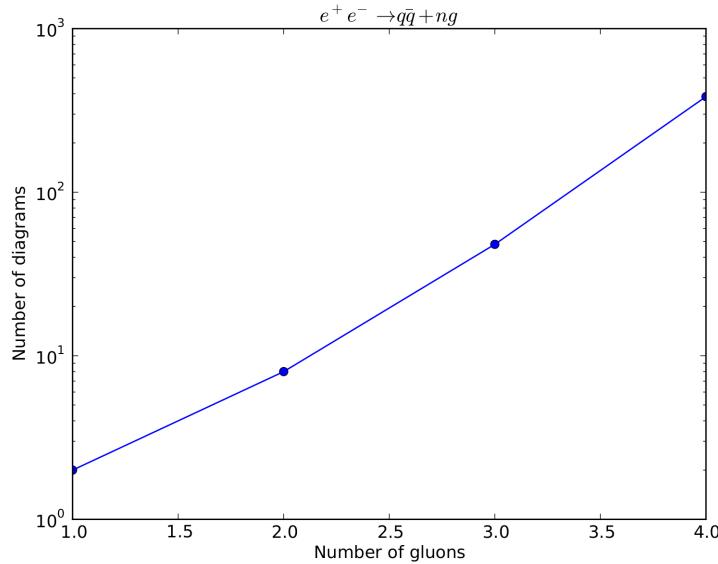
- $f_a^h(x, \mu)$  are the parton density functions (PDF) that depend on the  $x$  fraction of the parton  $a$ 's energy (Bjorken variable) respect to the  $h$ , and on the  $\mu_F$  factorization scale, which has been introduced in the Eq. XX.
- $\hat{\sigma}_{ab \rightarrow n}$  is the partonic cross section of the process  $ab \rightarrow n$ . The total differential cross section is given by the product of the corresponding square matrix element,  $|\mathcal{M}_{ab \rightarrow n}|^2$ , with the incident particle flow  $1/(2\hat{s}) = 1/(2x_a x_b s)$ , where  $\sqrt{s}$  is the energy of the system's center of mass.

- The matrix element  $|\mathcal{M}_{ab \rightarrow n}(\Phi_n, \mu_F, \mu_R)|^2$  can be written as the sum on all Feynman diagrams,

$$\mathcal{M}_{ab \rightarrow n} = \sum_i \mathcal{F}_{ab \rightarrow n}^{(i)}. \quad (3.3)$$

- $d\Phi_n$  it is the phase space differential for  $n$  particles in the final state.

The phase space will not be all physical space but will contain cuts for two reasons: the first is that the cuts will reflect the geometry and acceptance of the detector; the second because it is necessary put a cut on the transverse impulse of the particles produced in the process to avoid divergences in the calculation of the cross section <sup>1</sup>. In general, the calculation of the matrix element would require the calculation of all the Feynman diagrams which grow in a factorial way (Fig. 3.2) with the number of particles in the final state.



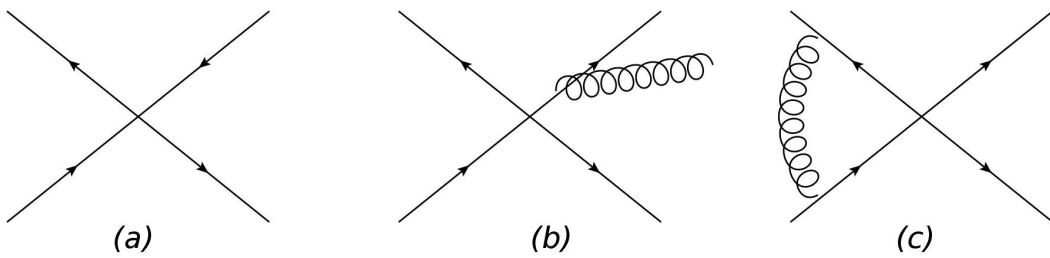
**Figure 3.2.** Trends in the number of Feynman diagrams as the number  $n$  of gluons increases in the process  $e^+e^- \rightarrow q\bar{q} + ng$ .

Usually the Monte Carlo events generators can compute the matrix element at the leading order for the Standard Model  $2 \rightarrow 1$ ,  $2 \rightarrow 2$  and  $2 \rightarrow 3$  [?] processes. However, if we stopped at the first perturbative order, we would have only a rough description of the process: in fact, subsequent orders involve important corrections both to the shape of the distributions and to the total cross section. LO is useful for a first study but and it is important to evaluate next-to-leading-order (NLO) <sup>2</sup>. The cross section calculated at the NLO is composed of three parts: by the LO part or part of Born, by the real and virtual part of the emission corrections (Fig. 3.3),

$$d\sigma^{NLO} = d\tilde{\Phi}_n [\mathcal{B}(\tilde{\Phi}_n) + \alpha_s \mathcal{V}(\tilde{\Phi}_n)] + d\tilde{\Phi}_{n+1} \alpha_s \mathcal{R}(\tilde{\Phi}_{n+1}), \quad (3.4)$$

<sup>1</sup>You can imagine having a singularity similar to that one has in scattering classic Coulomb.

<sup>2</sup>For some particularly important processes, for example  $gg \rightarrow H$ , the next-next-to-leading-order (NNLO) calculations are even available.



**Figure 3.3.** Examples of Feynman diagrams (a) Born, (b) real, (c) virtual.

where  $\mathcal{B}$ ,  $\mathcal{R}$  and  $\mathcal{V}$  denote the Born, the real and the virtual part respectively. The integral must be made on the  $n$  or  $n + 1$  final state particles and on the Bjorken variables related to the incident partons. Suppose, in the Born approximation, the process  $2 \rightarrow 2$ . If you want to go to the next order, NLO, you have to keep the element with an additional parton in the final state, the  $2 \rightarrow 3$  process, and virtual correction with a loop in the  $2 \rightarrow 2$  process. It should be noted that the cross-section for processes of the type  $2 \rightarrow 3$  is divergent when the energy of one of the partons tends to zero (soft divergence) or when two parts are collinear (collinear divergence).

## 3.2 Parton shower

In a collision between partons a charge of color is accelerated, so there will be bremsstrahlung emission. When studying a process of the type  $2 \rightarrow n$ , where  $n$  represents the number of partons in the final state, the LO matrix elements (called tree-level) will have divergences in the collinear and soft case. In particular, the processes that suffer from this type of divergence are  $q \rightarrow qg$ ,  $\bar{q} \rightarrow \bar{q}gq$ ,  $g \rightarrow gg$ : the first are similar processes to  $e \rightarrow e\gamma$  in QED, while the third is due to the fact that QCD is not an Abelian theory. The process  $g \rightarrow q\bar{q}$  does not have this type of divergence. The divergences of the tree-level matrix element can be removed by introducing the virtual corrections into the calculation, but they will be in the next order; these calculations are therefore particularly complex and they are only possible for a limited number of processes. The parton shower [?] algorithms offer an alternative and simple way to eliminate the collinear and soft divergences through:

- an iterative structure that combines the three states suffering from divergences in a single multi-partonic state,
- the introduction of the form factor of Sudakov.

The incoming or outgoing partons, which are far (temporally) from hard process, are called on-shell indeed the module of their four-momenta is equal to the mass at rest. However, closer you get to the interaction, the partons can be in a state called off-shell, where the module of the their four-momenta does not correspond to the mass at rest due to the uncertainty principle ( $\Delta E \Delta t \sim \hbar$ ). For this reason they are able to emit other partons and the energy of emitted partons is higher if they are closer to the scattering. If the emission occurs before the scattering, it is called initial state radiation (ISR), while after the interaction it is called final state radiation (FSR).

Each partition is characterized by a “virtuality scale”  $Q^2$  that corresponds roughly to a shower temporal scale. It is important to stress that different definitions are available for  $Q^2$ ; however regardless of the chosen convention, the  $Q^2$  scale increases as it approaches in the hard process, so then in the ISR, and decreases away, so in the FSR. If we take the FSR, the evolution starts at a  $Q_{max}^2$  scale that is related to the hard process and it ends when a limit scale is reached,  $Q_0$ , which will be on the order of 1 GeV.

The most common choice used is to set  $Q^2 = p^2 = E^2 - |\vec{p}|^2$ . With this convention in a process of type  $a \rightarrow bc$ , in FSR case,  $Q^2 > 0$ , that is of type time-like, and it will decrease until the limit scale  $Q_0$  is reached. The ISR case is more complicated: in this case  $a$  and  $b$  (supposed off-shell) have  $p^2$  of type space-like, then redefines  $Q_i^2 = -m_i^2$  in order to guarantee the increasing order of  $Q^2$ , i.e.  $Q_b^2 > Q_a^2$ . In contrast,  $c$  will not participate in hard process and will have  $p^2 > 0$  and therefore its shower will evolve like that of the FSR.

**Final State Radiation** In the parton shower approach, the final radiation state is modeled through a series of divisional processes of the type  $a \rightarrow bc$ . This is evident from the process  $q\bar{q}g$ , Fig. 3.3 (b), where the first order matrix element corrections correspond to the emission of a gluon. The evolution of the shower is described by two parameters: the fraction of energy carried by one of the two outgoing partons,  $z = E_b/E_a$ , and the order variable  $t$ . As we said, a possible choice for  $t$  is the virtuality  $Q_a^2$  of the incoming parton. In the collinear limit the probability of division  $d\mathcal{P}_{a \rightarrow bc}$ , in  $z$  and  $t = \ln(Q^2/\Lambda^2)$  is:

$$d\mathcal{P}_{a \rightarrow bc} = \sum_{bc} \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc} dt dz, \quad (3.5)$$

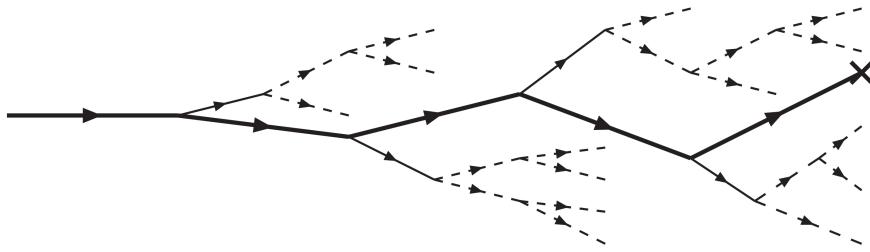
where  $dt = \frac{dQ^2}{Q^2}$ ,  $\alpha_{abc}$  it is the coupling constant that regulates the division process and  $P_{a \rightarrow bc}$  is the kernel splitting; these are universal functions and are valid in the collinear limit:

$$\begin{aligned} P_{q \rightarrow qg} &= \frac{4}{3} \frac{1+z^2}{1-z}, \\ P_{g \rightarrow gg} &= 3 \frac{(1-z(1-z))^2}{z(1-z)}, \\ P_{g \rightarrow q\bar{q}} &= \frac{n_f}{2} (z^2 + (1-z)^2), \end{aligned} \quad (3.6)$$

where  $n_f$  is the quarks flavour number *quark*. However the probability evaluated is larger than the unity indeed it suffers from the same divergences of the matrix element at the LO. The expression 3.5 is evaluated in the collinear approximation. In particular, there are two types of divergences: collinear, due to the dependency of type  $1/Q^2$ , and soft which corresponds to the limit  $z = 1$ .

To solve this problem, in the parton shower approach, the probability of dividing  $t$  and  $t + dt$  is evaluated; this is obtained with the integration of Eq 3.5 over  $z$  in the intervals  $[z_{min}(t), z_{max}(t)]$ :

$$d\mathcal{P}_{a \rightarrow bc} = \left( \sum_{bc} \int_{z_{min}(t')}^{z_{max}(t')} \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc} dt dz \right) dt. \quad (3.7)$$



**Figure 3.4.** Evolution of the initial state. The bold line corresponds to the part that will undergo the hard process (represented by a cross). Thin lines represent the partons that can not recombine, while the dashed lines are fluctuations that may or may not recombine.

As in other physical situations<sup>3</sup> the probability of something happening at  $t$  is given by the probability that this happens between  $t$  and  $t + dt$ , multiplied by the probability that this has not already occurred between the initial instant  $t_0$  and  $t$ . In this case then the division probability at  $t$  is:

$$d\mathcal{P}_a^{\text{FSR}}(t) = d\mathcal{P}_a \cdot \exp \left( - \sum_{bc} \int_{t_0}^t dt' \int_{z_{\min}(t')}^{z_{\max}(t')} \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc}(z) dz \right), \quad (3.8)$$

where  $t_0$  is the shower starting scale. The exponential term is called Sudakov factor and it represents the probability of non-division. If you want to interpret it in terms of Feynman diagrams, this represents the virtual corrections of LO matrix element. This total process can be combined to have more emissions at different steps: this will result in a partons shower which will be ordered in decreasing  $Q$ . Finally, the description given by parton shower is correct if you have collinear jet and it fail in configurations where there are well separated partons.

**Initial State Radiation** The the initial state radiation evolution is much more complicated than the final state. Indeed the quark and the gluons are emitted and absorbed continuously, inside the incoming proton. The initial stare radiation is already present during the hard scattering. The ISR simulation could start from the on-shell parton before the interaction and after could evolve to higher and higher  $Q^2$  scales until the hard process. However, this approach is very inefficient because the interest process is particularly rare and because it has the same probability to happen as in nature. In the event generators a different approach is used: first the hard process is produced and then it we try to rebuild back what may have happened. This procedure is called backward evolution, Fig. 3.4.

It is necessary to evaluate the probability for the process of type  $a \rightarrow bc$ , that a parton  $b$  has been produced by the parton  $a$ . For this reason the partonic density function is introduced. This evolves according to the DGLAP [?] equation,

$$\frac{df_b(x, t)}{dt} = \sum_{ac} \int_x^1 \frac{dx'}{x'} f_a(x', t) \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc} \left( \frac{x}{x'} \right), \quad (3.9)$$

---

<sup>3</sup>For example radioactive decay.

where  $f_{a,b}(x, t)$  are the parton PDFs  $a, b$ , that has  $x$  fraction of the incident and scale proton momenta  $t = \ln(Q^2/\Lambda^2)$ , instead  $P_{a \rightarrow bc}$  is the kernel splitting function. In the backward evolution the probability that the parton  $b$  has been generated from  $a$  in the interval between  $t$  and  $t - dt$  is given by:

$$d\mathcal{P}_b(t) = \frac{df_b(x, t)}{f_b(x, t)} = |dt| \sum_{ac} \int \frac{dx'}{x'} \frac{df_a(x', t)}{f_b(x, t)} \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc} \left( \frac{x}{x'} \right), \quad (3.10)$$

while the probability of non-division between the scale  $t_{max}$  and  $t < t_{max}$  is:

$$S_b(x, t, t_{max}) = \exp \left( - \int_t^{t_{max}} dt' \sum_{ac} \int \frac{dx'}{x'} \frac{df_a(x', t')}{f_b(x, t')} \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc} \left( \frac{x}{x'} \right) \right), \quad (3.11)$$

Finally then the probability of combining  $b$  in  $a$  is given in the range between  $t$  and  $(t - dt)$  from:

$$\begin{aligned} d\mathcal{P}_b^{\text{ISR}}(t) &= - \frac{dS_b(x, t, t_{max})}{dt} dt \\ &= \sum_{ac} \int \frac{dx'}{x'} \frac{df_a(x', t)}{f_b(x, t)} \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc} \left( \frac{x}{x'} \right) \cdot S_b(x, t, t_{max}) dt \end{aligned} \quad (3.12)$$

In this case the form factor Sudakov is different respect to FSR as it contains the PDFs. This means that the parton shower results do not depend only on the algorithm but also on the PDFs used.

**Resummation** When calculating an observable of the QCD in a perturbative way, the expansion in terms of  $\alpha_S$  contains terms of the type  $\alpha_S^n L^k$  ( $k < 2n$ ), where  $L = \ln(q_{cut}/s)$ , being  $q_{cut}$  the cut on resolvable emission. When we consider small values of  $q_{cut}$  the logarithm of the perturbative expansion becomes large and the perturbative series diverges. The main perturbative order of the expansion is  $n$  only if the successive terms of the series are negligible, however this is not guaranteed if there are high value of  $L$ . It is therefore necessary to consider the terms that have a high value of the logarithm. The study of these terms is called resummation and is done by putting the terms together in the perturbation series according to their degree of divergence:  $\alpha_S^n L^{2n}$  are the leading log (LL) terms,  $\alpha_S^n L^{2n-1}$  are the next-to-leading log (NLL) terms, and so on. At the end all  $\alpha_S$  orders terms are added. For many processes calculations are available at the NLL. The parton shower reproduces the effects of resuming approximately at the NLL.

**Merging among ME and PS** The two different approaches for the matrix element calculation and for the parton shower have advantages and disadvantages. Regarding the ME we have:

- the LO matrix element calculations can be performed exactly in the cases where there are many jet (of the order of six) in the final state,
- a good description of separate partons is performed,

- the perturbative calculations are correct,
- however, the cross section diverges in the collinear and soft case, so an exhaustive description of the internal structure of jet is not possible.

On the other hand the PS:

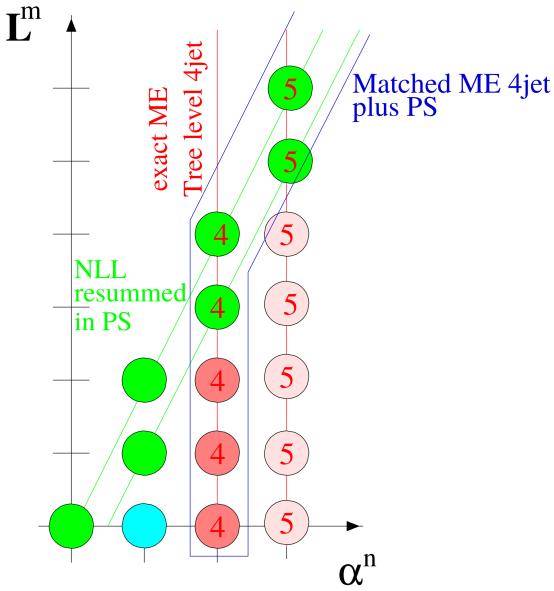
- it is a universal approach that produces a realistic configuration of the partons,
- the divergences, in the collinear limit, are treated with the introduction of the form factor of Sudakov. So we have an appropriate description of the jet evolution
- however, the method fails when describing separate partons, since the collinear approximation in this case can not be valid.

Clearly the two methods are complementary and their merging is desirable. There are different approaches that combine ME with PS. The main difficulty is to cover the total phase space without overlaps or holes: we want to describe a process in which there are  $n$  well separated partons in the final state, using the LO matrix element LO but also including the large logarithms resummation (LL, NLL) which is typical of the PS. A schematic description of the combination for four jet is given in Fig. 3.5. On the horizontal axis are the  $\alpha_S$  coupling orders, while on the vertical axis the logarithm. The PS describes the LL ( $m = 2n$ ) and the NLL ( $m = 2n - 1$ ) the green balls (e.g. in the case of  $n = 2$ ,  $m = 4, 3$  the two balls in green and marked as “4”). The balls that describe the event with four jet, combining the ME with the PS, are green, the blue and three red marked with the “ 4 ”. The difficulty arises because the ME describes exactly all the balls marked with the “ 4 ”: so if we simply sum up the two approaches we would have double counts of the green spheres called “ 4 ”. The main approaches to merge the ME and PS are:

- parton shower reweight: the basic idea is to start from the process to the lowest order and then re-evaluate the output of the PS as if it had been produced by the ME. This approach does not change the cross section, which remains at the lowest order, but improves the population of the phase space [?, ?].
- CKKW prescription: the phase space is divided into two zones using  $k_{\text{perp}}$  which is a measure of the cut  $Q_0^2$ : the region in which the jet is produced is filled with the ME, that of evolution with the PS [?, ?].
- The MLM prescription, which is also very widespread, is based on the same principle, but is implemented in a different way.

### 3.3 Multiple Interaction

Incident protons participating in the interaction are composed of a large number of partons (quark and gluons) that can interact independently with each other in addition to hard-process. The total cross section for the QDC process  $2 \rightarrow 2$  is dominated by the  $t$  process, so the cross section diverges as  $dp_{\text{perp}}^2/p_{\text{perp}}^4$  for  $p_\perp \rightarrow 0$  [?]. So when simulating a real event, in addition to the hard event, characterized



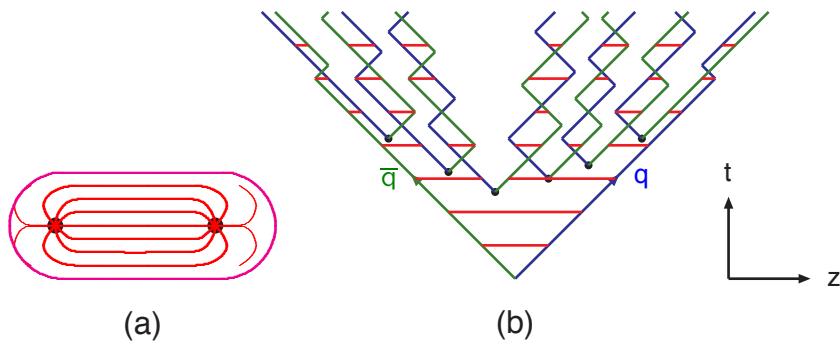
**Figure 3.5.** Merging among ME and PS.

by having large transverse transverse pulses, we must also take into account the additional collisions at small  $p_{\perp}$ . If these occur independently then a Poisson distribution is expected,  $P_n = \langle n \rangle^n \exp(-\langle n \rangle)/n!$ . However, conservation of energy and impulse means that interactions are not effectively independent, thus suppressing the possibility, for  $p_{\perp} \rightarrow 0$ , of having a high number of interactions. It should also be noted that in order to eliminate the divergence it is necessary to introduce a cut-off value of the transverse pulse, below which no collisions are generated.

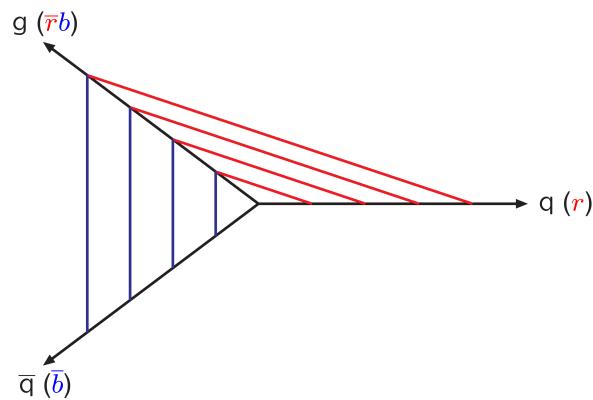
### 3.4 Hadronization

In this context, the process of atomization is a particular model, used in event generators, which describes the transition from the final partonic state to the final hadron state, which is an observable experimental. It is important to underline that this transition is treated in a phenomenological way and not by a rigorous approach. The two most important classes for tuning are the string model and the cluster model. The difference is that the former transforms the partonic systems directly into hadrons, while the second takes an intermediate step where it groups the objects to a scale of  $\sim 1$  GeV.

**String Model** The most complete and complete “string model” is that of Lund: we know from QCD that there is a linear confinement force between the partons that increases with distance. Consider, as an example, the final state in which there are two quark,  $q\bar{q}$ . As the partons move away the color flow tube is “stretched” between  $q$  and  $\bar{q}$ , Fig. ?? (a). The transverse dimensions of the tube are those typical of hadron, therefore of about 1 fm. If the tube is assumed to be uniform, the potential increases linearly,  $V(r) = \kappa r$ , with  $\kappa \approx 1$  GeV /fm, string constant.



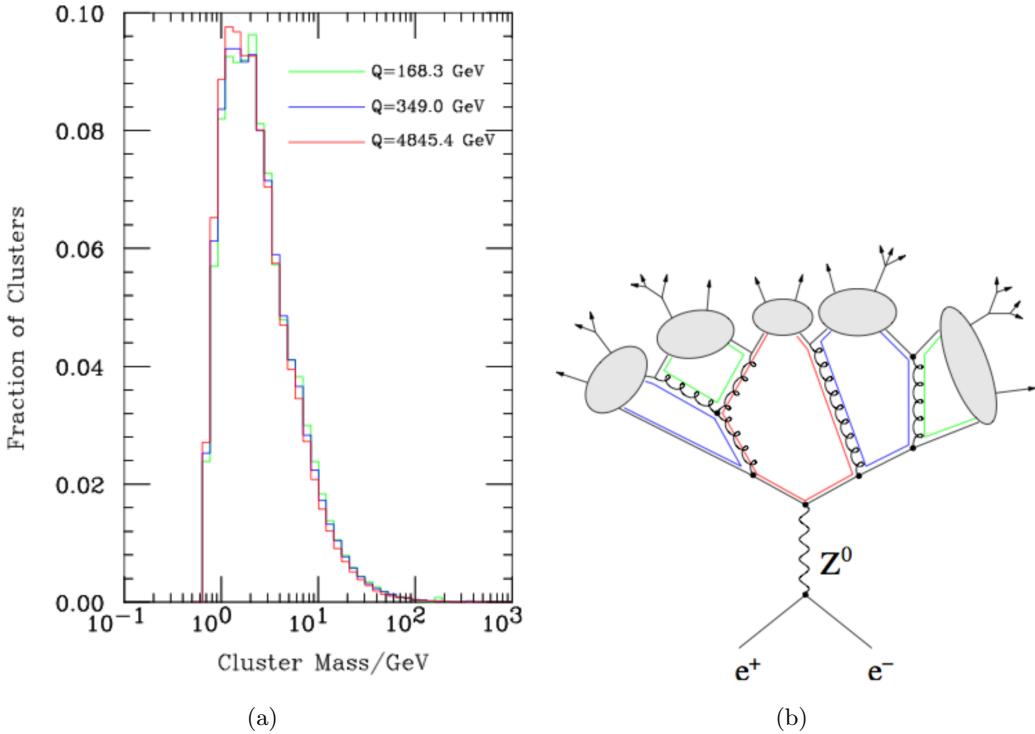
**Figure 3.6.** (a) The flow tube between a quark and an antiquark moving away. (b) Motion and breaking of a system string.



**Figure 3.7.** Motion of the string in the case  $q\bar{q}g$ .

At short distances it would be necessary to introduce an additional Coulomb term,  $\sim \frac{\alpha_s}{r}$ , however in the Lund model this negligible term is assumed. As the quark and antiquark move away from the creation vertex, the potential energy accumulated in the string increases until it breaks, giving rise to a pair  $q'\bar{q}'$ . So the system is divided into two new color singlets  $q\bar{q}'$  and  $q'\bar{q}$ . These two systems will move away in turn by repeating the process just described. The evolution of the system in space-time is represented in ?? (b). At the end of the process you will have a serious pair of  $q_i\bar{q}_i$  pairs, each of which will form a hadron. For now, only the case  $q\bar{q}$  has been considered. However, if more partons come from the interaction, the string model becomes more complicated. For an event in which there is an additional gluon,  $q\bar{q}g$ , the string is stretched between  $q$  and  $g$  and between  $g$  and  $\bar{q}$ , Fig. 3.7.

**Cluster Model** This model of atomization is based on the pre-confining property of the parton shower: the invariant mass of a single pair of opposite-colored partons is the same at any  $Q^2$  scale. This distribution has its maximum at a mass that is about cutoff of parton shower and decreases rapidly towards zero, Fig ?? (a). In the model, the parton shower gluons are represented by pairs of color-anticolor lines connected to the vertex. Each color line, near the cutoff, is connected to another colorless line present at the same scale. At this point the contiguous color / anticolor



**Figure 3.8.** (a) Parton shower structure in the cluster model. (b) Invariant mass distribution for singlets.

lines are interpreted, in the non-perturbative limit, as quark-antiquark pairs which give rise to mesons, which are observable objects in the final state. This mechanism is represented in ?? (b).

### 3.5 Hadronic Decays and Electromagnetic Radiation.

In the phasing phase, unstable hadrons can be produced which decay into other particles. So the final state of the event is the result of the convolution between the hadron and the decay. The information necessary for the simulation of the unstable particles of the decay is generally taken from the “Particle Data Book” (PDG) [?] which provides the properties (e.g. average life) of a large number of particles. In general, in an event generator, it is necessary to choose which hadrons to include in the simulation and then select the possible decay channels. In addition to hadronic decays, it is also necessary to simulate the emission of electromagnetic radiation. The most common approach adopted is to use algorithms similar to those used to simulate the emission of QCD in *parton shower*.

### 3.6 Jets Reconstruction

After the partonization and the decay of the unstable particles it is possible to estimate the quadrupole of the partons generated in the textit hard process by

the direction and energy of the textit jet that are reconstructed starting from the particles in the final state cite bib: run2jet, mass: in: dijet. The reconstruction of textit jet is entrusted to specific algorithms; these introduce the variable distance,  $d_{ij}$ , between two objects (particles or pseudo- textit jet) defined by,

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad (3.13)$$

where  $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  and  $k_{ti}$ ,  $y_i$  e  $\phi_i$  are the transverse momentum, la rapidity e the azimuthal angle of  $i$  respectively.  $R$  is the radial parameter. The distance between a  $i$  and a beam object is also introduced ,  $d_{iB} = k_{ti}^{2p}$

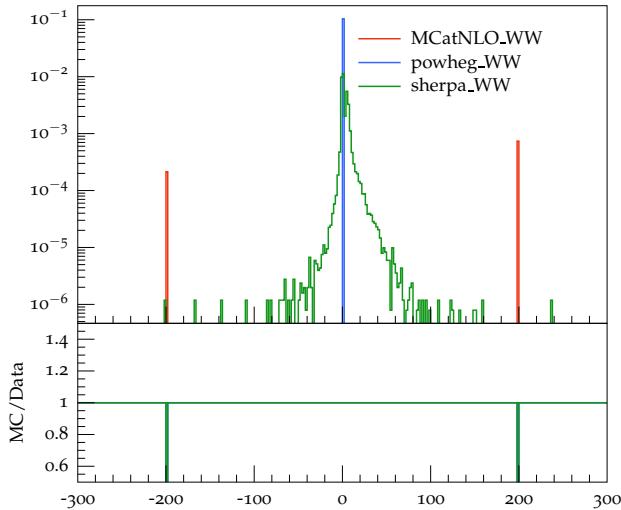
The algorithms proceed by calculating the distance  $d_{ij}$  and between all the pairs of particles  $i$ ,  $j$  identifying the minor. For the two particles with a smaller distance, the quadripulse are added. You also evaluate  $d_{iB}$  for every  $i$  and if it is less than the distance  $d_{ij}$  with all other particles  $j$ ,  $i$  is considered a jet and is removed from list of objects present in the event. Finally the distances are recalculated and this whole procedure is repeated until there are no more objects to be added. The value of  $p = -1$  defines the algorithm anti- $k_t$  [?], which is the one used, while the free parameter  $R$  has been set equal to 0.5.

### 3.7 Main Monte Carlo generators

For high energy physics, different Monte Carlo generators are available. Each of these has different methods for combining the ME with the PS. Here we will focus in particular on MADGRAPH interfaced with PYTHIA, POWHEG -anch'interface interfaced with PYTHIA - and SHERPA.

**Madgraph\_aMC@NLO** The MADGRAPH [?] approach is very ambitious, in fact the purpose of this generator is to calculate the cross section at the NLO including both real and virtual contributions in the calculation. The hard process is produced by the ME method while the soft emissions are combined with the PS. The first step is to compute the ME NLO corrections for a  $n$  parton process, including  $n + 1$  partons coming from the real corrections and  $n$  from the virtual ones. Next we evaluate how the parton shower populates the phase space at  $n + 1$  parton, excluding the Sudakov form factor at this stage. To get the true state where  $n + 1$  is present, MADGRAPH subtracts the PS expression from the  $n + 1$  state of the ME. The expression of the PS without a factor of Sudakov and of the ME are in agreement in the limit soft and collinear, so the singularities are deleted thus obtaining a finite value for the cross section in the case of  $n$  and  $n + 1$  partoni. A technical problem is that in the collinear limit there is no certainty that the ME always overhangs the PS. This problem is solved by introducing a fraction of negative-weight events, Fig. 3.9. Finally, the parton shower is applied, which includes the Sudakov factor and thus allows a finite and correct result to be obtained at the NLL.

**POWHEG** The idea behind POWHEG [?] is to generate the hardest radiation first, and then pass the event to the parton shower generator. In parton shower generators,



**Figure 3.9.** Weight distribution, for different Monte Carlo generators, with cross section normalization of  $1 \text{ fb}^{-1}$ .

the production, ordered in a transverse pulse, of the harshest radiation is always the first; so POWHEG simply replaces this with issuing to the NLO. In POWHEG events are produced with a positive and constant weight (Fig. 3.9).

**PYTHIA8** PYTHIA 8 [?] is a generator that can calculate the ME for processes with two particles or partons in the final state, but above all it generates the parton shower and the subsequent synchronization. The parton shower is ordered in a transverse impulse,  $p_T$ , and the first issue is corrected with the method of the reshuffle. For the use of the atomization, use the Lund model.

**SHERPA** SHERPA [?] is a Monte Carlo generator that as PYTHIA8 provides a complete description of hadronic collisions, from the calculation of the matrix element, up to the angle. The parton shower includes both QCD and QED emissions, i.e. photons. It can calculate the ME for the main processes (eg  $gg \rightarrow H$ ) at the NLO and combine the ME with the PS. The code is written completely in C ++ language.

### 3.8 Monte Carlo sample in High Mass Analysis

Concerning the simulated samples, several different Monte Carlo (MC) generators were used. In the simulation, ‘lepton’ includes also  $\tau$ . In order to perform the resonance search in a large part of the mass spectrum, several signal samples for the gluon-gluon fusion and the vector boson fusion mechanisms have been generated corresponding to different Higgs boson masses in the range between 200 GeV and 3 TeV. The signal lineshape for each mass point corresponds to the one expected for a SM Higgs boson at that mass. All signal samples, presented in Table , have been simulated with POWHEG v2 [3, 4, 5], designed to describe the full NLO

properties of these processes. In particular, for Higgs produced via gluon fusion [6], and vector-boson-fusion (VBF) [7], the decay of the Higgs boson into two W boson and subsequently into leptons was done using JHUGen v6.2.8 [8] for samples up to 300 GeV of mass and with v6.9.8 above that mass. The signals which correspond to a Higgs boson mass of 125 GeV have been simulated accordingly and are treated as backgrounds in the analysis, including the associated production with a vector boson ( $W^+H$ ,  $W^-H$ ,  $ZH$ ) [9], and gluon fusion produced  $ZH$  (ggZH). For associated production processes the Higgs boson decay was done via PYTHIA 8.1 [10].

The WW production, irreducible background for the analysis, was simulated in different ways. POWHEG v2 [11] was used for  $q\bar{q}$  induced WW in different decays. The cross section used for normalizing WW processes produced via  $q\bar{q}$  was computed at next-to-next-to-leading order (NNLO) [?]. In order to control the top quark background processes, the analysis is performed in jet bins. The jet binning enhances the importance of logarithms of the jet  $p_T$ , spoiling the convergence of fixed-order calculations of the  $qq \rightarrow WW$  process and requiring the use of dedicated resummation techniques for an accurate prediction of differential distributions [12, 13]. Since the  $p_T$  of the jets produced in association with the WW system is strongly correlated with its transverse momentum,  $p_T^{WW}$ , the simulated  $qq \rightarrow WW$  events are reweighted to reproduce the  $p_T^{WW}$  distribution from the  $p_T$ -resummed calculation.

Gluon fusion produced WW was generated, with and without Higgs diagrams, using MCFM v7.0 [14]. A  $t\bar{t}$  sample dilepton sample was also generated using POWHEG v2. The WW and  $t\bar{t}$  samples produced specifically for this analysis are presented in Table 3.1.

Process	Dataset Name	Events	$\sigma \times BR$ [pb]
$t\bar{t} \rightarrow WW bb \rightarrow 2l2\nu b\bar{b}$	TTTo2L2Nu_13TeV-powheg	5M	87.31
$q\bar{q} \rightarrow WW \rightarrow 2l2\nu$	WWTo2L2Nu_13TeV-powheg	2M	12.178
$q\bar{q} \rightarrow WW \rightarrow l\nu qq$	WpWmJJ-QCD-noTop_13TeV-powheg		
$gg \rightarrow WW \rightarrow 2l2\nu$	GluGluWWTo2L2Nu_MCFM_13TeV	500K	0.5905

**Table 3.1.** Simulated samples for  $t\bar{t}$  and WW production.

For the Drell-Yan backgrounds we use two different sets of samples. For the opposite flavor analysis (Sec 5.4), selecting events with an electron and a muon, a dedicated sample in which only the  $Z/\gamma^* \rightarrow \tau\tau \rightarrow e\mu\nu\nu$  decay is simulated. For the same flavor analysis (Sec. 5.5), in which pairs of electrons or muons are selected, a soup of different HT binned DY samples is used. A detailed study about this soup is given below in Sec. 3.8.1.

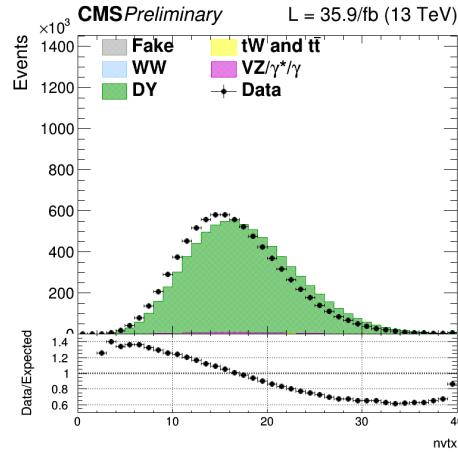
All processes are generated using the NNPDF3.0 [15, 16] parton distribution functions (PDF) for NLO generators, while the LO version of the same PDF is used for LO generators. All the event generators are interfaced to PYTHIA 8.1 [10] for the showering of partons and hadronization, as well as including a simulation of the underlying event (UE) and multiple interaction (MPI) based on the CUET8PM1 tune [17].

For all processes, the detector response is simulated using a detailed description of the CMS detector, based on the GEANT4 package [18]. The MC samples used are part of the RunIISummer16MiniAODv2 campaign with the global tag:

*PUMoriond17\_80X\_mcRun2\_asymptotic\_2016\_TrancheIV\_v6*

and the CMSSW version used is *8\_0\_26\_patch1*.

The simulated samples are generated with distributions for the number of pileup interactions that are meant to roughly cover, though not exactly match, the conditions expected for the different data-taking periods. In order to factorize these effects, the number of true pileup interactions from the simulation truth (as stored in the PileupInfo collection in the Monte Carlo) is reweighted to match the data. The re-weighting is propagated automatically to both the in-time pile up and the out-of-time one. In Figure 3.10, the effect of this reweighting on a sample enriched in Drell-Yan events is shown. In order to select this sample, events with two electrons with  $p_T > 25$  GeV for the leading one and  $p_T > 13$  GeV for the trailing one, are selected only if  $|m_{\ell\ell} - m_Z| < 10$  GeV.



**Figure 3.10.** Distributions of the number of vertices in a Drell-Yan enriched sample ( $Z \rightarrow ee$ ) in data

The pileup histogram for reweighting is calculated using the *pileupCalc* tool as described in [19].

Different sources and calculations are used to obtain the cross sections for the different processes at 13 TeV. For Higgs signals, the cross sections used are the ones reported by the LHC Higgs Cross Section Working Group [20], computed at NNLO and NNLL QCD and NLO EW for gluon fusion, and at NNLO QCD and NLO EW for the rest of the production modes. The branching fractions are the ones reported in [21].

The cross section used for normalizing  $q\bar{q}$ -produced WW processes was computed at next-to-next-to-leading order (NNLO) [?]. The leading-order (LO) cross section for ggWW is obtained directly from MCFM. For gluon fusion, the difference between LO and NLO cross sections is significantly big. A scale factor of 1.4 is theoretically calculated [?] and applied to the gg → WW background.

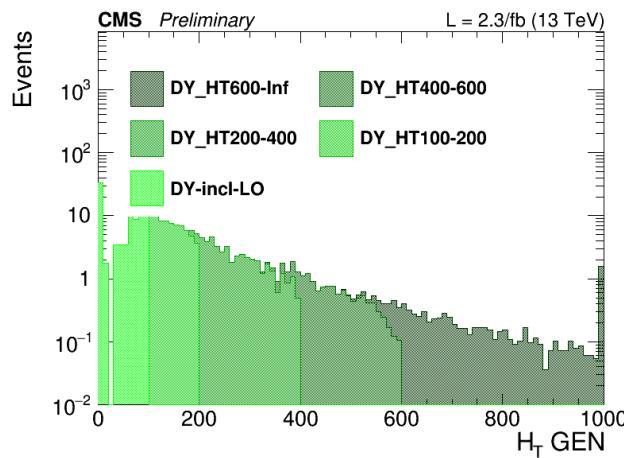
The cross sections of the different single top processes are estimated by the LHC Top Working group [22] at NLO. The  $t\bar{t}$  cross section is also provided by the LHC Top Working group [23], and it is computed at NNLO, with NNLL soft gluon resummation.

Drell-Yan (DY) production of  $Z/\gamma^*$  is generated using aMADGRAPH [24] and the cross section is scaled using a LO to NNLO k-factor equal to 1.23. Other multiboson processes, such as WZ, ZZ, and VVV ( $V=W/Z$ ), are generated with amc@NLO and normalized to the cross section obtained at NLO in generation. The cross sections for the remaining processes were directly obtained using the *GenXSecAnalyzer* tool [25] or from the Twiki presented in Ref. [26].

All processes are generated using the NNPDF2.3 [15, 16] parton distribution functions (PDF) for NLO generators, while the LO version of the same PDF is used for LO generators. All the event generators are interfaced to PYTHIA 8.1 for the showering of partons and hadronization, as well as including a simulation of the underlying event (UE) and multiple interaction (MPI) based on the CUET8PM1 tune [17].

### 3.8.1 The DY sample

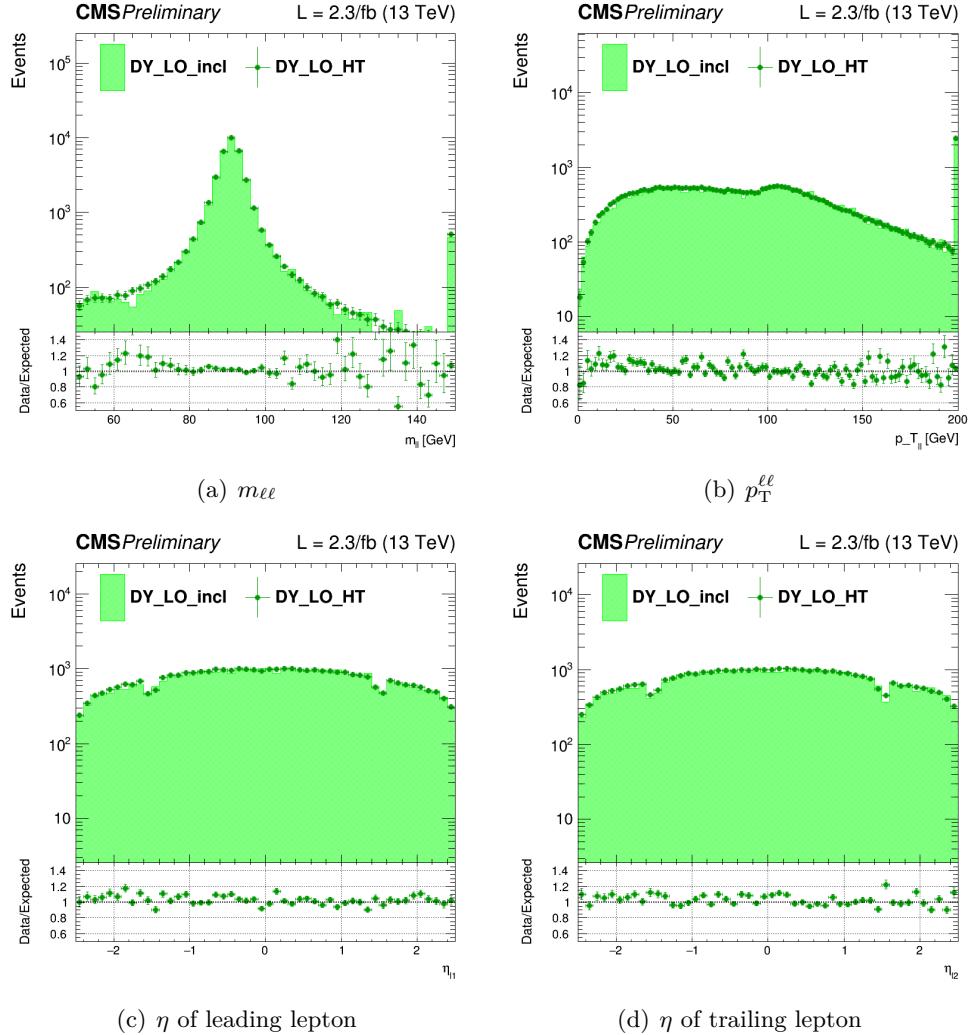
Given the lack of MC statistics in the LO inclusive DY sample the  $H_T$ -binned samples are used. This helps increasing the MC statistics especially in the VBF category of the same flavor analysis, which is characterized by large values of  $H_T$ . The LO inclusive sample is used for events with  $H_T < 100$  GeV and it has been merged to the other samples selecting events with  $H_T$  below 100 GeV using the parton level information. The cross sections of those samples have been scaled applying the LO to NNLO k-factor. In Fig. 3.11 the  $H_T$  distribution of the sample after the merging is reported, showing a smooth transition between different  $H_T$  samples.



**Figure 3.11.**  $H_T$  distribution for the merged DY sample.

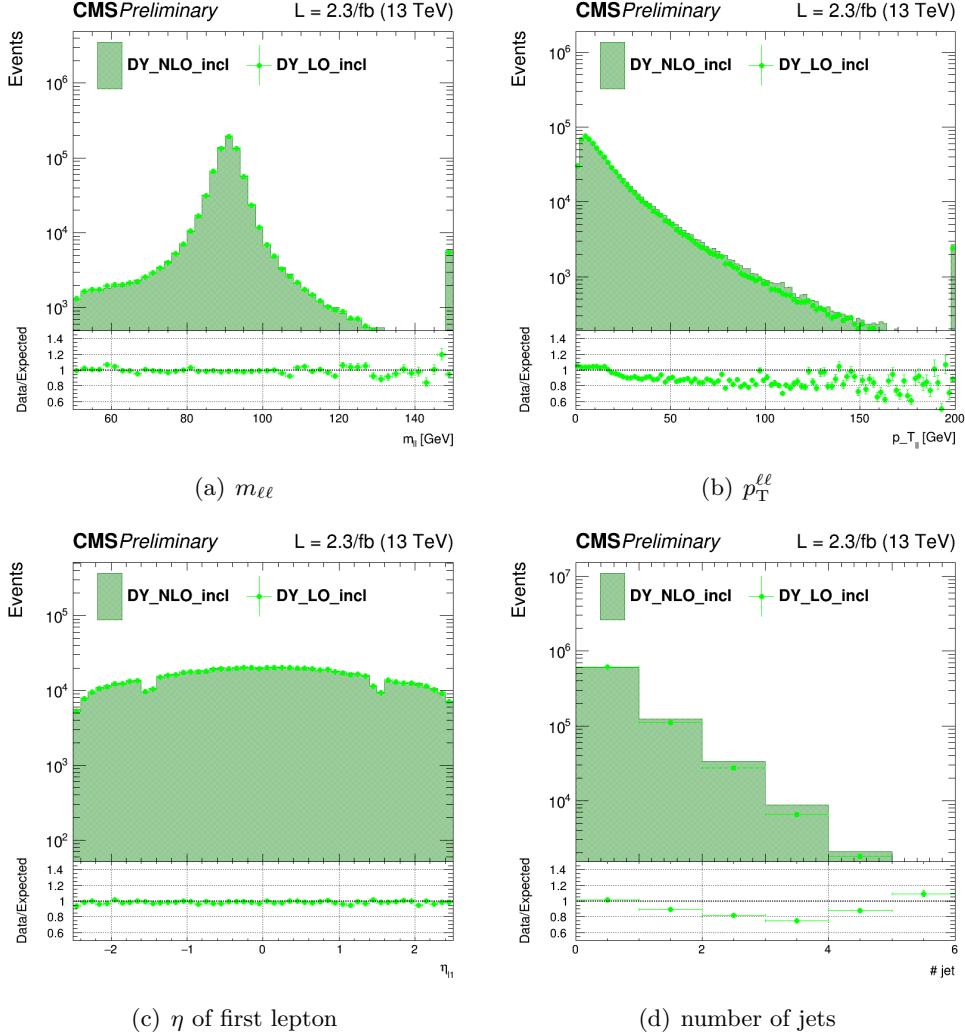
To further check the correct behaviour of the  $H_T$  binned samples we compared them to the inclusive LO sample, selecting only the events with a generator level  $H_T$  above 100 GeV. The comparison is done in a control region with two same flavor leptons with  $p_T > 20$  GeV and  $m_{\ell\ell} > 50$  GeV, showing very good agreement between the two samples. The distributions of some variables are shown in Fig. 3.12

To check the differences between the LO inclusive sample and the NLO sample simulated with MC@NLO, the two samples have been compared in a same flavor



**Figure 3.12.** Comparison between the inclusive LO DY sample and the  $H_T$  binned samples.

control region and some variables of interest are shown in Fig. 3.13. The control region is defined requiring two same flavor leptons with  $p_T > 20 \text{ GeV}$  and with  $m_{\ell\ell} > 50 \text{ GeV}$ .



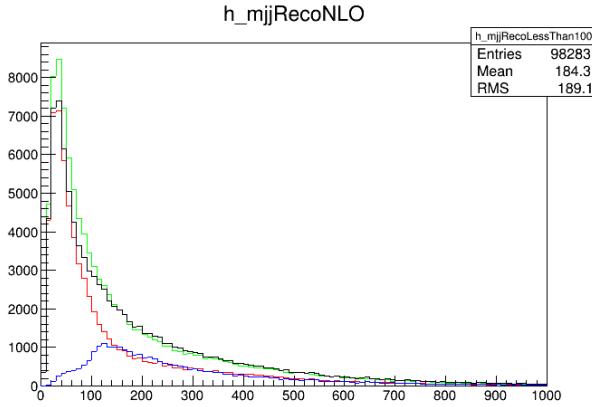
**Figure 3.13.** Comparison between the LO and NLO DY samples.

### 3.8.2 The WW sample

In the analysis two different WW Monte Carlo samples are merged: the “ $WW \rightarrow 2l2\nu$  NLO” and the “WW plus 2 jet” LO (WpWmJJ-QCD-noTop in Tab. 4).

The second sample, “WW plus 2 quark”, contains final state with two quarks or a gluon-quark system: only the final state with two quarks interferes with the signal. To avoid double count between the two samples a cut on di-jet mass at gen-level,  $m_{jjGenLev}$ , is applied. In particular the sample “ $WW \rightarrow 2l2\nu$  at NLO” is used for  $m_{jjGenLev} < 100 \text{ GeV}$  and the “WW plus 2 quark” for  $m_{jjGenLev} > 100 \text{ GeV}$ .

The distribution for the reco di-jet mass is shown in Fig.3.14. In particular the red distribution correspond the “ $WW \rightarrow 2l2\nu$  NLO” sample with a cut of  $mjj_{GenLev} < 100$ , the blue distribution to “ $WW$  plus 2 quark” with  $mjj_{GenLev} > 100$  GeV. The sum of the red and blue distributions is shown in black. There is a good agreement between with the black distribution and the “ $WW \rightarrow 2l2\nu$  NLO” without any  $mjj_{GenLev}$  distribution , in green.



**Figure 3.14.** Distribution for  $m_{jj}$  at RECO level for the merged WW sample.

## Chapter 4

# Event Reconstruction



## Chapter 5

# High mass resonances searching

### 5.1 Introduction

The search for a new resonance  $X$  is described in this chapter.

The main production mode for the Higgs boson particle over the all mass spectrum is the gluon-gluon fusion (ggH) process. At a center-of-mass energy of 13 TeV the ggH cross section for a Higgs boson mass ( $m_H$ ) of 125 GeV is 43.92 pb [20], that is almost one order of magnitude larger than the second process in terms of cross section at that mass, VBF, with 3.748 pb [20]. The gluon-gluon fusion cross section decreases with  $m_H$  but the VBF/ggH cross section ratio increases with the mass, making the VBF production mechanism more and more important as  $m_H$  approaches to high values.

The signal samples are interpreted in terms of the EWK singlet model described in Sec 5.3 below. The Higgs boson width and lineshape is reweighted at generator level according to the parameters defined in the model. The interference effects between the ggH signal, the ggWW background and SM Higgs boson, that are expected to slightly change the lineshape of the signal distribution, have been fully taken into account, as detailed in Sec. 5.3. A similar treatment is also applied for the interreference between the VBF high mass signal, the VBF SM Higgs and the quark initiated WW+2 quarks backgrround. The interference between the  $W^+W^- \rightarrow 2\ell 2\nu$  and  $ZZ \rightarrow 2\ell 2\nu$  is negligible due to the different phase space characteristic of these processes.

The analysis strategy for the high mass search with 2016 data in the  $W^+W^- \rightarrow 2\ell 2\nu$  decay channel is similar to the previous high mass analysis with 2015 data [27], but has several improvements.

The analysis is divided in two parts:

- the opposite-flavour final state,  $e^\pm\mu^\mp$ ,
- the same-flavour final state,  $e^+e^-$  and  $\mu^+\mu^-$ .

In the opposite-flavour final state four different jets-categories are defined: the 0-jet, the 1-jet, the 2-jet non VBF and finally the VBF. The 2-jet non-VBF category is new with respect to previous analysis with 2015 data.

In the same-flavour final state only the VBF category is considered. Indeed, the

only the VBF selection cuts are sufficiently tight to reduce the overwhelming Z+jets background to a manageable level.

## 5.2 Discriminating variable

This analysis is a shape analysis, meaning that after applying selection cuts we do not simply count events, but rather we fit a data histogram of a discriminating variable with the sum of signal and background templates, and extract the signal yield from the fit. The variable with the best discriminating value would be the invariant mass of the four lepton, which is not possible to reconstruct in the WW channel due to neutrinos.

For the Higgs boson in WW analysis, the shape analysis is based on two-dimensional templates of  $m_{\ell\ell}$  versus  $m_T^H$ , where the transverse mass  $m_T^H$  variable is defined as

$$m_T^H = \sqrt{2p_T^{\ell\ell}E_T^{\text{miss}}(1 - \cos\Delta\phi(\ell\ell, \vec{p}_T^{\text{miss}}))} \quad (5.1)$$

where  $\Delta\phi(\ell\ell, \vec{p}_T^{\text{miss}})$  is the azimuthal angle between the dilepton momentum and  $\vec{p}_T^{\text{miss}}$ .

However  $m_T^H$  (and also  $m_{\ell\ell}$ ) is not very sensitive to the signal mass hypothesis, so a new variable  $m_T^I$  defined as the visible mass,

$$m_T^I = \sqrt{(p_{\ell\ell} + E_T^{\text{miss}})^2 - (\vec{p}_{\ell\ell} + \vec{p}_T^{\text{miss}})^2} \quad (5.2)$$

has been introduced to discriminate better the high mass  $X$  signals generate at different masses. The distribution of the variables defined above are shown in Fig. 5.1, where it is visible the better power of  $m_T^I$  in discriminating different mass hypotheses respect the other variable.

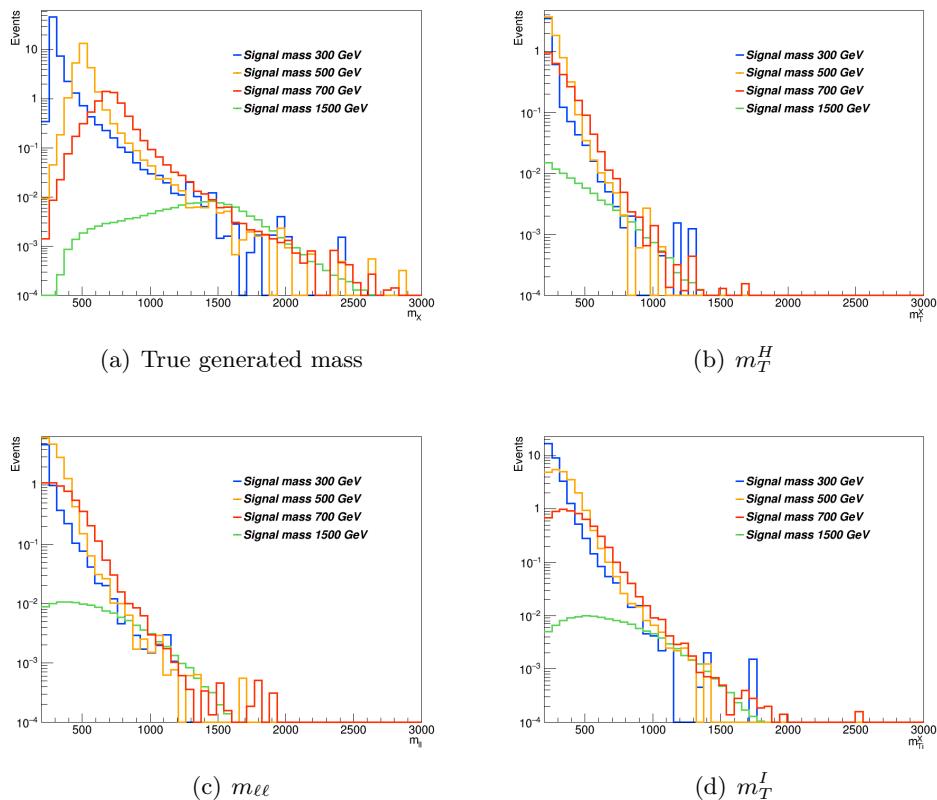
## 5.3 Signal interpretation

The signal is interpreted in terms of the electroweak singlet model, representing a scalar mixing with the 125 GeV Higgs boson. This model relies on two parameters: the scale factor of the couplings of the high mass resonance with respect to the SM,  $C'$ , and the branching fraction of the electroweak singlet to non-SM decays modes,  $BR_{\text{new}}$ . The electroweak singlet signal strength,  $\mu'$  and the modified width,  $\Gamma'$ , are related with the parameters in the model by the following equations:

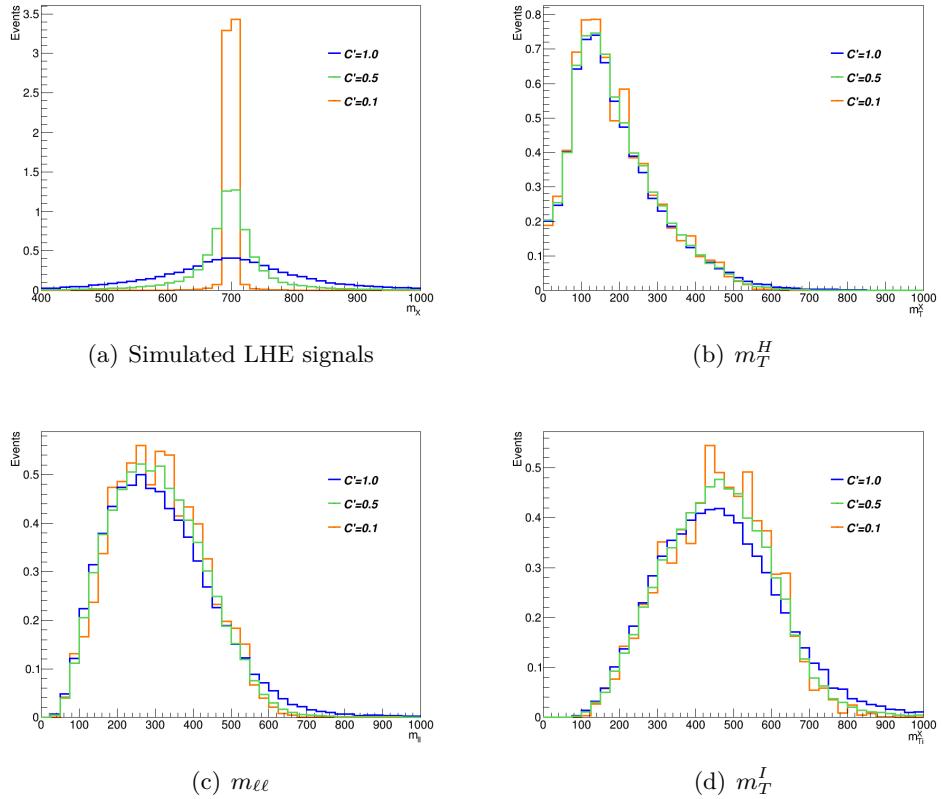
$$\mu' = C'^2 \cdot (1 - BR_{\text{new}}) \quad (5.3)$$

$$\Gamma' = \Gamma_{\text{SM}} \cdot \frac{C'^2}{1 - BR_{\text{new}}} \quad (5.4)$$

The available Higgs signal samples for different mass hypothesis have been reweighted according to this model. At the moment only the  $BR_{\text{new}} = 0$  hypothesis has been investigated while we tested different  $C'$  values. In Fig. 5.2 are shown the  $m_{\ell\ell}$  and  $m_T$  templates corresponding to a Higgs boson mass of 700 GeV for three different  $C'$  values:  $C' = 1$ , corresponding to the SM Higgs decay width,  $C' = 0.5$ , corresponding to  $\Gamma' = 2.5 \cdot 10^{-2} \Gamma_{\text{SM}}$ , and  $C' = 0.1$ , corresponding to  $\Gamma' = 10^{-2} \Gamma_{\text{SM}}$ . A value of  $BR_{\text{new}} = 0$  is considered in all cases. We note that the signal shape is not very sensitive to different  $C'$  values.



**Figure 5.1.** Distributions of the generated mass,  $m_T^H$ ,  $m_{\ell\ell}$  and  $m_T^I$  variables for different  $X$  mass hypothesis.



**Figure 5.2.** Distributions of the signals, the  $m_T^H$ , the  $m_{\ell\ell}$  and the  $m_T^I$  variables at generator level for different values of  $C'$ , without any selection.

### Study of the Interference effects

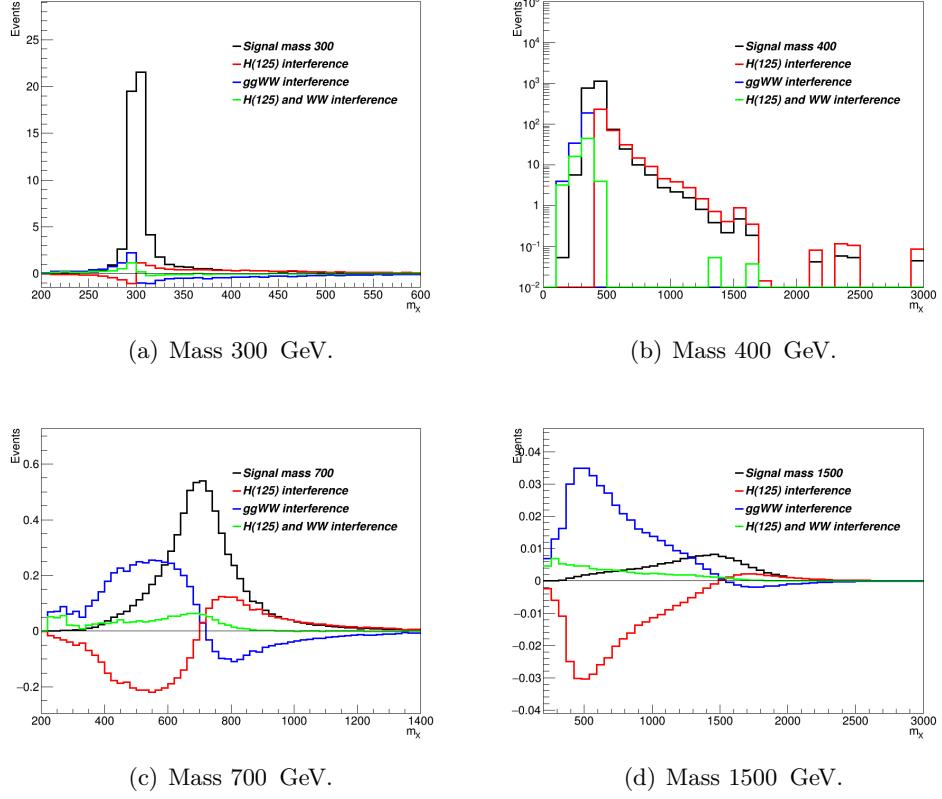
When a resonance  $X$ , with a non negligible width is considered, it is important to take into account also the interference effects both with the WW background , with same initial and final state, and with the Higgs boson off-shell tail.

In this analysis we take into account the interference effects between the new signal  $X$  produced in gluon-gluon fusion and in vector-boson-fusion. The effect of the various interference terms are shown in 5.3 and 5.4 for the two different production mechanism, gluon-gluon fusion and vector-boson.fusion. The contribution of the interference of  $X$  with WW background and with Higgs boson have opposite sign and partially cancel out. This cancellation effect is different for different resonance masses. The interference contribution is thus non negligible and is included in the fit.

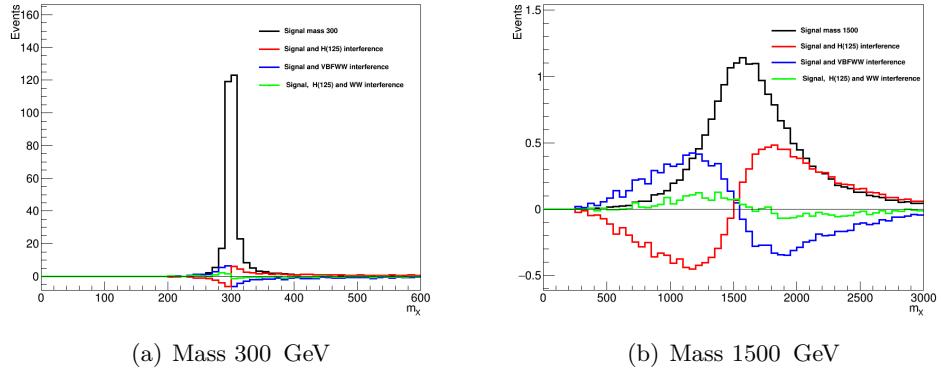
To prevent possible negative probability distribution function of the interference, during the fit the signal yield is computated as,

$$Yield = \sqrt{\mu} \times (S + B + I) + (\mu - \sqrt{\mu}) \times (S) + (1 - \sqrt{\mu}) \times (B) \quad (5.5)$$

where  $S$  is the signal,  $B$  the background and  $I$  the interference.



**Figure 5.3.** Distribution of for the  $X$  mass resonance, produced via gluon-gluon fusion for different masses. In black the high mass signal. In red the interference between the high mass signal and the Higgs boson. In blue the interference between the high mass signal and the background. In green the total interference i.e. high mass signal, Higgs bison and background.



**Figure 5.4.** Distribution of for the  $X$  mass resonance, produced via vector-boson-fusion fusion for different masses. In black the high mass signal. In red the interference between the high mass signal and the Higgs boson. In blue the interference between the high mass signal and the background. In green the total interference i.e. high mass signal, Higgs bison and background.

## 5.4 Opposite Flavor final state

In this section the analysis for the opposite-flavour final state  $W^+W^- \rightarrow \mu^\pm e^\mp 2\nu$  is described.

### Signal region

The events are requested to pass single or double lepton triggers, and exactly one electron and one muon are requested to be reconstructed in the event. One of the two leptons is requested to have a  $p_T$  greater than 25 GeV, the other is requested to have  $p_T$  greater than 20 GeV and both leptons are requested to be well identified and isolated, to reject non-prompt leptons and leptons coming from QCD sources. To suppress background processes with three or more leptons in the final state, such as ZZ, WZ, Z $\gamma$ , W $\gamma$  or triboson production, no additional identified and isolated lepton with  $p_T > 10$  GeV should be reconstructed. The low dilepton invariant mass region dominated by QCD production of leptons is not considered in the analysis and  $m_{\ell\ell}$  is requested to be higher than 50 GeV to reduce the SM Higgs boson ( $m_H=125$  GeV) contamination. A moderate MET cut is applied  $\text{MET} > 20$  GeV due to the presence of neutrinos in the final state searched for. Since a High mass signal is searched for, an  $m_T^I > 100$  GeV is applied. A cut on the transverse momentum ( $p_T^{\ell\ell} > 30$  GeV) and on the  $m_T^H > 60$  GeV are applied against  $DY \rightarrow \tau\tau$  background. Finally, against the top background, all jets above 20 GeV are requested not to be identified as b-jets according to the cMVAv2 tagger, loose WP. This is the full selection, defined as the “WW OF selection” :

- Two isolated leptons with different charge and flavor ( $\mu^\pm e^\mp$ );
- $p_T$  of the leading lepton  $> 25$  GeV;
- $p_T$  of the trailing lepton  $> 20$  GeV;
- Third lepton veto: veto events if a third lepton with  $p_T > 10$  GeV;
- $m_{\ell\ell} > 50$  GeV, to reduce H(125) contamination;
- MET  $> 20$  GeV;
- $m_T^I > 100$  GeV;
- $p_T^{\ell\ell} > 30$  GeV;
- $m_T^H > 60$  GeV;
- no b-tagged (cMVAv2 loose WP) jets with  $p_T > 20$  GeV;

Events passing the “WW OF selection” are categorized according to the jet multiplicity, counting jets above 30 GeV, to enhance the sensitivity, especially against the top background.

- **0 jet**, no jets are required in the event;

- **1 jet**, exactly 1 jet is required in the event;
- **2 jet**, exactly 2 jets are required in the event and in addition the condition  $\Delta\eta_{jj} < 3.5$  **or**  $m_{jj} < 500$  GeV;
- **VBF**, exactly 2 jets are required in the event and in addition the condition  $\Delta\eta_{jj} > 3.5$  **and**  $m_{jj} > 500$  GeV;

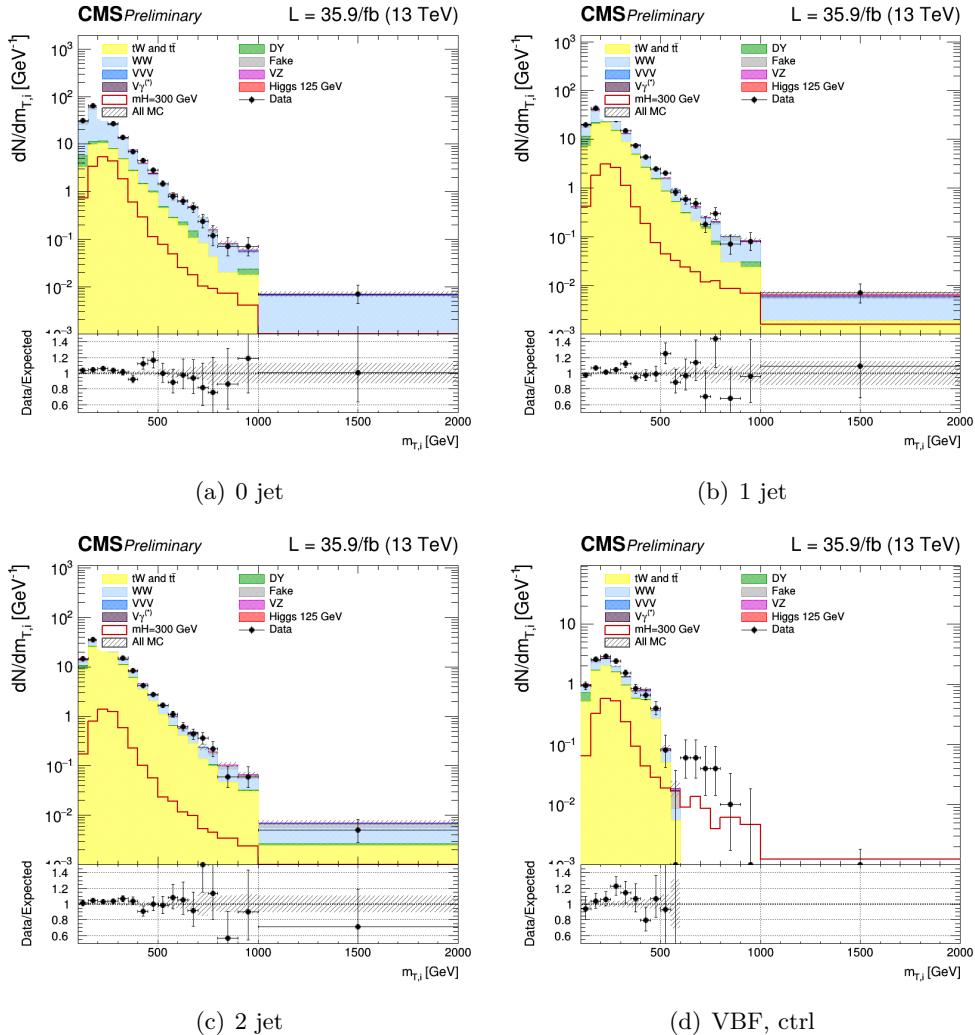
where the 2 jet and VBF regions are mutually exclusive by construction.

To extract high mass boson signals in these four categories, the strategy is followed: the  $m_T^I$  distribution is fitted as the sum of signal and background templates. Different binnings have been chosen for the  $m_T^I$  distributions in the different categories. The binning was chosen to have at least 10 top Monte Carlo events in each bin of the template. The chosen bins are:

- **0/1/2 jet**, [100,150,200,250,300,350,400,450,500,550,600,650,700,750,800,900,1000,2000]
- **VBF**, [100,150,200,250,300,350,400,500,700,1000,2000]

where the first number represents the lower edge of the first bin while the other numbers represent the upper edges. The last bin is an overflow bin.

The  $m_T^I$  distributions for the signal regions are presented in the four categories in Figs. 5.5.



**Figure 5.5.** Unblinding distributions  $m_T^I$  in the signal region for 0, 1, 2 and VBF categories. The signal hypothesis corresponding to  $m_X$  of 300 GeV.

## Background estimation

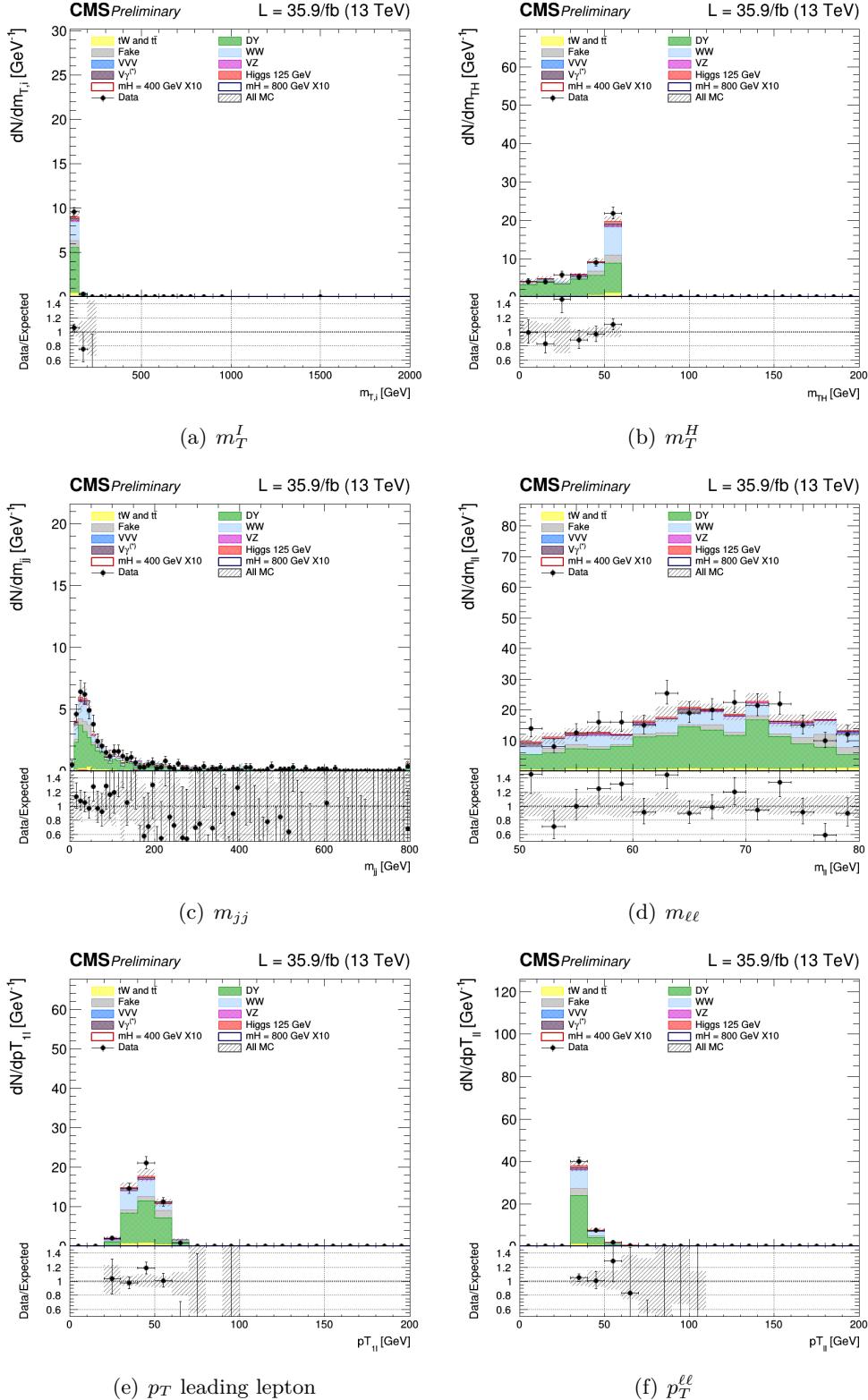
The main background processes that affect this signature arise from non-resonant WW production and from top production, including  $t\bar{t}$  pairs and single top production (mostly  $tW$ ), and are estimated using data. Instrumental backgrounds arising from non-prompt leptons in W+jets production and mis-measurement of  $E_T^{miss}$  in Drell-Yan events are also estimated from data. The contribution from  $W\gamma^*$  is estimated partly from data. The contribution of other sub-dominant backgrounds is obtained directly from simulated samples. The different data-driven background estimations are explained in the following subsections. More precisely top and DY backgrounds normalizations have been extracted directly from data-simulation comparison in specific control regions enriched in either one or the other background separately for the 0, 1, 2 and VBF jet categories, using the rateParam feature of the combine package [ref].

### Drell-Yan $\tau\tau$ control region

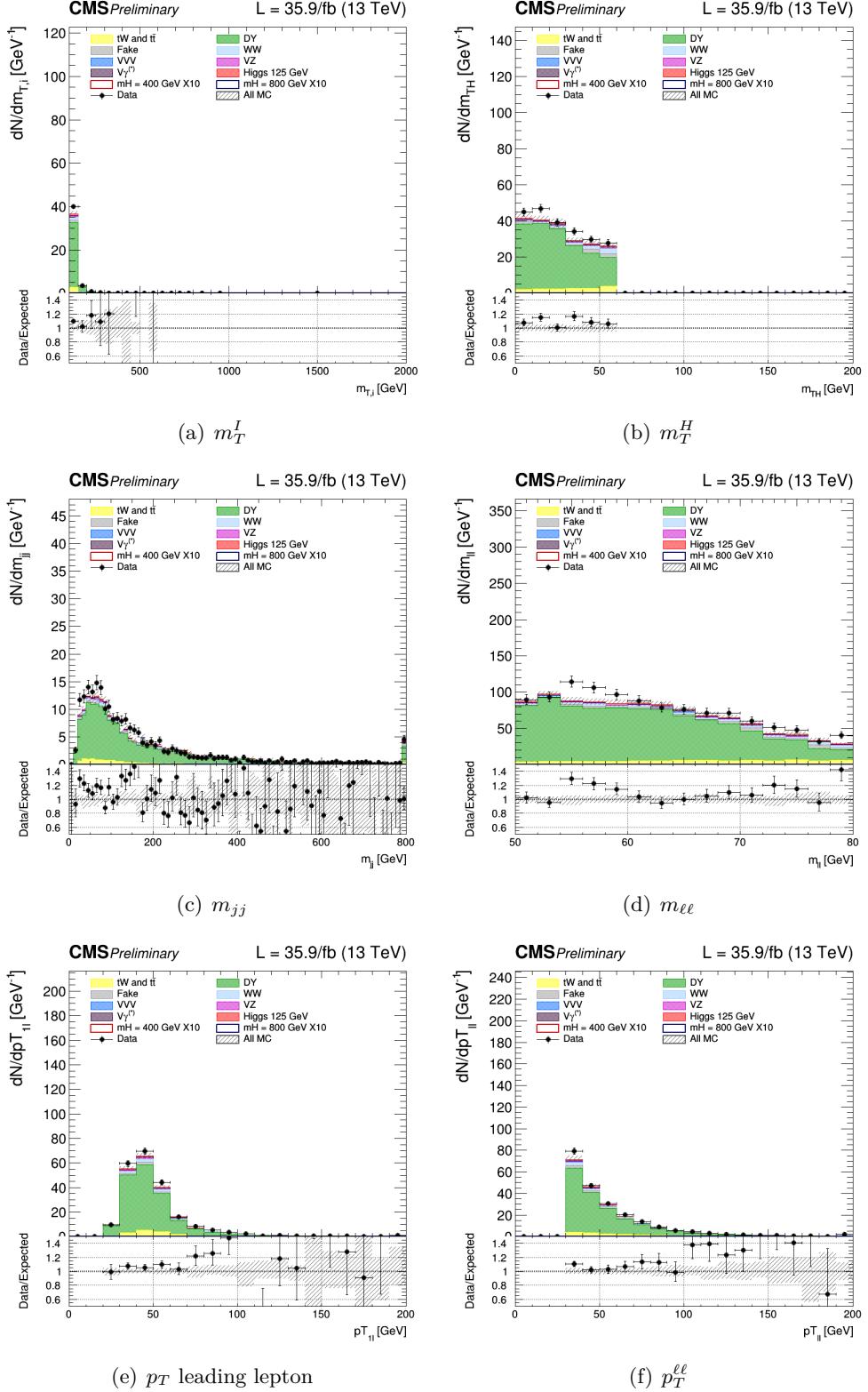
To normalize the Drell-Yan  $\tau\tau$  background to the data, control regions have been defined, as close as possible to the signal region, but enriched in  $Z \rightarrow \tau^+\tau^-$ . In particular, the “WW OF selection” is used with inverted  $m_T^H$  cut, i.e.  $m_T^H < 60$ . In addition a cut on the invariant mass of the two leptons  $50 \text{ GeV} < m_{\ell\ell} < 80 \text{ GeV}$  is requested to exclude possible contribution from non-prompt leptons (low limit) and from  $t\bar{t}$  (high limit).

For each signal category, a corresponding Drell-Yan  $\tau\tau$  control regions is defined. We thus have 4 total Drell-Yan  $\tau\tau$  control regions, for 0 jets, 1 jets, 2 jets and VBF.

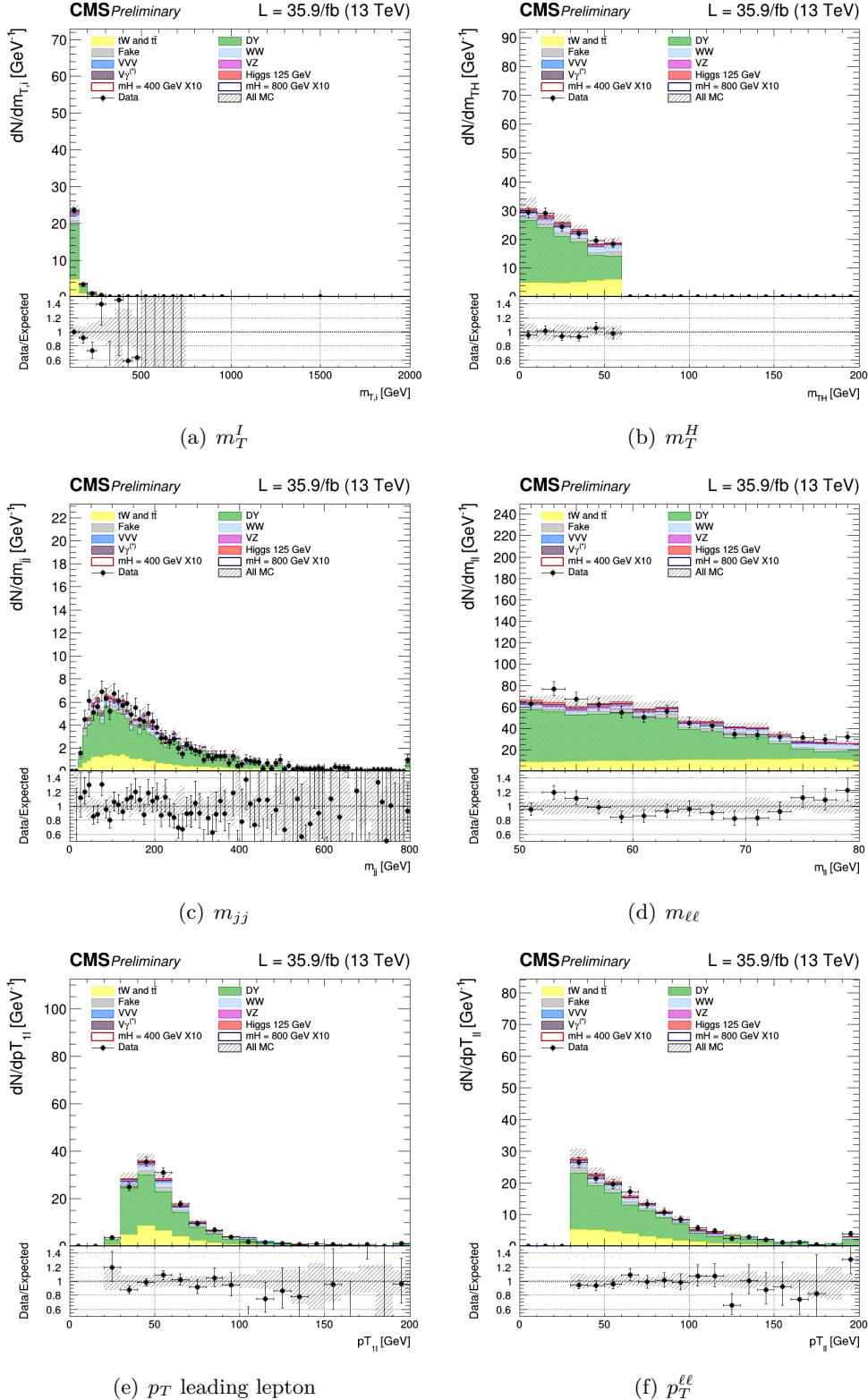
The control plots for several variables in a Drell-Yan enriched phase space for the four jets categories are shown in Figs. 5.6, 5.7, 5.8, 5.9. In general there is a good agreement between data and MC.



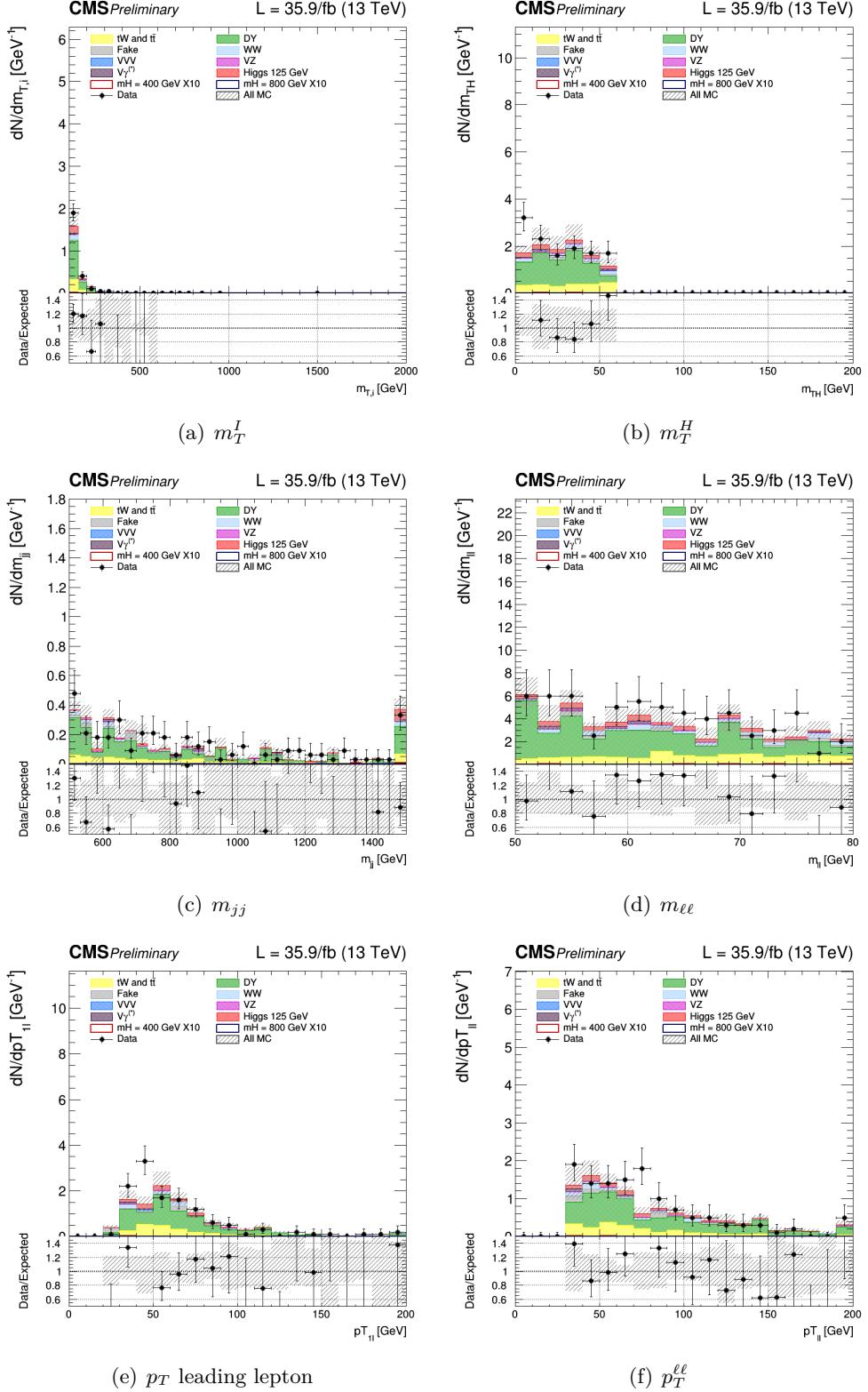
**Figure 5.6.** Control plots for several variables in a Drell-Yan enriched phase space for events with 0 jet.



**Figure 5.7.** Control plots for several variables in a Drell-Yan enriched phase space for events with 1 jet.



**Figure 5.8.** Control plots for several variables in a Drell-Yan enriched phase space for events with 2 jet.



**Figure 5.9.** Control plots for several variables in a Drell-Yan enriched phase space for events for VBF.

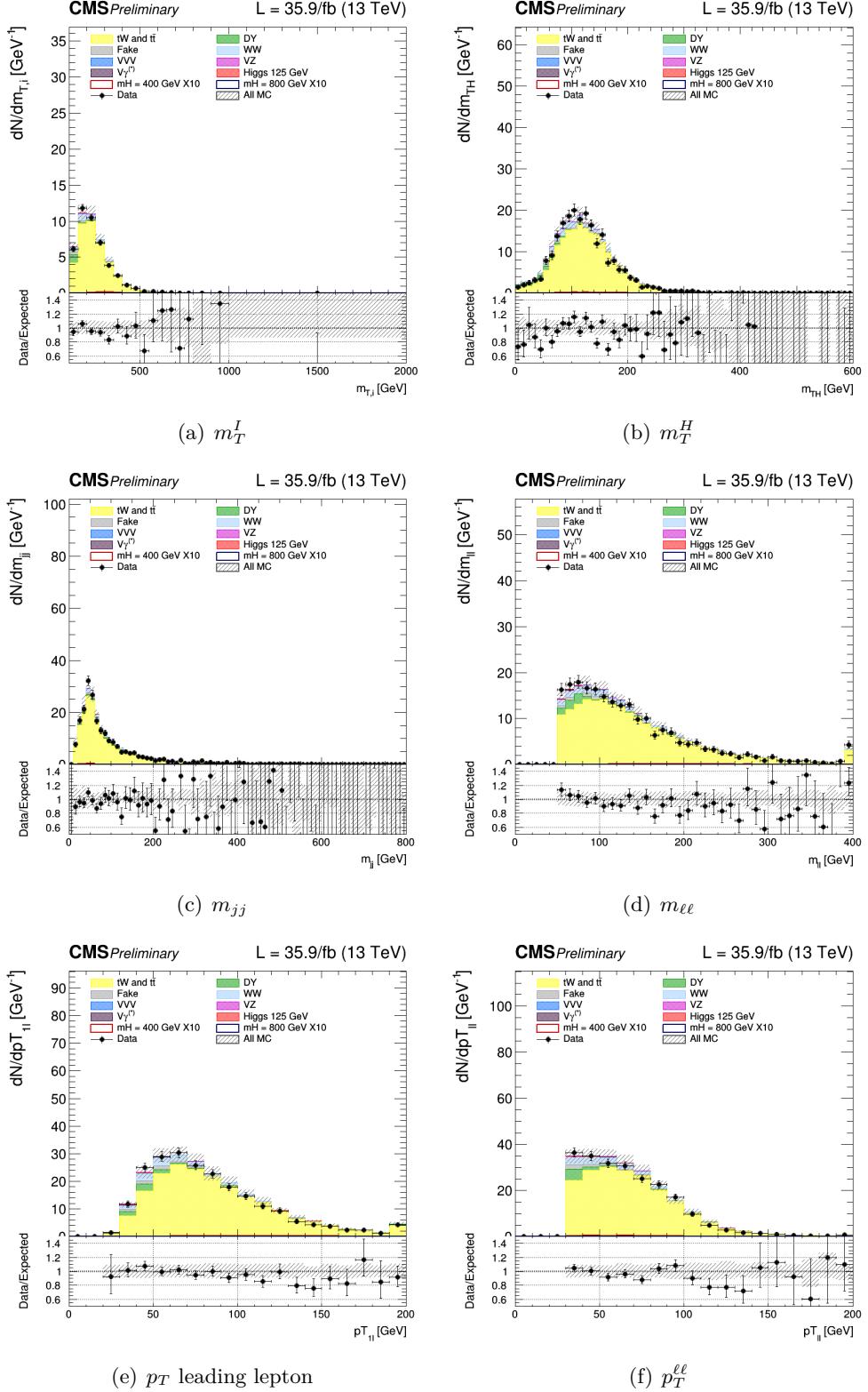
## Top control region

Similarly to the Drell-Yan  $\tau\tau$  case, control regions are defined for the Top background, and they are used to normalize the top background to data. The “WW OF selection” is used with inversion of the veto on b-jets. In particular the following conditions are imposed to select a top enriched control region for each of the 4 signal regions:

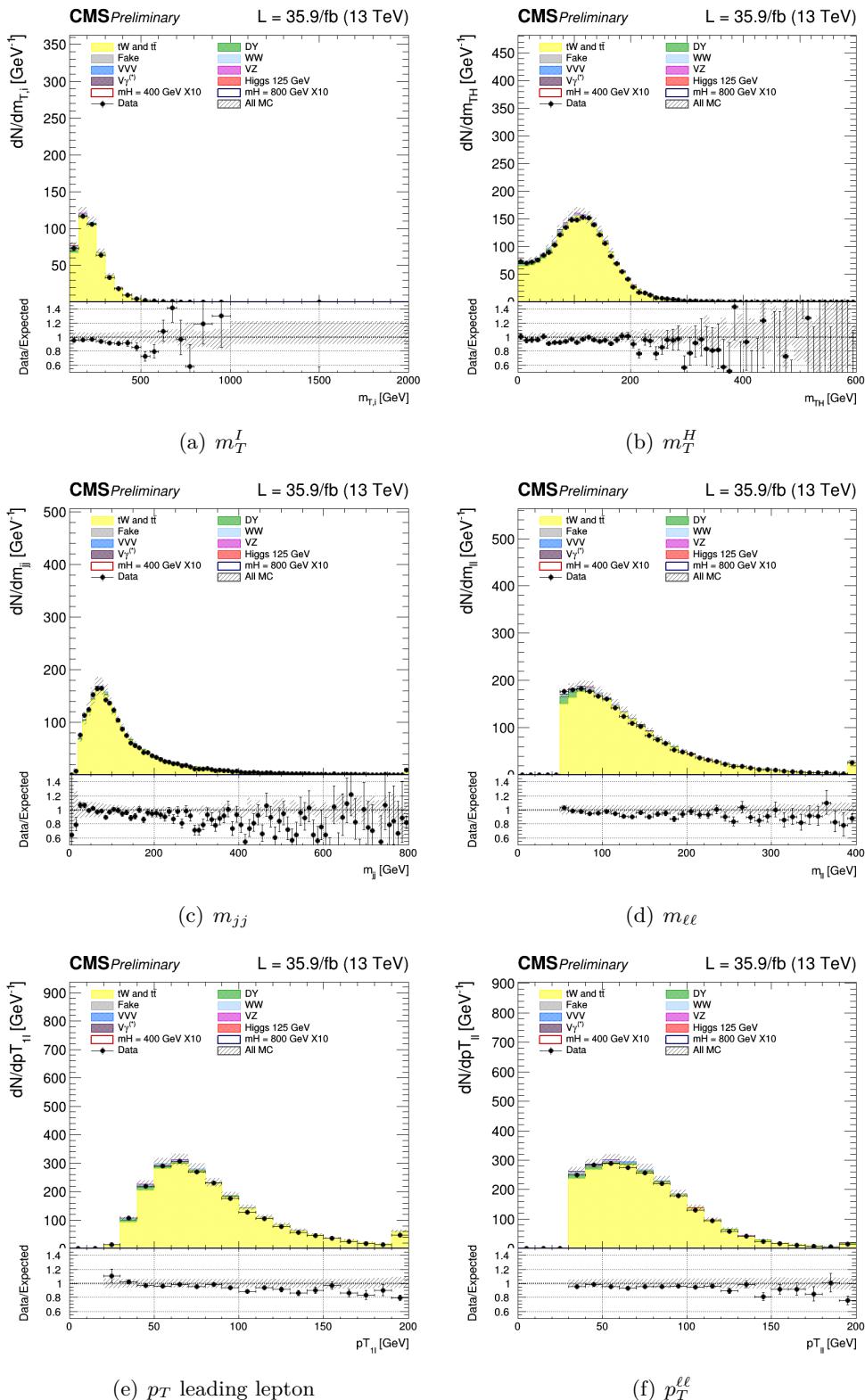
- **0 jet**, at least one b-tagged jet with  $20 < p_T < 30$  GeV is required;
- **1 jet**, exactly one b-tagged jet with  $p_T$  above 30 GeV is required;
- **2 jet**, exactly 2 jets with at least one of them b-tagged and in addition the condition  $\Delta\eta_{jj} < 3.5$  **or**  $m_{jj} < 500$  GeV;
- **VBF**, exactly 2 jets with at least one of them b-tagged and in addition the condition  $\Delta\eta_{jj} > 3.5$  **and**  $m_{jj} > 500$  GeV.

A jet is considered b-tagged if its cMVAv2 score is above the threshold defining the loose working point.

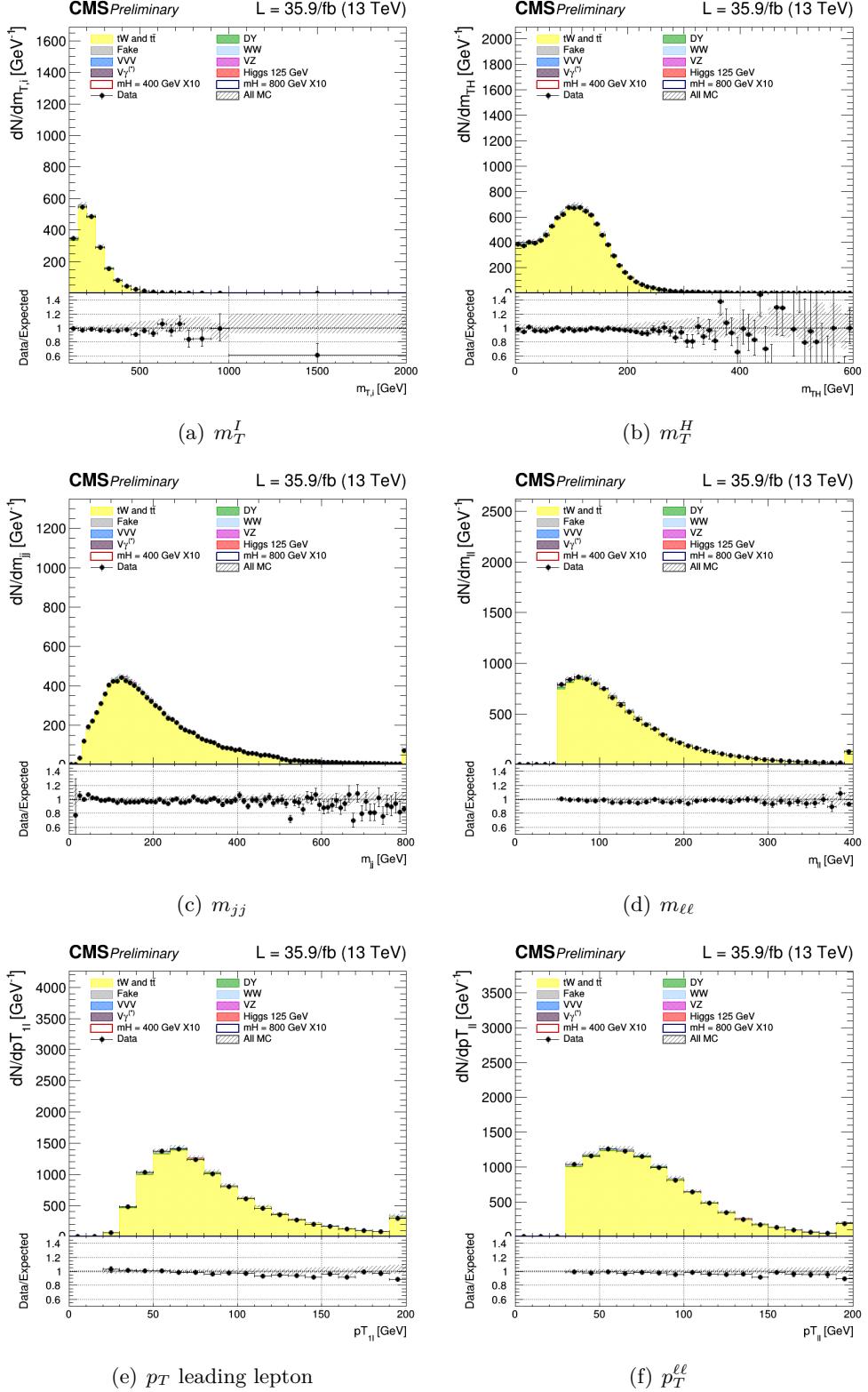
The control plots for several variables in a top enriched phase space for events are shown in the Fig. Figs. 5.10, 5.11, 5.12, 5.13. The last bin in the distribution is the overflow.



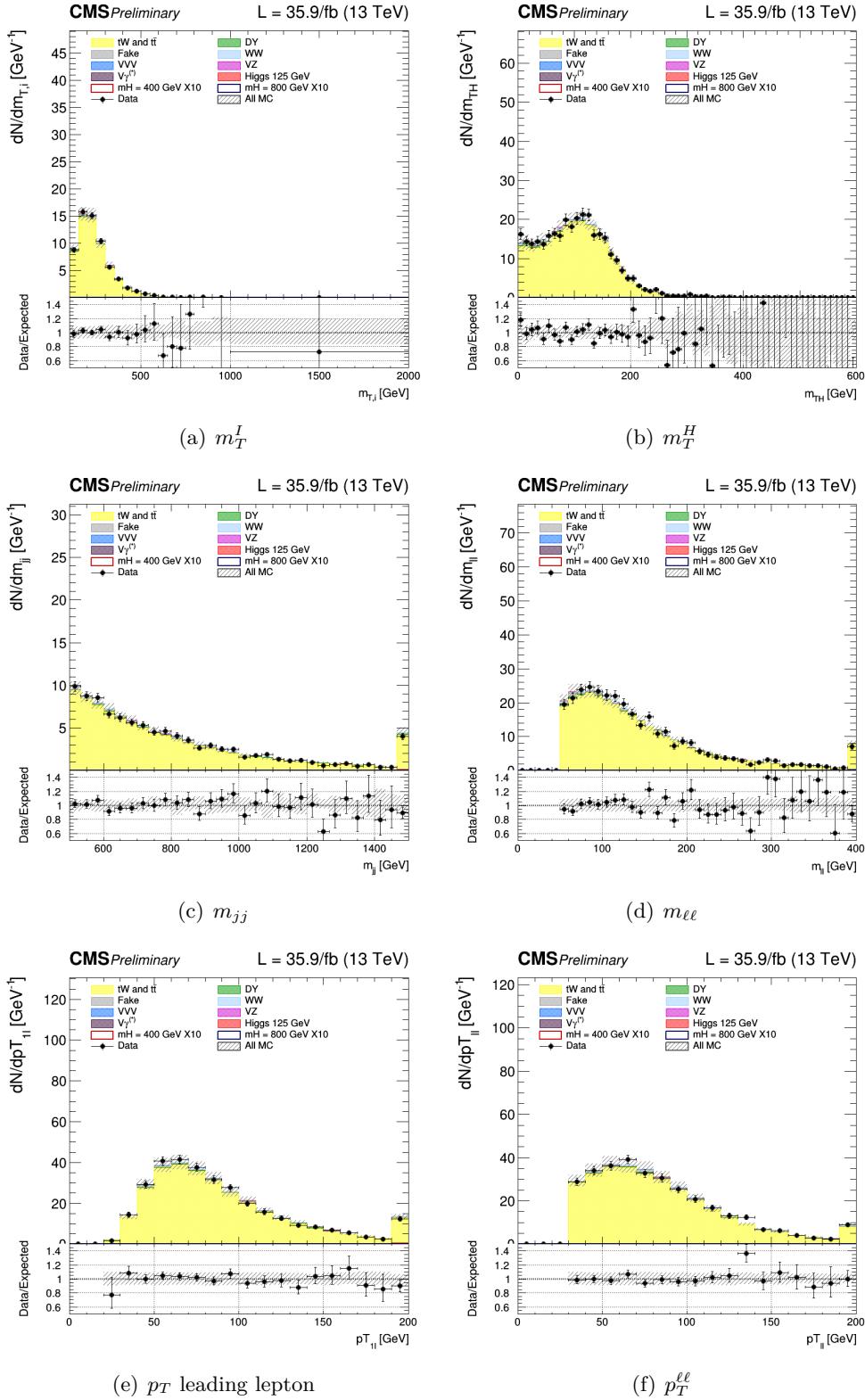
**Figure 5.10.** Control plots for several variables in the Top enriched phase space for events with 0 jet.



**Figure 5.11.** Control plots for several variables in the Top enriched phase space for events with 1 jet.



**Figure 5.12.** Control plots for several variables in the Top enriched phase space for events with 2 jet.



**Figure 5.13.** Control plots for several variables in the Top enriched phase space for events in VBF region.

## 5.5 Same Flavor final state

The analysis if the same-flavour final state  $W^+W^- \rightarrow \mu^\pm\mu^\mp 2\nu$  and  $W^+W^- \rightarrow e^\pm e^\mp 2\nu$  is described.

### Signal region

Events are requested to pass single or double lepton triggers and all the physics objects definitions are the same as in the OF analysis. The final state consists of two well identified electrons or two muons with  $p_T > 20$  GeV, opposite charge, and large missing transverse energy from the undetected neutrinos.

In addition to the backgrounds described for the OF final state, the background from  $DY \rightarrow \mu^+\mu^-$  and  $DY \rightarrow e^+e^-$  is very large in this final state. Indeed, due to this very large background, the SF analysis only targets the VBF topology, where the DY background is suppressed by the tight jet requirements. In addition, an invariant mass of the two leptons larger than 120 GeV is requested. The full selection, defined as the “WW SF selection”, is :

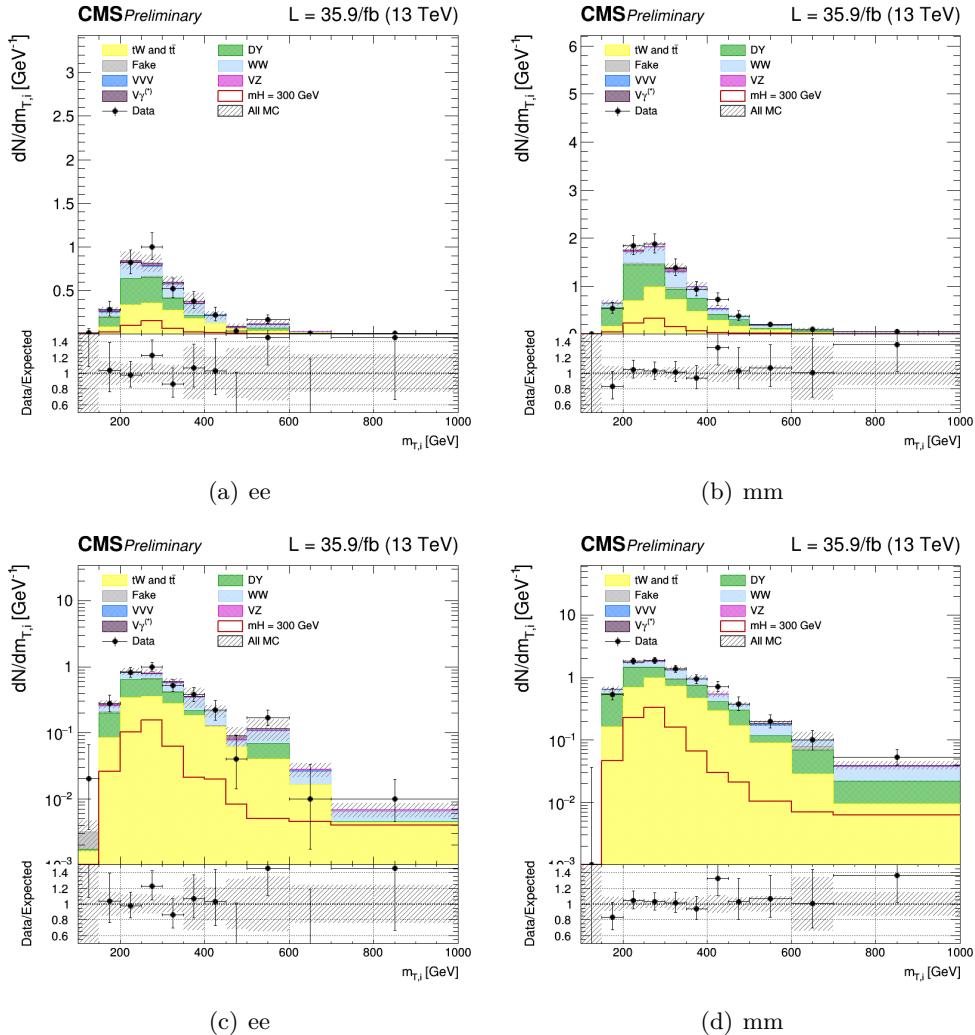
- Two isolated leptons with same flavor and opposite charge ( $\mu^\pm\mu^\mp$  and  $e^\pm e^\mp$ );
- $p_T$  of the leading and trailing lepton  $> 20$  GeV;
- Third lepton veto: veto events if a third lepton with  $p_T > 10$  GeV;
- $m_{\ell\ell} > 120$  GeV
- $p_T^{\ell\ell} > 30$  GeV;
- MET  $> 50$  GeV;
- $m_T^I > 100$  GeV;
- At least 2 jets non b-tagged (according to cMVAv2 loose WP) with  $p_T > 30$  GeV.
- $\Delta\eta_{jj} > 3.5$ ;
- $m_{jj} > 500$  GeV;;

Similarly to the opposite-flavour analysis, the signal is extracted from a template fit of the  $m_T^I$  distribution. The  $m_T^I$  distributions has the following binning:

- **VBF**, [100,150,200,250,300,350,400,450,500,600,700,1000];

where the first number represents the lower edge of the first bin while the other numbers represent the upper edges. The last bin is an overflow bin. The binning has been chosen in order to have at least 10 expected Top-backgrounds event and at least 10 expected Drell-Yan events in each bin of the template.

The distributions for the signal region of  $m_T^I$  variable is shown in Fig. 5.5



**Figure 5.14.** Unblinding distributions  $m_T^I$  in the signal region for  $ee$  and  $\mu\mu$  categories in linear and log scale. The signal hypothesis corresponding to  $m_X$  of 300 GeV.

### Drell-Yan control region

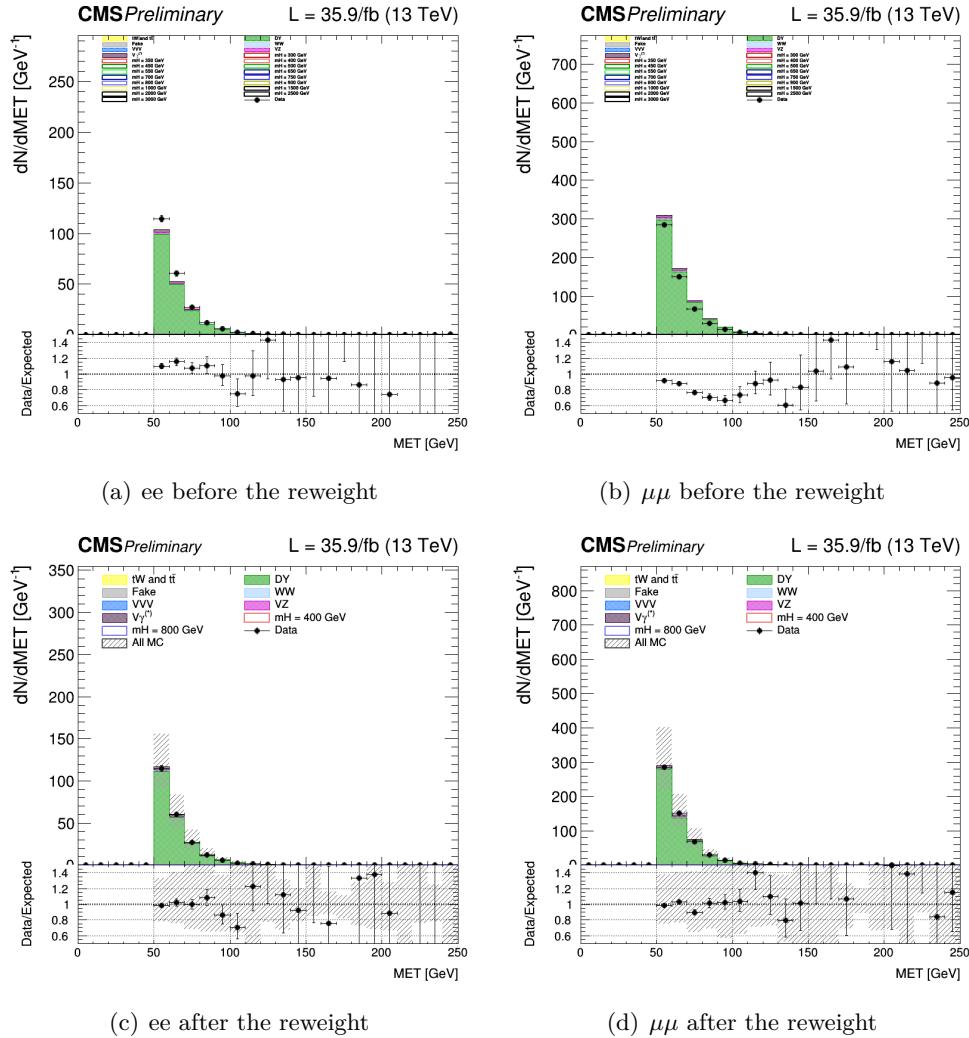
The main background for the SF analysis is the DY. A control region has been defined, as close as possible to the signal one to be used for the normalization of the DY background, separately for electrons and muons.

The control region is defined by the “WW SF selection”, except for the  $m_{\ell\ell}$  requirement which is changed to  $70 \text{ GeV} < m_{\ell\ell} < 120 \text{ GeV}$  to include the Z boson.

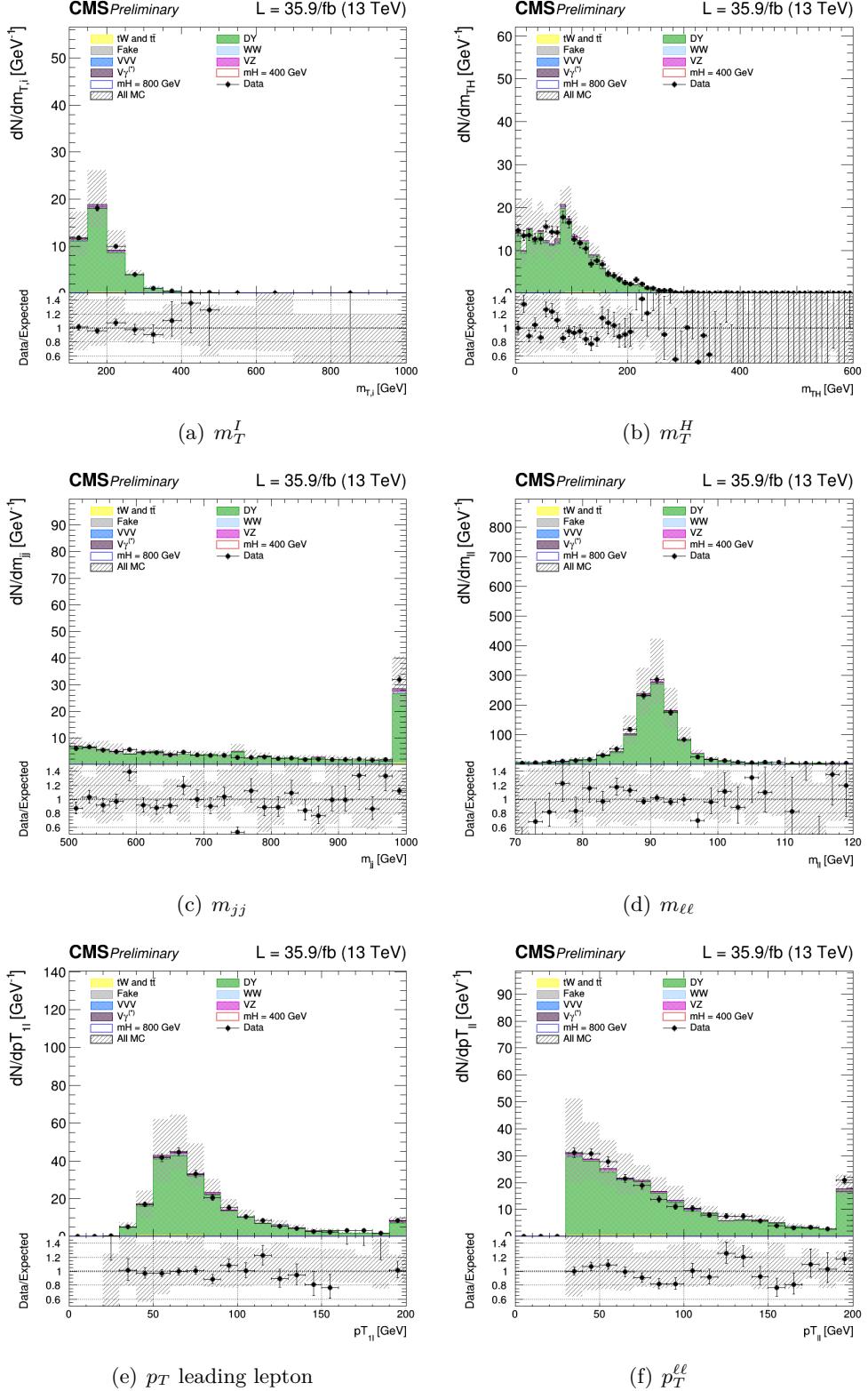
The missing transverse energy distribution in the data shows discrepancies respect to Monte Carlo simulation in ee and  $\mu\mu$  Drell-Yan control regions. A correction is applied reweighting all the simulated samples with a weight per event which depends on the MET value. The weight is evaluated as the ratio between data, one subtracted all backgrounds except the DY, and the Drell-Yan itself, in each bins of the distribution, separately for ee and  $\mu\mu$  categories. The weight is assumed to be linear as function of the MET value.

This kind of reweighting allows to correct for shape differences between data and MC, , Fig. 5.15.

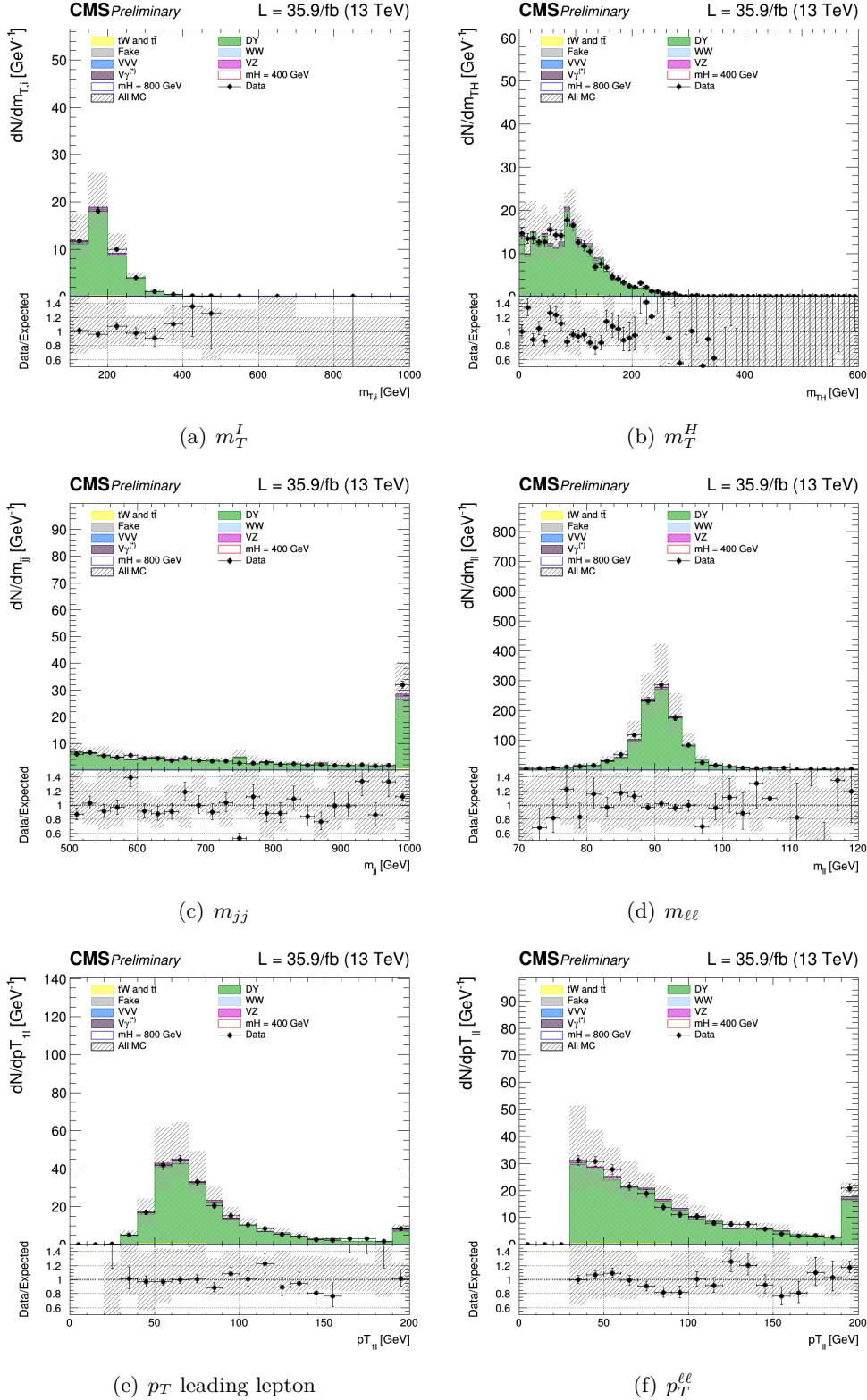
The control plots for several variables in a Drell-Yan enriched phase space for the ee and  $\mu\mu$  are shown in Figs. 5.17 for the dielectron case and Figs. 5.18 for the dimuon case. In general there is a good agreement between data and MC.



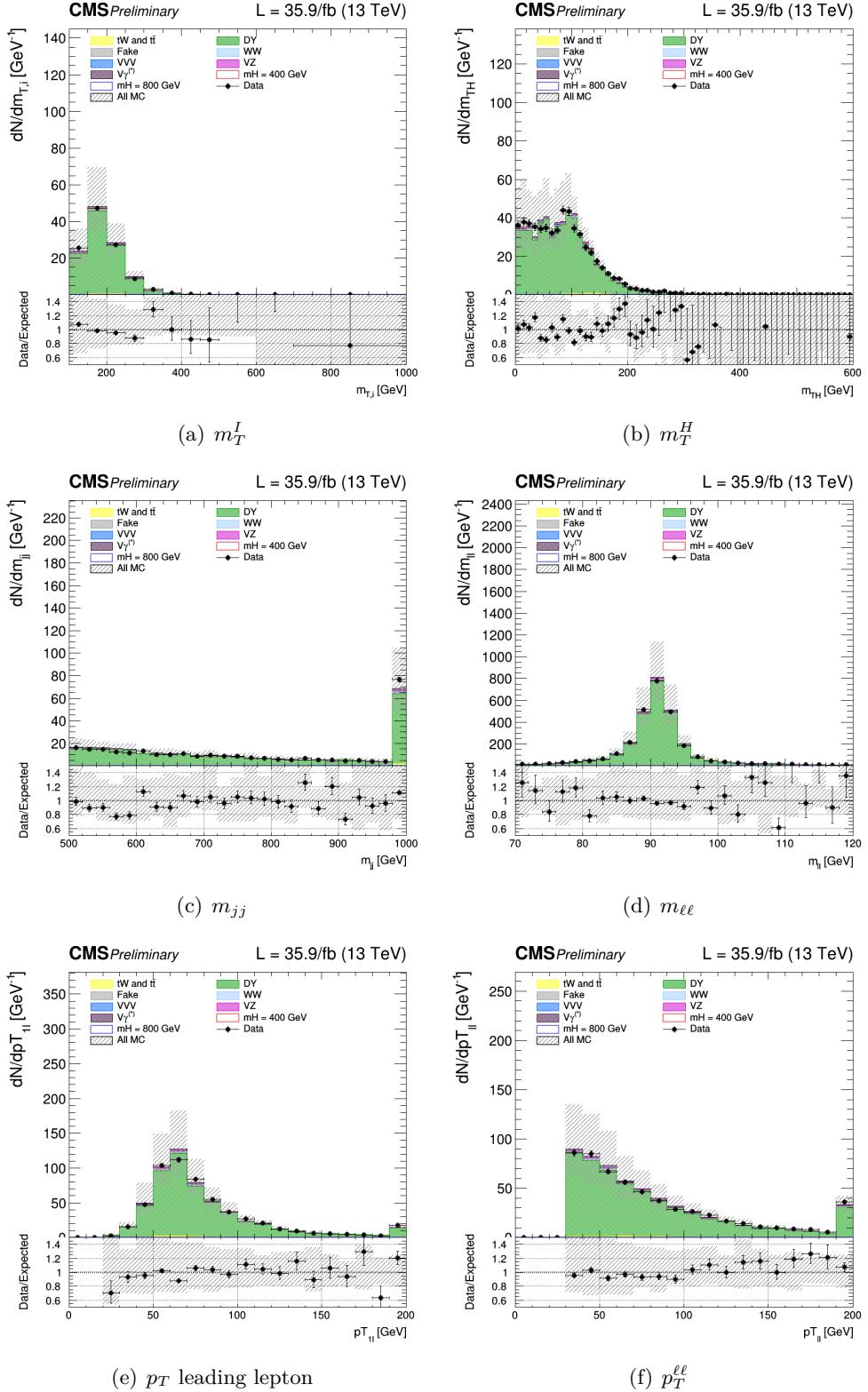
**Figure 5.15.** MET control plots for Drell-Yan fot ee categories in *a* and for  $\mu\mu$  in *b* before the reweight. In *c* and *d* the same distribution after the correction.



**Figure 5.16.** Control plots for several variables in a Drell-Yan enriched phase space for ee.



**Figure 5.17.** Control plots for several variables in a Drell-Yan enriched phase space for ee.

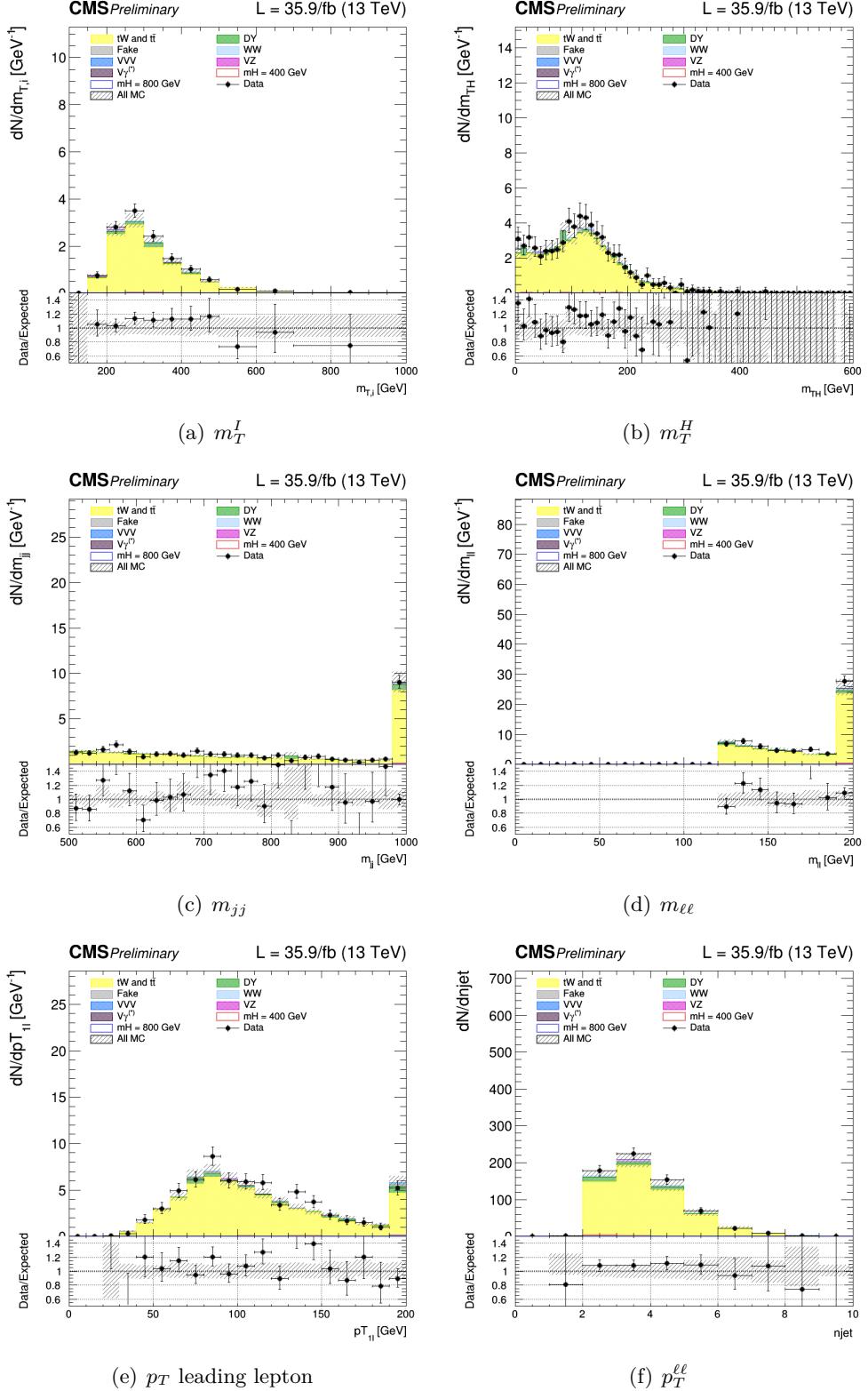


**Figure 5.18.** Control plots for several variables in a Drell-Yan enriched phase space for  $\mu\mu$ .

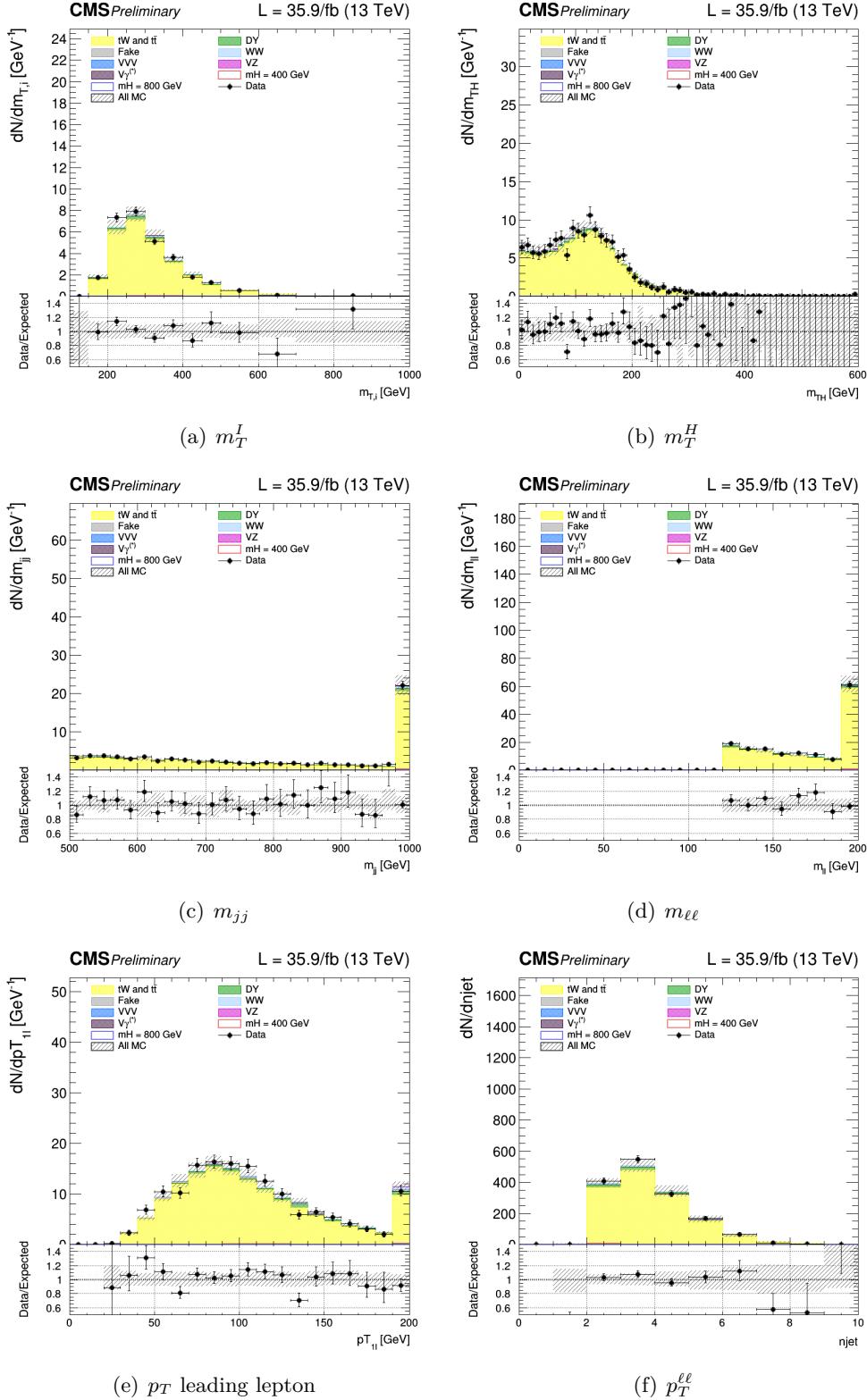
### Top control region

A top-enriched control region is defined to normalize the top backgrounds, separately for electrons and muons. The “WW SF selection” is required with the inversion of the b-tagging requirement, i.e. the two jets are both requested to be b-tagged according to cMVAv2 loose WP.:

The control plots for several variables in a top enriched phase space for events are shown in the Figs. 5.19 for the dielectron case and 5.20 for the dimuon case. Good agreement is observed between data and MC.



**Figure 5.19.** Control plots for several variables in a Top enriched phase space for ee.



**Figure 5.20.** Control plots for several variables in a Top enriched phase space for  $\mu\mu$ .

## 5.6 2HDM and MSSM interpretations

In this section the interpretation of this analysis in a Two-Higgs-doublet model (2HDM) and in some scenarios of the Minimal Supersymmetric extension to the Standard Model (MSSM) is described.

### 5.6.1 Introduction to 2HDM and MSSM

The 2HDM is a well motivated extension of the SM. It contains two Higgs doublets, from which a total of five Higgs bosons are predicted: Two CP-even bosons  $h$  and  $H$ , a CP-odd boson  $A$  and two charged bosons  $H^\pm$ . In most theories,  $h$  exhibits the features of the SM Higgs boson, while  $H$  is a CP-even Higgs boson at a higher mass. In this study, limits are calculated on the production cross section of the Higgs boson  $H$  multiplied with the branching fraction of the decay into two  $W$  bosons.

The 2HDM comprises many free parameters. Two of these are of particluar interest:

- $\tan \beta$ : The ratio  $\frac{v_u}{v_d}$  of the vacuum expectation values of the two Higgs doublets.
- $\alpha$ : The mixing angle of the two scalar Higgs bosons  $h$  and  $H$ .

The quantity  $\cos(\beta - \alpha)$  is also of interest, as the coupling of the heavy scalar Higgs boson  $H$  to two vector bosons is proportional to this factor. In the decoupling limit, which occurs at  $\cos(\beta - \alpha) = 0$ , all couplings become SM-like.

A 2HDM of type-2 is considered in this analysis. Here up-type quarks couple to one doublet, while down-type quarks and leptons couple to the other doublet. The MSSM is a type-2 2HDM. On tree level only two parameters are left free. By convention, these parameters are chosen to be  $\tan \beta$  and  $m_A$ , the mass of the pseudoscalar Higgs boson. The exclusion limits can be set in a two-dimensional plane as a function of these two parameters. Due to higher order diagrams additional free parameters occur. Benchmark scenarios are then used in order to constrain these parameters. Here two MSSM scenarios are used: The  $m_h^{mod+}$  scenario and the hMSSM scenario [?].

The analysis follows the same steps as described in sections 5.4 and 5.5.

### 5.6.2 Statistical inference

The necessary model predictions for these scenarios are provided by the LHC Higgs Cross Section Working Group [28]. For both MSSM scenarios the ggF cross sections have been computed with SusHi (v.1.4.1)[29]. These cross sections include NLO supersymmetric QCD corrections and NNLO QCD corrections for the top quark contribution in the effective theory of a heavy top quark, as well as electroweak effects by light quarks. The masses of the Higgs bosons, their mixing, the branching fractions and the effective Yukawa couplings in the  $m_h^{mod+}$  scenario are all calculated with FeynHiggs (v.2.10.2)[30, 31, 32, 33, 34]. For the hMSSM scenario the branching fractions are obtained from HDECAY (v.6.40)[35, 36]. The results for general 2HDM are obtained using the ggF cross sections computed with SusHi (v.1.5.0) and the branching fractions from 2HDMC (v.1.7.0)[37]. The VBF cross sections are calculated using an approximation. The BSM Higgs production cross sections

for VBF, which are provided for different masses by the LHC Higgs Cross Section Working Group [38], are taken and multiplied by  $\cos^2(\beta - \alpha)$ , resulting in VBF cross sections for a heavy CP-even Higgs boson.

The exclusion limits obtained for the MSSM scenarios are displayed in the  $m_A$ - $\tan\beta$  plane. A fine grid is chosen in this plane, and for each point of this grid a maximum likelihood fit is performed after the  $m_A$  and/or  $\tan\beta$  dependent values of the model, such as cross sections and masses of the Higgs bosons are calculated. These fits are done using the asymptotic method. Performing a maximum likelihood fit in this manner is equivalent to a hypothesis test, where the signal hypothesis is tested against the SM-and-background hypothesis. The signal hypothesis for a combination of  $m_A$  and  $\tan\beta$  is excluded at 95 % confidence level. In the two-dimensional plane this limit is determined from interpolation between the points of the grid. The limits in the more general 2HDM are obtained in the same way, although a different parameter is chosen in place of  $m_A$ .

## 5.7 Systematic uncertainties

Systematic uncertainties are introduced as nuisance parameters in the fit and can affect the normalization and the shape of the different contributions.

Statistical uncertainties from MC simulated events are taken into account. Systematic uncertainties are represented by individual nuisance parameters with log-normal or shape-based distributions. The uncertainties affect the overall normalization of the signal and backgrounds as well as the shape of the predictions across the distribution of the observables. Correlations between systematic uncertainties in different categories and final states are taken into account. Systematic uncertainties play an especially important role in this analysis where no strong mass peak is expected due to the presence of undetected neutrinos in the final state. Below we describe in detail sources and quantities of systematics in this analysis and their effects on the signal and background processes. A list of the most important background uncertainties is given below.

### Background normalization uncertainties

One of the most important sources of systematic uncertainty is the normalization of the backgrounds that are estimated on data control samples whenever is possible. The signal extraction is performed subtracting the estimated backgrounds to the event counts in data. The amount of uncertainty depends on the considered background:

- jet-induced background: normalization and kinematic shapes are derived from a data control region and both normalization and shape systematic uncertainties are considered. A conservative 30% uncertainty on the fake rate is assumed correlated across the different analysis regions. The contribution to the uncertainty in the signal region due to the limited electron statistics in the background enriched control regions is about 10%, while the contribution due to the limited muon statistics 3%.
- WW background: The normalization of the WW background is performed independently in each jet multiplicity via the rateParam feature of combine. A WW electroweak (VBS) sample is used in addition to the standard WW sample in the phase spaces with at least two jets, where its contribution becomes non negligible. The uncertainty in the cross section for this process is evaluated using the variations of the renormalization and factorization QCD scales, as well as the PDF variations, and amounts to 11%.
- $t\bar{t}$  and tW backgrounds: Top events are estimated with b-tagging in data control regions. The two top background enriched control regions are defined as additional categories in the fit while the kinematic shapes are taken from the simulation corrected for the b-tagging discriminant scale factors. The top normalization is correlated between the top control region and the Higgs signal categories separately in the different jet multiplicities, and these normalizations are left unconstrained using the rateParam feature of combine.
- Drell-Yan background: The Drell-Yan background enters the different flavor analysis via the leptonic decays of the  $\tau$  leptons from  $Z\gamma^* \rightarrow \tau\tau$ . In the

different flavor analyses the normalization of these background is controlled via the rateParam feature of combine and with a dedicated control region in each jet bin category.

- $W\gamma^*$  background: The kinematic shape of this background is predicted by simulation, normalized to its data-driven estimate, and constrained within the respective uncertainty, which is 25%.
- WZ : The kinematic shapes of this backgrounds are predicted by simulation and normalized to their theoretical predictions in the different and same flavour analysis.
- $Z\gamma^*$  : The kinematic shapes of this backgrounds are predicted by simulation and normalized to their theoretical predictions in the different and same flavor analysis.
- ZZ: The kinematic shapes of this backgrounds are predicted by simulation and normalized to their theoretical predictions in the different and same flavor analysis.

## Experimental uncertainties

Effects from experimental uncertainties are studied by applying a scaling and/or smearing of certain variables of the physics objects, followed by a subsequent recalculation of all the correlated variables. This is done for MC simulation, to account for possible systematic mismeasurements of the data. All experimental sources except luminosity are treated both as normalization and shape uncertainties. For background with a data-driven normalization estimation, the shape uncertainty is considered only. The following experimental systematic sources have been taken into account.

- Luminosity: The uncertainty determined by the CMS luminosity monitoring is 2.3% for 13 TeV data.
- Lepton trigger systematics: Lepton trigger systematics are of the order of less than 1%. These uncertainties are computed by varying the tag selection as well as the Z window in the tag and probe method used to compute the corresponding scale factors.
- Lepton reconstruction and identification efficiency: The lepton reconstruction and identification efficiencies are measured with the tag and probe method in data. To correct for the difference in the lepton identification efficiencies between data and MC, data/MC scale factors dependent on  $p_T$  and  $\eta$  are applied to the MC. The resulting uncertainty in the signal region is 1% for electrons and 2% for muon.
- Muon momentum and electron energy scale: Uncertainties on both the scale and resolution individually amount to 0.6-1% for electrons and 0.2% for muons.

- MET miss modelling: The MET miss measurement is affected by the possible mismeasurement of individual particles addressed above, as well as the additional contributions from the pile-up interactions. The effect of the missing transverse momentum resolution on the event selection is studied by propagating each component of the MET uncertainty to the absolute value and direction of MET.
- Jet energy scale (JES) uncertainties: We estimate this uncertainty applying the official jet uncertainties on the JES and compute the variation of the selection efficiency. JES uncertainty affects the rates in the signal region at the level of 10%.
- b-jet misidentification modelling: The uncertainties on the selection of non-b jets is taken into account by looking at the b-jet misidentification efficiency. The uncertainties on these scale factors need to be taken into account, and are of the order of a few percent.

### Theoretical uncertainties

- PDF and higher-order corrections (renormalization and factorization scale): PDF uncertainties and the missing knowledge on higher-order corrections, evaluated by means of scale variation, directly affect the cross section, as well as the acceptance of a simulated process. The uncertainties that arise from using different PDF sets were obtained by reweighting events with different PDF sets.
- Underlying event and parton shower modelling: The underlying event (UE) and parton shower (PS) modelling uncertainties are estimated by comparing samples interfaced with different parton showers (Pythia vs Herwig) and UE tunes
- Single top tW and tt ratio: The ratio between the single top and top pair cross section is varied by the uncertainty on the ratio between their cross sections, calculated considering scale variations,
- A QCD and PDF scales for the signal samples at different masses. The uncertainties are taken from the Yellow Report 3 and the same values are used both for the large width hypothesis and for different values of  $C'$ . The effect of QCD and PDF scale uncertainties on the analysis selection has also been taken into account.
- The categorization of events based on jet multiplicity introduces additional uncertainties related to higher order corrections. These uncertainties are associated to the ggH production mode and are evaluated independently following the recipe described in [39] and are 5.6% for the 0-jet and 13% for the 1-jet and 20% for the 2-jet and VBF categories.

The top background shape is estimated from simulation and corrected using a data driven b-tagging scale factor. The normalization is measured in a top quark

enriched control region obtained inverting the b-veto requirement of the signal region. Three control regions are defined, one for each jet bin category. A nuisance parameter is added to take into account the effect of the parton shower uncertainty on the top background.

The DY background shape is also estimated from simulation and analogously to the Top background, the DY normalization is measured with a data driven technique in three control regions enriched in DY events.

A dedicated nuisances for MET reweighting in DY control region is introduced in SF analysis. It is evaluated separately for ee and  $\mu\mu$  categories. The uncertainty is quoted as maximum and minimum best-fit lines of the linear fit.

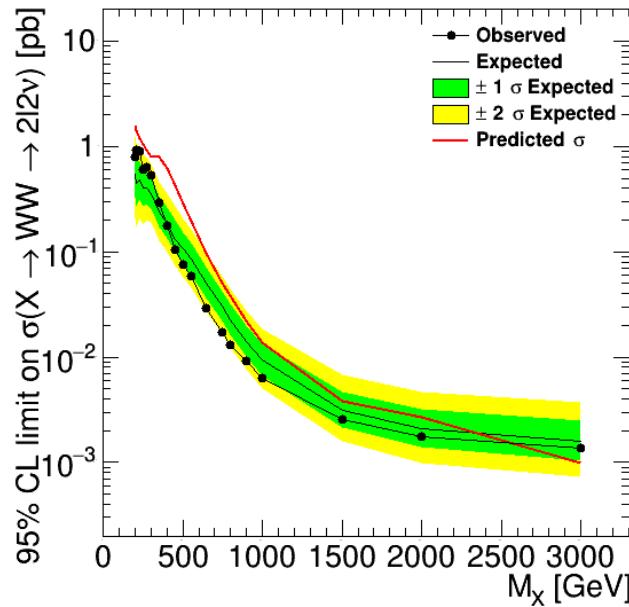


## Chapter 6

# Results and Interpretation

### EWS results

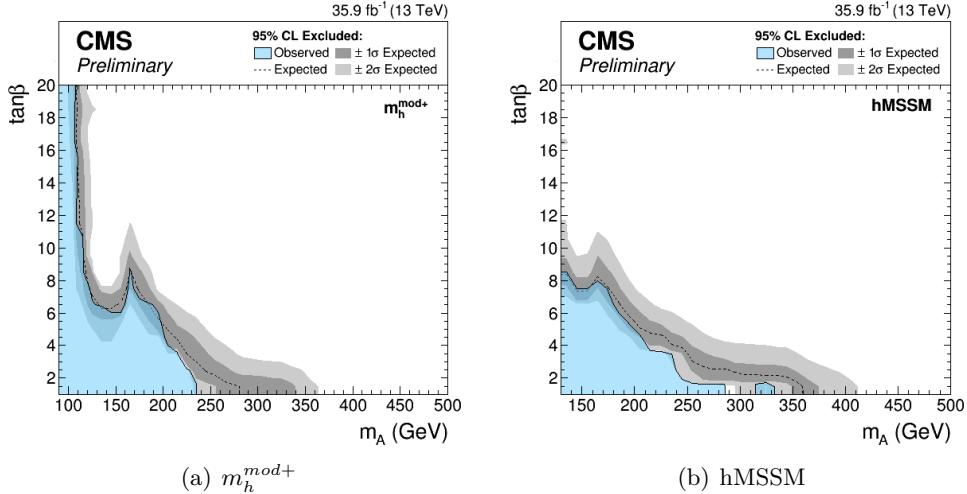
The final binned fit is performed using the  $m_T^I$  histogram for all signals and the number of events for the backgrounds. For the oppiste-flavour and same-flavour analysis, for every categories and for every mass point from 200 GeV up to 3 TeV the significance and the 95% CL upper exclusion limit are calculated. The expected final limit from the combination of the OF and SF analysis are shown in fig. ???. This limit represent a considerable improvement respect to the high mass search done with 2015 data and the expected limits is compatible with the ATLAS results for a similar analysis (CERN-EP-2017-214; arXiv:1710.01123).



**Figure 6.1.** 95% CL exclusion limits, on the production ggH and VBF cross section times branching fraction as a function of the mass for the **combination** of the two analysis OF and SF, in the full mass range. The red line represent the predicted cross-section for EW high mass bosons.

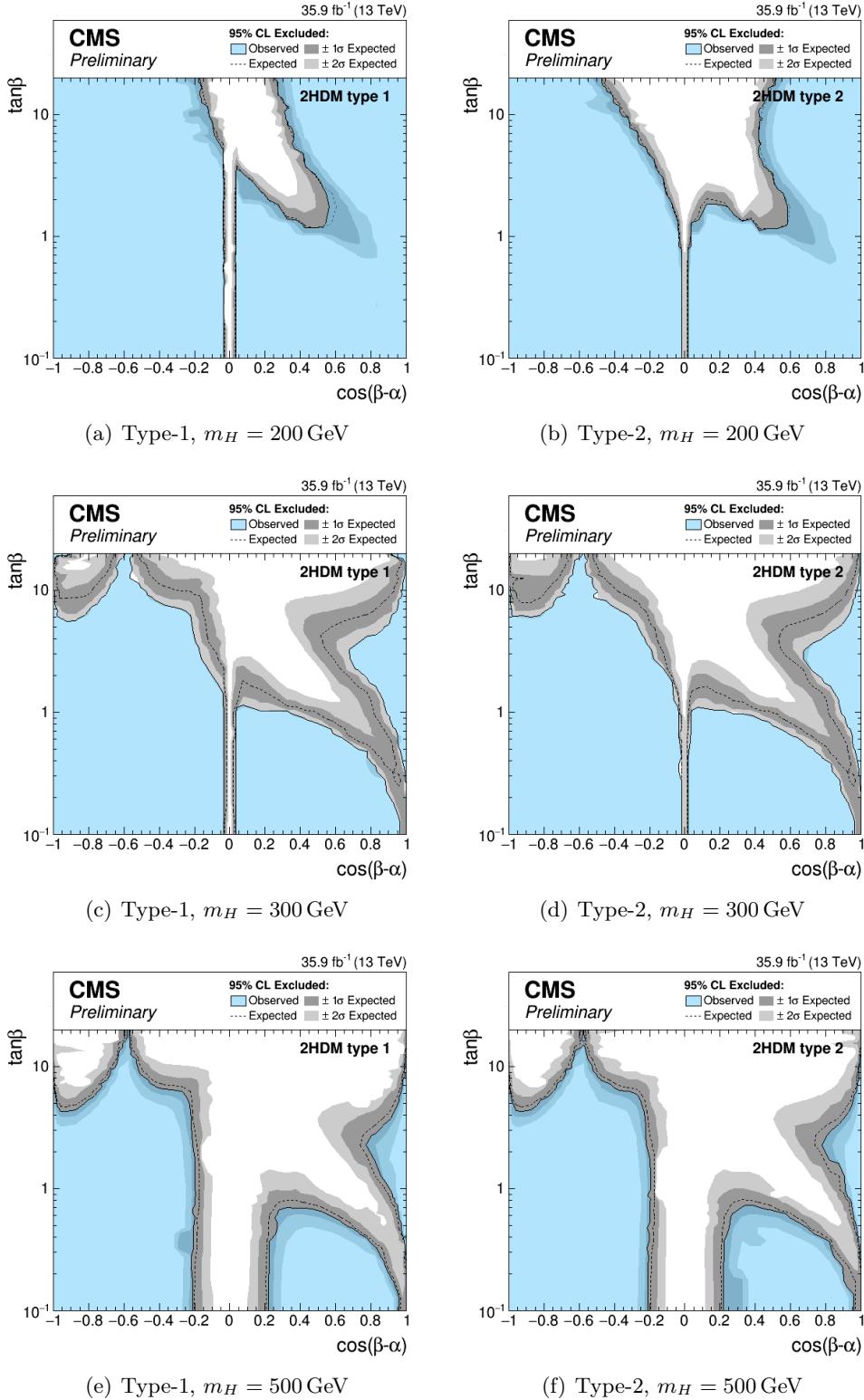
## 2HDM results

In Fig. 6.2 the exclusion limits are shown for the  $m_h^{mod+}$  scenario on the left and the hMSSM scenario on the right. The dashed line marks the limit, while the green area shows the side of the limit that is excluded. The bands surrounding the limit indicate the  $\pm 1, 2\sigma$  contours. For both scenarios the region at low values of  $m_A$  and  $\tan \beta$  is excluded. These results complement well with the exclusion limit given by the MSSM  $H \rightarrow \tau\tau$  analysis, where the sensitivity is lower for low  $m_A$  and  $\tan \beta$ .

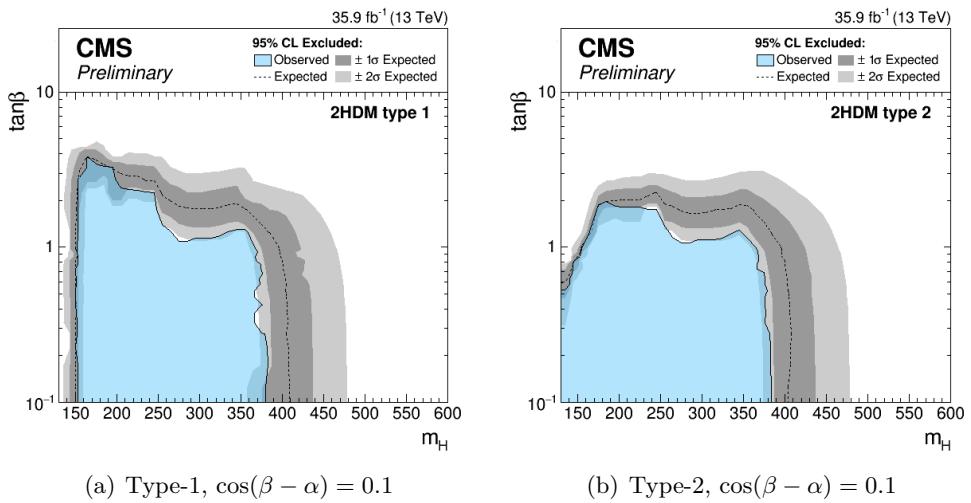


**Figure 6.2.** 95% CL exclusion limits on the MSSM  $m_h^{mod+}$  scenario (left) and the hMSSM scenario (right).

In Fig. 6.3 and 6.4 the exclusion limits are shown for a 2HDM. The limits in 6.3 for both a type-1 and type-2 2HDM is displayed in a  $\cos(\beta - \alpha)$ - $\tan \beta$  plane, in which the neutral heavy Higgs boson masses are  $m_H = m_A = 200, 300, 500$  GeV and the convention  $\sin(\beta - \alpha) > 0$  is used. The plots in Fig. 6.4 show the limit in the  $m_H$ - $\tan \beta$  plane. Here it is again assumed that  $m_H = m_A$  and  $\sin(\beta - \alpha) > 0$ , but here the relationship between  $\beta$  and  $\alpha$  is  $\cos(\beta - \alpha) = 0.1$ . The exclusion limits seen here are larger compared to those produced in the similar analysis by ATLAS. A possible reason may be the choice of the discriminating variable  $m_T^I$  or the different categorization.



**Figure 6.3.** 95% CL exclusion limits on a 2HDM with  $\cos(\beta - \alpha)$  on the x-axis. Limits are shown for a type-1 and type-2 2HDM for different masses  $m_H = 200, 300, 500$  GeV.



**Figure 6.4.** 95% CL exclusion limits on a 2HDM with  $m_H$  on the x-axis. Limits are shown for a type-1 and type-2 2HDM for  $\cos(\beta - \alpha) = 0.1$ .

# Appendix A

## Special commands



# Bibliography

- [1] T. Sjostrand. Monte Carlo generators. 2006.
- [2] A. Buckley et al. General-purpose event generators for LHC physics. *Phys. Rept.*, 504:145–233, 2011.
- [3] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [4] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [5] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [6] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP*, 04:002, 2009.
- [7] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 02:037, 2010.
- [8] A. V. Gritsan *et. al.* S. Bolognesi, Y. Gao. Jhugen. <http://www.pha.jhu.edu/spin/>.
- [9] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano.  $HW^\pm/HZ + 0$  and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO. *JHEP*, 10:083, 2013.
- [10] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.
- [11] Tom Melia, Paolo Nason, Raoul Rontsch, and Giulia Zanderighi. W+W-, WZ and ZZ production in the POWHEG BOX. *JHEP*, 11:078, 2011.
- [12] Patrick Meade, Harikrishnan Ramani, and Mao Zeng. Transverse momentum resummation effects in  $W^+W^-$  measurements. *Phys. Rev.*, D90(11):114006, 2014.
- [13] Prerit Jaiswal and Takemichi Okui. Explanation of the  $WW$  excess at the LHC by jet-veto resummation. *Phys. Rev.*, D90(7):073009, 2014.

- [14] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Bounding the Higgs width at the LHC: Complementary results from  $H \rightarrow WW$ . *Phys. Rev.*, D89(5):053011, 2014.
- [15] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, and Juan Rojo. Parton distributions with QED corrections. *Nucl. Phys.*, B877:290–320, 2013.
- [16] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO. *Nucl. Phys.*, B855:153–221, 2012.
- [17] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. 2015.
- [18] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [19] 2015 pileup json files. [https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData#2015\\_Pileup\\_JSON\\_Files](https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData#2015_Pileup_JSON_Files).
- [20] Sm higgs production cross sections at  $\sqrt{s} = 13\text{-}14$  tev. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt1314TeV>.
- [21] J R Andersen et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013.
- [22] Nlo single-top channel cross sections. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec>.
- [23] Nnlo+nnll top-quark-pair cross sections. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO>.
- [24] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [25] How to compute cross sections with the genxsec analyzer. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HowToGenXSecAnalyzer>.
- [26] Summary table of samples produced for the 1 billion campaign, with 25ns bunch-crossing. <https://twiki.cern.ch/twiki/bin/view/CMS/SummaryTable1G25ns>.
- [27] Search for high mass Higgs to WW with fully leptonic decays using 2015 data. Technical Report CMS-PAS-HIG-16-023, CERN, Geneva, 2016.
- [28] Lhc hxswg for bsm higgs (wg3). <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG3>.

- [29] Robert V. Harlander, Stefan Liebler, and Hendrik Mantler. SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM. *Comput. Phys. Commun.*, 184:1605–1617, 2013.
- [30] S. Heinemeyer, W. Hollik, and G. Weiglein. FeynHiggs: A Program for the calculation of the masses of the neutral CP even Higgs bosons in the MSSM. *Comput. Phys. Commun.*, 124:76–89, 2000.
- [31] S. Heinemeyer, W. Hollik, and G. Weiglein. The Masses of the neutral CP - even Higgs bosons in the MSSM: Accurate analysis at the two loop level. *Eur. Phys. J.*, C9:343–366, 1999.
- [32] G. Degrassi, S. Heinemeyer, W. Hollik, P. Slavich, and G. Weiglein. Towards high precision predictions for the MSSM Higgs sector. *Eur. Phys. J.*, C28:133–143, 2003.
- [33] M. Frank, T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein. The Higgs Boson Masses and Mixings of the Complex MSSM in the Feynman-Diagrammatic Approach. *JHEP*, 02:047, 2007.
- [34] T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein. High-Precision Predictions for the Light CP -Even Higgs Boson Mass of the Minimal Supersymmetric Standard Model. *Phys. Rev. Lett.*, 112(14):141801, 2014.
- [35] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension. *Comput. Phys. Commun.*, 108:56–74, 1998.
- [36] A. Djouadi, M. M. Muhlleitner, and M. Spira. Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface). *Acta Phys. Polon.*, B38:635–644, 2007.
- [37] Johan Rathsman and Oscar Stal. 2HDMC - A Two Higgs Doublet Model Calculator. *PoS*, CHARGED2010:034, 2010.
- [38] Bsm higgs production cross sections at  $\sqrt{s} = 13$  tev. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBSMAt13TeV>.
- [39] Radja Boughezal, Xiaohui Liu, Frank Petriello, Frank J. Tackmann, and Jonathan R. Walsh. Combining Resummed Higgs Predictions Across Jet Bins. *Phys. Rev.*, D89(7):074044, 2014.