Measurement of the Higgs boson transverse momentum spectrum in the WW decay channel at 8 TeV and first results at 13 TeV

Lorenzo Viliani
of University of Florence

PhD Thesis

Abstract

The cross section for Higgs boson production in pp collisions is studied using the H $\rightarrow$ W$^+$W$^-$ decay mode, followed by leptonic decays of the W bosons, leading to an oppositely charged electron-muon pair in the final state. The measurements are performed using data collected by the CMS experiment at the LHC with pp collisions at a centre-of-mass energy of 8TeV, corresponding to an integrated luminosity of 19.4fb^{-1}. The Higgs boson transverse momentum (p_{T}) is reconstructed using the lepton pair p_{T} and missing p_{T}. The differential cross section times branching fraction is measured as a function of the Higgs boson p_{T} in a fiducial phase space defined to match the experimental acceptance in terms of the lepton kinematics and event topology. The production cross section times branching fraction in the fiducial phase space is measured to be 39 ± 8 (stat) ± 9 (syst)fb. The measurements are compared to theoretical calculations based on the standard model to which they agree within experimental uncertainties.

Contents

Chapter 1

Electroweak and QCD physics at LHC

Chapter 2

The CMS experiment at the LHC

2.1 The Large Hadron Collider

2.2 The CMS experiment

2.3 The CMS trigger system

2.4 Objects definition and event reconstruction

sec:Objects

Chapter 3

Higgs boson properties in the H → WW decay channel

3.1 Higgs boson measurements at LHC

The discovery of a new boson consistent with the standard model (SM) Higgs boson has been reported by ATLAS and CMS Collaborations in 2012. The discovery has been followed by a comprehensive set of studies of properties of this new boson in several production and decay channels and no evidence of deviation from the SM expectation has been found so far. The CMS studies in the H → WW → $2\ell2\nu$ decay channel include the measurement of the Higgs properties, as well as constraints on the Higgs total decay width and gauge bosons anomalous couplings.

3.2 Higgs boson measurements in the H → WW decay channel

Chapter 4

Measurement of the Higgs boson transverse momentum at 8TeVusing H → WW → 2ℓ2ν decays

4.1 Introduction

In this chapter the measurement of the transverse momentum spectrum of the Higgs boson, produced in proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 8\text{TeV}$, is reported. This measurement can be used to directly inspect the perturbative QCD theory in the Higgs sector. In particular the p_T^H variable is sensitive to the Higgs production mode and the differential distribution in this variable can be used to inspect the effects of the top quark mass in the gluon fusion top loop. Moreover, any observed deviation from the SM expectation, especially in the tail of the p_T^H distribution, could be a hint of physics beyond the SM.

Similar measurements have already been performed by CMS and ATLAS experiments in the ZZ and $\gamma\gamma$ Higgs decay channels. The measurement reported here is the first measurement of the Higgs p_{T} spectrum in the WW decay channel.

The cross section has been measured in a fiducial phase space defined using generator level variables in order to mimic the experimental acceptance and reduce the systematic uncertainties on the procedure of extrapolating the results in a larger phase space.

The Higgs transverse momentum has been reconstructed calculating the vector sum of the

dilepton system transverse momentum plus missing transverse energy

$$\vec{p}_\mathrm{T}^\mathrm{H} = \vec{p}_\mathrm{T}^{\ell\ell} + \vec{p}_\mathrm{T}^\mathrm{miss} \tag{4.1}$$

The signal has been extracted subtracting all backgrounds by means of a binned Maximum Likelihood fit and has been then corrected for the efficiency of the analysis selections and for the detector resolution effects using an unfolding procedure.

The differential measurement has been performed in six bins of p_T^H with variable widths, chosen to have approximately the same purity in each bin, as explained in section 4.3.

4.2 Datasets, Triggers and MC samples

This analysis is largely built on top of the already published H $\rightarrow$ WW measurments [**Chatrchyan:2013**] in terms of code, selections and background estimates for both the gluon fusion (ggH) [**AN-2013-022**] and the vector boson fusion (VBF) [**AN-13-097**] prodcution mechanisms.

4.2.1 Datasets and triggers

The datasets used for the analysis correspond to 19.4fb^{-1}at $\sqrt{(s)} = 8$ TeVof integrated luminosity composed of the following CMS data taking periods during 2012: 2012A (892 pb^{-1}), 2012B (4440 pb^{-1}), and 2012C (6898 pb^{-1}) and 2012D (7238 pb^{-1}). Data have been vchecked and validated and only data corresponding to good data taking quality are considered. The e$^\pm\mu^\mp$ final state is considered in this analysis. The following five Primary Datasets have been used for the signal extraction: SingleElectron, SingleMu and MuEG (Muon-ElectronGamma).

For the data samples, the events are required to fire one of the unprescaled single-electron, single-muon or muon-electron triggers. A full description of these triggers in given in [**AN-2012-228**] for 8 TeVdata. Although identification and isolation criteria are also applied, a brief overview of the HLT transverse momentum (p_T) criteria on the leptons is given in Table 4.1. While the HLT lepton p_T thresholds of 17 and 8 GeVfor the double lepton triggers accommodate the offline lepton p_T selection of 20 and 10 GeV, the higher p_T thresholds in the single lepton triggers help partially recovering double lepton trigger

inefficiencies as a high p_T lepton is on average expected due to the kinematic of the Higgs decay.

Table 4.1: Highest transverse momentum thresholds applied in the lepton triggers at the HLT level. Double set of thresholds indicates the thresholds for each leg of the double lepton triggers.

Trigger Path	7 TeV	8 TeV
Single-Electron	$p_T > 27$ GeV	$p_T > 27$ GeV
Single-Muon	$p_T > 15$ GeV	$p_T > 24$ GeV
Muon-Electron	$p_T > 17$ and 8 GeV	$p_T > 17$ and 8 GeV
Electron-Muon	$p_T > 17$ and 8 GeV	$p_T > 17$ and 8 GeV

No trigger requirement is made on the simulated events but the combined trigger efficiency is estimated from data and applied to all simulated events. The detailed trigger efficiencies and the weighting procedure can be found in Appendix C of [**AN-2013-022**] [**AN-2013-052**]. The average trigger efficiency for signal events that pass the full event selection is measured to be about 96% in the $e\mu$ final state for a Higgs boson mass of about 125GeV.

4.2.2 Monte-Carlo samples

Several Monte Carlo event generators are used to simulate the signal and background processes:

- The POWHEG program [**powheg**] provides event samples for the $H \rightarrow WW$ signal for the Gluon Fusion (ggH) and VBF production mechanisms, as well as $t\bar{t}$ and tW processes.

- The $q\bar{q} \rightarrow WW$, Drell-Yan, ZZ, WZ, $W\gamma$, $W\gamma^*$, tri-bosons and W+jets processes are generated using the MADGRAPH 5.1.3 [**madgraph**] event generator.

- The VH process is simulated using PYTHIA 6.424 [**pythia**].

For leading-order generators samples, the CTEQ6L [**cteq66**] set of parton distribution functions (PDF) is used, while CT10 [**ct10**] is used for next-to-leading order (NLO) ones. Cross section calculations [**LHCHiggsCrossSectionWorkingGroup:2011ti**] at next-to-next-

to-leading order (NNLO) are used for the $H \to WW$ process (POWHEG NLO generator is tuned to reproduce NNLO accuracy on the on-shell Higgs p_T spectrum and scaled to NNLO inclusive cross-section), while NLO calculations are used for background cross sections. For all processes, the detector response is simulated using a detailed description of the CMS detector, based on the GEANT4 package [**Agostinelli:2002hh**].

Minimum bias events are superimposed on the simulated events to emulate the additional pp interactions per bunch crossing (pile-up). The number of pile-up events simulated in the MC samples (in the same bunch crossing, in time, or in the previous or following one, out of time pile-up) have been generated poissonianly sampling from a distribution similar to what is expected from data. These samples are reweighted to represent the pile-up distribution as measured in the data. For a given range of analyzed runs, the mean number of pile-up interactions per bunch crossing is estimated per luminosity block using the instantaneous luminosity provided by the LHC, integrated over the entire run range and normalized. This distribution is then used to reweight the simulated pile-up distribution. The average number of pile-up events per beam crossing in the 2011 data is about 10, and in the 2012 data it is about 20.

We checked that the contribution of the ttH production mechanisms is negligible in each bin of p_T^H (below 1%) and has been neglected. In figure 4.1 is shown the relative fraction of the four different production modes in each bin of p_T^H.

4.3 Analysis Strategy

The Higgs boson transverse momentum is measured in a fiducial phase space, which is defined at generator level requiring

- Exactly two status 3 leptons, an electron and a muon, originated from the H $\to$ WW $\to$ $2\ell2\nu$ decays, with opposite charge, with $|\eta| < 2.5$ and > 20 GeV and > 10 GeV.

- Generator level invariant mass of the two leptons $m_{\ell\ell} > 12$ GeV.

- Vector sum of the two status 3 leptons $p_T^{\ell\ell} > 30$ GeV.

- Generator level transverse mass $\sqrt{(p_T^{\ell\ell}+)^2 - (p_T^{\vec{\ell\ell}} +)^2} > 50$ GeV.

Figure 4.1: Relatve fraction of ggH, VBF, VH and ttH in each bin of the Higgs boson transverse momentum.

Experimentally, the Higgs boson transverse momentum is reconstructed as the vector sum of the lepton momenta in the transverse plane and MET.

$$\vec{p}_T^H = \vec{p}_T^{\ell\ell} + \vec{p}_T^{miss} \tag{4.2}$$

Compared to other differential analysis of the Higgs cross section, such as those in the ZZ and $\gamma\gamma$ decay channels, this analysis has to cope with the limited resolution due to the entering the transverse momentum measurement. The effect of the limited resolution has two main implications on the analysis strategy:

- the choice of the binning in the transverse momentum spectrum needs to be reasonable when compared to the resolution. A detailed explanation of how the binning is defined is given in Sec. 4.5.

- Non negligible bin migration effects are present, and an unfolding procedure needs to be applied, not only to correct for selection efficiencies, as in ZZ and $\gamma\gamma$, but also to correct for bin migration effects. This is explained in Sec. **??**.

A detailed description of the fiducial region definition and about its optimization is given in appendix **??**.

The selection is essentially based on the one in the published analysis [**Chatrchyan:2013iaa**] with one noticeable difference being the fact that in this analysis we do not make categories in the number of jets. The reason for this choice is that the number of jets is strongly correlated with the transverse momentum, so making an inclusive analysis in the number of jets allows the dropping of most of the uncertainties related to the signal modeling of the number of jets produced in association with the Higgs boson. A detailed description of the selection is shown in Sec. 4.4.

The estimation of the backgrounds is different, to some extent, with respect to the one of the published . This is mainly due to the absence of the jet binning. The techniques used to assess the backgrounds in each bin are discussed in Secs. **??**, **??**, **??**.

Concerning the signal extraction, this analysis is again based on the already published analysis, although we fit the signal component in each of the transverse momentum bins, using two dimensional templates in the $m_{\ell\ell}$, plane. The signal extraction is discussed in Sec. **??**.

Finally an unfolding procedure is needed to extract the differential distribution in a fiducial phase space. This is discussed in detail in Sec. **??**.

4.4 Event reconstruction and selections

4.4.1 Event reconstruction

The muons, electrons, jets and missing transverse energy () reconstruction and criteria are described in in details in [**AN-2012-194**]. The following criteria are only a brief summary:

- **Muons:** *GlobalMuon* (with $\chi^2/ndof < 10$, at least one good muon hit and at least two muon segments in different muon stations) or *TrackerMuon* (provided it satisfies

the "Tracker Muon Last Station Tight" selection). Several cut-based identification criteria are applied as well as the particle flow (PF) Isolation. In 2012, the PF Isolation is replaced by an MVA algorithm.

- **Electrons:** *GSF Electons*. A MVA identification criteria is applied as well as an MVA algorithm.

- **Jets:** *Anti-k_T PF jets* (with R=0.5 and applying L1, L2 and L3 jet energy corrections, including Pile-Up jet corrections from Fastjet method). Only jets and $|\eta| < 4.7$ are considered. A specific Pile-Up MVA-based rejection algorithm is applied.

- **:** The is reconstructed the *PF Algorithm* or considering *only tracks* originating from the same vertex as the two leptons. In addition, the minimum of the projections of these two to the closest lepton direction if they are in the same hemisphere, otherwise of their original values () is used in the analysis.

4.4.2 Event selection

Unlike the main analysis, this analysis is inclusive in number of jets, so we do not have to define different jet multiplicity categories. The event selection consist of several steps. The first step is to select WW -like events applying a selection that is heavily based on the main analysis selection except for few different cuts explained below. The WW -like event preselection consists of the following set of cuts:

1. **Lepton preselection**:

   - at least two opposite-sign and opposite-flavour ($e\mu$) leptons reconstructed in the event;

   - $|\eta| < 2.5$ for electrons and $|\eta| < 2.4$ for muons;

   - $p_\text{T} > 20$ GeV for the leading lepton. For the trailing lepton, the transverse momentum is required to be larger than 10 GeV.

2. **Extra lepton veto**: the event is required to have two and only two opposite-sign leptons passing the lepton selection.

3. **preselection**: particle flow is required to be greater than 20GeV.

4. **Di-lepton mass cut**: $m_{\ell\ell} > 12 \text{GeV}$ in order to reject low mass resonances and QCD backgrounds.

5. **Di-lepton p_T cut**: $p_T^{\ell\ell} > 30 \text{GeV}$.

6. **projected selection**: required to be larger than 20 GeV.

7. **Transverse mass**: $m_T^H > 60 \text{GeV}$ to reject Drell-Yan to $\tau\tau$ events.

In addition to the WW -like preselection other cuts are applied in order to reduce the top background ($t\bar{t}$ ans single-top), which is one of the main backgrounds in this final state. We operate two different selections depending on the number of jets with $p_T > 30$ GeVin the event. This is done to suppress the top background both in the low p_T^H region, where 0-jets events have the biggest contribution, and for higher values where also larger jet multiplicity events are important. The selection for 0-jets events relies on a soft muon veto, which rejects events with non-isolated soft muons (likely belonging to b-jets), and on a soft jets (with $p_T < 30$ GeV) anti b-tagging requirement. The latter requirement exploits the Track Counting High Efficiency tagger (TCHE) to reject soft jets that are likely to come from b quarks hadronization. These are exactly the same requirements applied in the 0-jets bin of the main analysis.

For events with a jet multiplicity greater or equal than one, we apply a different selection with respect to the main analysis. In this case we exploit the good b-tagging performances of the *JetBProbability* tagger to reject all the jets with $p_T > 30$ GeVthat are likely to come from a b quark. This jet veto relies on a cut on the *JetBProbability* tagger discriminant as has been also done in the VH () analysis [**CMS˙PAS˙HIG˙13-017**]. Any jet with a discriminant value below 1.4 is identified as a non b-jet. The analysis selection requires no b-tagged jets with $p_T > 30$ GeV.

A cut-flow plot is reported in figure 4.2 showing the effect of each selection on top of Monte Carlo samples. In the first bin, labelled as *No cut*, no selection has been applied and the bin content correspond to the total expected number of events with a luminosity of 19.46 fb^{-1}. All the events in this bin have at least two leptons with a loose transverse momentum cut of 8 GeV. In the following bin the lepton cuts are applied, including the requirement to have two opposite-sign and opposite-flavour leptons and the extra lepton veto. Then are progressively reported all the other selections, showing the effect of each

cut on backgrounds and signal. For each selection is also reported the expected signal over background ratio which after the full selection reach a maximum value around 3%.

The selection efficiency is shown in Fig. 4.3 (a). The efficiency denominator is the number of events that pass the acceptance, while the numerator is the number of events that pass both the selection and the acceptance, in each $p_{\mathrm{T}}^{\mathrm{H}}$ bin. The fake rate, defined by the ratio of signal events that pass the selection but are not within the acceptance, divided by the total number of events passing both the selection and the acceptance is shown in Fig. 4.3 (b). For both the selection efficiency and the fake rate the signal samples included
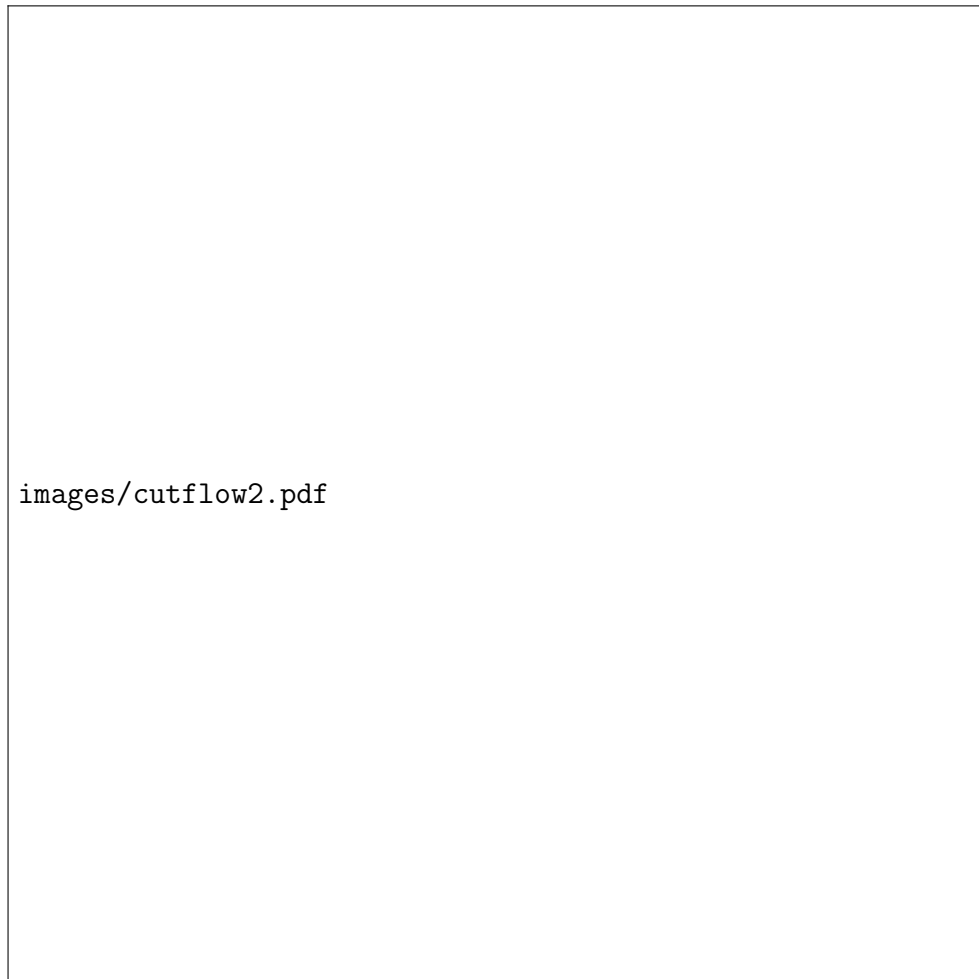
images/cutflow2.pdf

Figure 4.2: Effect of single selections on MC samples. The signal (red line) is multiplied by 100 and superimposed on stacked backgrounds. In each bin, corresponding to a different selection, is reported the expected number of events in MC at a luminosity of 19.46 fb^{-1}.
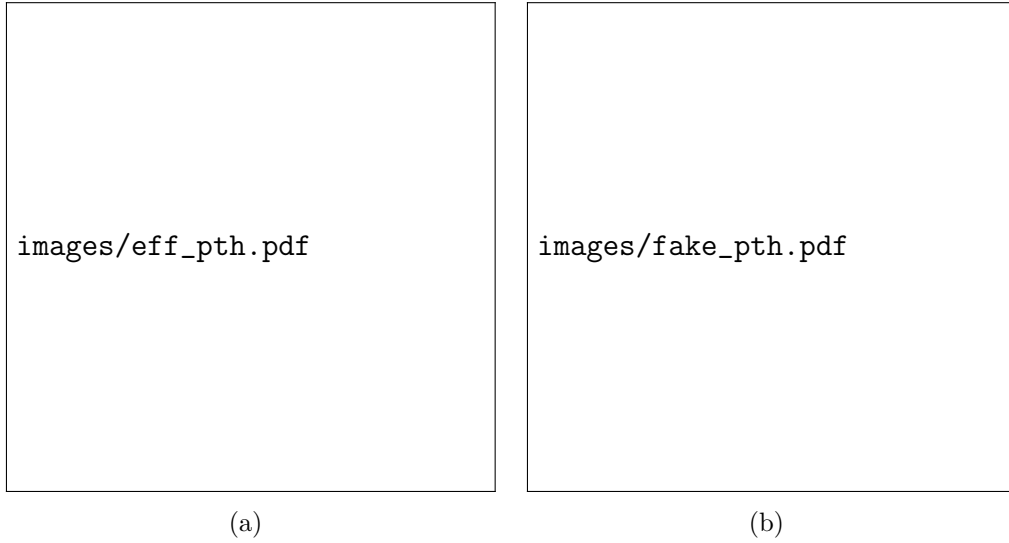
<div align="center">(a)   (b)</div>

Figure 4.3: Efficiency of the full selection (a) and fake rate (b) as a function of p_T^H.

correspond to the *ggH*, *VBF* and *VH* production mechanisms. The overall efficiency and fake rate are: $\epsilon = 0.362 \pm 0.005$ and $fake\ rate = 0.126 \pm 0.004$, where the errors are only statistical.

If we define a 4π acceptance, requiring just that the Higgs decays to WW and then to $2l2\nu$, the efficiency is $\epsilon = 0.03960 \pm 0.00033$.

4.5 Binning of the p_T^H spectrum

4.6 Background estimation

Chapter 5

First H → WW results at 13 TeV

5.1 Higgs boson search at 13TeV

5.2 Search for a high mass resonance in the WW decay channel at 13TeV

Chapter 6

Conclusions