

Measurement of the top-quark mass using
 J/ψ -in-jet candidates at the CMS experiment

Summer Student Program 2012

Maria Giulia Ratti

Contents

1	Introduction	2
1.1	Motivation	2
1.2	$t\bar{t}$ production and decay at the LHC	2
1.3	J/ψ final states in dilepton channel	4
2	Analysis	5
2.1	Data samples and event selection	5
2.2	J/ψ mass peak	6
2.3	$m_{J/\psi L}$ distribution in data	7
2.4	Simulation-based calibration	7
2.5	Measurement of the top mass	9
2.6	Conclusions	10

Chapter 1

Introduction

1.1 Motivation

The top quark is the heaviest known particle and its mass is one of the fundamental parameters of the Standard Model. The best top mass value now available comes from the combined results of the measurements performed by D0 and CDF experiments at the Tevatron:

$$m_t = 173.18 \pm 0.94 \text{ GeV} \quad (1.1)$$

At the LHC several methods are being used to reach such a precision and possibly do better; most of them require the full reconstruction of the top event. At high luminosities, the increased pile up will make the full reconstruction more difficult and the precision on the mass will be dominated by the systematics.

An alternative method was suggested back in 1999 by Avto Karchilava [1] and studied at simulation level for the CMS experiment [2]: it would make use of final states of $t\bar{t}$ with J/ψ -in-jet candidates. The extremely clean signature of J/ψ events makes this method appealing for precision measurements - not only for the top mass - but at the same time rather challenging, because of the very low branching ratio of $t\bar{t}$ final states needed for the measurement.

In this project, such approach has been studied and a preliminary measurement for the top mass has been obtained without accounting for systematics.

1.2 $t\bar{t}$ production and decay at the LHC

At the LHC, top production is dominated by $t\bar{t}$ pair production in pp collisions. The total approximate production cross section at $\sqrt{s} = 7 \text{ TeV}$ is $\sim 165 \text{ pb}$. The main production mechanisms are shown in figure 1.1:

- gluon gluon fusion ($\sim 85\%$)
- $q\bar{q}$ annihilation ($\sim 15\%$)

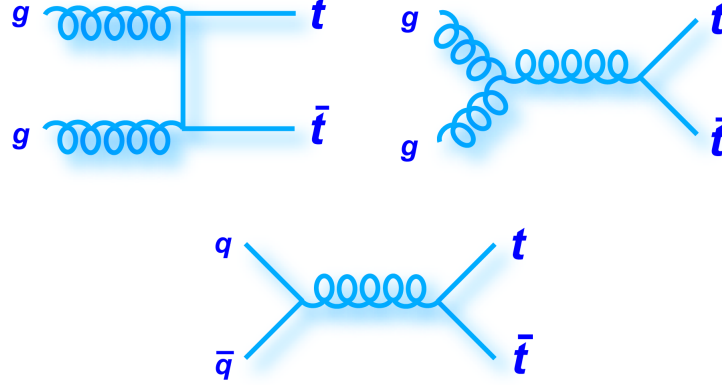
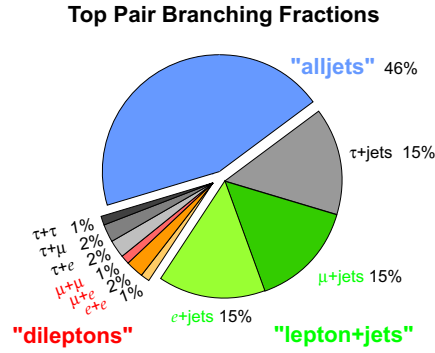
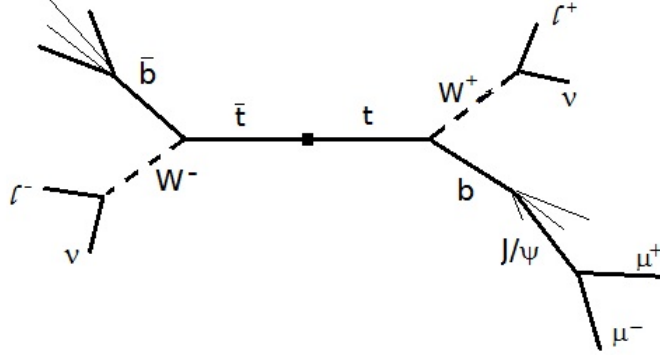
Figure 1.1: $t\bar{t}$ production Feynman diagrams

Figure 1.2: Top pair branching fractions

In the Standard Model, the top quark decays almost exclusively into W boson and b quark: this means that the W decay modes ($l^+(l^-) + \nu(\bar{\nu})$, $q\bar{q}$) define the $t\bar{t}$ final states. Three are the possible decay channels:

- all-hadronic (6 jets)
- semileptonic (1 lepton + 4 jets)
- dileptonic (2 leptons + 2 jets)

The branching ratios of the different decay modes are shown in figure 1.2: the dilepton channel is the one with lowest rate (9%), but also lowest background (mainly Drell-Yan process). My study is limited to dilepton channel only, but a similar analysis could be performed in the semileptonic channel as well.

Figure 1.3: Dilepton final state with J/ψ

1.3 J/ψ final states in dilepton channel

For this analysis the interesting final state of $t\bar{t}$ is a dilepton state that includes a J/ψ meson ($c\bar{c}$) coming from one of the b-quarks (see figure 1.3); the J/ψ then decays into two opposite-sign muons¹.

The variable used to extract the top mass is the invariant mass of the two opposite-sign muons coming from the J/ψ and the lepton coming from the same top. I define it as $m_{J/\psi L}$. It is expected that this variable is highly enough correlated to the top mass: because of the relatively high mass of the J/ψ , most of the momentum of the b-quark is carried by the J/ψ itself[2]. For this reason, it is possible to avoid the reconstruction of the other decay products of the b-hadronization.

The correlation of the top mass with the $m_{J/\psi L}$ is used to measure the top mass through a template method, i.e.: distributions of $m_{J/\psi L}$ are generated for different assumed values of the top mass and fitted with a convenient parametrization; the value of the reconstructed $m_{J/\psi L}$ is then obtained as a function of the input top mass. This function is later fitted to give the calibration of the $m_{J/\psi L}$ and used to measure the top mass when the reconstructed $m_{J/\psi L}$ in data is compared to the calibration.

¹Other two possible decays $J/\psi \rightarrow e^+e^-$ and $J/\psi \rightarrow \tau^+\tau^-$ are not considered in this analysis.

Chapter 2

Analysis

2.1 Data samples and event selection

The whole analysis is built on top of the event reconstruction centrally performed by CMS via the CMS software. The data samples used are the full dataset at $\sqrt{s} = 7$ TeV recorded by the CMS detector in the year 2011, for a resulting integrated luminosity of 4.9 fb^{-1} . For simulations, the matrix-element generator MadGraph is used in association with the tool for generation of high-energy collisions PYTHIA (Fall11 sample) with a simulated top mass of 172.5 GeV.

The event selection required for this analysis includes several steps:

- the term *preselection* is used to indicate a number of triggers and typical event cleanings that are performed in the standard $t\bar{t}$ reference selection. These include typical requirements on the primary vertex size, on the quality of the tracks, on the rejection of anomalous HCAL noise and beam scraping event veto. In addition to these, specific cuts are applied to leptons and jets: two hard isolated leptons (ee , $e\mu$, $\mu\mu$) are required with a $p_T > 20$ GeV, $|\eta| < 2.4$. Jets are particle-flow jets, reconstructed with anti- k_T algorithm, $R=0.5$, $p_T > 30$ GeV, $|\eta| < 2.4$; only one jet is required.
- I refer to *Dimu-in-jet* selection to indicate the events with opposite-sign dimuon candidates in jets.
- Finally, I call *J/ψ window* the events that pass the further selection criterion of having opposite-sign dimuons with an invariant mass between 2.8 and 3.2 GeV.

In the table 2.1 the number of events in data and simulations for these different steps are shown.

The Monte Carlo yields are given scaled to the data and the $t\bar{t}$ fraction of signal events is then estimated as the ratio between $t\bar{t}$ signal events from simulations and the number of events in data.

	Data	$t\bar{t}$ sig (MC)	$t\bar{t}$ bkg (MC)	$t\bar{t}$ fraction
Preselection	3462315	21613	142	1%
Dimu-in-jet	525	206	1	39%
J/ψ window	38	18	0	48%

Table 2.1: Event yields in data and simulations

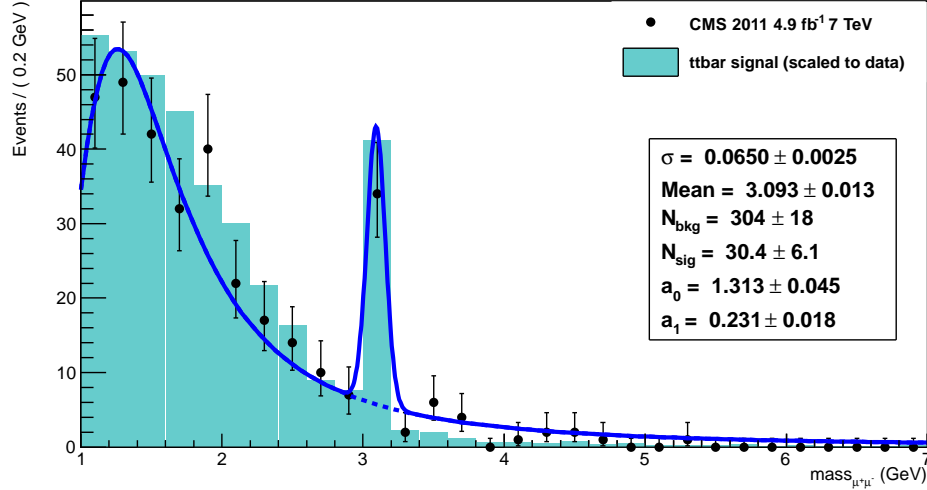


Figure 2.1: Invariant mass distribution of dimuons-in-jet

No b-tagging requirement has been explicitly applied both for statistical reasons (only few events pass all the selections) and because it is believed that requiring a J/ψ -in-jet is an implicit b-tagger.

Other cuts that are usually performed in the dilepton channel to increase the $t\bar{t}$ fraction of signal events in data have been left aside (such as missing transverse energy requirement, Z veto, QCD veto, ...) due to lack of statistics.

2.2 J/ψ mass peak

The fundamental ingredient of this analysis is the J/ψ meson produced in the b-hadronization. As mentioned before, we reconstruct the J/ψ as dimuons (opposite-sign muons) coming from the same jet. The distribution of events with dimuons-in-jet is shown in figure 2.1: a clean peak at the J/ψ mass is present. Both data and $t\bar{t}$ signal (scaled to data) are shown: within the statistic uncertainties they agree. The $t\bar{t}$ background is not shown, but it does not change the picture significantly.

A fit is performed on data to estimate the position of the J/ψ mass peak and the number of J/ψ events. It is an unbinned, extended, maximum likelihood fit (performed with RooFit) where the background has been modelled with a Landau

distribution and the signal with a Gaussian.

The extracted position of the peak 3.093 ± 0.013 GeV is compatible with the reference PDG value (2012): 3.096916 ± 0.000011 GeV.

The extracted number of J/ψ candidates (first time seen in $t\bar{t}$ data) is 30.4 ± 6.1 . However, for the purposes of this analysis all events in the 2.8-3.2 GeV window (as suggested in [2]) are selected for the reconstruction of the invariant mass $m_{J/\psi L}$ for a total number of 38.

2.3 $m_{J/\psi L}$ distribution in data

The $m_{J/\psi L}$ distribution is obtained after matching the selected J/ψ candidates with one of the leptons. In principle, it would be best to match the J/ψ with the correct lepton, but this is not done here due to statistical reasons and both the possible assignments are considered: the “correct” and the “wrong” one. It is expected that this is translated into a weaker correlation between $m_{J/\psi L}$ and the top mass; for further discussion on this see the following paragraph.

In figure 2.4 the distribution for data and simulation (again scaled to data) are shown. Within the huge statistic uncertainty on data, the two distributions are compatible.

2.4 Simulation-based calibration

As said, the measurement of the top mass is obtained by “calibrating” the extracted value of $m_{J/\psi L}$ to the top mass scale. For this purpose, a Monte Carlo-based calibration is needed and this is done through the template method described above.

The simulated distributions of $m_{J/\psi L}$ for different top masses are shown in figure 2.2. Due to very low statistics in the official top mass samples, the distributions for 161.5 GeV and 184.5 GeV have been produced privately with a different generator: Powheg+Pythia in FastSim (3M of generated $t\bar{t}$ events for each mass point)¹.

Each distribution is fitted with a Gaussian and the reconstructed value of the $m_{J/\psi L}$ is extracted as the mean of this Gaussian. The fit is unbinned.

The choice of the Gaussian is done for convenience: in [2] a 4th-order polynomial is suggested and the maximum taken as extracted value of $m_{J/\psi L}$. In [1] the approach is slightly more complex and the contribution from the correct and wrong assignments are separated: the correctly-assigned distribution is fitted with a Gaussian and the wrongly-assigned distribution with a 3rd-order polynomial.

Both the suggested approaches were tried, but the Gaussian fit in the range of 20-120 GeV was preferred for stability of the fit.

¹It is expected that the different generator does not have a significantly different behaviour or at least that the final measurement of the top mass is not highly influenced because of the dominating statistic uncertainty in this analysis. For more precise studies, the different behaviour of generators for this measurement should be investigated.

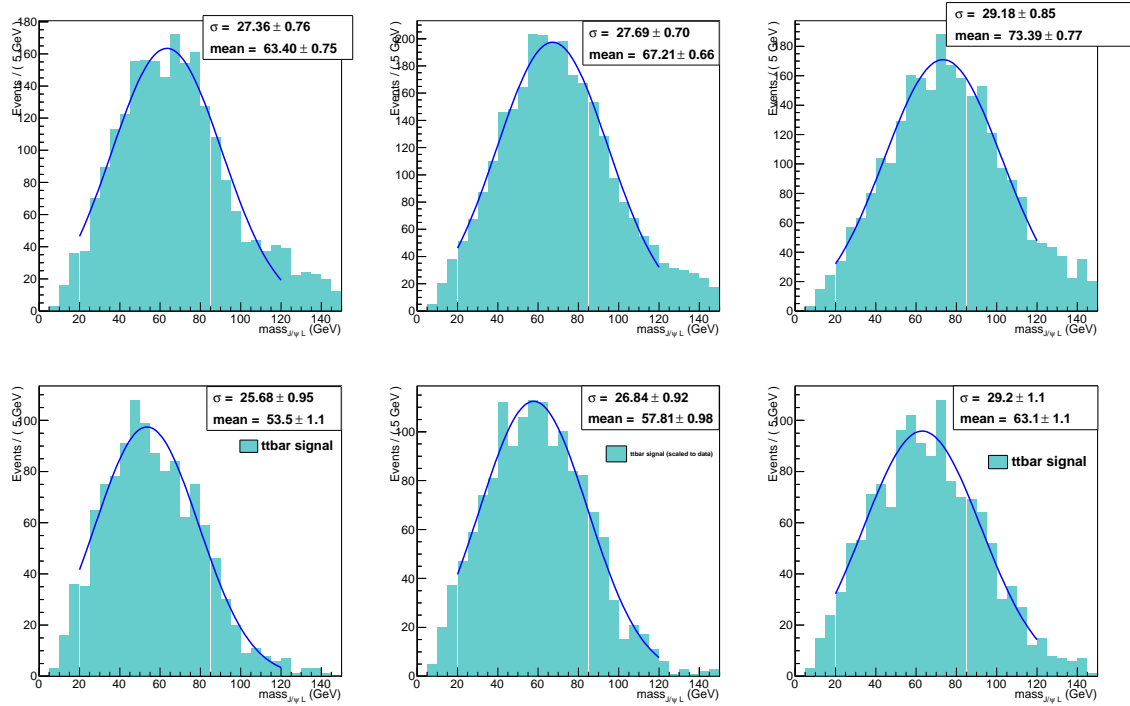


Figure 2.2: Simulated invariant mass distributions of $J/\psi + \text{lepton}$ for any-lepton assignment (top panels), and lower-angle-lepton assignment (bottom panels), for different assumed values of the top mass: from the left to the right 161.5 GeV, 172.5 GeV, 184.5 GeV.

m_t (GeV)	$m_{J/\psi L}$ (GeV)
161.5	63.40 ± 0.75
172.5	67.21 ± 0.66
184.5	73.39 ± 0.77

Table 2.2:

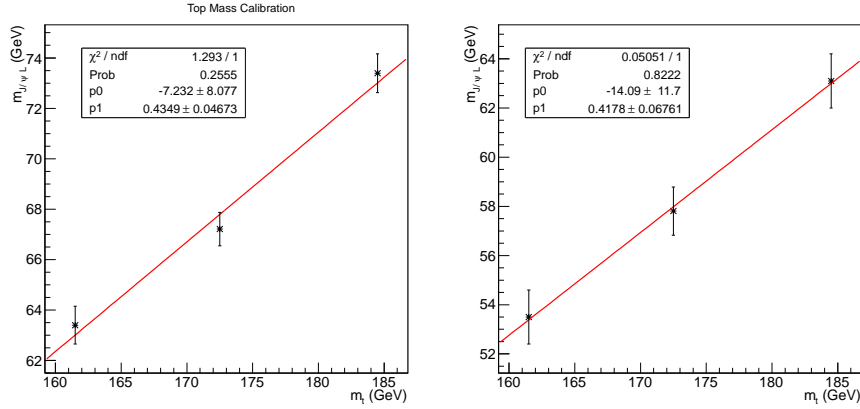


Figure 2.3: Simulation-based calibrations for any-lepton assignment (on the left) and lower-angle-lepton assignment (on the right)

The final extracted values of the $m_{J/\psi L}$ for the different simulated top masses are summarized in table 2.2 and shown in the right panel of figure 2.3.

A nice linear correlation is seen but the uncertainties on the fit parameters are high.

An approach where only one lepton is assigned to the dimuons has been tried. A simple algorithm that selects the lepton with the lowest angle with respect to the J/ψ has been implemented. The number of entries for the $m_{J/\psi L}$ distributions is then reduced of a factor of two, which is translated in increased statistic uncertainties on the extracted value of the $m_{J/\psi L}$ and on the fit parameters.

2.5 Measurement of the top mass

To finally obtain the measurement, the data distributions of $m_{J/\psi L}$ for both any-lepton and lower-angle-lepton assignments are fitted, the mean value of the Gaussians extracted and then scaled with the simulation-based calibration (see figure 2.4). The measured values of the top mass for the any-lepton assignment is:

$$m_t^{any} = 166 \pm 11(\text{stat}) \pm 26(\text{calib}) \text{ GeV}$$

which is compatible within the huge total statistic uncertainty (around 28 GeV) with 1.1. I have separated the uncertainty in two components: the first coming from the

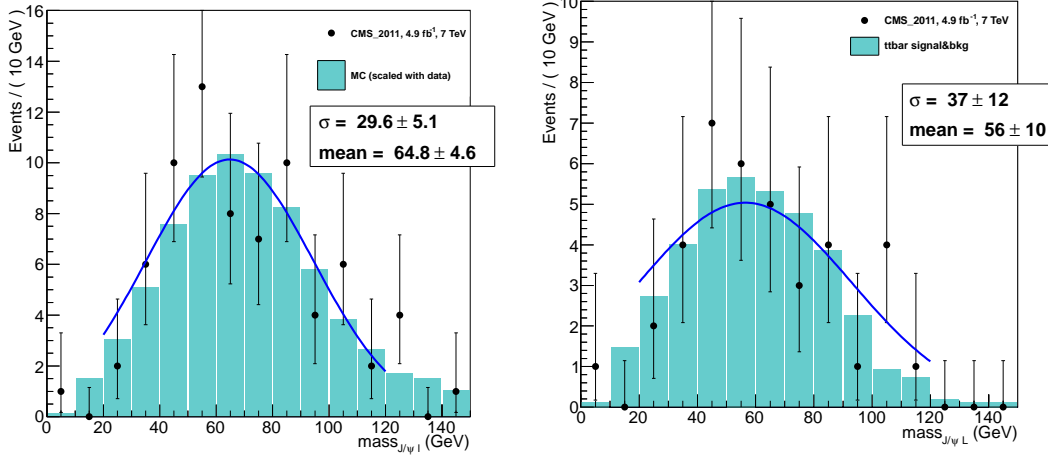


Figure 2.4: Fitted $m_{J/\psi L}$ distributions in data for any-lepton assignment (on the left) and lower-angle-lepton assignment (on the right)

error on the extracted value of the $m_{J/\psi L}$, the second one from the error on the fit parameters.

For the lower-angle lepton assignment the measured top mass is:

$$m_t^{\text{lower-angle}} = 168 \pm 28(\text{stat}) \pm 36(\text{calib})$$

with an even larger statistic uncertainty.

2.6 Conclusions

This analysis has shown for the first time J/ψ -in-jet candidates in the CMS $t\bar{t}$ data. A preliminary calibration of the $m_{J/\psi L}$ correlation with the top mass has been given, without accounting for systematics. A very preliminary result of the top mass has been reached with a total statistic uncertainty of 28 GeV. The precision in this measurement is limited by the lack of statistics, both in data and in simulations. This study should be improved with a more precise calibration: this could be done with larger Monte Carlo samples and with accurate studies on how to choose the correct lepton. Also, 2012 data samples should be included in the analysis.

Bibliography

- [1] Avto Kharchilava, Top Mass determination in leptonic final states with J/ψ , arxiv: hep-ph/9912320.
- [2] R. Chierici, A. Dierlamm, *Determination of the top mass with exclusive events $t \rightarrow Wb \rightarrow l\nu J/\psi X$* , CMS-NOTE-2006-058.