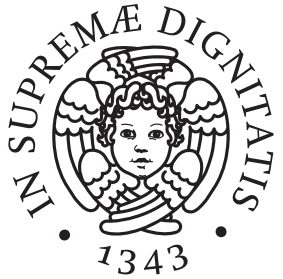


Università degli Studi di Pisa

Final Report

Validation Studies of the Matrix Element Method for the Top Quark Mass Measurement at CDF

Andrea Malara

supervised by
Costas Vellidis, FNAL

October 10, 2016

Abstract

The purpose of this report is to describe the results I obtained during the Summer School at Fermilab. In particular I studied the pull distribution in the background matrix element integration and the signal transfer functions and parton reconstruction efficiencies, trying to validate either the integration technique or the "new" transfer functions and efficiencies.

Contents

1 Introduction to top quark physics

1.1 Standard Model of the particles physics

As it is well known, the Standard Model (SM) is the theory which explains three of the four forces actually known through the bosons and also explains the structure of matter through the fermions. It includes 12 fermions of $\frac{1}{2}$ spin, which are the constituents of matter, and 4 bosons of spin 1, which carry the three forces (see Figure 1.1) plus the additional recent Higgs boson, which gives masses to all fermions. Fermions are further classified into two types: leptons and quarks, which carry color charge and so they can also interact strongly. For each of the various fundamental constituents, its symbol and its electric charge in unit of elementary charge $|e|$ of the electron are given in Table 1.1. The leptons which carry integral electric charge include electron e, muon μ and tau lepton τ: they can interact via all forces except from strong interaction. The neutral leptons are called neutrinos, denoted by the generic symbol ν and they only feel weak and gravitational interactions. A different flavor of neutrino correspond to different flavor of charged lepton. The quarks carry fractional charges, of $+\frac{2}{3}|e|$ or $-\frac{1}{3}|e|$. In Table 1.1 the masses increase from left to right, both for leptons and quarks. Particles in higher generations (having higher masses) are unstable; in order to observe and study these particles, we need to produce them via collisions with other stable particles. While leptons exist as free particles, quarks seem not to do so. It is a peculiarity of the strong forces between the quarks that they can be found only in combinations, not individually. This phenomenon is known as confinement. The quarks are the only fundamental particles interaction through all the possible forces.

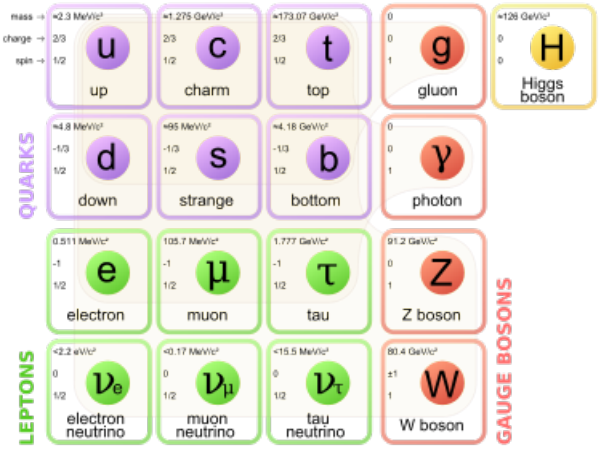


Figure 1.1: The building blocks of matter: six quarks (violet) and six leptons (green). Four gauges bosons represent(red) the three forces. Last discovery: Higgs boson (yellow).

	Generation			
	I	II	III	
Particle	Flavour			Charge
Quarks	u	c	t	$+\frac{2}{3}$
	d	s	b	$-\frac{1}{3}$
Leptons	e	μ	τ	-1
	ν_e	ν_μ	ν_τ	0

Table 1.1: The three generations of fundamental fermions. The charge is quoted in units of the absolute value of the electron charge.

Gauge Boson	Interaction	Mass	Charge	Spin
Gluon, g	Strong	0	0	1
Photon, γ	Em	0	0	1
$W^{\pm}$	Weak	80.4 GeV/c^2	±1	1
Z^0	Weak	91.2 GeV/c^2	0	1
H		125.1 GeV/c^2	0	0

Table 1.2: SM bosons: mediators of the fundamental forces between interacting elementary particles and the Higgs boson which gives masses to all fermions.

1.2 Top quark physics

The top quark can be produced either via the strong interaction leading to a production by $t\bar{t}$ pair (Figure 1.2) or via the electroweak interaction leading to the single top production (Figure 1.3). The pair production is the dominant mode at hadron colliders. At the $\sqrt{s} = 1.96$ TeV center-of-mass energy of the Tevatron, the top quark is expected to be produced from either $q\bar{q}$ annihilation (85%) or gg fusion (15%). The cross section of the process can be calculated perturbatively in QCD Next-to-Next-to-Leading Order (NNLO) and the measured value is $\sigma_{t\bar{t}} = 7.02 \pm 0.63$ pb (assuming $m_t = 175$ GeV/c^2). The $t\bar{t}$ cross section is extracted using the following formula:

$$\sigma_{t\bar{t}} = \frac{N_{signal} - N_{background}}{A(m_t)\epsilon_{t\bar{t}}(m_t)L} \quad (1)$$

where N_{signal} and $N_{background}$ are the number of observed data events and predicted number of background events. L is the recorded integrated luminosity, A denotes the acceptance, determined by MC and defined as the ratio of

events reconstructed in the detector to all events produced in the collisions, and ϵ is the product of all efficiency corrections determined from calibrations based on the comparison of control data with MC.

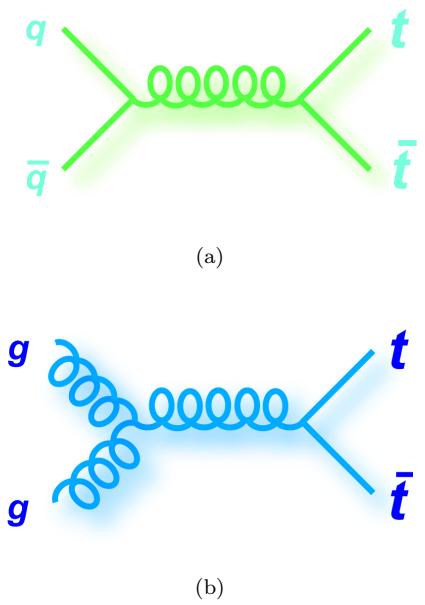


Figure 1.2: Example of the Feynman diagrams of $t\bar{t}$ production through quarks annihilation (a) and gluon fusion (b) [1].

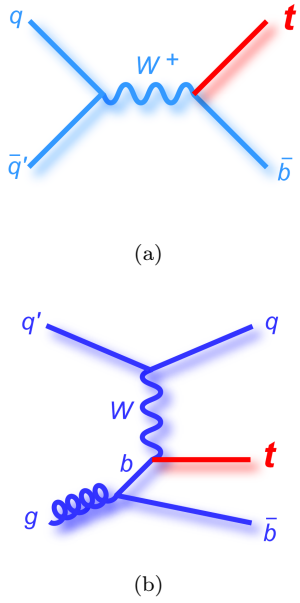


Figure 1.3: Example of the Feynman diagrams of single top weakly production via s-channel (a) or t-channel (b)[1]. The s-channel has been observed only at the Tevatron so far.

The $t\bar{t}$ production process is explained by the parton model of the hard scattering process: two partons, one of each colliding proton and anti-proton, take part in the interaction, carrying unknown momentum fractions z_1 and z_2 of the incoming hadron momenta. The Parton Distribution Function $f_i(z_j, Q^2)$, for a parton with flavor i participating in the hard scattering interaction with momentum fraction z_i, are unknown and can be extracted from data. Given the PDFs, the total cross section of $t\bar{t}$ production can be approximated by summing over all possible parton interactions $\left(p_A p_B \to t\bar{t}\right)$:

$$\sigma_{t\bar{t}} = \sum_{A,B} \int \mathrm{d}z_1 \, \mathrm{d}z_2 \left[f_A(z_1, Q^2) f_B(z_2, Q^2) \right] \sigma \left(p_A p_B \to t\bar{t} \right) \tag{2}$$

The top quark decays weakly into a W boson and a down-type quark (d, s, b) . As the decay rates of the top into down-type quarks are proportional to the squared CKM matrix element $|V_{t(d;s;b)}|^2$ (see Equation (3)), and considering that $|V_{td}|$ and $|V_{ts}|$ are two order of magnitude lower than $|V_{tb}|$. A recent measurement of the magnitude of the top-to-bottom quark coupling has been obtained from the combination of CDF and DØ results in the measurement of cross sections for single-top quark production, returning the value $|V_{tb}| \approx 1$. Direct measurements of $|V_{tb}|$ not relying on this assumption can be made, but this yields a much weaker constraint of $|V_{tb}| > 0.74$.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} * & * & * \\ * & * & * \\ 10^{-2} & 10^{-2} & 0.9991 \end{pmatrix} \tag{3}$$

So, the fraction of top quark decay branching ratios is:

$$R = \frac{Br(t \to Wb)}{Br(t \to Wq)} = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2} \approx 1 \tag{4}$$

Consequently, when measuring the top mass, we can assume that a $t\bar{t}$ pair will always decays into a pair of W bosons and a pair of b quarks.

On the other hand, W boson can decay either leptonically into any charged lepton-neutrino pair of the same family (with nearly family-independent branching ratio), or hadronically into quark anti-quark pair of different flavors. So the $t\bar{t}$ final states are classified according to the W boson decay modes.

- When the two W bosons coming from the top and the anti-top decay hadronically, the final state is called the **all-jets** or **full hadronic** final state. This decay channel has the largest branching ratio (∼55%) but also low signal to background ratio (S/B), due to the largest QCD background coming from multijet events.
- The **lepton+jets** channel occurs ∼38% of the time when one W boson decays leptonically and one hadronically. This mode has reasonable statistics and reasonable background (Figure 1.4).
- When two W bosons decay leptonically, we have **dilepton** channel: here lepton denotes electron or muon; even if the request of two leptons makes the signal to background ratio very large, this channel suffers from low statistics at the Tevatron (∼7%). Moreover this measurement is complicated by the two non observable neutrinos in the final state.

Particular attention is required for τ events, where at least one W boson decays into a $\tau\nu_\tau$ final state. In fact, when W decays into $l\nu_l$, with $l = e, \mu$, the result is an unambiguous experimental signatures, whereas τ lepton presence is more complicated as tau quick decay can be misidentified as a narrow jet, an isolated track, or an electron or muon:τ is challenging to identify in an hadronic environment. Our measurement is performed in the *lepton+jets* channel, as it offers the best balance between a clean signature and a good branching ratio. However, the missing neutrino measurement requires to simulate top-quark decay and product together with detector response; the price is the systematic uncertainties introduced by the finite detector resolution of the simulation model. Moreover, we also include events of W decay into τ, in which the τ decays into an electron or muon. Top quark pairs decay modes and their branching ratios are illustrated in Figure 1.5.

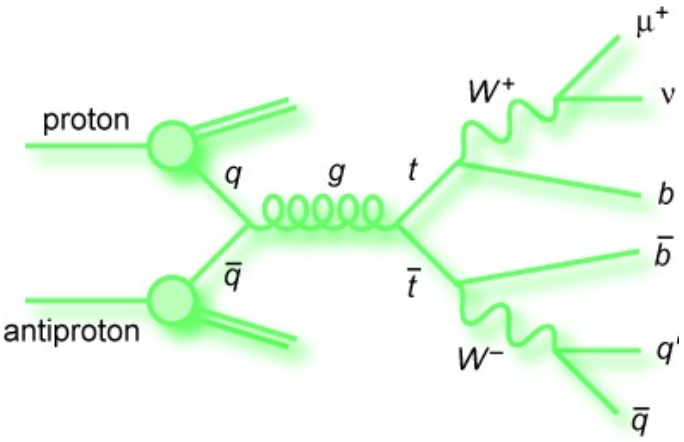


Figure 1.4: Example of the Feynman diagrams of $t\bar{t}$ production through annihilation of quarks and *lepton+jets* decay channel[see 1].

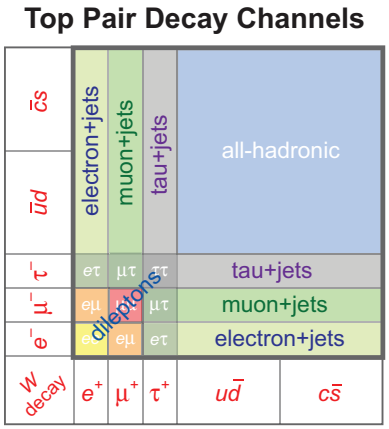


Figure 1.5: The graphic represents all the $t\bar{t}$ decay modes. The areas are proportional to the branching ratios[1].

1.3 Role of the top quark mass in SM and beyond SM particle physics

The Standard Model does not predict any of the particles' masses, so the only way to determine them is by direct measurement. The top quark is the heaviest known fundamental particle, with a measured mass of 173 GeV/c^2. Being heavier than a W boson (see Table 1.2), it is the only quark that can decays semi-weakly, i.e., into a real W boson and a b quark, with a very short lifetime ($\tau_\tau \approx 0.3 \cdot 10^{-24}$ s), despite the decay being mediated by the weak force ($\tau_{weak} > 100 \cdot 10^{-12}$ s). This quick decay means that, unlike lighter quarks, the top quark does not hadronize ($\tau_{QCD} \sim 3 \cdot 10^{-24}$ s) before decaying but rather decays as a free quark. Consequently, this means that the properties of the top quark (mass, spin, charge, couplings, production cross section, decay branching ratios, etc.) can be measured directly through reconstruction of its decay products.

In addition, it is the only quark whose Yukawa coupling to the Higgs boson is order of unity. Its peculiar feature provides a unique laboratory where our understanding of the strong interactions, both in the QCD perturbative

(pQCD) and non perturbative regimes, can be tested: top physics is mainly described by pQCD, but top is a colored and unstable particle so non-perturbative effects enter through the back door[1].

The $t\bar{t}$ pair cross section $\sigma_{t\bar{t}}$ is a quantity that allows to test the SM prediction. This measurement enables the possibility to search for new physics that can manifest itself in anomalous cross section values for different top decay channels. $t\bar{t}$ events are also an important background for Higgs boson searches.

An accurate knowledge of top properties can bring key information on fundamental interactions at the electroweak breaking scale and beyond. In fact top mass value is close to electroweak symmetry breaking scale, due to top-quark mass high value; so this, together with W and H precision physics, provides strong lever for testing the internal consistence of SM. The top-quark and W-boson mass values constrain the Higgs-boson mass, as shown in Figure 1.6.

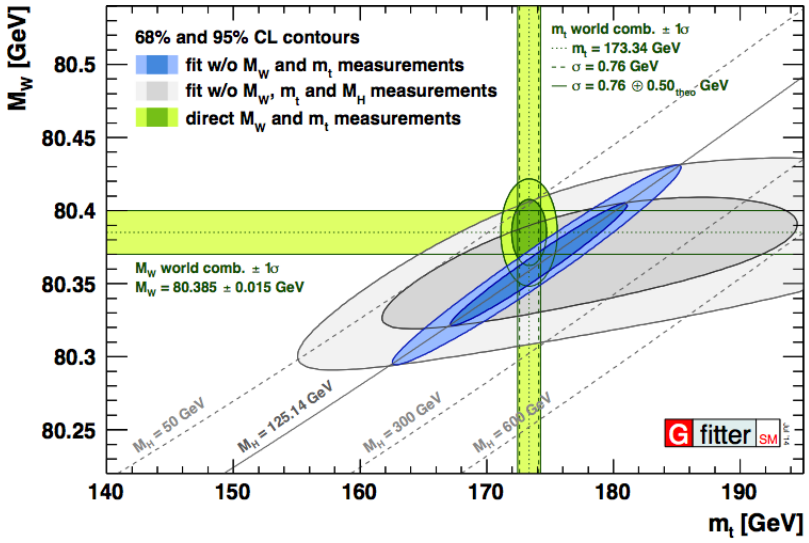


Figure 1.6: In this plot, the diagonal gray lines indicate the predicted value of the Higgs mass, given W-boson and top-quark mass value and various electroweak parameters. The horizontal and vertical bands indicate 1-σ confidence regions of the measured W-boson and top-quark mass values. The green oval contour are 1-σ and 2-σ confidence areas for joint W-boson and top-quark masses. The blue contours are 1-σ and 2-σ confidence areas for W-boson and top-quark masses as predicted from electroweak parameters and Higgs-boson measured mass. The agreement between the experimental measurement and the predictions indicate a self-consistency of the SM [2].

The EW vacuum stability depends crucially on the precise top mass value, as shown in Figure 1.7: higher top mass value eventually leads to scenario of metastable or unstable Universe. The potential, a function of the Higgs field, depends on two terms: one determines the Higgs-boson mass and the other is a self-interaction term that is sensitive to the value of the top-quark mass. For some combination of top quark mass and Higgs boson masses the potential minimum in which the Higgs field currently sits in not the absolute minimum of the potential and quantum tunnelling to a lower-energy state is permitted. In such cases, particles physicists speak of metastability of the electroweak vacuum: the universe is in a state that may endure for a very long time, but not forever. Unfortunately, systematic and theoretical uncertainties limits the precision of top-quarks mass measurements obtained at the hadron colliders, and to improve the measurement precision, the community will need an electron-positron collider capable of producing top quarks.

1.4 Status of the top quark mass measurement

The last world combination of measurements of the top quark mass has been performed in 2014 [3] between the results of CDF and DØ at the Tevatron and the ATLAS and CMS experiments at the LHC. The resulting combined measurement, with a precision of 0.4% is:

$$m_t = 173.34 \pm 0.27(stat) \pm 0.71(syst) \text{ GeV}/c^2 \quad (5)$$

[1]Non-perturbative QCD affects top quark physics through the showering and hadronization processes of the lighter quarks produced in top quark decays. The showering process accounts for multiple soft gluon radiation and the hadronization process accounts for the final formation of colorless hadrons from the colorful partons produced in the perturbatively treated top quark production and decay.

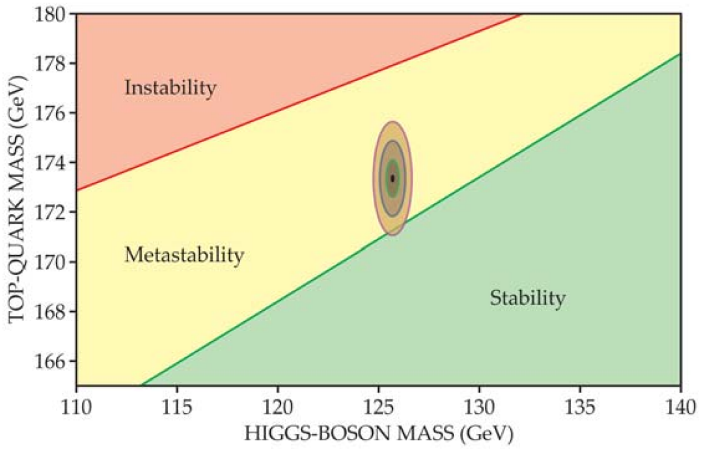


Figure 1.7: The Universe is metastable, long-lived but not eternal, for certain combination of top-quark and Higgs-boson masses. Given the current determination of the top-quark mass the Higgs-boson mass of about 125 GeV is close to the boundary of stability, but a definitive answer will require a much more precise measurement of the top-quark mass. The three ellipses represent 1-σ, 2-σ and 3-σ confidence areas for the mass determinations.

Figure 1.8 summarizes the input measurements and the results of the combination. However, this scenario is now incomplete because it does not include two more recent and very precise results:

– DØ final measurement in lepton+jets: $m_t = 174.98 \pm 0.76$ GeV/c^2 [4].

– CMS measurements in all channels: $m_t = 172.44 \pm 0.48$ GeV/c^2 [5].

The DØ latest result used matrix element technique in lepton+jets final states of the full Run II sample, while CMS result is based on the proton-proton data $\sqrt{s} = 8$ TeV, in the combined lepton+jets, all-jets and dilepton decay channels, by using the ideogram method. These results represent the most precise measurements ever obtained for the top quark mass. However, there is a significant discrepancy between them of $\sim 3\sigma$. With this new top mass measurement, CDF aims at improving considerably the systematic uncertainty as well as the statistical weight of the events and the statistics itself, exploiting the full CDF Run II dataset, reaching a total error of less than 0.5%. Hopefully, the new measurement will help solving the current tension between the two latest measurement listed above.

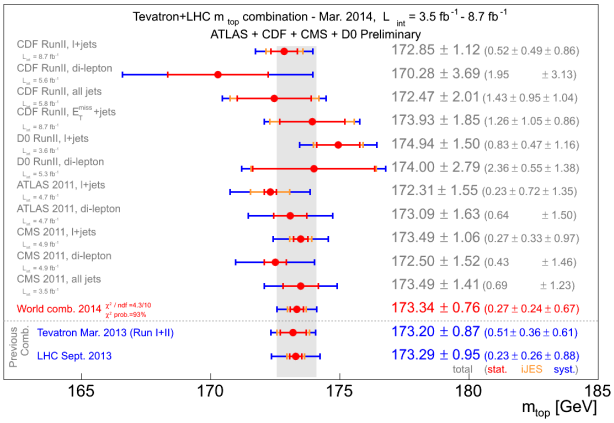


Figure 1.8: Combination of the results for top mass measurements, compared (lowest lines) to Tevatron and LHC m_t values. The red lines correspond to the statistical uncertainty while the blue lines show the total uncertainty. While in the Tevatron experiments the statistical error dominates, at the LHC the error is still mostly systematic. This combination refers to results until 2014, but DØ and CMS 2014 results are not shown.

2 Description of the present analysis

2.1 Event selection

The signature of the lepton+jets events is constituted by a high p_T charged lepton [2], large missing transverse energy [3], accounting for the escaping neutrino, and at least four jets coming from quark hadronization. Jets are divided into *tight* and *loose* jets; the former are required to have $E_T > 20$ GeV and pseudorapidity $|\eta| < 2.0$ while the latter need $E_T > 12$ GeV and $|\eta| < 2.4$. Moreover *lepton+jets* events are further divided into sub-samples based on the number of identified b jets:

- no b-tagged jet (0-tag),
- one b-tagged jet (1-tag),
- and two or more b-tagged jets (2-tag).

In the 0-tag events, we require exactly four *tight* jets. We divide the 1-tag and 2-tag samples into sub-samples based on the number of *tight* jets: the "tight" sub-category requires exactly 4 *tight* jets and any number of *loose* jets, while "loose" sub-category requires 3 *tight* jets and at least one *loose* jet plus eventually any other kind of jets. Thus the analysis uses five categories: 0-tag, 1-tagL, 1-tagT, 2-tagL, and 2-tagT, where L and T subscripts represent loose and tight selection, respectively. The selection requirements are summarized in Table 2.1.

	0-tag	1-tagT	1-tagL	2-tagT	2-tagL
Lepton E_T	> 20	> 20	> 20	> 20	> 20
Lepton $\|\eta\|$	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0
$\not\!\!E_T$	> 20	> 20	> 20	> 20	> 20
3 jets E_T	> 20	> 20	> 20	> 20	> 20
3 jets $\|\eta\|$	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
4^{th} jets E_T	> 20	> 20	> 12	> 20	> 12
4^{th} jets $\|\eta\|$	< 2.0	< 2.0	< 2.4	< 2.0	< 2.4
Extra jets		Any loose	Any loose or ≥ 1 tight	Any loose	Any loose or ≥ 1 tight

Table 2.1: Event selection category with respective requirements depending on the number of identified b jets and on the number of *tight* and *loose* jets.

For signal events we use a large variety of Monte Carlo (MC) simulated sample in implementing our method. In order to model transfer functions, and to calculate the acceptance, we use $t\bar{t}$ signal samples generated by PYTHIA6.2 at leading-order (LO), in the range of top masses between 157.5 GeV and 187.5 GeV. For the calibration procedure we will use $t\bar{t}$ samples accurate at Next-To-Leading-Order (NLO), simulated by using POWHEG with the parton showering performed by PYTHIA6.4. Since the calibration step needs high accuracy, we use a NLO generator and we also enlarge the range of masses used in this procedure with respect to the previous analysis (top mass in *all hadronic* channel) by producing additional MC $t\bar{t}$ samples with top mass in the range 157.5 GeV to 187.5 GeV.

[2] In our analysis we only use electrons which are reconstructed from clusters in the Central Electromagnetic calorimeter (CEM), covering a region in pseudorapidity up to $|\eta| < 1.0$, with $p_T = E_T > 20$ GeV: note that the particle 4-momentum is reconstructed assuming massless electrons. We also select one isolated muon with $p_T > 20$ GeV and $|\eta| \leq 0.68$ in the case of muon observed in central muon detector (CMU), or $0.65 \leq |\eta| \leq 1.0$, in the case of muon detected in central muon extension detector (CMX).

[3] Since for neutrinos the probability of interacting in the detector material is negligible, they escape detection causing an energy imbalance in the observed event. The transverse energy vector must be null in the final state as it is in the initial state of the events, and can serve to sense escaping neutrinos [6]. The missing transverse energy is defined as follows:

$$\not\!\!E_T = -\sum_i E_i^T \hat{n}_i > 20 \text{ GeV} \tag{6}$$

where E_i^T is the transverse energy measured in the i-th tower of the calorimeter and $\hat{n}_i$ is the projection of the versor pointing from the event vertex to the i-th calorimeter tower in the plane perpendicular to the beam axis.

2.2 Sample composition

The W+jets background contains both events with mistagged light flavor jets and events with properly tagged heavy flavor jets. In fact a certain fraction of the events passing our selection cuts can mimic the signature of the signal of interest, even though they are not true $t\bar{t}$ events, but rather background. Since the background events do not contain any useful information on the top mass (the single top events, of course, do contain some information but are still not useful if reconstructed as $t\bar{t}$ events), it is necessary to adopt a strategy to minimize their effect on the reconstructed top mass, so we need to know the expected contributions of each of the above background types to the overall observed total [6].

The main contribution to the background in the *lepton+jets* decay channel comes from events in which a W boson is produced together with a number of hadronic jets. Such processes could create light flavor quarks, but one of the jets is mistagged as a b jet. There is also a background of non-W multi-jet events due to QCD decay. There are also *Diboson* event, in which W and/or Z are produced; in these events first boson decay in two charged lepton, with one lepton unobserved, and the other boson decay hadronically. Some example of the Feynmann diagrams of these events are displayed in Figure 2.1 and Figure 2.2.

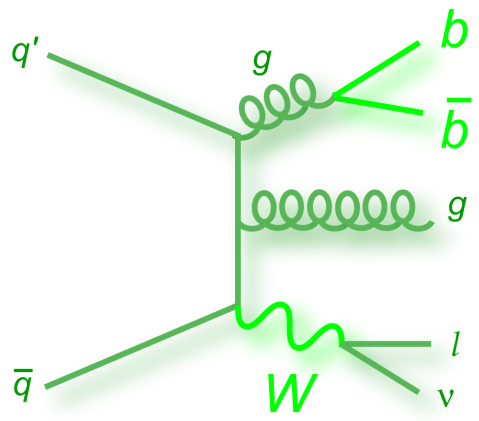


Figure 2.1: Example of the Feynman diagrams of *W+jets* process [1].

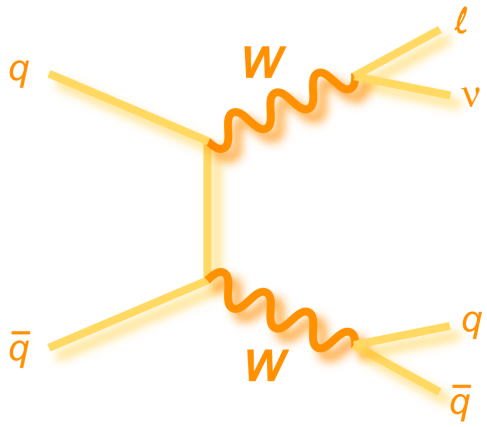


Figure 2.2: Example of the Feynman diagrams of *Diboson* process [1].

For the determination of the selected sample composition we refer to a widely used method known in internal CDF terminology as "*Method II For You*". The basic principle of Method II is to use MC simulation to estimate the background contributions; for channels where the MC is not known it employs a data-based approach. In general, we observe a good agreement between the number of total expected events and the number of observed events. Only in the 2-tagT category we note a substantial difference of $\sim 20\%$. The total signal-to-background ratio, obtained by summing up the overall signal and background and calculating the ratio, is $S/B \sim 1$. All these events refer to a luminosity of 9 fb^{-1}.

	0-tag	1-tagL	1-tagT	2-tagL	2-tagT	All
W+ h.f	697	357	161	34	21	1269
W+ l.f	1581	171	77	3	2	1834
Z+ jets	169	25	14	2	1	212
Diboson	166	31	18	3	2	220
Single top	14	17	8	7	5	50
QCD	623	120	60	1	6	811
Background	3251	720	338	49	37	4395
Signal	960	999	1086	331	425	3801
Total	4211	1719	1424	380	462	8196
S/B	0.3	1.4	3.2	6.8	10.6	0.9
Observed	4474	1711	1434	365	375	8359

Table 2.2: Expected and observed sample composition.

MC samples for background are produced by several generators, as listed below.

- We use ALPGEN+PYTHIA and PYTHIA, respectively to simulate and shower both *W + jets* and *Z + jets* events.

- The single top samples are simulated with a top mass of 172.5 GeV using the MADGRAPH5 package along with PYTHIA for the parton shower and hadronization.

- For modeling *diboson* events we use PYTHIA

- As MC is not able to adequately model the non-W QCD background, we use sample exclusive to our signal sample: practically these events are taken by the inversion of the selection cuts that discriminate events with bad electrons.

3 Introduction to the matrix element method

3.1 Motivation

The Matrix Element (ME) method is a powerful tool in experimental particle physics as it provides a superior statistical sensitivity in evaluating parameters, especially at hadron colliders to measure the top mass. The statistical sensitivity of this method is due to the completeness of the information exploited in each event. This is because the computation of the matrix element checks the full kinematics of the event, instead of using information from only a few variables.

The ME method, as well as other techniques, can be used to determine several unknown parameters[4] at the same time in one measurement, thus also allowing for a reduction of systematic uncertainties. Together with the top quark mass, a simultaneous additional measurement of the jet energy scale (JES) is incorporated in the ME method. For each selected event, the likelihood to observe it is calculated as a function of the assumed top quark mass. For this purpose, all possible reactions yielding to final states that could have led to the observed event are considered. An integration is performed over the multi-dimensional volume in the space of final state momenta defined by the finite detector resolution (In particular, we do not integrate over lepton momenta, assumed to be measured with "infinite" resolution).

The theoretical assumptions about the process under study (parton distribution function, matrix element, transfer functions) are incorporated into the data analysis in the most efficient manner. Some widely used data analysis methods, like the template technique, introduce implicit assumptions about the shape of detector resolution functions.

Despite all the advantages mentioned, the matrix element approach is not necessarily the obvious first choice among various high energy data analysis techniques. A practical challenge associated with the use of the ME technique is that the likelihood calculated is only an approximation of the true likelihood. This situation arises because of finite detector resolution, corrections to the fixed-order matrix element, and neglected information. Therefore, to properly calibrate the applied method, one is forced to perform a large number of "pseudo-experiments" (PE) which can be very expensive from the standpoint of computing time.

3.2 Description of the method

The key concept of the ME method is the fact that the quantum-mechanical probability for top pair production and decay depends strongly on the top mass through the kinematic and topological information. It is based on maximization of a suitable likelihood function (Equation (7)), given by the product of all the events of a combination ($a(f_{sig}, f_{back})$ and $b(f_{sig}, f_{back})$) of the likelihood both of signal and background[5]. You can notice as the factors depend on the fraction of signal and background respectively (f_{sig} and f_{back}) which have to be determined by

[4]Both theoretical parameters describing the physics processes measured as well as experimental parameters describing the detector response can be determined.

[5]Obtaining a consistent m_t estimate is guaranteed by an important property of the Breit-Wigner shape of the top quark mass distribution (which limits the m_t resolution that can be achieved in a single even by a perfect detector): the slope of its log-likelihood is bounded. This means that a contaminating event inconsistent with the model can only make a finite contribution to the determination of top quark mass. In case of Gaussian distribution, even a single observation can give arbitrarily large contributions to the average. On the contrary, the maximum likelihood estimate of m_t is guaranteed to be robust and can tolerate certain deficiencies in the background modelling.

calibration with simulated measurements (PE) . We would consider f as a function of m_t and Δ_{JES} if we want to take into account different expected sample composition for different m_t and Δ_{JES} values. Moreover the signal likelihood depends on both m_t and Δ_{JES} parameters while the background one depends only on the Δ_{JES}; both they have to be determined by calibration with simulated measurements. [6].

$$L_{tot} = \prod_{i=1}^{N} \left[a\left(f_{sig}, f_{back}\right) L_i^{sig}\left(m_t, \Delta_{JES}\right) + b\left(f_{sig}, f_{back}\right) L_i^{back}\left(\Delta_{JES}\right) \right] \tag{7}$$

where

$$JES = \frac{p_T^{MC-jet}}{p_T^{Cal-jet}} = 1 + \Delta_{JES} \cdot \sigma_{p_T}^{Cal-jet} \tag{8}$$

Here Δ_{JES} parametrizes the uncertainty in our knowledge of the jet energy scale (JES) and it is a shift applied to all jet energies in units of the jet-dependent systematic error. By introducing Δ_{JES} into the likelihood, we can use the information contained in W boson decays to constrain Δ_{JES} and reduce error due to this uncertainty - a method known as the *in situ* calibration of the jet energy.
Equation (9) is the formal expression of the signal likelihood.

$$L_i^{sig}\left(m_t, \Delta_{JES}\right) = \frac{1}{\sigma\left(m_t\right)} \frac{1}{A\left(m_t, \Delta_{JES}\right)} \sum_{j=1}^{24} w_{ij} P^{sig}\left(\vec{x}_i | m_t, \Delta_{JES}\right) \tag{9}$$

- $\sigma(m_t)$ refers to the total cross section;
- $A\left(m_t, \Delta_{JES}\right)$ refers to acceptance, defined as as the ratio between the number of selected events in the specific category and the number of the total lepton+jets generated events. The acceptance is calculated for each category varying m_t value in the range $157.5 - 187.5$ GeV with steps of 1 GeV and Δ_{JES} parameter between -3 and +3 with steps of 0.2;
- the topological information of the process is contained in the sum, which is over all the possible jets coming from different quarks and the weights take into account the efficiency in the identification of the right jet.

The formula for the P^{sig} is given by equation (10):

$$\begin{aligned} P^{sig}\left(\vec{x}_i | m_t, \Delta_{JES}\right) = \int \epsilon\left(\vec{x}_i | \vec{y}_i, \Delta_{JES}\right) T\left(\vec{x}_i | \vec{y}_i, \Delta_{JES}\right) \left| M_{2p\to l\nu_l+4p}^{t\bar{t}}\left(m_t, \vec{y}_i\right) \right|^2 \\ \times \frac{f\left(z_1, Q^2\right) f\left(z_2, Q^2\right)}{z_1 z_2} \Big|_{Q^2=4m_t^2} \mathrm{d}z_1\, \mathrm{d}z_2 J(\vec{y}_i | \vec{y'}_i)\, \mathrm{d}\Phi\left(\vec{y_i'}\right) \end{aligned} \tag{10}$$

where M represents the matrix element of the production of $t\bar{t}$ and the decay of the W bosons in lepton+jets. There is also a flux factor $(z_1 z_2)$ which normalize all things, and all this functions are integrated in a 19 dimensions phase space. This formal expression implies the Jacobian J for the change of the variables, while the Transfer Function (TF) assume a efficiency factor ϵ, referred to the event detection system and reconstruction algorithm, to be normalized [7]. The 19 integration variables (y_i') we choose are:

- The two masses squared m_t^2 of the two top quarks in the event.
- The two masses squared M_W^2 of the two W bosons in the event.
- The logarithm $\ln(p_1/p_2)$ of the ratio of the jet momenta in the hadronic W decay.
- The two components of the transverse momentum p_T^{ttbar} of the top quark pair.
- The momenta of the four quarks (2 b and 2 from the hadronic W decay) in the (η, ϕ, m) space (12 elements).

A 20^{th} variable is added, the charged lepton momentum, when the leptonic W decays into a τ, which in turn decays into an electron or muon and neutrinos [8]. Moreover ($\vec{y}$) refers to parton-level quantities used in the matrix element while ($\vec{x}$) refers tp the reconstructed quantities observed in our detector.

This formula refers to the signal; a similar formula is constructed for the background: the differences lies on the non dependence on the top mass and on different event of course. In fact the matrix element M refers to a process without top quarks and, because there are no intermediate particles in the background, like the top quarks in the signal sample, whose kinematic integration is sensitive to final state angular variables, we only integrate over a limited number of variables:

- The momenta of the four jets (4 elements).
- The two components of the transverse momentum p_T^W.
- The W squared mass M_W^2.

This gives a total of only 7 integration variables in the case of the *W+jets* event probability, while the integration over M_W^2 retains the dependence on Δ_{JES}.

3.3 Transfer functions

The transfer functions are one of the most important components of the matrix element method, as they depend on jet fragmentation properties convoluted with detector response: in other words, they connect the parton-level quantities (y) used in the matrix element with the reconstructed quantities observed in our detector (x). In principle, transfer functions can be used for all of the objects observed in our detector: jets, leptons, and $\not\!\!E_T$. However, we only apply our transfer functions to the jets, as we assume that the lepton momentum is well-measured in our detector[6] and we do not use the value of the measured $\not\!\!E_T$ in our likelihood calculation at all, since it does not contain sufficient information to be useful in our top mass calculation. The transfer functions are built by taking Monte Carlo events, matching the jets to the partons. This give the probability T(x|y) desired[7].

There are two version of the TFs.

$$\boxed{T_{old}} = F_1\left(\frac{p_T^j}{p_T^p}; p_T^p, \eta_p, m_p\right) \times F_2\left(\Delta\eta_{j-p}, \Delta\phi_{j-p}; p_T^p, \eta_p, m_p\right) \tag{11}$$

$$\boxed{T_{new}} = F_3\left(\frac{p_T^j}{p_T^p}, \Delta R_{j-p}; p_T^p, \eta_p, m_p\right) \qquad \text{where} \quad \Delta R_{j-p} = \sqrt{\left(\Delta\eta_{j-p}\right)^2 + \left(\Delta\phi_{j-p}\right)^2} \tag{12}$$

We factorize the old transfer functions into separate parts for the momentum and for the angles. The momentum transfer functions are constructed as a function of the variable $r = \frac{p_T^j}{p_T^p}$, while the angular transfer functions are built as a function of $\Delta\eta_{j-p} = \eta_j - \eta_p$ and $\Delta\phi_{j-p} = \phi_j - \phi_p$; both of them depend on the parton variables p_T^p, m_p and η_p. The new TFs are functions of the momentum and the variable ΔR_{j-p}.

The difference between old and new TF are summarize in the following statements:

T_{old}	T_{new}
– Derived from Pythia6.2.	– Derived from Pohweg + Pythia6.4.
– Only LO.	– Extra parton emission at NLO resolved by jet-to-parton matching.
– Angular variables factorised as $\Delta\eta_{j-p}$ vs $\Delta\phi_{j-p}$.	– Angular decomposition is made through the Jacobian: $\Delta R_{j-p} \to (\Delta\eta_{j-p}, \Delta\phi_{j-p})$.
– T_{old} are constructed for tight event categories.	– T_{new} include also loose event categories.

[6] Essentially the transfer function for lepton momentum is a delta function.

[7] Note that the forward-backward asymmetry does not affect TFs because they are symmetric in angle.

They refer to different MC generators: PYTHIA6.2 standalone for the old one while POWHEG+PYTHIA6.4 for the new ones; this means that NLO correction are added to the new ones, instead of only LO in the old one. Both TF sets are subject to jet-to-parton matching; it eliminate situations where the partons and jets are scrambled, which can happen with or without an extra parton emission in the "hard" process described by the ME and such situations give badly behaving likelihood. Jet-to-parton requires the jet axis to lie within a cone of radius $|\Delta R_{j-p}| \leq 0.4$ with axis along the momentum vector of the parton.

As I mentioned above we have different variable for new and old TF, so the decomposition from ΔR_{j-p} to $(\Delta\eta_{j-p}, \Delta\phi_{j-p})$ is made through the Jacobian. Finally the old TF are constructed only for 1-tagT and 2-tagT (*tight* category), while the new ones include also 0-tag, 1-tagL and 2-tagL. (*loose* category)

3.4 Description of pMC and qMC integration

Usually we have to evaluate the integral of this complicate function (f) in a $s = 19$ dimension phase-space $(\mathrm{d}\vec{x})$ of volume V. To do that, it is convenient to approximate the integral with the average of the function itself evaluated on a finite number (N) of point $(\vec{x_i})$, with the result of getting an error (ϵ)given by the precision of the approximation. Now the problem is how to choose these points.

$$\int_{[0,1]^s} f(\vec{x})\,\mathrm{d}\vec{x} \approx \frac{V\left([0,1]^s\right)}{N}\sum_{i=1}^{N} f(\vec{x}_i) \qquad \text{error} \quad \epsilon \equiv \left|\int_{[0,1]^s} f(\vec{x}) - \frac{1}{N}\sum_{i=1}^{N} f(\vec{x}_i)\right| \tag{13}$$

The common method is to use pseudo random sequences of point uniformly distributed and this method is called pseudo-MC (pMC). The other way is to use Quasi-Random Monte Carlo sequences (qMC) [9], that are generated by choosing those points in the integration space. They both use importance sampling algorithms to identify intervals of the phase space where the likelihood function is most significant and then sample more points in those intervals and fewer in the rest of the phase space. The importance sampling is random in pMC and deterministic in qMC, but in both methods it is based on finding the important intervals for integration (those where the function is most significant). The errors in precision of these techniques depend on the number of integration points N as:

$$\text{pseudo-Monte Carlo:} \quad \epsilon_{_{pMC}} \propto \frac{1}{\sqrt{N}} \qquad \text{quasi-Monte Carlo:} \quad \epsilon_{_{qMC}} \propto \frac{(lnN)^s}{N} \tag{14}$$

The difference in precision between the two methods is shown in Figure 3.1 and Figure 3.2. Notice that both pMC and qMC are of the same speed for the same number of points, because almost all the CPU time is spent in evaluating the likelihood function at the integration points. But qMC is more accurate (the expected integration error is much smaller), because it is smarter in finding importance intervals (some of the sampled points in pMC are always wasted because they are random). This difference allows for using qMC with fewer points for the same integration accuracy, that effectively saves time. In our analysis we will use Quasi–Monte Carlo (Quasi-MC) integration technique which significantly reduces the time required to integrate an event, allowing us to improve the precision and accuracy of our result; but I have done some studies with Pseudo-Monte Carlo (pMC) integration technique.

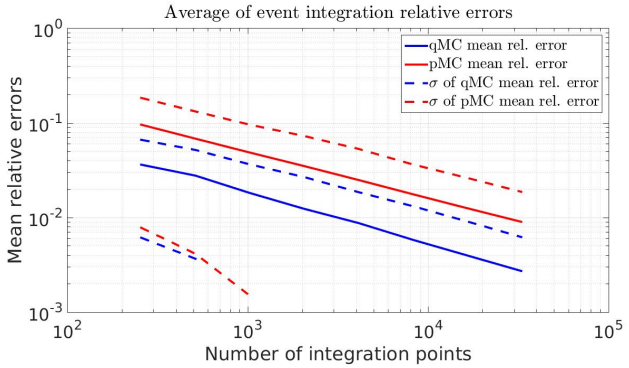


Figure 3.1: Trend of the relative computational error in pMC and qMC technique.

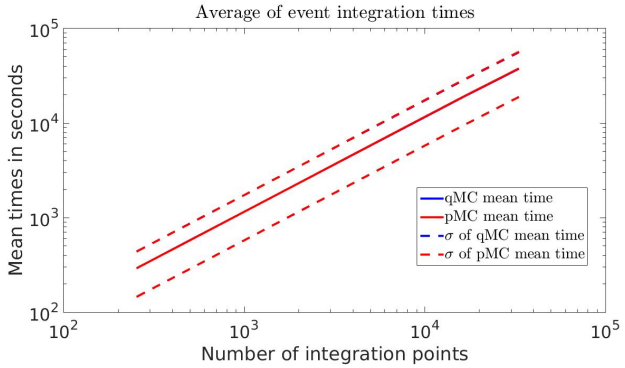


Figure 3.2: Trend of the computational time in pMC and qMC technique.

4 Description of the analysis tool kit

The integration for our analysis is performed through a complex computation framework, known as "CAF attack framework" not too different, in principle, from products like SETI@home or folding@home. FermiGrid is the actual Central Analysis Farm (CAF) of Fermilab that you are using for your "CAF attack framework". It has the significant advantage over most CDF analyses in that it is completely independent of CDF software and data handling, which means you can run it pretty much anywhere you can get CPU time. It consists in three main parts organized in a hierarchy: one director, one or more masters and several workers. The director coordinates the master and the workers. The masters serve the events to workers. The workers run on the FermiGrid and are responsible for the actual integration. At the beginning, the worker connect first to the director which directs it to an available master, then to a master which will continue serving events until all events in the CAF attack are completed; it's not uncommon for workers to be killed by the CAF for running out of time while in the middle of the integration. The worker performs the integration and sends the event's results back to the master to be stored. The process is iterated till completion of the project. A project is a set of files to be run by a single director with a given configuration set for the integration, defining the matrix elements, the transfer functions, the range of the $(m_t, \Delta_J ES)$ grid to scan over, the convergence target and so on. The integration for a single event ends when one of the following three conditions is reached:

- the maximum allowed time to process one event (usually set for 24 h).
- the maximum number of integration points (usually set for 1024).
- the required integration precision (usually set for 10^{-10}).

For the signal probability the integration is performed on a grid consisting of 61 points in Δ_{JES} from -3.05 to 3.05, and 40 points in 1 GeV intervals in m_t in the range $157.5 - 187.5$ GeV. For the W + jets we must study only the dependence on Δ_{JES}, which is done in the same range and step-spacing as for the signal.

5 Study of the pull distribution for the W+jets integration

To test the pMC we created the distribution of the variable δ_i defined in equation (15), where μ and σ are respectively the mean and the standard deviation of different integration [10].

$$\delta_i = \frac{x_i - \mu}{\sigma} \tag{15}$$

We created a pull for background samples: notice that for the moment the acceptance described above in not included yet. I built this pull with the following algorithm:

- Fixed the specific event, over 1000 different events.
- Fixed the specific Δ_{JES}.
- Used 10 different integration results, changing the integration seed each time.
- Create the mean and the standard deviation with these 10 different results.
- Create pull with these 10 results.
- Sum the pull for different Δ_{JES}.
- Sum the pull for different events.

We expect that this distribution asymptotically behave as a Gaussian normal distributed. The result obtained is good enough. In fact we obtain the following value:

- $\mu = -0.116 \pm 0.002$
- $\sigma = 1.185 \pm 0.001$

The fit is displayed in Figure 5.1 to show how close the pull histogram follows the expected Gaussian distribution. The fit's values do not match perfectly with a standard Gaussian; negative skewness is also present. It is plausible to expect that these differences will be reduced as statistics increase for the pull construction: 10 different integration trials are not sufficient.

We also make a profile of the mean (Figure 5.2) and the standard deviation (Figure 5.3) while Δ_{JES} parameter is varying: the result is pretty good because they have respectively the maximum and the minimum for central Δ_{JES}, but their expected behaviour in the neighbourhood of the maximum for the mean or the minimum for the standard deviation profile is parabolic, not quartic, but it is not observed so because the pull distribution is not exactly a Gaussian.

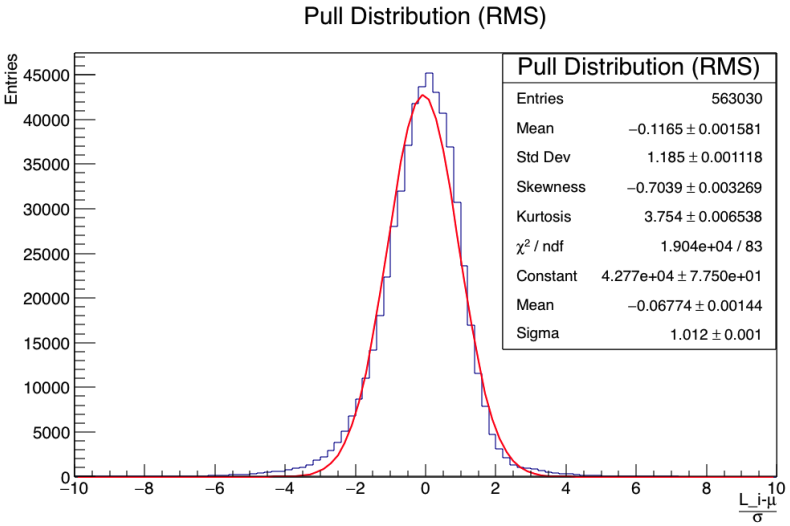


Figure 5.1: Result of the pull distribution. The fitting curve is a Gaussian.

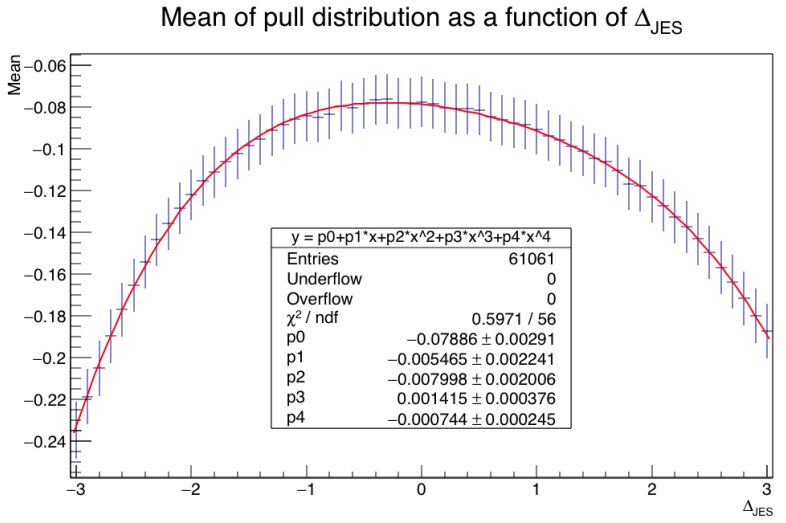


Figure 5.2: Profile of the mean of the pull for different Δ_{JES}. The fitting curve is a quartic polynomial.

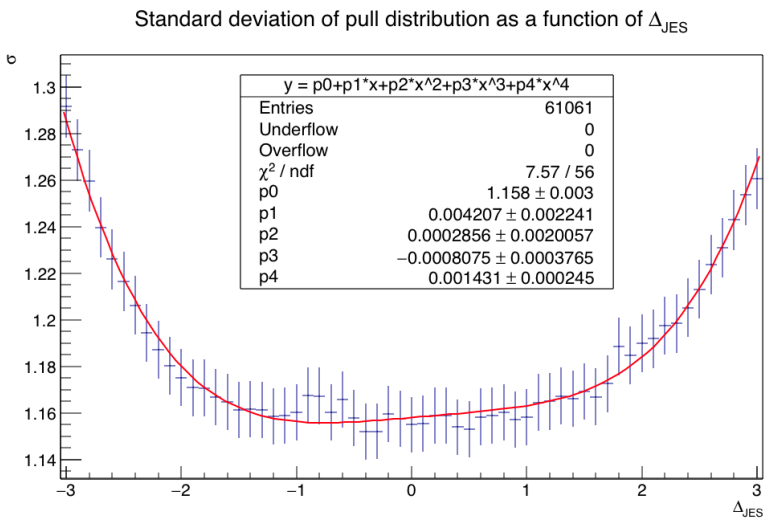


Figure 5.3: Profile of the sigma of the pull for different Δ_{JES}. The fitting curve is a quartic polynomial.

6 Comparative study of the transfer functions derived from LO and NLO MC

The T_{old} are projected on $\Delta\phi_{j-p}$ axis of a 2D histograms of $\Delta\eta_{j-p}$ vs $\Delta\phi_{j-p}$. The T_{new} are projected on $\Delta\phi_{j-p}$ axis of a 2D histograms of $\frac{p_T^j}{p_T^p}$ vs $\Delta\phi_{j-p}$. They both depend on m_p, p_T^p, η_p, Δ_{JES} and the parton type ($isB = 0$ for light quarks or $isB = 1$ for b-quarks). Notice that the case with $m_p = 0.5$ contains only plots with $isB = 0$, because b quark bare mass is about 5 GeV.

The purpose of this analysis is to compare the old TFs (made with PYTHIA6.2) with the new ones (made with POWHEG+PYTHIA6.4). The kinematic variables use for the large scan are shown in the following tables. On the left (Table 6.1), there are the wide values of kinematics among all the possible values to scan coarsely a number of extreme possibilities; on the right (Table 6.2), there are values of kinematics that are the most representative of $t\bar{t}$ decays at the Tevatron: the expected $\langle p_T^j \rangle$ for light-flavor jets is $\sim$ 37 GeV and for heavy-flavour jets is $\sim$ 54 GeV. An approximate estimate for the corresponding $\langle p_T^p \rangle$ is 40 and 60 GeV, respectively, and thus we choose these values to display the TFs. We also choose $m_p = 10$ GeV and $m_p = 20$ GeV as moderate parton off-shell mass values between extreme values of 0.5 and 50 GeV used in the "wide kinematics" scan. As a reminder, our range of Δ_{JES} is generally $-3 \leq \Delta_{JES} \leq +3$, so we choose $\Delta_{JES} = -2$, 0, +2 as indicative values, whereas $\Delta\eta_{j-p}$ is bounded by the fact that the probability of having a jet at an angle with the original parton drops rapidly with that angle and thus the TFs become flaky at large values of $\Delta\eta_{j-p}$.

Wide kinematics

m_p	0.5	5	10	20	50
p_T^p	5	25	40	60	100
η_p	-2	-1	0	+1	+2
Δ_{JES}	-2	0	+2		
$\Delta\eta_{j-p}$	-0.2	0	+0.2		
isB	0	1			

Table 6.1: Kinematics values for "Wide kinematics" scan.

Central kinematics

m_p	10	20	
p_T^p	40	60	
η_p	-1	0	+1
Δ_{JES}	-2	0	+2
$\Delta\eta_{j-p}$	-0.2	0	+0.2
isB	0	1	

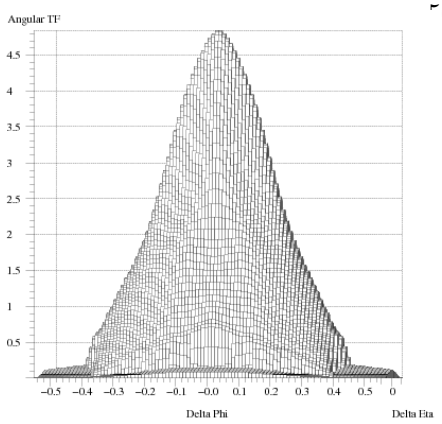
Table 6.2: Kinematics values for "Central kinematics" scan.

6.1 Discussion of the results

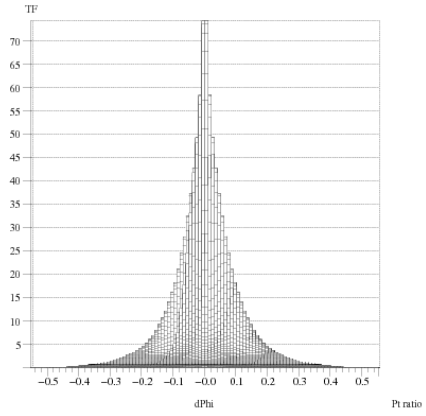
The most prominent difference concerns the width of the angular distributions, which are broader in the old TFs but they show longer tails in the new TFs (Figure 6.1). Prominent is also the difference in the slope of the $\frac{p_T^j}{p_T^p}$

density on the low side, which is steeper in the new TFs (Figure 6.2). These differences grow as the parton mass increases and, at the same time, the parton p_T decreases. In certain cases (Figure 6.3) the angular distribution in the new TFs shows a deep kink in the middle ($\Delta\phi_{j-p} = 0$ bin). In other cases the angular distribution in the new TFs appears flat but wavy in an unphysical way (Figure 6.4).

m_p	p_T	η_p	isB	Δ_{JES}
50	100	0	0	0



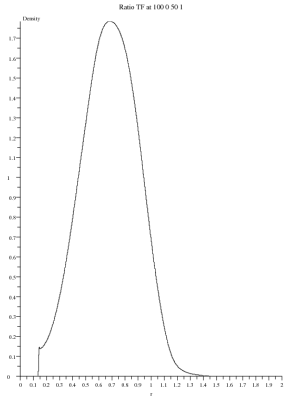
$\Delta\phi$ axis of $\Delta\eta$ vs $\Delta\phi$ plot for T_{old} with $\Delta\eta_{j-p} = 0$.



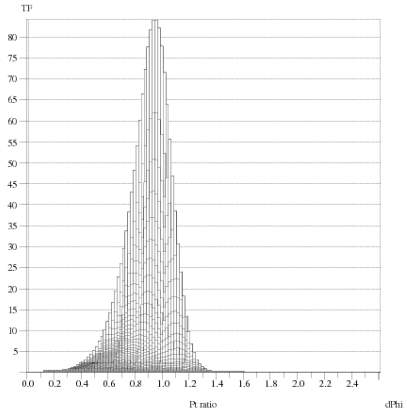
$\Delta\phi$ axis of $\frac{p_T^j}{p_T^p}$ vs $\Delta\phi$ plot for T_{new} with $\Delta\eta_{j-p} = 0$.

Figure 6.1: Comparison of width for angular distribution of TF.

m_p	p_T	η_p	isB	Δ_{JES}
50	100	0	1	0



$\frac{p_T^j}{p_T^p}$ plot for T_{old} with $\Delta\eta_{j-p} = 0$.



$\frac{p_T^j}{p_T^p}$ plot for T_{new} with $\Delta\eta_{j-p} = 0$.

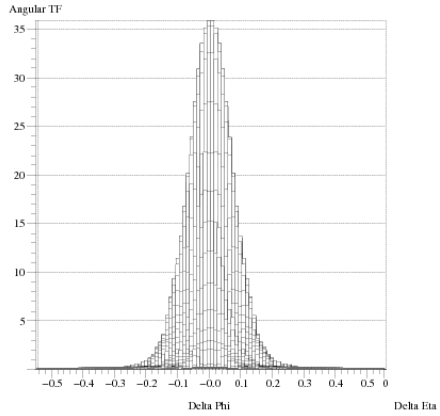
Figure 6.2: Comparison of slope for ratio density.

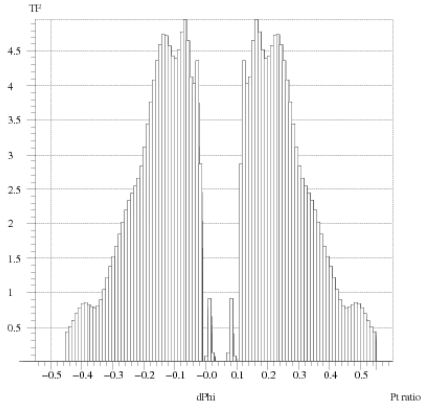
I have done also a comparison between the jet efficiencies made with the previous MC samples. Now both efficiencies are displayed as 2D histograms of p_T^p vs Δ_{JES}, given m_p, η_p and the parton type ($isB = 0$ for light quarks or $isB = 1$ for b-quarks). The values $\eta_p = -2$ and $\eta_p = -1$ are chosen to show the symmetry of the efficiencies about $\eta_p = 0$.

isB	0	1				
η_p	-2	-1	0	+1	+2	
m_p	0.5	1.5	5	10	20	40

A noticeable difference between the two efficiencies is detected only for $p_T^p < 10$ GeV, where the old efficiency dies off smoothly, whereas the new one does not go to zero for $p_T^p = 0$ (Figure 6.5). This is tentatively explained

m_p	p_T	η_p	isB	Δ_{JES}
5	5	0	0	0



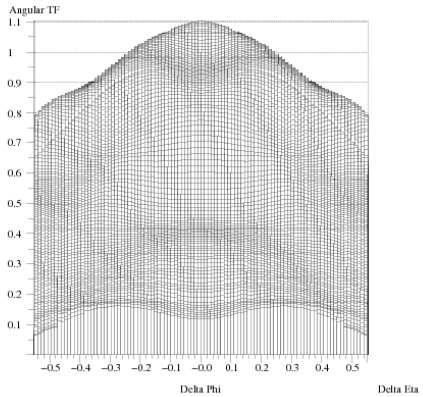


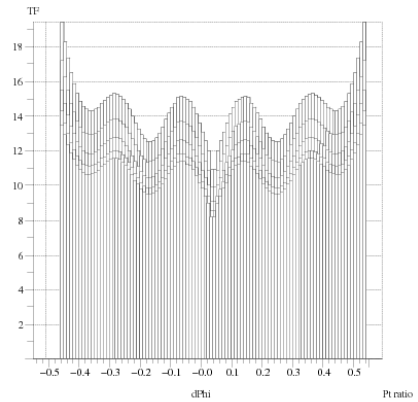
$\Delta\phi$ axis of $\Delta\eta$ vs $\Delta\phi$ plot for T_{old} with $\Delta\eta_{j-p} = 0$.

$\Delta\phi$ axis of $\frac{p_T^j}{p_T^p}$ vs $\Delta\phi$ plot for T_{new} with $\Delta\eta_{j-p} = 0$.

Figure 6.3: Comparison of "kink" in the angular distribution.

m_p	p_T	η_p	isB	Δ_{JES}
50	5	2	0	0





$\Delta\phi$ axis of $\Delta\eta$ vs $\Delta\phi$ plot for T_{old} with $\Delta\eta_{j-p} = 0$.

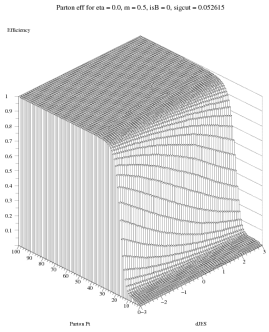
$\Delta\phi$ axis of $\frac{p_T^j}{p_T^p}$ vs $\Delta\phi$ plot for T_{new} with $\Delta\eta_{j-p} = 0$.

Figure 6.4: Comparison of wavy behaviour in the angular distribution.

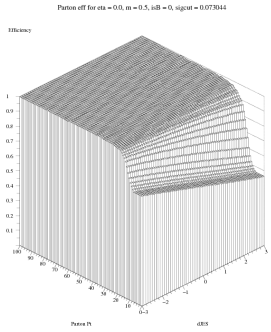
by the extra parton emission in POWEG, which can give partons with arbitrarily small p_T^p. The finite efficiency at $p_T^p = 0$ effect is more prominent at $\eta_p = 0$ (compare Figure 6.5 with Figure 6.6) , where "hard" gluon emission from the primary $t\bar{t}$ event is more likely, and gets reduced with increasing m_p, as the rate of time-like gluon radiation is expected to behave. Nearly no difference between the old and new efficiencies is observed for $isB = 1$, since the process of generating a nearly zero-momentum b quark at NLO in the primary $t\bar{t}$ event (via a hard gluon splitting) has a tiny acceptance (Figure 6.7).

We thinks that all of these differences are mainly caused by different MC used to derive the TFs are also possible. This scenario assumes that the parton-jet matching does not work efficiently, i.e. there are still events where a parton from extra emission is misidentified as a top decay product.

$m_p = 0.5$ $isB = 0$ $\eta_p = \quad 0$



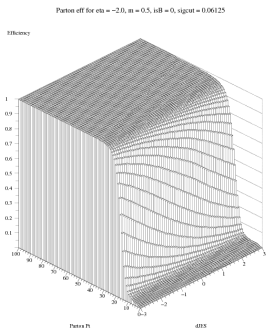
Efficiency plot for ϵ_{old}.



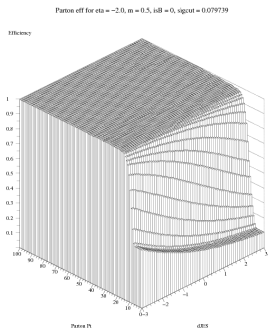
Efficiency plot for ϵ_{new}.

Figure 6.5: Comparison of jet efficiencies.

$m_p = 0.5$ $isB = 0$ $\eta_p = -2$



Efficiency plot for ϵ_{old}.



Efficiency plot for ϵ_{new}.

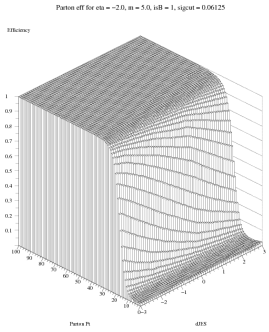
Figure 6.6: Comparison of jet efficiencies.

7 Summary and analysis prospects

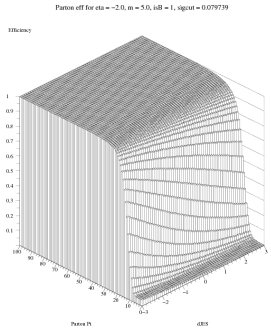
To sum up, we noticed a discrepancy between TFs, so next step is to solve these differences. Also because problem in TFs for signal events may hint problems in background TFs because of similar construction, so this is a way for a better understanding of background TFs. About the pull, the acceptance need to be included in background pulls and together with more statistic this could lead to more precise pulls. Moreover further steps in the development of the analysis framework are:

- Correct the signal TFs.
- Test the background TFs.
- Finalize the qMC integration technique, using scrambled Sobol's sequences which should accelerate the integration convergence.

$m_p = 5.0$ $isB = 1$ $\eta_p = -2$



Efficiency plot for ϵ_{old}.



Efficiency plot for ϵ_{new}.

Figure 6.7: Comparison of jet efficiencies.

- Test the final likelihood (including acceptance and cross section normalization).
- Calibration procedure with realistic pseudo-experiments (PE).
- Run PE with Poisson average equal to the expected candidate events.
- Find average m_t, bias, expected σ_m, pull width for the PE ensemble.
- Derive a combined signal+background likelihood as defined in equation 7 (essentially the factors a and b).
- Derive a calibration parametrization for dependences not included in our likelihood.
- Conduct validation tests (linearity, pull tests ...) of the complete likelihood framework
- Estimation of systematic uncertainties.

References

[1] D0 Collaboration et al. Useful diagrams of top signals and backgrounds. *Public web page. URL http://www-d0. fnal. gov/Run2Physics/top/top_ public_ web_ pages/top_ feynman_ diagrams. html*, 2009.

[2] GFitter Collaboration et al. Results for the global electroweak standard model fit, 2014.

[3] CDF ATLAS. Cms, d0: First combination of tevatron and lhc measurements of the top-quark mass (2014). *arXiv preprint arXiv:1403.4427*.

[4] Victor Mukhamedovich Abazov, Abbott, et al. Precision measurement of the top-quark mass in lepton+ jets final states. *Physical Review D*, 91(11):112003, 2015.

[5] Vardan Khachatryan, Sirunyan, et al. Measurement of the top quark mass using proton-proton data at (s)= 7 and 8 tev. *Physical Review D*, 93(7):072004, 2016.

[6] CECILIA TOSCIRI. Approaching the cdf top quark mass legacy measurement in the lepton+ jets channel with the matrix element method. 2016.

[7] Igor Volobouev. Matrix element method in hep: Transfer functions, efficiencies, and likelihood normalization. *arXiv preprint arXiv:1101.2259*, 2011.

[8] Paul Joseph Lujan. *Precision Measurement of the Top Quark Mass in the Lepton+ Jets Channel Using a Matrix Element Method with Quasi–Monte Carlo Integration.* PhD thesis, University of California, Berkeley, 2009.

[9] William J Morokoff and Russel E Caflisch. Quasi-monte carlo integration. *Journal of computational physics*, 122(2):218–230, 1995.

[10] Luc Demortier and Louis Lyons. Everything you always wanted to know about pulls. *CDF note*, 43, 2002.