Top physics at the Tevatron Run II

Boris Tuchming
CEA Saclay, Dapnia/Spp, 91191 Gif-Sur-Yvette, France

The $p\bar{p}$ collider Tevatron, with its high center of mass energy of 2 TeV is presently the unique top quark factory in the world. Thousands of top-quarks have been produced during Run IIa. This gives the opportunity to study the production and the properties of the heaviest known fundamental particle. This article will summarize a sample of recent top quark physics results obtained at the Tevatron.

1. Introduction

The top quark is the heaviest known fundamental fermion of the Standard Model (SM) with a mass of $\simeq$ 175 GeV/c^2, close to the electroweak breaking scale. This unexpected large mass delayed for several years its discovery. It occurred only 11 years ago at the $p\bar{p}$ collider Tevatron [1, 2], during the so called Run I (1992-1996).

The high mass of the top quark implies a large coupling to the Higgs boson ($Y_{top} \simeq 1$). This makes think it could play a particular role in the electroweak breaking mechanism and may be the gatekeeper to physics beyond the SM. It also causes large effect to the W boson propagation through virtual quantum effects, so that precision measurements of the top mass together with the W mass put strong constraints to the yet to be discovered Higgs boson. Besides, this quark has a very short lifetime and decays before it hadronizes. Therefore the bare quark properties are transfered to his decay products and are not hidden in the spectroscopy that we observe for the other heavy flavor quarks.

Top physics can be studied at several levels. First the production mechanism can be examined, for example by looking at the production kinematics or searching for a resonant production. Secondly the particle properties, can be studied, such as its mass, width, lifetime, spin and charge. Finally the decay properties and the couplings can be studied by looking at the decay products, their kinematics or searching for rare decays.

As the top discovery is recent and only a few events were recorded during Run I, top physics still need to be studied in details. The only existing top factory at this time, is the Tevatron, running at 1.96 TeV since the beginning of Run II in spring 2001. One of the main physics goal of the Tevatron experiments, CDF and DØ, is the study of the top physics.

As of summer 2006, both experiments have about 1 fb^{-1} of data available for physics analyzes, which is almost 8 times more than what was collected during Run I. Analysis presented in the following are based on smaller subsets, from 300 to 750 pb^{-1}. This is still a small fraction of what is expected to be recorded by the end of Run II, in 2009: between 4 and 8 fb^{-1} [3].

THUPL05

2. Production of top quarks at the Tevatron

2.1. Pair production

At the Tevatron, the top quarks are mainly produced by pair. The quark anti-quark annihilation contribute at the level of $\simeq$ 85% while the gluon-gluon annihilation is only $\simeq$ 15%. The main diagrams for these processes are displayed in Figure 1. The cross-section is predicted to be 6.7 pb [4, 5], the level of uncertainty being $\simeq$ 12%.

Typically for $\simeq$ 1 fb^{-1} of data and for a channel with a branching ratio of 30% and an acceptance of 15%, 300 events per experiment are expected to be observed.
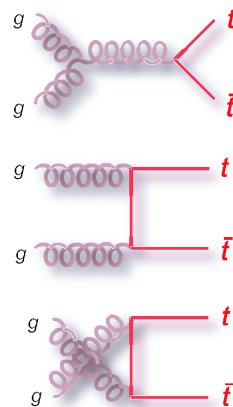

Figure 1: Main diagrams for top pair production.

2.2. Electroweak single top production

Single top quarks can be produced via the electroweak coupling to the W boson. The main diagrams for these processes are displayed in Figure 2. Since they have a different final state, the $s-$channel and $t-$channel are usually distinguished. The $s-$ and the $t-$ channels have respectively cross sections of the order of 0.88 pb and 1.98 pb which is quite low. Given that the background is quite high (there is only one

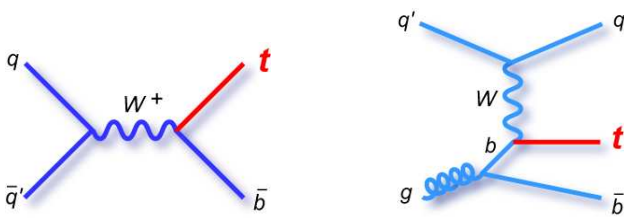top as a signature), this process has not been observed yet and ongoing analysis are searching for it.

Figure 2: Diagrams for single top production. The $s-$ and $t-$channels have different final states.

3. Top pair production

3.1. Signatures

Within the Standard Model, the top decays into a W boson and a b quark almost 100% of the time. The different channels arise from the possible decays of the pair of W bosons: either $W \to e\nu$ ($\simeq 10.7\%$), or $W \to \mu\nu$ ($\simeq 10.7\%$), or $W \to \tau\nu$ ($\simeq 10.7\%$), or $W \to q\bar{q}'$ ($\simeq 69.0\%$).

The main channels are then:

- "lepton + jets" channels ($\simeq 30\%$) correspond to events for which one W decays hadronically and the other one decays into electrons or muons. These channels have a moderate yield and a moderate background arising from W+jets production, Z+jets production, or QCD processes.

- "di-lepton" channels ($\simeq 4.5\%$) correspond to events for which the two W decay into electrons or muons. These channels are very pure but have a small yield. The background is due to di-bons events (mainly WW+jets production), Z+jets and also QCD processes.

- "all jets channels" ($\simeq 44\%$) correspond to events where both W bosons decay hadronically. The yield is high, but the background arising from QCD multi-jet production is very large.

- "tau channels" ($\simeq 22\%$) arise from events where at least one of the W decays into τ. As the τ decays are hard to identify, especially in a hadronic environment as at Tevatron, these channels are weakly exploited.

Because of the high mass of the top quark, the decay products have high momenta and large angular separations. Reconstructing and identifying the production of top quarks demands reconstruction and identification of high transverse momenta (p_T) electrons, muons, jets, and the measurement of the missing transverse energy (E_T). Good momentum resolution for these objects is required and the jet energy scale (JES) has also to be known with a good precision.

Identifying the b-jets is an effective way of improving the purity of the selections. The b-tagging makes use of the presence of secondary vertices and tracks with high impact parameters involved by the decays of b-hadrons. Typically for $p_T = 50$ GeV, a 50% efficiency per b-jet is achieved while the mistag rate is about $0.5 - 1\%$.

Both CDF and DØ analyze their data using various methods, in the different channels. These numerous analysis are updated as the amount of data increases. Only a few examples and the typical selections are discussed in the following.

3.2. Lepton+jets channels

The signature consists of a central spherical energetic events, with a high p_T lepton and a high $\not{E}_T$. Four or more jets can be reconstructed. Among them two jets are expected to arise from b-quark.

The cuts to preselect the events are typically: high p_T lepton trigger (lepton+jets trigger at DØ), isolated lepton with $p_T > 20$ GeV/c, $\not{E}_T > 20$ GeV, 4 jets with $p_T > 15$ GeV/c. Usually the scalar transverse energy H_T defined as the sum of transverse momenta of the reconstructed objects of the events, $H_T = \sum_i |p_{T\,i}|$ can be used to improve purity, as $t\bar{t}$ events are expected to have a higher H_T than the background.

To discriminate between signal and background after these preselections, two main approaches are employed: either the tagging of b-jets or topological cuts which maximize the use of the events kinematics. Each of these approaches is also a cross-check of the other one.

The acceptance is typically at the level of 10 to 20% (for $N_{jets} \geq 3$), it is determined with the simulation (MC), corrected for DATA/MC weights measured on independent samples.

The backgrounds arise mainly from $W + n$ jets, $Wjjj$, $Wbbj$, $Wccj$ (here j stands for light partons, as opposed to heavy flavor, b or c). These backgrounds are determined with the MC, but as the yield suffers from high theoretical uncertainties, these backgrounds are normalized to the data. Non-W background is mainly QCD where one jet fakes a lepton, it is determined on the data. The WW+jets background is determined using the MC.

3.2.1. Lepton+jets with topological analysis at CDF

This analysis is based on an artificial neural network (ANN) which combines seven kinematic variables to discriminate against the main background, W+jets. A likelihood fit of the ANN output distribution allows to determine the $t\bar{t}$ content of the selected events as quoted in Table I. The ANN output distribution
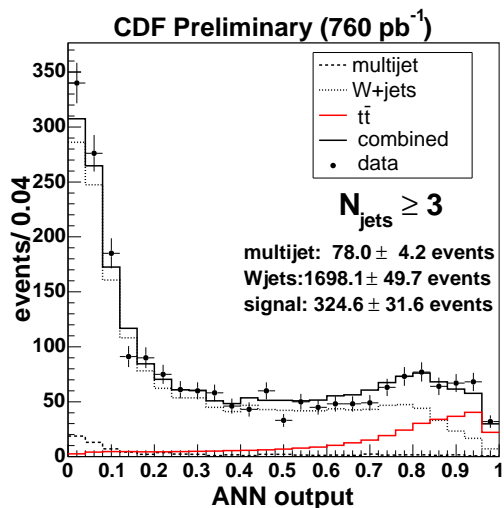
Figure 3: distribution of the *ANN* output of the selected events for the topological lepton+jets analysis at CDF.

Sample	Events	Fitted $t\bar{t}$	$\sigma(t\bar{t})$
$W+ \geq 3$-jets	2102	324.6 ± 31.6	$6.0 \pm 0.6 \pm 0.9$ pb
$W+ \geq 4$-jets	461	166.0 ± 22.1	$5.8 \pm 0.8 \pm 1.3$ pb

Table I Results of the CDF topological letpon+jets analysis.

for selected events is shown in Figure 3. The measurement based on 760 pb^{-1} of data reads:

$$\sigma(t\bar{t}) = 6.0 \pm 0.6 \pm 0.9 \text{ pb.}$$

The main systematic uncertainties for this measurement arise from the JES (8.3%), the W+jets modeling (10.2%) and the luminosity measurement (5.8%).

3.2.2. Lepton+jets with *b*-tagging at CDF

A secondary vertex tag is employ to identify *b*-jets in this analysis based on 695 pb^{-1}. Events can have either ≥ 1 b-tag or ≥ 2 b-tag. The jet multiplicity distribution for one b-tag events is shown in Figure 4. The cross-sections are measured from the 3rd and 4th jet bin in ≥ 1 b-tag events after the cut $H_T > 200$ GeV:

$$\sigma(t\bar{t}) = 8.2 \pm 0.6(stat) \pm 1.0(syst) \text{ pb.}$$

The main systematic uncertainties for this measurement arise from the b-tagging uncertainty (6.5%) the luminosity measurement (6.0%) and the parton density function uncertainty (5.8%).

The $t\bar{t}$ yield can be also extracted from the jet multiplicity of ≥ 2 b-tag events, but given the large uncertainty this result is only a cross-check for the previous measurement:

$$\sigma(t\bar{t}) = 8.8^{+1.2}_{-1.1}(stat)^{+2.0}_{-1.3}(syst) \text{ pb.}$$

THUPL05

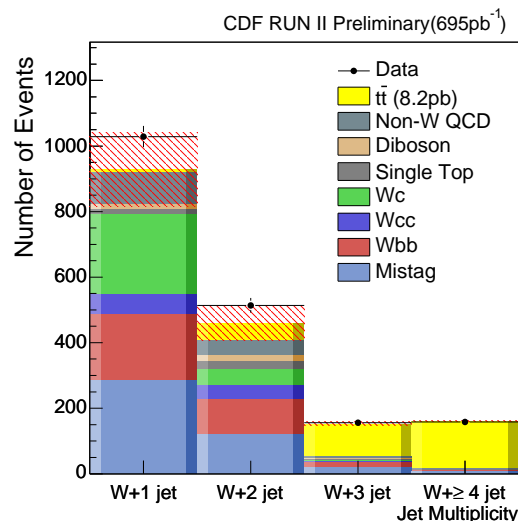Figure 4: Distribution of the jet multiplicity in lepton+jets data at CDF after requiring 1 b-tag and $H_T > 200$ GeV.

3.3. Di-leptons channels

The signature consists of a central spherical energetic events, with two high p_T lepton and high $\not{E}_T$. Two ore more jets can be reconstructed. Among them two jets are expected to arise from *b*-quark.

The cuts to preselect the events are typically: high p_T lepton trigger (or di-lepton), presence of isolated leptons with $p_T > 20$ GeV/c, $\not{E}_T > 20$ GeV, 2 jets with $p_T > 15$ GeV/c and $H_T > 120 - 200$ GeV. This selection is so pure that the b-tagging is actually not employed.

The acceptance is typically at the level of 10 to 15% and is determined with the simulation (MC), corrected for DATA/MC weight factors.

The backgrounds arise mainly from $Z \rightarrow \mu\mu$+jets, $Z \rightarrow ee$+jets, $Z \rightarrow \tau\tau$+jets, WW+jets These backgrounds are determined with the MC. Non-W background consists mainly of QCD events in which jets are faking leptons, it is determined from the data.

The results for the CDF analysis in 750 pb^{-1} of data reads:

$$\sigma(t\bar{t}) = 8.3 \pm 1.5(stat) \pm 1.0(syst) \pm 0.5(lumi) \text{ pb.}$$

For this analysis the main systematic uncertainties arise from the lepton-id (4%), the JES (3.1%) and the $t\bar{t}$ modeling (5%).

3.4. Fully hadronic channel

The fully hadronic channels suffers from a large QCD background and a large combinatoric when one tries to match the jets to the W systems or the $t \rightarrow Wb$ system. The requirement of 2 b-tag jets in the

events helps in improving the signal over background ratio and reducing the combinatorics. But the background is still very large as shown in Figure 5 where the invariant mass spectrum of the reconstructed top system is displayed, before and after background subtraction. DØ with 360 pb^{-1} of data measures

$$\sigma(t\bar{t}) = 12.1 \pm 4.9(stat) \pm 4.6(syst) \text{ pb}.$$

The main uncertainties arise from the background subtraction (25%), the JES (15%) and the b-tagging efficiency uncertainty (18%).

3.5. Pair production summary

The results of the different channels in both CDF and DØ are displayed in Figure 6. The combined CDF result reads:

$$\sigma(t\bar{t}) = 7.3 \pm 0.5(stat) \pm 0.6(syst) \pm 0.4(lumi) \text{ pb}.$$

It is worthwhile to notice that the level of uncertainty is 12%, as low as the theoritical uncertainty. The accuracy is also no more limited by statistics. Thus it can be said that the Tevatron has entered a precision era regarding the $t\bar{t}$ cross-section measurement.

4. Measurement of the top mass

Measuring the top mass is a complex experimental task. The events are complicated since the $t\bar{t}$ production yields six elementary particles in the final state. Even if the $b-$ jets have been identified, there is still a large combinatorics when trying to assign reconstructed jets to initial partons.

Besides, these six particles correspond to different type of objects (jets, b-tagged jets, leptons, $\not{E}_T$), which have to be well identified and measured. As a consequence, the mass measurement depends on the simultaneous understanding of most of the different subdetectors (tracking system, electromagnetic calorimetry, hadronic calorimetry, muon system) of CDF or DØ.

Another difficulty comes from the presence of background. Its contamination is sizable so that it can bias the measurement if it is not well understood and under control.

Other sources of difficulty are encountered when reconstructing the kinematic information in events with $\not{E}_T$. This is particularly true in the di-leptons channels for which two neutrinos carry out unmeasured energy.

Finally, one of the largest source of experimental difficulty is the parton energy measurement. Resolution effects are typically $85\%\sqrt{E_T}$ and go down with increasing statistics. The uncertainty on the jet energy scale is however an irreducible systematics. Recently, an improved method have been set up to reduce this uncertainty: using an *in-situ* calibration of the JES, a part of its uncertainty scales down with increasing statistics as well. This method will be discussed in the following sections.

4.1. Methodology

A lot of different analysis have been performed both at DØ and CDF. They are different from each other depending on:

- The channel, lepton+jets, or di-leptons of fully hadronic;

- the event selection, for example requiring 0 b-tag, 1 b-tag or 2-btag;

- the jet multiplicity requirements;

- the method to derive a mass measurement.

The various methods can be classified in two families:

- The template methods consist in choosing a given kinematic observable, creating signal templates at different M_{top} thanks to the MC, creating template for background events, and using a likelihood fit to determine the best signal template; this method makes use of the kinematic information only.

- The matrix method consists in building a likelihood function based upon the PDF, the matrix elements of $t\bar{t}$ process, and the transfer functions which relate detector measurements (eg: jets p_T) to the top decay products (eg quarks p_T). Unmeasured quantities are integrated out and a maximum likelihood fit allows to derive the top mass. This method needs also to be calibrated using MC events.

The different analysis do not suffer from the same sources of uncertainty. This allows to check their consistency and precision is gained by combining the results. Only three examples (the two most sensitive and one original analysis) are briefly described in the following.

4.2. Lepton+jets template method at CDF

The Lepton+jets template analysis at CDF has been performed on 680 pb^{-1} of data. Four exclusive selections with different purities are performed, namely the 0-btag selection, the 1 b-tag loose selection, the 1 b-tag tight selection and the 2 b-tag selection.
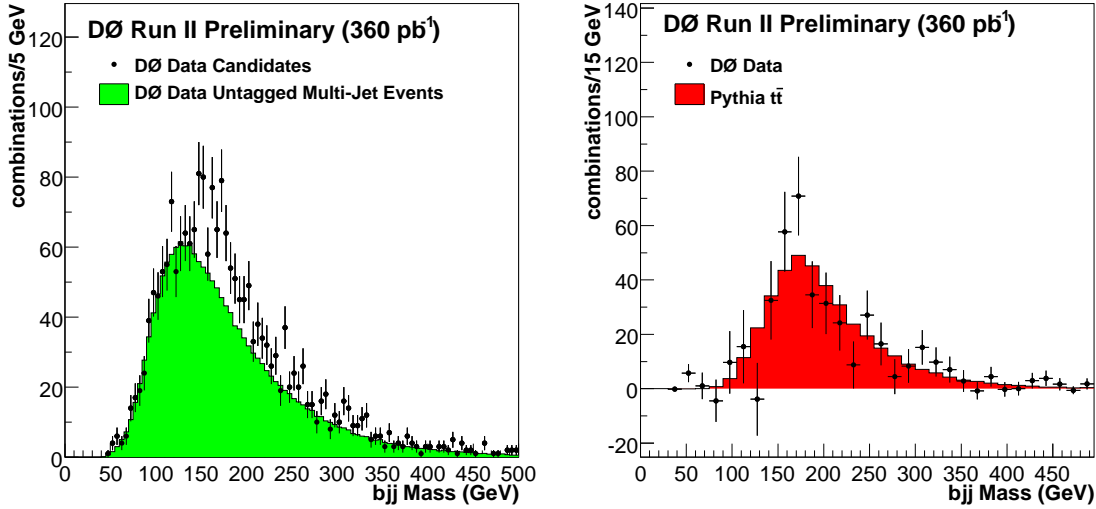
Figure 5: Invariant mass of top system (bjj), before and after background subtraction in the DØ fully hadronic cross-section analysis.
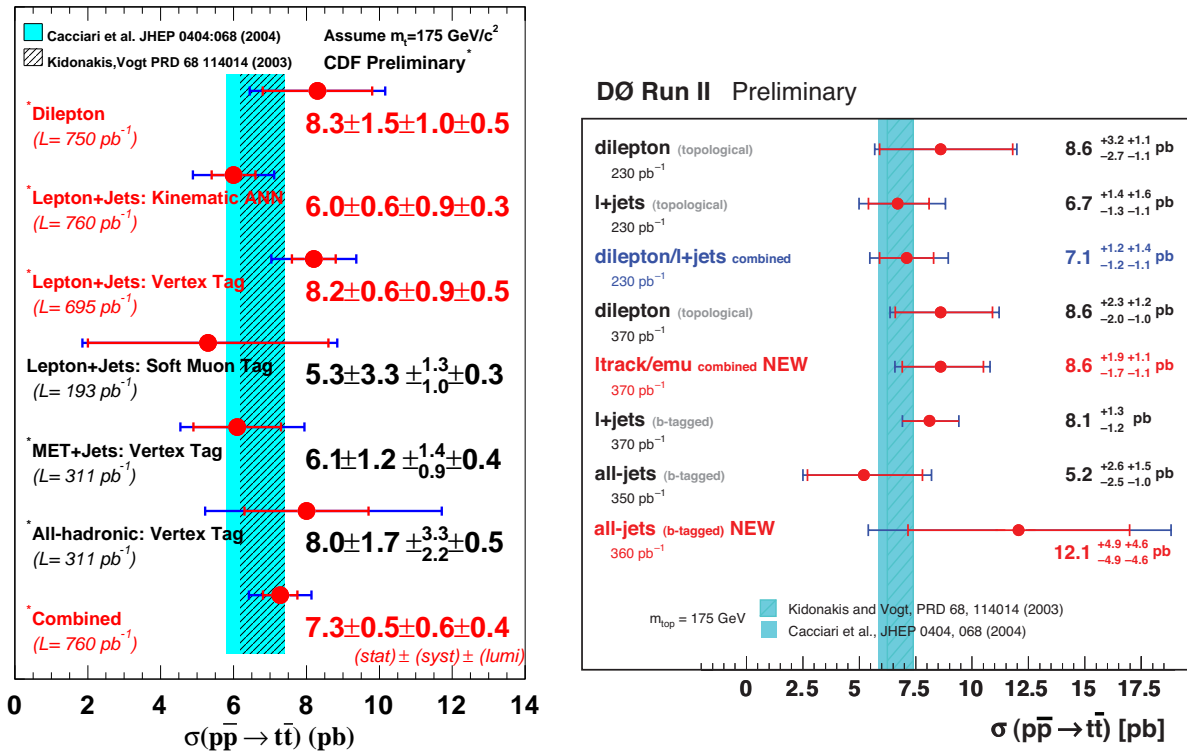

Figure 6: cross-sections measured at CDF and DØ in the various channels

In this channel, the six decay products (if eg the W^+ decays semi-leptonically) are expected to be: a lepton ℓ^+, a neutrino ν, two b-quarks $b\bar{b}$ and two light quarks q and q'. A kinematic fit allows to choose among the various combinations, making use of the b-tag, which reconstructed jets correspond the the qq' system, which jets correspond to the hadronic top, $qq'\bar{b}$, and which jet should be associated to the "leptonic top", $b\ell^+\nu$. Thus for each event a top mass, $M_{qq'\bar{b}} = M_{l\nu b}$, and a W mass, $M_{qq'}$, are reconstructed.

Templates for these two quantities are prepared from the MC after varying the top mass hypothesis,
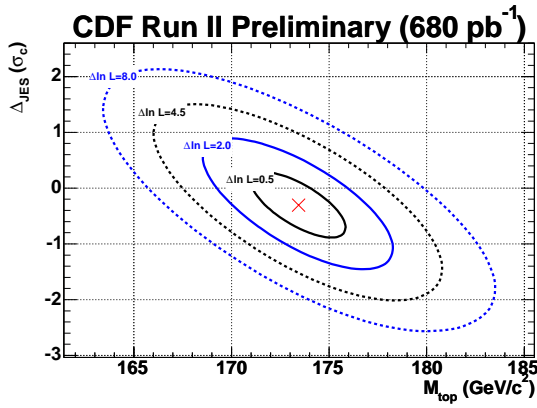
THUPL05

Figure 7: Likelihood contour in plane $(M_{top}, \Delta JES)$ for the CDF template method.

Systematic Source	$\Delta M_{top}(\text{GeV/c}^2)$
b-jet energy scale	0.6
Residual JES	0.7
Background JES	0.4
ISR	0.5
FSR	0.2
PDF	0.3
Generators	0.2
Background Shape	0.5
b-tagging	0.1
MC statistics	0.3
total	1.3

Table II Systematic uncertainties for the top mass measurement in the lepton+jets channel at CDF.

.

M_{top}, and ΔJES which is defined as the deviation from the usual JES calibration in unit of uncertainty on this calibration.

By comparing the templates to the data, a simultaneous fit of $(M_{top}, \Delta JES)$ is performed. Results of this fit are displayed in Figure 7. The results of this fit are the world best single measurement, at the time of PIC06. They read

$$M_{top} = 173.4 \pm 2.5(stat + \Delta JES) \pm 1.3(syst) \text{ GeV/c}^2.$$

The measured ΔJES is

$$\Delta JES = -0.3 \pm 0.6(stat + M_{top}).$$

This signifies that the systematic uncertainty due to the JES as been scaled by 0.6 thanks to the in-situ calibration and the W mass constraint. A further decrease is expected in the future with more statistics.

The systematics on the mass measurement are summarized in Table II.

THUPL05

4.3. Lepton+jets matrix element method at DØ

The DØ collaboration has analyzed 380 pb^{-1} of data using a matrix element method in the lepton+jets channels. Three exclusive samples are selected, namely a 0 b-tag, 1 b-tag and 2 b-tag samples. The signal probability for a set of measurements (jets, leptons) $\vec{x}$ is built:

$$P(\vec{x}, M_t, JES) = \qquad\qquad (1)$$
$$\frac{1}{\sigma} \int d\sigma(\vec{y}, M_t) \ f_1(\tilde{q}_1)f_2(\tilde{q}_2) \ W(\vec{x}, \vec{y}, JES)d\tilde{q}_1 d\tilde{q}_2,$$

where $\tilde{q}_i$ are the initial parton energies, $f_i(\tilde{q}_i)$ are the proton (and anti-proton) parton density function, W is the transfer function from true kinematic quantities $\vec{y}$ to measured $\vec{x}$ and $d\sigma$ is the differential cross-section.

In Equation 1, JES is an additional scaling factor to the externally calibrated Jet Energy Scale. It is left as a free parameter, so that it can be simultaneously fitted, together with the top mass M_t. Thus an in-situ calibration is performed, mostly relying on the constraint arising from the $W \rightarrow q\bar{q}'$ system.

In a similar way, a probability is built for background events P_{bkg} and the final likelihood reads:

$$\mathcal{L} = \prod_{evt}(f_{top}P(\vec{x}, M_t, JES) + (1 - f_{top})P_{bkg}(\vec{x}, JES))$$
$$\qquad\qquad (2)$$

The Likelihood function is shown in Figure 8.
The results of the fit reads:

$$M_{top} = 170.6^{+0.4}_{-4.7}(stat + \Delta JES) \pm 1.4(syst) \text{ GeV/c}^2.$$

The JES in-situ calibration gives:

$$JES = 1.027^{+0.033}_{-0.030}(stat + M_{top}).$$

it indicates a 2.7% correction relative to the externaly calibrated JES, well compatible with its uncertainty ($\simeq 3$ to 4%). The error on the fit reads $^{+0.033}_{-0.030}$. When compared to the externally calibrated JES uncertainty this indicates a small improvement thanks to the in-situ calibration. But this improvements will be more significant with a higher statistics.

The systematics on the mass measurement are summarized in Table III.

4.4. Decay length method

This method employed at CDF is very different from the other ones, as it does not rely on the energy measurements of outgoing particles. Thus it is an important cross-check with reduced impact from the JES uncertainty.
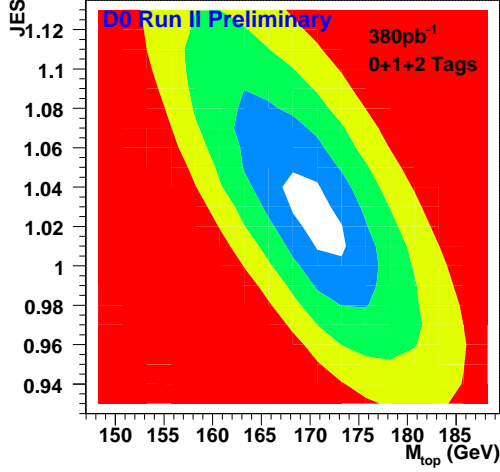
Figure 8: Likelihood contour in plane (M_{top}, JES) for the DØ matrix element method. The different colored region correspond to steps $\Delta\mathcal{L} = 0.5$.

Systematic Source	$\Delta M_{top}(\mathrm{GeV/c}^2)$
b-fragmentation	0.7
b-response	0.8
MC calibration	0.45
W background	0.3
QCD background	0.3
total	1.4

Table III Systematic uncertainties of the top mass measurement at DØ in the lepton+jets channel with the matrix element method.

.

The method makes use of the B hadron decay length and the boost of the b-quark in top decays, directly related to the top mass:

$$\gamma_b = \frac{M_t^2 + M_b^2 - M_W^2}{2M_t M_b} \simeq 0.4 \frac{M_t}{M_b}. \qquad (3)$$

In a lepton+jets sample selected from 695 pb^{-1} of data, the decay length of the b-hadrons are reconstructed and compared to template simulated with different top mass hypothesis. The result reads:

$$M_{top} = 183.6^{+15.7}_{-13.9}(stat) \pm 0.3(JES) \pm 5.6(syst)\,\mathrm{GeV/c}^2.$$

As expected, the impact of JES uncertainty is very small, but the result is statistically limited using the present data sample. However such kind of analysis will become competitive at the end of Run II, when ten times more data will be recorded.
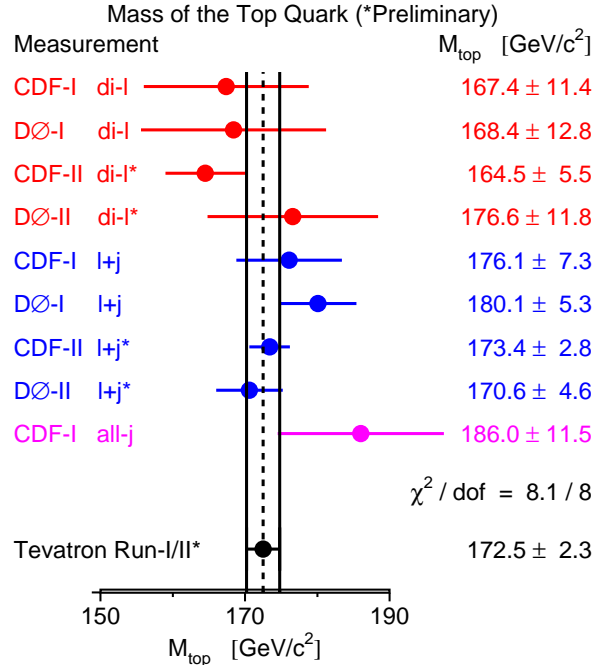
THUPL05

Figure 9: Measurements of the top mass at CDF and DØ.

4.5. Combined results

Nine measurements performed at CDF and DØ have been combined to achieve a 1.3% accuracy:

$$M_{top} = 172.5 \pm 1.3(stat) \pm 1.9(syst)\,\mathrm{GeV/c}^2.$$

The various measurements are summarized in Figure 9.

The chi square of the combination is $\chi^2/dof = 8.1/8$ which demonstrate a good consistency.

It should be stressed that the Run IIa goal was to obtain a 2.5 GeV/c^2 combined uncertainty using $\simeq$ 2 fb^{-1} of data. This goal has been exceeded. The 1% precision is in sight for the near future.

4.6. Impact on the Higgs Boson Mass

The top mass is a free parameter in the standard model. But its mass can be related to various electroweak observables thanks to radiative corrections. Thus the measurement of the top mass provides a consistency check of the SM that can be used to probe for new physics.

In particular the top mass plays an important role in the radiative corrections to the W boson propagator: $\Delta M_W \alpha M_t^2$. As the Higgs boson mass also influences the W boson mass through radiative corrections, $\Delta M_W \alpha \log(M_H)$, the combined measurements of the W and top masses yield constraints onto the SM Higgs boson mass.