Studies of top pair production in the fully hadronic channel at LHC with CMS

Claudia Ciocca

University of Bologna
CMS Collaboration

DIS 2006
XIV International Workshop on Deep Inelastic Scattering
Tsukuba, 20 April 2006

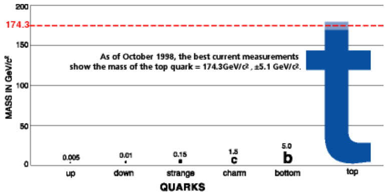
The top quark

Discovery of top quark: 1995 at Tevatron

$I(J^P) = 0(\frac{1}{2}^+)$ charge=$\frac{2}{3}e$

$m_t = 174.3 \pm 5.1$ GeV

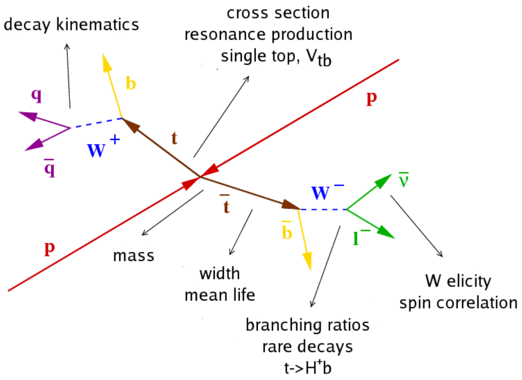
$\Gamma_t \approx 1.5$ GeV $\tau_t \sim 10^{-24}$s



The top quark completed the 3-generation structure of the Standard Model and opened up the new field of the top quark physics. The top quark is distinguished from other quarks:

- decay before hadronization:
  - spin and momentum trasferred to decay products
- large mass, intriguingly close to the scale of electroweak symmetry breaking:
  - role of m_t in EW symmetry breaking mechanism
- unique place to search physics beyond the SM:
  - anomalies in top quark production and decay

The phenomenology of top quark in a slide



$t\bar{t}$ pairs:

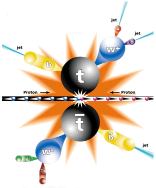
- production rate
- mass
- W elicity
- QCD tests
- new physics in $X \to tt$
- anomalous couplings

single top:

- $|V_{tb}|$
- polarization
- QCD tests
- new physics

Top physics at LHC

LHC will be a real "top factory": production of $\sim$ 8 million $t\bar{t}$ pairs per year (L = 10 fb^{-1}) and of $\sim$ 3 million single top.
NLO cross section for $t\bar{t}$ production at LHC = 830 pb



Precision physics measurements:

- top quark properties in production and decay (production rate, mass)
- QCD tests
- EW precision observables / radiative corrections $\propto m_t^2$

New physics:

- resonances $X \rightarrow tt$
- anomalous couplings in production ($gt\bar{t}$, Wtb, FCNC) and decay ($t \rightarrow H^+ b$, FCNC decays)

Tool for calibration:

- jet energy scale
- b-tagging

CMS experiment and simulation setup

Besides the construction of the CMS detector, big efforts are in the preparation for data analysis which will be documented in the upcoming CMS Physics TDR. Complete overview of the strategy to explore the rich physics program offered by LHC:

- detector performance, operational procedures and reconstruction tools, software and data issue
- physics capabilities and analysis tools

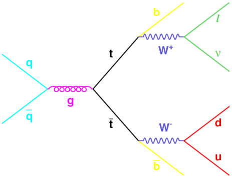
Event generation and simulation:

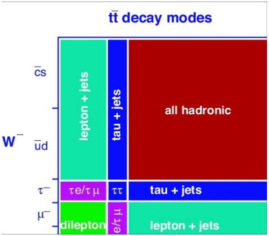
- large number of Monte Carlo generators available (general purpose PYTHIA, HERWIG, CompHEP, ALPGEN and dedicated TopRex)
- complementary QCD tools
- full detector simulation based on GEANT 4
- fast simulation FAMOS

Systematics from theory and detector:

- ISR/FSR, minimum bias and underlying event, PDF
- accurate understanding of QCD dynamics required
- jet energy scale and b-tagging efficiency (also influenced by jet structure as fragmentation model and shape)

Top production and decay





Dominant top pair production mechanisms at LHC are gluon-gluon fusion (90%) and quark-antiquark annihilation (10%).

Within the SM the top quark decays almost esclusively via:

$$t \longrightarrow Wb \quad \sim 100\%$$

Final state topology of $t\bar{t}$ events depends only from the W boson decay mode:

$$t\bar{t} \longrightarrow l\nu l\nu bb \quad 5\% \quad (e+\mu)$$
$$t\bar{t} \longrightarrow l\nu qqbb \quad 30\% \quad (e+\mu)$$
$$t\bar{t} \longrightarrow qqqqbb \quad 46\%$$

The fully hadronic final state

Nominal 6-jet topology: $t\bar{t} \to WWb\bar{b} \to qqqqb\bar{b}$

Signature based analysis:

- quarks (light jets)
- b quark jets (displaced vertices)

+ **largest branching fraction**
+ **kinematics that can be fully reconstructed**
- **very large background from QCD multijet production which makes the isolation of the signal rather challenging**
- **difficulties for triggering (multijet trigger thresholds only up to 4 jets and high E_T thresholds)**

Improvements in the signal-to-background ratio are possible by requiring the presence of b-quark jets and by selecting central and very high-energy kinematic configurations.

A specific multijet trigger that use b-tagging information has been devised for this analysis and an optimized kinematical selection has been applied.

Trigger selection

The trigger selection uses the inclusive jets trigger which consider multijets with different E_T thresholds depending on the number of jets, up to 4 jets, and a special inclusive b-jets trigger implemented according the follow criteria:

1. the b-tagging requirement is combined with an inclusive jets trigger with tuned E_T thresholds: 350, 150, 55 GeV (single jet, 3-jet, 4-jet)
2. b-tagging based on pixel and regional track and vertex recontruction on the two most energetic jets

The trigger requires multijet (n-jet) or a b-tagged jet among the two highest-E_T jets (b-jet).

Channel	QCD $50 < \hat{p}_T < 470$ GeV				
	Production	L1	HLT		
			b-jet	n-jet	n-jet + b-jet
Rate [Hz]	49k	3.3k	19.4	6.3	23.2
$\sigma\epsilon$ [pb]	25M	1.7M	9.7k	3.2k	11600

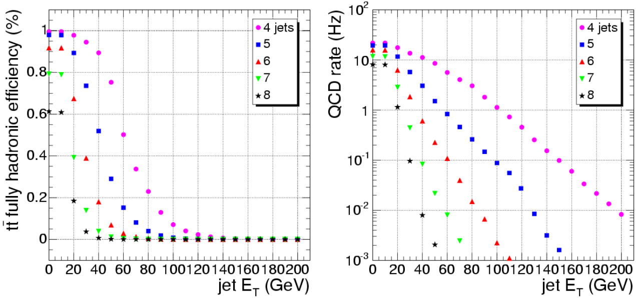
Channel	$t\bar{t} \to$ qqqq $b\bar{b}$				
	Production	L1	HLT		
			b-jet	n-jet	n-jet + b-jet
Rate [Hz]	0.76	0.43	0.11	0.03	0.13
$\sigma\epsilon$ [pb]	380	220	60	20	64
Efficiency (%)	100	57.2	14.9	4.4	16.8

The b-jet stream significantly improves the efficiency of the n-jet straem (15%).

S/B = 1/180 $\epsilon = 16.8\%$

Kinematical selection

The $t\bar{t}$ fully hadronic efficiency (factorizing out the trigger efficiency) and the QCD rate are shown as a function of jet E_T for different values of the minimum number of jets considered:



A discriminant selection is needed in order to improve the signal to background ratio. Different choices on the minumum number of jets and jet E_T are possible.

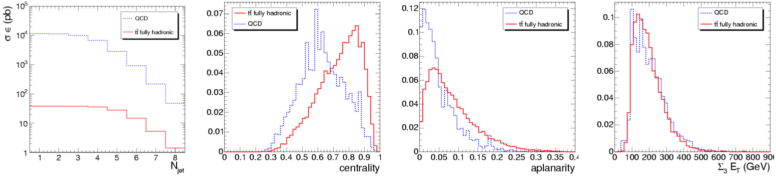
Kinematical selection

The optimal kinematical selection is based on the best statistical significance of the signal achievable for an integrated luminosity of L = 1 fb^{-1}

$$S/\sqrt{S+B} \qquad \longrightarrow \quad = (\sigma/\Delta\sigma)_{stat}$$

Kinematical selection steps:

- $6 \leq N_{jets} \leq 8$ topology with $|\eta| < 2.4$ and $E_T > 30$ GeV
- Kinematical variables:
  - centrality > 0.68
  - aplanarity > 0.024
  - $\sum_3 E_T > 148$ GeV



Selection summary

A b-tagging is then applied considering selection criteria of at least 1 b-jet and 2 b-jet.

Selection chain in cascade ($t\bar{t}$ and QCD effective cross sections, S/B and signal efficiency achieved at each step):

Selection	Requirement	$\hat{\sigma}$ $t\bar{t}$ [pb]	$\hat{\sigma}$ QCD [pb]	S/B	ϵ $t\bar{t}$ (%)
Trigger	HLT b-jets + n-jets	64	11600	1/180	16.8
Kinematical	$6 \leq N_{jet} \leq 8$	59	7900	1/130	15.5
	$E_T \geq 30$ GeV	25	930	1/37	6.6
	$centrality \geq 0.68$	16.8	324	1/19	4.4
	$aplanarity \geq 0.024$	15.3	251	1/16	4.0
	$\sum_3 E_T \geq 148$ GeV	15.2	234	1/15	4.0
b-tagging	1 b-tag	14.5	148	1/10	3.8
	2 b-tag	10	54	1/5	2.6

S/B = 1/10 (1/5) respectively for 1 (2) b-tag
ϵ = 3.8% (2.7%)

Cross section

$$\sigma = \frac{n-b}{\epsilon L}$$

$$(\Delta\sigma)_{stat} \approx \frac{(\Delta n)_{stat}}{\epsilon L} = \frac{\sqrt{n}}{\epsilon L} = \sigma \frac{\sqrt{n}}{n-b}$$

n = candidate events = $N(t\bar{t})$ + N(QCD)
b = expected background = N(QCD)
$n - b = N(t\bar{t})$
= surviving events
ϵ = total signal efficiency
L = integrated luminosity

The signal efficiency relative to the total inclusive $t\bar{t}$ events sample, to be used in the calculation of the total $t\bar{t}$ production cross section, becomes 2.3% (1.6%), respectively for the request of 1 (2) b-tag.

Requirement	L = 1 fb^{-1}				
	$t\bar{t}$ inclusive events	QCD events	ϵ (%)	$(\Delta\sigma)_{stat}$ [pb]	$(\Delta\sigma/\sigma)_{stat}$ (%)
1 b-tag	19000	148000	2.3	$\sim$ 18	$\sim$ 2
2 b-tag	13000	54000	1.6	$\sim$ 17	$\sim$ 2

Systematic uncertainties

In addition to the statistical uncertainties, significant contributions due to the systematic uncertainties on the signal efficiency (fragmentation model, PDF, ISR/FSR, jet energy scale, b-tagging), the expected background, and the integrated luminosity, are expected to be the delicate part of the cross section measurement. The fine-tuning of the optimized selection will be derived considering these uncertainties.

FAMOS

Dominating at Tevatron

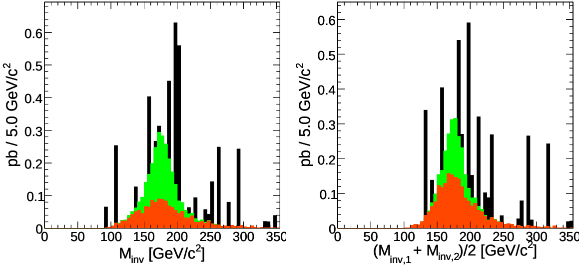
	$\Delta\sigma/\sigma$ (%)
Pile Up	10.0
Underlying Event	4.1
Fragmentation	1.9
PDF	4.2
IS/FS Radiation	7.9
Jet Energy Scale	9.3
b-tagging	5.0
Background	5.0
Integrated Luminosity	5.0
Statistical Uncertainty (1 fb^{-1})	2.0

$$\Delta\sigma/\sigma = 2\%(\mathit{stat}) + 18\%(\mathit{syst}) + 5\%(\mathit{luminosity})$$

Mass determination

The kinematical selection described including the demand for the two b-tags, forms the basis for a selection of fully hadronic $t\bar{t}$ events suitable for a kinematical top-mass reconstruction.

An additional cut of the form 100 GeV/c $< p_T <$ 300 GeV/c is applied on the two leading jets.



Invariant mass distribution of the recontrusted and rescaled top and for the average of both tops, for background and signal. Only one top chosen for kinematical mass determination to recover signal statistics from the intrinsic background.

Top mass uncertainties

Extracted top mass (input mass= 175 GeV/c^2, L = 1 fb^{-1}):

$$m_t = 175.0 \pm 0.5(\text{stat}) \pm 4.2(\text{syst})\ \text{GeV}/c^2$$

Systematic uncertainties:

	Δm_t (GeV/c^2)
Pile Up	0.4
Underlying Event	0.6
PDF	1.4
IS/FS Radiation	2.3
Fragmentation	0.9
Jet Energy Scale	2.3
b-Tagging	0.3
Background	2.0

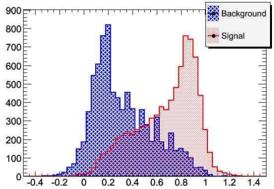
Kinematical selection based on a neural net

A more refined kinematical selection can be based on a neural net exploiting the same kinematical variables considered so far. Such approach is attempted in order to to investigate the possibility of improving the S/B ratio and/or the efficiency.
The previous kinematical selection which will be called “early".

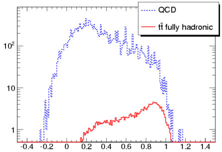
The most effective among the studied neural net configurations is the one referring to the $t\bar{t}$ and QCD samples after a cut of $E_T > 25$ GeV and consists of 6 input nodes:

- E_T 1st jet
- E_T 6th jet
- Centrality
- Aplanarity
- $\sum_3 E_T$
- Sphericity

Output of the neural net training:

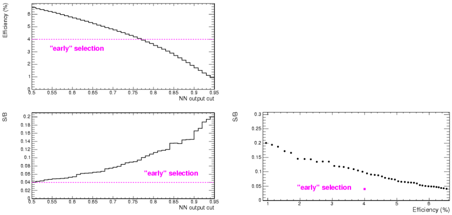


Kinematical selection based on neural net



Output distributions of the neural net applied to the whole sample of $t\bar{t}$ and QCD events.

The performance of the neural net is described by plots of efficiency and expected S/B ratio as functions of the neural net cut and of S/B ratio as a function of the efficiency.



With respect to the “early" selection, the request for a neural net output, NNcut > 0.77, improves the S/B ratio from 1/15 to 1/6 with same efficiency (4%).

Kinematical selection based on neural net

Selection	Requirement	$\hat{\sigma}$ $t\bar{t}$ [pb]	$\hat{\sigma}$ QCD [pb]	S/B	ϵ $t\bar{t}$ (%)
Trigger	HLT b-jets + n-jets	64	11600	1/180	16.8
Kinematical	$6 \leq N_{jet} \leq 8$	59	7900	1/130	15.5
	$E_T \geq 25$ GeV	33	1650	1/50	8.7
	neural net	15.2	91	1/6	4.0
b-tagging	1 b-tag	14.5	61	1/4	3.8
	2 b-tag	10.1	20	1/2	2.7

Improved S/B ratios, 1/4 (1/2) respectively for 1 b-tag (2 b-tag), can be achieved using the neural net, with same efficiencies.

Selection	L = 1 fb^{-1}				
	$t\bar{t}$ events	QCD events	ϵ (%)	$\Delta\sigma_{stat}$ [pb]	$(\Delta\sigma/\sigma)_{stat}$ (%)
1 b-tag	19000	61000	2.3	12	1.4
2 b-tag	13000	20000	1.6	11	1.3

A gain in terms of systematic uncertainty on jet energy scale, 7.6% instead of 9.3%, is also provided by the use of the neural net.

Conclusions

Top production at CMS with LHC will be an excellent place to understand the intrinsic properties of the top quark and its interactions, make precision electroweak measurements and search for new physics. Big efforts on accurate understanding of detector, backgrounds and systematics. The Physics TDR will be the occasion to show the CMS physics potential.

Particulary challenging is the $t\bar{t}$ fully hadronic decay channel which has required:

1. the definition of a specific trigger (jets + b-tagging)
2. the introduction of an optimized kinematical selection to discriminate signal from background based on cuts or on neural net
3. techniques for b-tagging for top sample isolation

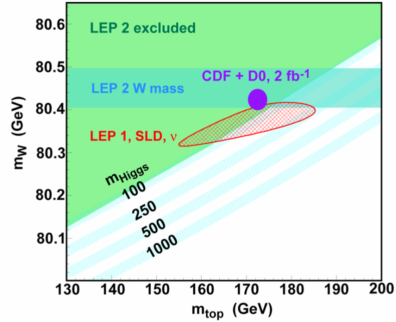
With these requirements is possible to isolate $t\bar{t} \rightarrow qqqqb\bar{b}$ sample with an efficiency of $\epsilon = 3.8\%$ (2.7%) and obtaining a signal to background ratio of $S/B = 1/10$ $(1/5)$ respectively for a request of 1 (2) b-tag which can be significantly improved by using a neural net, $S/B = 1/4$ $(1/2)$.

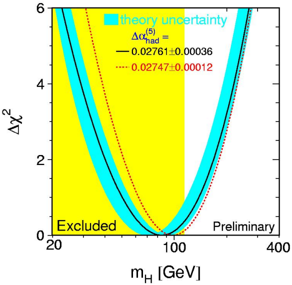
This will allow to perform precision measurements of the cross section and top mass. Measurements and interpretations will be dominated by experimental and theoretical systematic uncertainties. The fine tuning of the selection will be derived considering these uncertainties.

Backup

Role of the top quark in EW precision physics

Top quark mass enters EW precision observables as an input parameter via quantum effects (loop corrections). As distinctive feature, the large m_t value gives rise to sizeable corrections (power of m_t). By comparing EW precision measurements with theoretical predictions is possible to derive indirect constraints on all SM parameters. This information is complementary to those gained from the direct particle production.



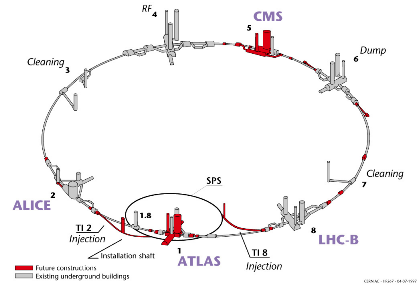


Tevatron ($\Delta m_t \approx 2$ GeV) $\rightarrow$ LHC ($\Delta m_t \approx 1$ GeV)

The Large Hadron Collider (LHC)

LHC is a machine of unprecedented collision energy and luminisity which promises to open up a new chapter in understanding the structure of matter at a deeper level exploring directly the TeV scale.

Layout of the LEP tunnel including future LHC infrastructures.



***pp* collider**

$\sqrt{s} = 14$ **TeV**

low (initial) luminosity $\mathcal{L} = 2 \times 10^{33}$ **cm**$^{-1}$ **s**$^{-1}$ $\rightarrow$ **high (design) luminosity (2010)** $\mathcal{L} = 10^{34}$ **cm**$^{-1}$ **s**$^{-1}$

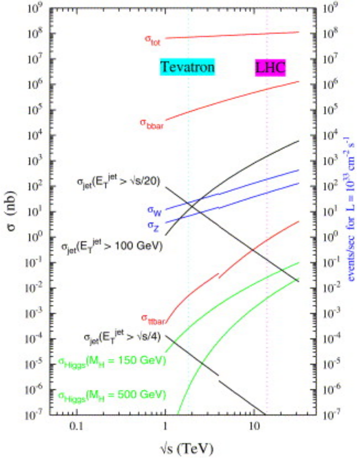
LHC Physics:

- EW symmetry breaking mechanism (Higgs?)
- physics beyond SM (SUSY?)
- SM precision physics

Detectors:

- CMS and ATLAS: general purpose detectors
- ALICE: heavy ions, QGP
- LHCb: b-physics, CP violation

Event production ad LHC

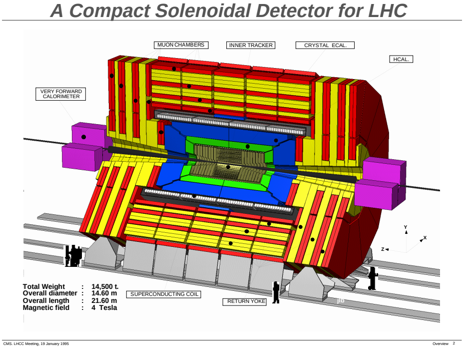


Enormous event production rate:
+ allows million events to be stored (main discoveries during first year of operations)
- impose stringent requierements on trigger and detectors permomance

Characteristics:

1. interaction rate = 10^9 Hz
2. dominant QCD jet production
3. *bunch crossing* 25 ns
4. $\sim$ 20 “minimum bias" events

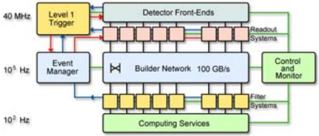
CMS (Compact Muon Solenoid)



- Tracker: charged tracks and vertex reconstruction
- Electromagnetic Calorimeter: isolated photon energy and electron identification
- Hadronic Calorimeter: jets position and energy
- Muon Detector: muon reconstruction

Trigger and Data Aquisition

A high performance trigger system is needed in order to reduce the event rate by about six order of magnitude at LHC.



HLT table:

Trigger object	Threshold (GeV)	Rate (Hz)
Isolated muon	19	25
Isolated electron	29	33
Single Jet, 3 Jet, 4 Jet	657, 247, 113	9
Jet + missing E	180,123	5
Inclusive tau jet	86	3
Electron + jet	19, 45	2
Inclusive b-jets	237	5

Kinematical variables

Kinematical variables, potentially able to separate signal from background, taken into account:

1. jet total transverse energy: $\sum E_T = \sum_{i=1}^{N_{jet}} E_T^i$
2. jet total transverse energy obtained removing the two most energetic jets : $\sum_3 E_T = \sum E_T - E_T^1 - E_T^2$
3. sfericity $= \frac{3}{2}(Q_1 + Q_2)$, Q_i being the 3 normalized eigenvalues of sfericity tensor $M_{ab} = \sum_j P_{ja} P_{jb}$
4. aplanarity $= \frac{3}{2} Q_1$
5. fraction of the hard scattering going in the transverse plane, called centrality, $= \sum E_T / \sqrt{\hat{s}}$
6. cosinus of the angle between leading jet and beam axis as computed in the jet frame $= cos\theta_1$
7. same for the next to leading jet = $= cos\theta_2$