Expectations for Top Physics at CMS IPM09, اصفهان (Isfahan)

University of California, Santa Barbara

Steven Lowette



Outline

- the top quark a peculiar particle
- from the Tevatron to the LHC
- top physics with first data
- the top quark as a calibration tool
- measuring top quark properties
- the top quark as a probe for new physics



Outline

- the top quark a peculiar particle
- from the Tevatron to the LHC
- top physics with first data
- the top quark as a calibration tool
- measuring top quark properties
- the top quark as a probe for new physics





The top quark

- SU(2) partner of the b-quark, heavy brother of the up and charm quark
 Fermion, electrical charge 2/3
- discovered at the Tevatron in 1995

completes the third generation of fermions in the Standard Model
still the only place where the top quark was studied directly

• heaviest fundamental particle known: $m_{top} = 173.1 \pm 0.6 \pm 1.1 \text{ GeV}$

→ precision of 0.7% !

pair production cross section ~ 7 pb

only a few ten-thousands tops have been produced so far

- the top can also be produced alone
 - → single top observation (5σ) came only less than 2 months ago from both DØ (arXiv:0903.0850) and CDF (arXiv:0903.0885)



Top pair production: the Tevatron versus the LHC

- production through the strong interaction
 - → Tevatron: ~ 7pb ; LHC (14TeV): ~ 830pb
- at the kinematic threshold for top pair production: $s_{hat} = x_1 x_2 s = (2m_1)^2$
 - > Tevatron: $x_1x_2 = 0.034$ -> <x> ~ 0.19-> valence quarks dominant> LHC (14TeV): $x_1x_2 = 0.0006$ -> <x> ~ 0.025-> gluon sea dominant

ttbar cross section is largest at the kinematic threshold

Tevatron: dominated by ~ 85% qq->tt
LHC (14TeV): the inverse: dominated by ~ 87% gg->tt



Tevatron measurement: gg->tt / pp->tt < 0.38 (95% CL) (CDF link)



$\begin{bmatrix} u \\ s \end{bmatrix} \begin{bmatrix} c \\ b \end{bmatrix}$

Single top production

- produced through the weak interaction
- three distinct production modes



LHC cross sections (14TeV)

| | t-channel | s-channel | tW-channel |
|----------|---------------------------|---------------------------|--------------------------|
| top | σ ^{NLO} = 153 pb | σ ^{NLO} = 6.6 pb | σ ^{NLO} = 60 pb |
| anti-top | σ ^{NLO} = 90 pb | σ^{NLO} = 4.1 pb | σ ^{NLO} = 60 pb |
| Tevatron | × 120 | × 10 | × 400 |

- processes with gluons in the initial state become dominant at LHC
 - > best possibilities in t-channel: large cross section + forward quark
 - tW-channel within reach
 - →s-channel very difficult





The top quark decay

Standard Model prediction

→ lifetime τ = 0.4 × 10⁻²⁴ s
 → decay before hadronization
 → BR(t->Wb) ~ 100% (|V₊| ~ 1)



- exploration of rare decays is a handle to new physics
- some limits set at Tevatron, often statistically limited

→ BR(t->Zq) < 3.7% (95% CL) (CDF link)

→ BR(t->H⁺b) < 35% (95% CL) (D0 link)

→ BR(t->Wb)/BR(t->Wq) > 0.79 (95% CL)

|V_{tb}| > 0.89 (95% CL) if 3x3 CKM is unitary (DO link)



Outline

- the top quark a peculiar particle
- from the Tevatron to the LHC
- top physics with first data
- the top quark as a calibration tool
- measuring top quark properties
- the top quark as a probe for new physics





At the LHC, everything is "Large"

- * "standard candles" are overwhelming
 → W's: ~200/s ; Z's: ~50/s
- top quark rate huge: reaching 1/s
 - → Tevatron: thousands

→LHC: millions!

| Process | Tevatron x-sec. | LHC x-sec. (14 TeV) | Tevatron to LHC |
|-----------------------|-----------------|---------------------|-----------------|
| top pairs | 7 pb | 830 pb | x 120 |
| single top s-channel | 0.9 pb | 11 pb | × 10 |
| single top t-channel | 2.0 pb | 250 pb | x 120 |
| single top tW-channel | 0.15 pb | 66 pb | x 400 |

- top quark physics at LHC allows precision measurements
- top quark events become standard candle themselves, providing control and calibration samples







- + $\sigma(\text{tt})$ grows more rapidly than QCD, W and Z cross sections as a function of the center-of-mass energy
- so backgrounds are expected negligible?
 - no! argument only works inclusively...
- example: W+jets
 - → σ(W) increases with a factor ~10
 → σ(W+4jets) with a factor ~100
- also LHC will need to control QCD, W+jets,...







A feeling of a 10 fb⁻¹ LHC analysis with top pairs







LHC Chamonix workshop aftermath

- + LHC will run at 10TeV c.o.m. energy
- year-long run in 2009-2010: deliver ~200/pb good collisions to experiments
- quark-initiated processes gain importance
- how does the picture change for top physics?

 $\frac{\sigma(t\bar{t})}{\sigma(t\bar{t})}$ at 14 TeV = 2.3



typical acceptances in ttbar found to be the same

what about the backgrounds? only really relevant for lepton+jets

 $\frac{\sigma(W)}{\sigma(W)}$ at 14 TeV = 1.4 but not clear how W+4 jets scales...

>probably pessimistic; Tevatron rejects more with similar selection

- bottomline: factor 2 loss in cross section for similar background rejection
- all results in the remainder of the talk are for 14TeV collisions



Outline

- the top quark a peculiar particle
- from the Tevatron to the LHC
- top physics with first data
- the top quark as a calibration tool
- measuring top quark properties
- the top quark as a probe for new physics





Rediscovery of the top quark

- much more statistics with 10/pb than used for the original top discovery
- but the detector and hence physics objects will not be well understood

reed simple, robust methods for event selection

- e.g. high thresholds to avoid not-well controlled noise
- MET and b-tagging will probably not be sufficiently understood
- a mass peak from lepton+jets is much easier to establish than a cross section deviation as a function of number of jets

>but also here no advanced kinematic methods can be employed to select which jets to combine, because JES will not be well known

no mass measurement, so mass doesn't need to be at the right value

top rediscovery with 10 pb⁻¹ in lepton+jets

CMS PAS TOP-08-005

- → hard isolated muon
- →4 hard jets
- → angular muon-jet separation
- \rightarrow top mass from 3 jets with highest ΣE_{T}
 - yields ~35% correct assignment







Intermezzo: on measuring cross sections

cross section measurements seem easy in principal

- → select events with some cuts
- subtract your expected background contribution
- count the obtained number of events
- > extract cross section folding in selection and trigger eff.'s and luminosity

event selection: not just a set of random cuts

- > needs to be simple -> avoid unnecessary uncertainties
- >but needs to efficiently reduce backgrounds and systematics

estimating backgrounds: the difficult part

- one should always aim for data-driven techniques
- defining and understanding control samples can be difficult, extrapolation to signal region must be proved to work, uncertainties must be assessed

translating back to cross section

- Iess difficult as long as you stick to default cuts -> efficiencies and uncertainties will be obtained for you
- \rightarrow otherwise there's a lot of work in this as well, for instance in the trigger



Top Physics with First Data



- the result of a cross section analysis can typically be summarized in 1 plot
 - # events as a function of # jets,
 comparing MC expectations to data
 proving signal backgrounds and
 - proving signal, backgrounds and control regions are well understood
 a lot of work in 1 number and 1 plot!

Cross section measurements in early data

- example: ttbar dilepton cross section with 10/pb
 - → ee, µµ and eµ channels each an analysis themselves
 - →Z->dilepton is the main background
 - → 0 and 1 jet bins crucial to prove understanding of backgrounds in control region
 - → expected statistical uncertainty O(10%)
 - > comparable systematics anticipated; work ongoing









Control of the QCD background

- example of one background study as part of a cross section measurement
- QCD multijet events can look like top events
 - when a real lepton within a jet looks as if it is isolated
 when weird jet or misreconstruction "fake" a lepton (esp. electrons)
- both effects are rare (but QCD cross section is huge!)

impossible to model reliably in simulations

need to estimate from data the amount of QCD passing the selection cuts

several methods being developed

- → "ABCD": use low MET and/or bad Iso to predict high MET, low Iso
- → similar: isolation vs p_{T,lepton} or d_{0,muon}
- →ad-hoc method of extrapolation of the Iso variable in the signal region
- If these methods have their drawbacks and limits, but can probably be sufficient for this small QCD contamination, allowing large uncertainties
- area of large activity! [also as part of the CMS V+jets group]





Cross section measurements with 100/pb



 differentiation of larger jet multiplicities is interesting for MC tuning and searches with multi-jets, like in SUSY



Outline

- the top quark a peculiar particle
- from the Tevatron to the LHC
- top physics with first data
- the top quark as a calibration tool
- measuring top quark properties
- the top quark as a probe for new physics





Jet energy scale calibration using the W-mass constraint

- reconstructed hadronic W mass in lepton+jets events is sensitive to jet energy scale of light quark jets
- imposing the W-mass it is possible to determine the JES corrections down to the parton level
- needed correction for true W mass scales linear with initial miscalibration



top quarks will be the most important JES calibration sample at the LHC





CMS PAS TOP-08-002

Extension to calibration of b-jet energy scale

- full kinematic fit imposes W and top mass on 3 well-paired jets
- miscalibrations are applied on all 3 jets to scan for the JES corresponding to the true W and top mass
- closure test on MC ok, no significant bias expected
- systematics
 - pile-up -> can be monitored via primary vertex count
 - effect combinatorial and non-ttbar
 background negligible
 - effect from jet energy resolution negligible
- expected performance
 - ~1% statistical uncertainty on JES for light and b-jets
- ${\mbox{ + this is inclusive, can be differentiated versus } p_{_{T}}, |\eta|}$
- detailed understanding of pile-up JES interplay crucial







B-Tagging Efficiency Measurement

CMS Note 2006/013

- b-jet-rich samples are selected in ttbar e+mu, mu+jets and e+jets events
 - Challenges in selecting a pure b-jet sample come from both process backgrounds and combinatorial background

extraction of b-tagging efficiencies from b-jet samples

 \Rightarrow apply any b-tag algorithm on a selected data sample -> fraction x_{tag} tagged

 \rightarrow estimate sample's b-tag purity \mathbf{x}_{b} and mistag probability ε_{0} from MC

from data, contributes $\epsilon_{\rm b} = \frac{1}{x_{\rm b}} [x_{\rm tag} - \epsilon_{\rm o}(1 - x_{\rm b})]$ statistical uncertainty from MC shown not to contribute additional systematics

from MC, systematics determined as function of likelihood ratio cut

systematics

- by far dominated by uncertainties on initial state gluon radiation
- → we will need LHC data or better Monte-Carlo to improve
- no systematics from misalignment -> this is in fact measured
- systematics decrease as the purity of selected samples is increased!





- + differentiation versus \textbf{p}_{τ} and η possible with 1fb^-1 of data
- three samples are exclusive and similar -> can be combined



- top events will provide the most important b-tagging efficiency calibration samples at the LHC
- new methods are in the works, especially to measure c-jet and possibly light jet efficiencies simultaneously



Outline

- the top quark a peculiar particle
- from the Tevatron to the LHC
- top physics with first data
- the top quark as a calibration tool
- measuring top quark properties
- the top quark as a probe for new physics



CMS Note 2006/066

Top mass strategies in the golden channel

- lepton + jets event selection: efficiency ~6%
 - I isolated lepton and 4 jets, of which 2 b-tagged
 extra cuts on kinematic fit, jet combination likelihood and signal/background separation likelihood

mass reconstruction performed for the hadronically decaying top

- → leptonic side used for determination of the best jet association
- estimators studied:
 - → Gaussian fit on the reconstructed top mass peak
 - → parameterized ideogram
 - an event-by-event pdf as a function of the top mass, with top mass uncertainty calculated from fit and assumed Gaussian

→ full scan ideogram

- the event-by-event top mass resolution function is now built by scanning with an additional top mass constraint in the kinematic fit
- details in the backup





Systematics for the top mass in the golden channel

• most of the following table consists of ambitious detector systematics

| | Standard Selection | | | |
|---|--------------------|-------------------|-----------------------|--|
| | Gaussian Fit | Gaussian Ideogram | Full Scan Ideogram | |
| | $-m_{v}$ | $-m_{*}$ | $-m_{v}$ | |
| | (GeV/c∸) | (GeV/c∸) | (GeV/c [∠]) | |
| Pile-Up (5% On-Off) | 0.32 | 0.23 | 0.21 | |
| Underlying Event | 0.50 | 0.35 | 0.25 | |
| Jet Energy Scale (1.5%) | 2.90 | 1.05 | 0.96 | |
| Radiation (pQCD) | 0.80 | 0.27 | 0.22 | |
| Fragmentation | 0.40 | 0.40 | 0.30 | |
| b-tagging (2%) | 0.80 | 0.20 | 0.18 | |
| Background | 0.30 | 0.25 | 0.25 | |
| Parton Density Functions | 0.12 | 0.10 | 0.08 | |
| Total Systematical uncertainty | 3.21 | 1.27 | 1.13 | |
| Statistical Uncertainty (10fb ⁻¹) | 0.32 | 0.36 | 0.21 | |
| Total Uncertainty | 3.23 | 1.32 | 1.15 | |

- main uncertainty from the b-jet energy scale
- to reach 1 GeV uncertainty we need a very good detector understanding

The more recent CMS studies indicate that this should indeed be possible





Top mass from the di-lepton decay channel

- event selection similar to cross section
- kinematics underconstrained, but indirect mass reconstruction possible
- most important systematic from jet energy scale (also neutrino modeling)
 > expected 4.2/2.9/1.0 GeV for 1/10/>10 fb⁻¹
- systematics dominated already for 1 fb⁻¹

Top mass from the fully hadronic decay

- all kinematics measured
- needs to trigger on b-jets
- jet pairing: 6 possible combinations after b-tag
- need to deal with multi-jet background
- systematics the main problem, again...
- needs very good detector understanding and advanced analysis techniques







CMS Note 2006/058

Top mass from $b \rightarrow J/\Psi \rightarrow \mu\mu$ decays

- + top mass is correlated to $m_{_{J/\Psi\ell}}$, for J/Ψ and lepton from same top
- "only" involving leptons

very good mass resolution

→ orthogonal systematics, mostly theoretical: fragmentation, $\Lambda_{_{\rm OCD}}$,...

• but low statistics this time! BR(tt->J/ Ψ {X) ~ 5.5 10⁻⁴ -> 4500 per 10fb⁻¹

> and don't forget trigger, reconstruction and pairing efficiencies

this mass determination only becomes useful at high luminosity

Top mass from b decay length

hep-ex/0501043

- top mass is correlated to decay length of the B-meson decay in the b-jet
- currently gaining importance at the Tevatron
- "only" involving tracks

→ orthogonal systematics, mostly theoretical: fragmentation, $\Lambda_{_{\rm OCD}}$,...

not explored yet for CMS! statistics no problem, best at low luminosity





CMS PAS in preparation

Constraining V_{tb} through the R = t->Wb / t->Wq ratio

test of the standard model (unitarity of the CKM matrix)

 \rightarrow sensitive to anomalous decays or a 4th fermion generation

- complement of the b-tagging efficiency measurement in ttbar events
- analysis strategy
 - standard event selection for ttbar lepton + 4 jets
 - > choose a working point for b-tagging
 - > determine the distribution of the number-of-b-tags (0 up to 4)
 - fit this distribution to the expectation, with R and mistag rate as free parameters and b-tag efficiency from an independent measurement
- background subtraction can be made data-driven
- main systematics expected from the b-tagging efficiency and the jet energy scale and resolution
- analysis under review, public results expected soon



Outline

- the top quark a peculiar particle
- from the Tevatron to the LHC
- top physics with first data
- the top quark as a calibration tool
- measuring top quark properties
- the top quark as a probe for new physics

The Top Quark as a Probe for New Physics



Top physics as integral part of a CMS discovery program

- top events live in some multi-dimensional space of event requirements
- SM backgrounds populate a separate, but not completely disjoint region
- in a simplified picture, new physics can
 - strongly overlap with the core ttbar region (BSM1)
 be only affected by the tail of ttbar (BSM2)
 be almost totally distinct (BSM3)



- first task: prove to ourselves that we understand the control regions
- then ask: is the ttbar sample consistent with the ttbar hypothesis
 opens up a broad search for ttbar physics early on



New physics in top-related differential distributions

- examples of such distributions include m_{tt} , $p_{T,top}$, H_{T}
- the Tevatron provides constraints up to relatively low scales
- at the LHC we can go to scales of several TeV
 - \rightarrow e.g. high-mass resonances in the m_{tt} spectrum, from MSSM, RS,...
 - these are not necessarily long-term studies
- at such scales we will need to deal differently with top quarks
 - The high boost will collimate the decay products of top quarks into socalled "top jets" -> can be reconstructed summing all energy in large cones
 - Ieptons lose their isolation -> needs different trigger approaches
- first studies are proceeding
 - Iast year a working group was started on highly-boosted top quarks in CMS
 - currently the use of a dedicated boosted-top jet-algorithm is under review

The Top Quark as a Probe for New Physics



Charged Higgs observability in top decays

charged Higgs H[±] bosons arise in models with
 2 Higgs doublets, like MSSM

fully described at tree level by
 2 parameters: tanβ and m_{μt}

channel of interest:

tt -> WbH[±]**b**, W -> ℓ_V

with H[±] -> τv (BR ~ 98%) and τ -> hadrons

- selection: 1 muon/electron, 1 tau jet, at least 3 jets, exactly 1 b-tag, hard MET cut
- systematics from b & tau tagging, ttbar knowledge,...

no data-driven methods yet though
analysis currently systematics dominated

+ the LHC will allow exploration up to the top mass for high tan $\!\beta$

 \rightarrow smooth transition to $m_{H_{+}} > m_{T_{+}}$





Search for an excess of same-charge top pairs CMS Note 2006/065

- corresponds to a search for a same-charge signal with ttbar kinematics
- fully leptonic final states considered: mu+mu, e+mu and e+e
 simple selections: two leptons and two b-jets
- wrongly charged leptons can be fakes or charge mis-id'd leptons

> mu+mu is the most powerful channel

- analysis looks for a deviation of
 - the R = $N_{++,--} / N_{+-}$ ratio
 - → 1fb⁻¹: R₁₁₁ = 0.0027 ± 0.0007 (stat)

most systematics cancel in ratio

- systematics from tau's and Z+jets uncertainties negligible, and none expected from knowledge of charge mis-id
- same-sign signal reach depends directly on integrated luminosity
- much better results can be obtained working in specific models







New physics in rare top decays

CMS Note 2006/093

• beyond the SM, rare top decays through loops can become sizable

| Decay | SM | two-Higgs | SUSY with R | Exotic Quarks | Exper. Limits(95% CL) | |
|--------------------------|---------------------|----------------|----------------|------------------------|-----------------------|-------------|
| t ightarrow gq | 5×10^{-11} | $\sim 10^{-5}$ | $\sim 10^{-3}$ | $\sim 5 	imes 10^{-4}$ | < 0.29 (CDF+TH) | |
| $t \rightarrow \gamma q$ | $5 	imes 10^{-13}$ | $\sim 10^{-7}$ | $\sim 10^{-5}$ | $\sim 10^{-5}$ | < 0.0059 (HERA) | |
| $t \rightarrow Zq$ | $\sim 10^{-13}$ | $\sim 10^{-6}$ | $\sim 10^{-4}$ | $\sim 10^{-2}$ | < 0.14 (LEP-2) | <0.037 (CDF |

• selections devised to isolate ttbar events with one top in a rare decay

→ need to select as little ttbar, single top,... as possible



• statistically limited in the beginning, but great potential in the long run





• top quarks provide a very rich physics program at the LHC

- → SM measurements in the top sector
- → calibrations
- >beyond the SM searches
- early day top physics will be exciting
 - many analyses become accessible in the first LHC year, some immediately
 - \rightarrow ideal playground to acquire understanding of the complete CMS detector

top quark physics should be seen as an essential part of an global CMS discovery program

> top physics is omnipresent in CMS, as a signal, as a background or as a playground for tools and calibrations









Backup















Backup











A few words on kinematic fitting

- using kinematic fit techniques, constraints (like mass, momentum balance,...) can be enforced on the reconstructed event topology
- here a linearized least-square method with Lagrange multipliers was used
 part of the CMSSW PhysicsTools packages
- example: top mass when W mass enforced in the t->Wb->qq'b decay



+ to obtain the same precision without the fit, 5 times more data is needed





More details on ideograms

- For each event an ideogram $I(\{p_i\}|m_{top})$ can be defined as a probability function to measure the observed parameters $\{p_i\}$ given a certain mass m_{top}
- method 1: this ideogram can be assumed Gaussian, with $\sigma_{m_{fit}}$ calculated from the fitted jet covariance matrices: $\Rightarrow I({p_i}|m_{ton}) \sim exp[-0.5 (m_{fit}-m_{ton})^2/\sigma_{m_{fit}}^2]$
- method 2: the ideogram can be scanned imposing different m_{top} hypotheses as an extra constraint in the fit:

→ I({ p_i }| m_{top}) ~ { $P_{m_{top}}(\chi_{fit}^2)$ } with m_{top} = 125, 135, ..., 225 GeV



inclusion of b-jet kinematics in the kinematic fit results in a better ideogram resolution



Backup



- to go from the experimental space m_{top} to the theoretical space M_{top}, the ideogram of an event is convoluted with a theoretically expected template function T(m_{top} | M_{top}):
 → L_i(M_{top}) = ∫ I({p_i}|m_{top}). T(m_{top} | M_{top}) dm_{top}
- where

$$\Rightarrow T(m_{top} | M_{top}) = P^{signal} \cdot [P^{comb} \cdot BW(m_{top} | M_{top}) + (1 - P^{comb}) \cdot B^{comb}(m_{top})]$$

$$+ (1 - P^{signal}) \cdot B^{channel}(m_{top})$$

- finally all the event-by-event likelihoods are combined
 → L_{tot}(M_{top}) = Π_i L_i(M_{top})
- and the minimum of

$$\chi^2$$
 = -2 ln L_{tot}

gives the best estimate for the top mass







Top mass results in the golden channel

• comparison estimators: Gaussian fit, Gaussian ideogram, scanned ideogram



| | Gaussian Fit | Gaussian Ideogram | Full Scan Ideogram |
|--|------------------|-------------------|--------------------|
| Bias (GeV/c ²) | -0.84 ± 0.59 | -4.35 ± 0.54 | -2.58 ± 0.31 |
| Pull | 0.82 | 1.01 | 1.01 |
| Expected uncertainty for 1fb ⁻¹ (GeV/c ²) | 1.01 | 1.14 | 0.66 |
| Expected uncertainty for 10fb^{-1} (GeV/c ²) | 0.32 | 0.36 | 0.21 |







CMS Note 2006/084

Single top t-channel cross section measurement

- best results at the LHC are expected in the process where a W is exchanged in the t-channel
 - > distinct signature of a forward light jet
 - Presence of a lepton, MET and b-jet from single top decay
 - no ambiguity in reconstructing the top
 - >backgrounds from ttbar, W+jets and QCD multi-jet

selection involves

- standard lepton, MET and b-jet cuts
- suppression of backgrounds with a forward jet requirement
- wide W transverse mass and top mass window cut
- suppression of the ttbar background in particular with a total transverse momentum requirement







expected number of events after 10fb⁻¹

| ⇒ signal | ttbar | Wbbj | Wjj | |
|----------|-------|------|-----|-----|
| → | 2389 | 1188 | 195 | 402 |

- QCD estimated <8% of total background
- uncertainties for 10fb⁻¹
 - → statistical uncertainty: 3%
 - > estimation main systematics (theory, b-tagging, JES): 8%
 - → luminosity: 5%
- no background estimations from data yet, starts from theory calculations

Other single top channels

- tW channel: accessible at 10fb⁻¹
 - >both dilepton and semi-leptonic decays possible; di-lepton cleanest
 - ttbar is the major background by far, several times larger than the signal
- * s channel: very challenging

>very large backgrounds from ttbar and single top itself

CMS Note 2006/086