UNIVERSITY OF LJUBLJANA
FACULTY OF MATHEMATICS AND PHYSICS
DEPARTMENT OF PHYSICS

Liza Mijović

TOP QUARK MASS MEASUREMENT WITH THE ATLAS DETECTOR

Doctoral Thesis Defence

ADVISER: Prof. Dr. Borut Paul Kerševan

Ljubljana, 2010

Overview and the Motivation of the Thesis

Overview

- **Thesis topic is the top quark mass measurement.**
  - Top (t) quark is the heaviest known Standard Model particle.
  - Top quark can be produced in proton-proton ($p - p$) collisions at the **L**arge **H**adron **C**ollider (**LHC**). The t mass can be measured with the **ATLAS** (**A** **T**oroidal **L**HC **A**pparatu**S**) detector.
- **Thesis focus is the systematic uncertainty of the top quark mass peak reconstruction.**

Motivation

- **Due to the large top quark production rates at the LHC the top quark measurement precision will (eventually) be limited by the systematic uncertainty. Systematic uncertainty is estimated from the computer simulations of the t production, decay and detection (Monte Carlo simulations).**
- Top quark production is appropriate for the MC generators comparisons and tuning.
- Due to the large mass of the t the top physics can be used to tune the MC generators at the high momentum scales. This is of relevance for the new particle searches.
- The top mass (physics): knowledge of the the modelling uncertainty is of importance for both the early reconstruction and the precision measurements.
- A sizable sample of t quarks is yet to be collected at the LHC, so the MC generators are used to simulate the t production, decay, detection.

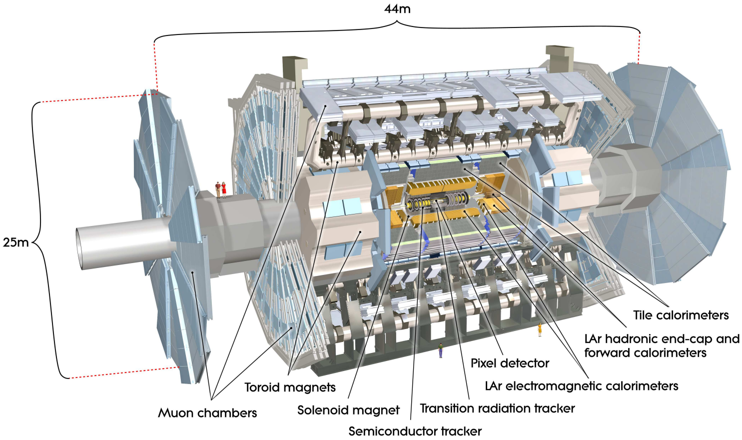
The LHC:

- the **L**arge **H**adron **C**ollider, CERN, Geneva, Switzerland.
- $p-p$ collisions at high center of mass energies (E_{CM}),
- design E_{CM}=14 TeV, current E_{CM}=7 TeV,
  for comparison: Tevatron, Fermilab, Chicago: E_{CM}=1.96 TeV.
- **CERN** stands for the **C**onseil **E**uropéen pour la **R**echerche **N**ucléaire.
- LHC tunnel: circumference: 27 kilometres, down to 175 metres underground.
- Protons are travelling with (99.999999% of) the speed of light ($c = 3 \cdot 10^8$ m/s).

- $p-p$ collision data collected by four experiments: ALICE, **ATLAS**, CMS, LHCb.
  - ATLAS = A Toroidal LHC Apparatus,
  - CMS = Compact Muon Solenoid,
  - LHCb = LHC-beauty,
  - ALICE = A Large Ion Collider Experiment.

The ATLAS Detector

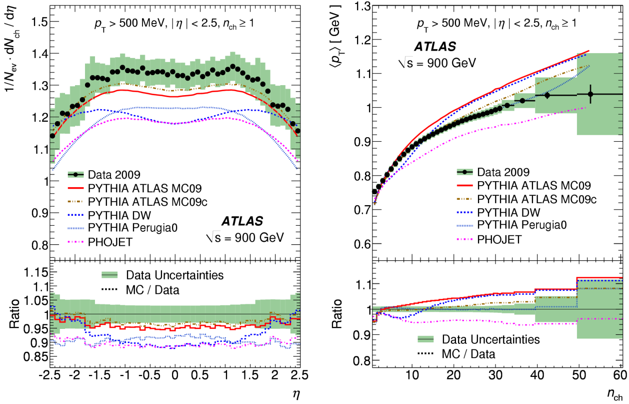
- The dimensions of the detector are 25 m in height and 44 m in length. The overall weight is 7000 tonnes [1].
- The z axis corresponds to the beam axis, $x - y$ plane is the plane transverse to the beam direction.
- θ=angle from the beam axis, pseudorapidity η: $\frac{p_z}{|\vec{p}|} = \cos\theta = \tanh(\eta)$.



[1] G. Aad et al. [ATLAS], JINST **3**, 2008.

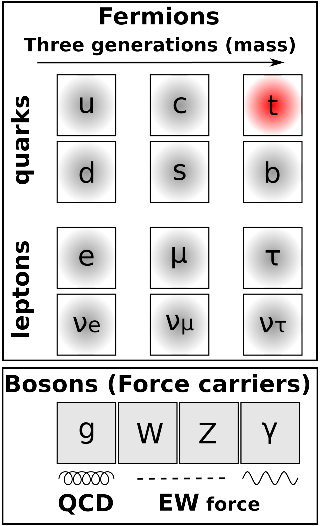
Data-taking with the ATLAS detector

- Current E_{CM} of collisions is 7 TeV.
- Figures: charged-particle distributions in $p - p$ interactions at $E_{CM} = 900$ GeV, data collected with the ATLAS detector in December 2009 [2].
- top quark observation in 7 TeV collisions:
  - has not yet been reported,
  - is anticipated in 2010.



[2] G. Add, L. Mijović et al. [ATLAS], Phys. Lett. **B688**, 2010.

The Standard Model and the Top (t) Quark



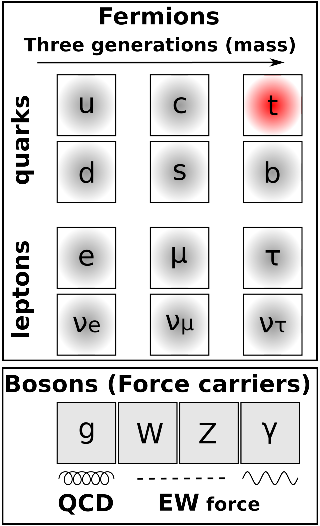
Facts about the t quark:

- Discovered: 1995:
  Tevatron, Fermilab, Chicago [a], [b].
- $\mathrm{m}_t = 171.2\ \mathrm{GeV} \pm 1.2\ \mathrm{GeV}$ (stat.) $\pm 1.8\ \mathrm{GeV}$ (syst.) [c]
  (units: c=1 $\Rightarrow$ [mass]=[Energy]).
- N.b.: $\delta \mathrm{m}_t / \mathrm{m}_t \sim 1\%$,
  $\sim$ precision needed @ the LHC.
- LHC vs Tevatron:
  Tevatron: $p - \bar{p}$, $E_{CM} \sim 2\ \mathrm{TeV}$.
  LHC: $p - p$, $E_{CM} \sim 7 - 14\ \mathrm{TeV}$.
  LHC: more statistics, but also
  LHC: harder to control syst.

---

[a] F. Abe et al. [CDF], Phys. Rev. D **74**, 1995.
[b] S. Abachi et al. [D0], Phys. Rev. Lett. **74**, 1995.
[c] C. Amsler et al. [PDG], Phys. Lett. **B667**, 2010.

The Standard Model and the Top (t) Quark



Experimental signature:

QCD confinement, hadronization $\Rightarrow$
jets=sprays of hadrons.
Hadron=colourless (q, $\overline{q}$) bound state.
Exception:
t decays before the hadronization.

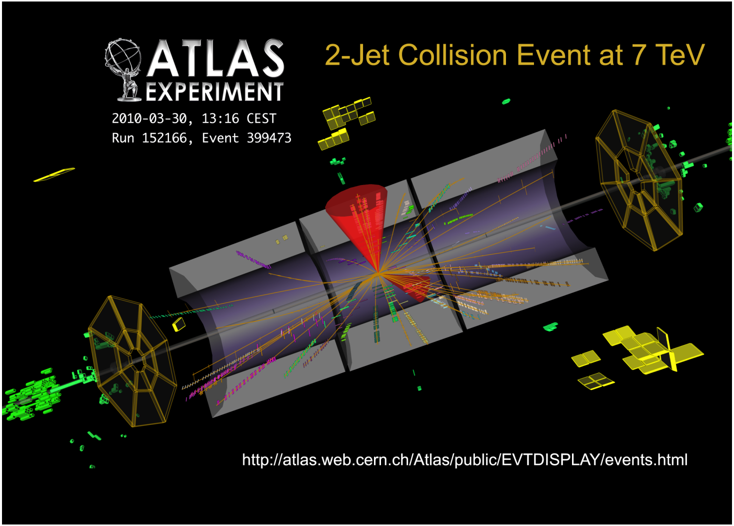
---

e, μ: easy to detect.
ν: escapes direct detection,
but contributes to $\not{E}_{\mathrm{T}}$.

W, Z: decay to quarks and leptons,
g: form jets.

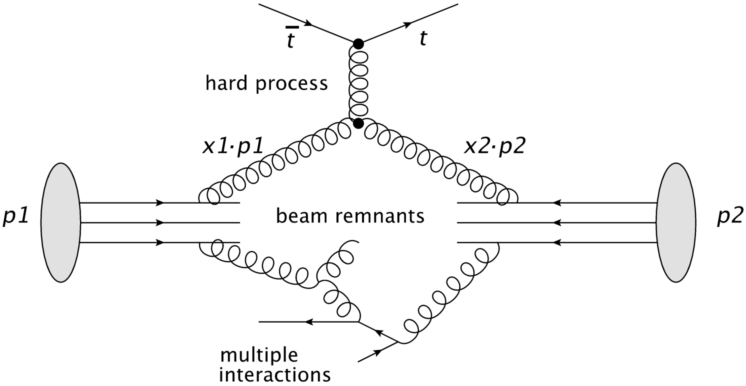
The Standard Model and the Top (t) Quark

QCD jets are an experimental signature of the quarks and gluons.



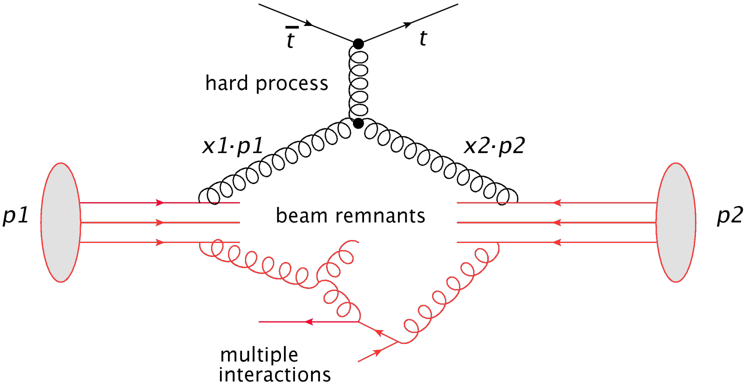
Top Quark Production in $p - p$ collisions

- Use $t\bar{t}$ production for the top quark mass measurement.
- Hard process: perturbative QCD calculation.
- Parton Distribution Functions, beam remnants, multiple interactions: effective model.



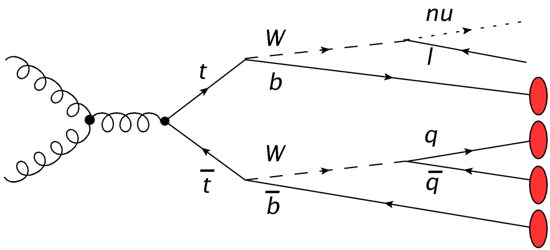
Top Quark Production in $p - p$ collisions

- Use $t\bar{t}$ production for the top quark mass measurement.
- Hard process: perturbative QCD calculation.
- Parton Distribution Functions, beam remnants, multiple interactions: effective model.



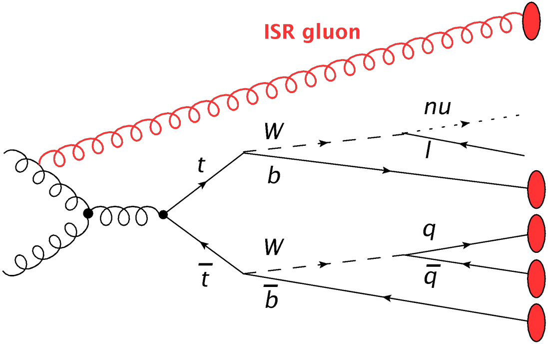
Top Quark Decay Observable Final State

- Hard process: perturbative QCD calculation.
- Final state quarks, gluons → jets: effective model.
- Initial State Radiation (ISR): effective model [adds ∼ 1.5 jets to the $t\bar{t}$ event].
- Final State Radiation (FSR): effective model.
- The Parton Shower (PS) picture is used for the QCD radiation modelling.
- The effective model contains free parameters to be tuned to the data.



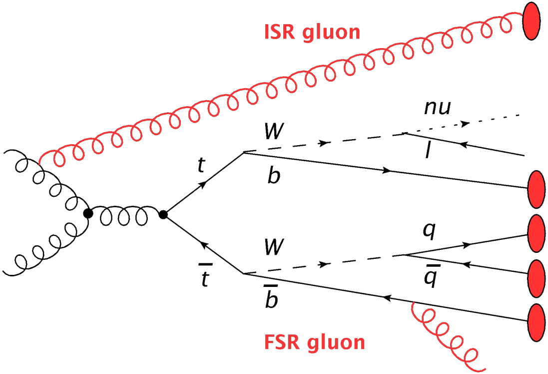
Top Quark Decay Observable Final State

- Hard process: perturbative QCD calculation.
- Final state quarks, gluons $\rightarrow$ jets: effective model.
- Initial State Radiation (ISR): effective model [adds $\sim$ 1.5 jets to the $t\bar{t}$ event].
- Final State Radiation (FSR): effective model.
- The Parton Shower (PS) picture is used for the QCD radiation modelling.
- The effective model contains free parameters to be tuned to the data.



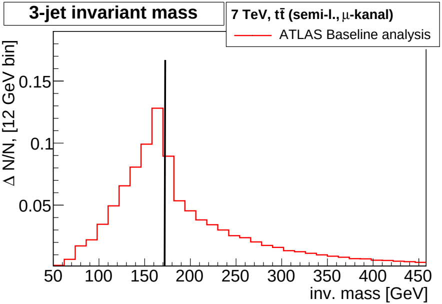
Top Quark Decay Observable Final State

- Hard process: perturbative QCD calculation.
- Final state quarks, gluons $\rightarrow$ jets: effective model.
- Initial State Radiation (ISR): effective model [adds $\sim$ 1.5 jets to the $t\bar{t}$ event].
- Final State Radiation (FSR): effective model.
- The Parton Shower (PS) picture is used for the QCD radiation modelling.
- The effective model contains free parameters to be tuned to the data.



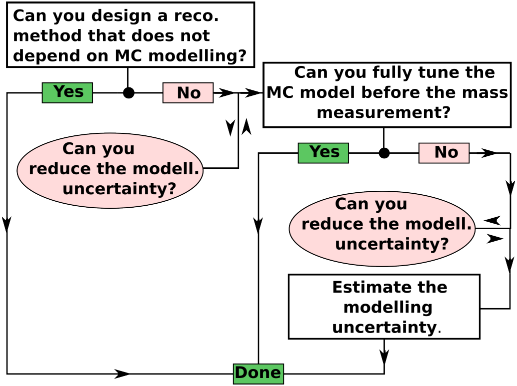
The t Quark Mass Peak Distribution

- The top quark properties: $\mathrm{m}_t = 172.5\ \mathrm{GeV}$, $\Gamma = 1.4\ \mathrm{GeV}$.
- This are the computer simulation inputs, used for the simulation of the **hard process**.
- For the top quark mass **measurement** the mass is extracted from the **experimentally observable quantities**.
- The aspects that need to be **modelled** (PDFs, Multiple Interactions, Fragmentation, Parton Shower . . .) **modify** the reconstructed t quark mass distribution.
- The imperfect reconstruction procedure also results in the broadening and the shape modulation of the reconstructed distribution.



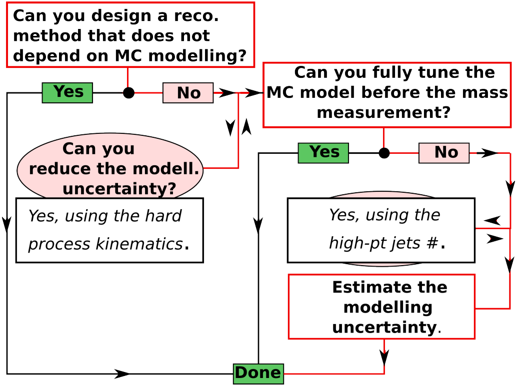
Monte Carlo Systematic Uncertainty of the m_t

- The effective models used in the Monte Carlo generators contain free parameters.
- The parameters are to be set by the comparisons to the experimental data (generator tuning).
- The hard event properties (top quark mass) need to be extracted from the observable quantities.
- The use of (many) effective models is needed for the simulation of the observable quantities.
- Q: What is the systematic uncertainty of the extracted m_t due to the use of effective models in the Monte Carlo generators?



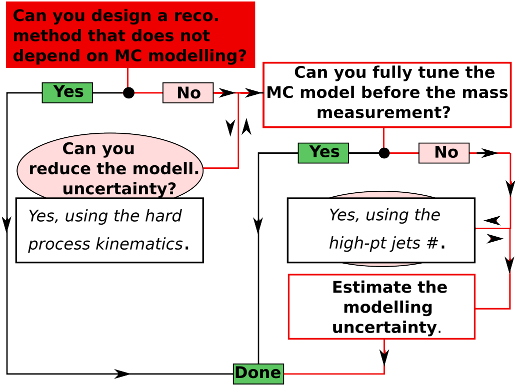
Monte Carlo Systematic Uncertainty of the m_t

- The effective models used in the Monte Carlo generators contain free parameters.
- The parameters are to be set by the comparisons to the experimental data (generator tuning).
- The hard event properties (top quark mass) need to be extracted from the observable quantities.
- The use of (many) effective models is needed for the simulation of the observable quantities.
- Q: What is the systematic uncertainty of the extracted m_t due to the use of effective models in the Monte Carlo generators?



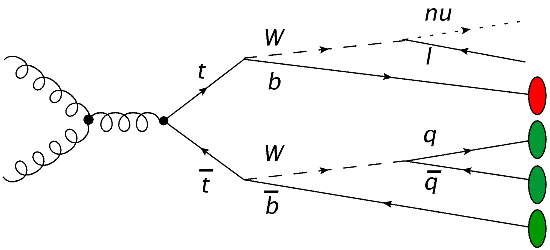
Monte Carlo Systematic Uncertainty of the m_t

- The effective models used in the Monte Carlo generators contain free parameters.
- The parameters are to be set by the comparisons to the experimental data (generator tuning).
- The hard event properties (top quark mass) need to be extracted from the observable quantities.
- The use of (many) effective models is needed for the simulation of the observable quantities.
- Q: What is the systematic uncertainty of the extracted m_t due to the use of effective models in the Monte Carlo generators?



Top Quark Mass Reconstruction

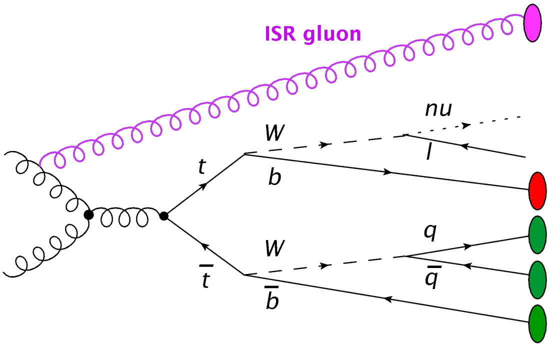
- Using the semi-leptonic decay channel.
- Initial State Radiation (ISR): extra jets ⇒ larger comb. background.
- Final State Radiation (FSR): jet broadening, energy loss, kinematics modulation.



jets,
t ($\bar{t}$) mass
reconstruction

Top Quark Mass Reconstruction

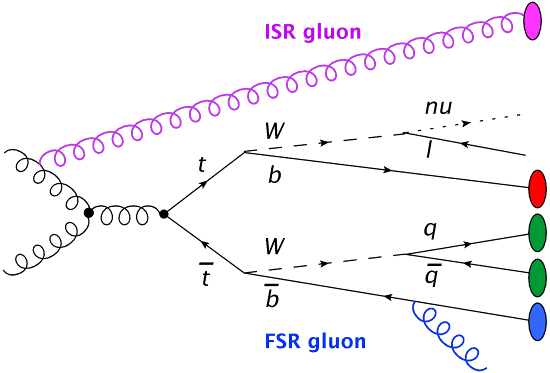
- Using the semi-leptonic decay channel.
- Initial State Radiation (ISR): extra jets ⇒ larger comb. background.
- Final State Radiation (FSR): jet broadening, energy loss, kinematics modulation.



jets,
t ($\bar{t}$) mass
reconstruction

Top Quark Mass Reconstruction

- Using the semi-leptonic decay channel.
- Initial State Radiation (ISR): extra jets ⇒ larger comb. background.
- Final State Radiation (FSR): jet broadening, energy loss, kinematics modulation.



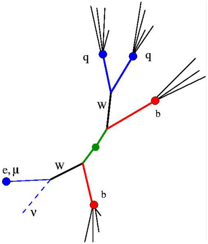
jets,
t ($\bar{t}$) mass
reconstruction

The ATLAS Baseline Analysis

A (published, [3]) strategy for the relatively early top physics analysis:

Pseudorapidity η is defined as:

$$\frac{p_z}{|\vec{p}|} = \cos\theta = \tanh(\eta) \ .$$



- Semi-leptonic $t\bar{t}$ event candidates selection ($t\bar{t} \to WWb\bar{b} \to (l\nu)(q\bar{q})b\bar{b} \to (l\nu)(\mathrm{jj})\mathrm{j}_{\,\mathrm{b}}\mathrm{j}_{\,\bar{\mathrm{b}}}$):
  - exactly one lepton with $\mathrm{p_T} > 20$ GeV and $|\eta| < 2.5$,
  - missing transverse energy of at least 20 GeV,
  - at least four jets with $\mathrm{p_T} > 20$ GeV and $|\eta| < 2.5$,
  - at least three jets with $\mathrm{p_T} > 40$ GeV and $|\eta| < 2.5$.
- selection of the three jets:

$$\mathrm{p_T}(\mathrm{j1} + \mathrm{j2} + \mathrm{j3}) = \mathrm{max.} \ .$$

- The analysis strategy evolves with E_{CM}, $\mathcal{L}$, studies and detector performance.
- Results shown for E_{CM}=10 TeV, E_{CM}=7 TeV and m_t=172.5 GeV.

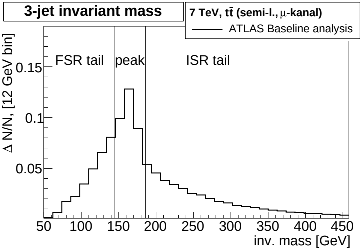
[3] G. Aad et al. [ATLAS], ATL-PHYS-PUB-2009-087, CERN, Geneva, 2009.

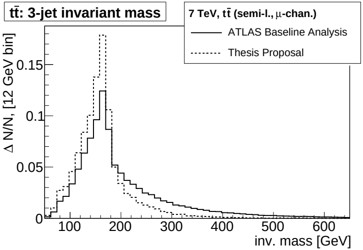
The t Mass Peak Reconstruction

- Baseline Analysis method: reconstruct hadronically decaying t from 3 jets:

$$\mathrm{p_T}(\mathrm{j1} + \mathrm{j2} + \mathrm{j3}) = \mathrm{max.}\ .$$

- $\Rightarrow$ FSR: low mass tail, ISR: high mass tail.
- Thesis: developed separation of QCD radiation and hard event jets.
- $\Delta\eta(\mathrm{j1} + \mathrm{j2} + \mathrm{j3}) = \min\ \Rightarrow$ less ISR combinatorial background, better mass reco.
- Figures: events pass Baseline analysis cuts, normalization: unit area.

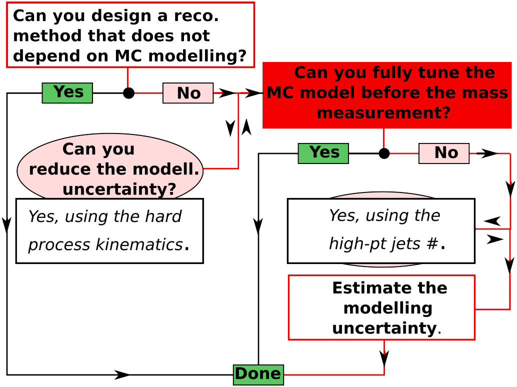




- Residual ISR/FSR dependence.
- Additional event rejection needed, optimal strategy depends on selection cuts, reco method, need uncertainty evaluation for a general case.

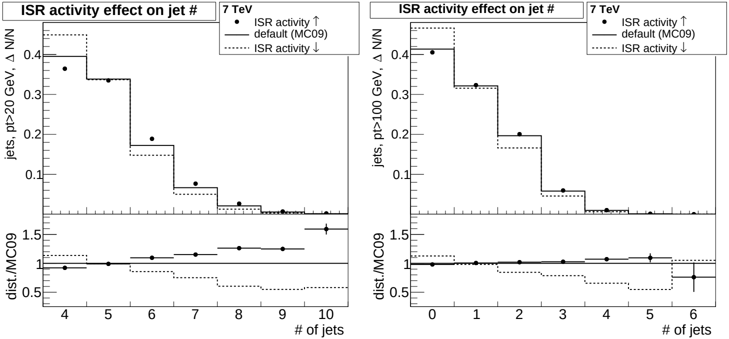
Monte Carlo Systematic Uncertainty of the m_t

- The effective models used in the Monte Carlo generators contain free parameters.
- The parameters are to be set by the comparisons to the experimental data (generator tuning).
- The hard event properties (top quark mass) need to be extracted from the observable quantities.
- The use of (many) effective models is needed for the simulation of the observable quantities.
- Q: What is the systematic uncertainty of the extracted m_t due to the use of effective models in the Monte Carlo generators?



ISR ↑,↓ : jet

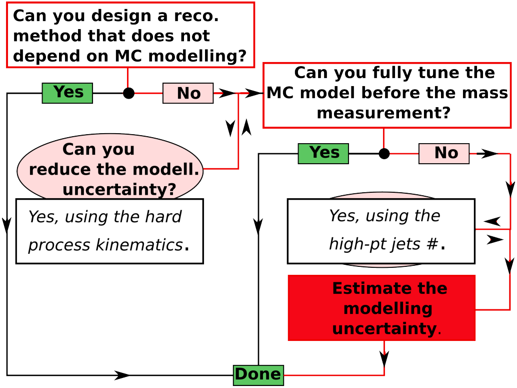
- ISR: much more activity @ LHC (wrt. Tevatron), additional resolved jets.
- Figures: events pass Baseline analysis cuts, normalization: unit area.



- FSR: activity @ LHC similar to Tevatron, small effect on jet #, use jet energy scale calibration.
- Jet p_T, η and some other distributions can be used for model comparisons.
- Likely scenario: mass peak reconstruction before the generators tuning.

Monte Carlo Systematic Uncertainty of the m_t

- The effective models used in the Monte Carlo generators contain free parameters.
- The parameters are to be set by the comparisons to the experimental data (generator tuning).
- The hard event properties (top quark mass) need to be extracted from the observable quantities.
- The use of (many) effective models is needed for the simulation of the observable quantities.
- Q: What is the systematic uncertainty of the extracted m_t due to the use of effective models in the Monte Carlo generators?



Estimating the Modelling Uncertainty

- The existing experimental data has already been used for the model development and default parameter values.
- Due to high energy scale of the $t\bar{t}$ process the existing experimental data (often) does not additionally constraint parameter values.

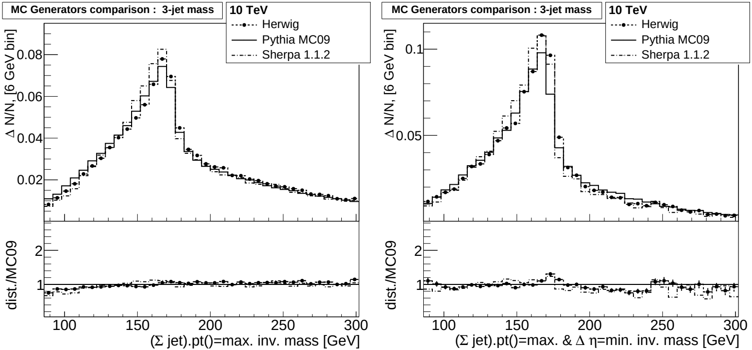
How to estimate the systematic uncertainty due to the use of effective models in MC generators:

- identify the relevant modelling aspects. Top mass ($t\bar{t}$) the ISR, FSR effects are dominant.
  Other sources of modelling uncertainty have also been treated in the Thesis.
- For each relevant modelling aspect:
  1. compare predictions of observable value obtained by different Monte Carlo generators models (implementing different assumptions),
  2. for the different models vary model parameters within their valid ranges. Determine the effects of the variations on the observable quantity of interest.
  3. The larger of the differences in points 1) and 2) correspond to Monte Carlo modelling uncertainty.

Showing results: reconstructed top quark peak position and width.
Also treated: jet # (tuning) and sel. efficiency effects (event selection bias).

1) Model comparison

- AcerMC (Matrix Element) + Pythia (ISR,FSR, hadronization), MC09 tune,
- MC@NLO (Matrix Element) + Herwig (ISR,FSR, hadronization),
- Sherpa (1.1.2).

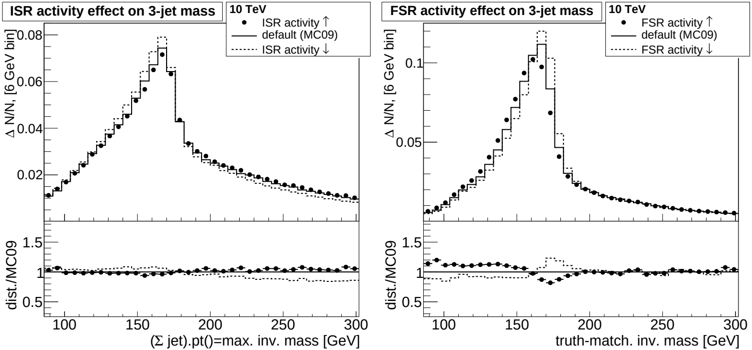


The anticipated measurement precision is $\delta m_t \sim 1\ \mathrm{GeV}$.
Also worked with:

- AcerMC (Matrix Element) + Pythia, virtuality-ordered parton shower,
- Sherpa 1.2.0 + 0,1 and 2 additional FS partons from the ME.

2) Parameter variations

- AcerMC (Matrix Element) + Pythia (ISR,FSR, hadronization), MC09 tune.
- Collective variation of ISR parameters: ISR activity ↑,↓.
- Collective variation of FSR parameters: FSR activity ↑,↓.



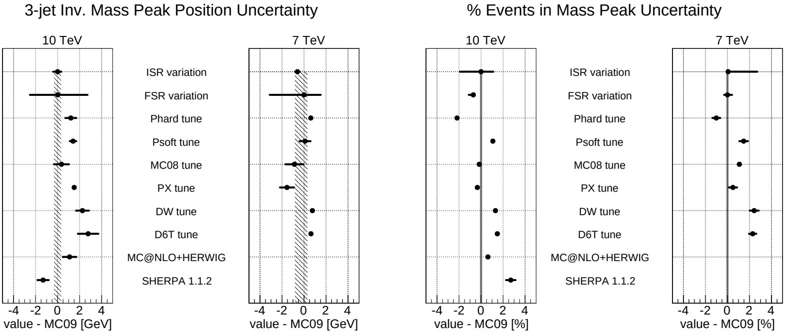
The anticipated mass measurement precision is $\delta m_t \sim 1\ \mathrm{GeV}$.

- x-checked with systematics samples developed by the Pythia author [4].
- ISR: can be tuned from jet # (studies done for individual generator parameters).
- FSR effects can be reduced using the jet energy scale calibration (Tevatron experience).

[4] P. Z. Skands, arXiv:hep-ph/0905.3418 .

3) The Uncertainty

MC09 reconstructed value: $m_t \sim 168 \ \mathrm{GeV}$. MC09 reconstructed % events ~ 40 %.



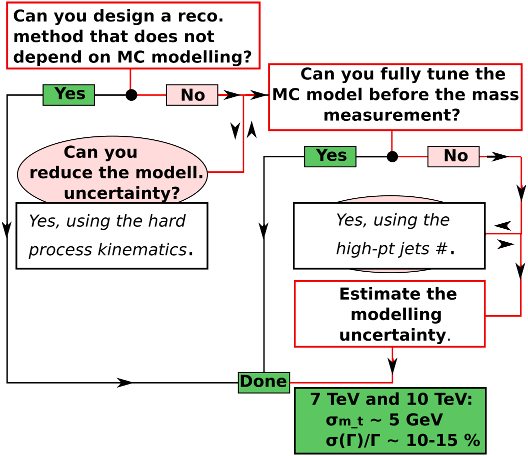
$$\sigma_{10 \ \mathrm{TeV}} \sim \sigma_{7 \ \mathrm{TeV}} \sim 5 \ \mathrm{GeV} \ .$$

$$\frac{\Delta\sigma_{10 \ \mathrm{TeV}}}{\sigma_{10 \ \mathrm{TeV}}} \sim \frac{\Delta\sigma_{7 \ \mathrm{TeV}}}{\sigma_{7 \ \mathrm{TeV}}} \sim 10 - 15 \ \% \ .$$

The central value is further modified by the detector effects.
Uncertainty estimation is expected to be relevant also at the sim.-level.

Modelling Uncertainty in $t\bar{t}$ Process: Results

- Q: What is the systematic uncertainty of the extracted m_t due to the use of effective models in the Monte Carlo generators?



Summary of the Results

- The systematic modelling uncertainty of the top quark mass reconstruction in the $t\bar{t}$ events has been evaluated.
- The dominant source of uncertainty are the ISR and the FSR modelling.
- LHC: more ISR than @ Tevatron, $\sim$ same level of FSR activity as @ Tevatron.
- The impact of the ISR jets contamination can be reduced by exploiting the small $\Delta\eta$ of the jets from the $t\bar{t}$ hard event.
- The high-pt jets $\#$ can be used to tune the generators ISR parameters and compare different MC models.
- **For the Monte Carlo generator Pythia the ISR and FSR parameter variations are studied in detail. The relevant model parameters are identified, the variation of their effect is evaluated. The tuning strategy is developed for the generator's initial state radiation parameters for the $t\bar{t}$ events.**
- The Parton Shower activity causes observable modulation of the reconstructed top quark mass distribution.
- **The peak position shifts of $\sim 5 \text{ GeV}$ can be caused by the different PS modelling. The mass measurement precision at the Tevatron is $\sim 1 \text{ GeV}$.**
- The uncertainty values have been evaluated for the ATLAS Baseline Analysis.
- **Defining systematic samples enables evaluation for other reconstruction methods and observables.** This uncertainty evaluation method has been used @ ATLAS so far.

Outlook

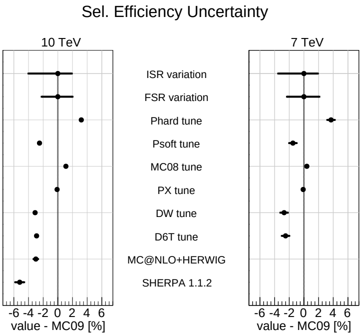
- The presented systematic modelling uncertainty figures have been obtained from the MC models. When the inputs from the experimental data are available the systematic modelling uncertainty is expected to be decreased.
- The comparison of the LHC $t\bar{t}$ process data to the MC distributions will enable evaluation of the Parton Shower models performance and the Parton Shower parameters tuning.
- Due to the high mass of the top quark the top quark mass measurement can be used to evaluate the performance of the MC generators for the high mass resonance measurement @ the LHC. For the processes that involve high energy scales the ISR(FSR) emissions result in the additional resolved jets in the final state. This is of relevance for the new particle searches at ATLAS. The $t\bar{t}$ process is also an important background for the new particle searches.
- Therefore I estimate the Parton Shower tuning in $t\bar{t}$ events should be of interest for the top group, Monte Carlo group and the ATLAS collaboration.

Thank You

Additional Slides

The Selection Efficiency Uncertainty Evaluation

- Events pass the Baseline analysis cuts.
- MC09 reconstructed value: $\sim 40\%$.

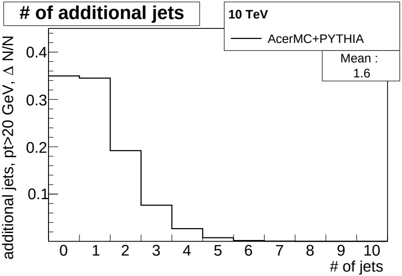


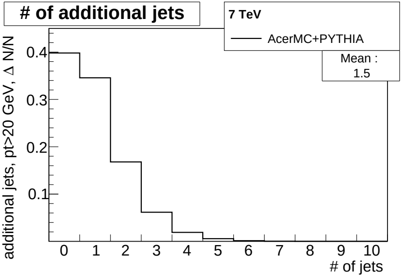
$$\sigma^2_{\mathrm{ISR+FSR}} = \sigma^2_{\mathrm{ISR}} + \sigma^2_{\mathrm{FSR}} \Rightarrow \sigma_{7\ TeV} = 7.2\%$$

$$\frac{\Delta\sigma_{7\ TeV}}{\sigma_{7\ TeV}} \sim 20\ \%$$

ISR/FSR Systematics : 10 TeV vs 7 TeV

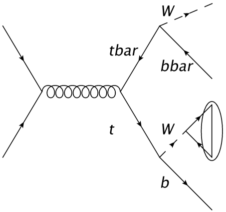
- 10 TeV vs 7 TeV: the expected trends observed, but **not a big change**;
  - $t\bar{t}$ still gg dominated at 7 TeV;
    % of gg extracted from official MC@NLO inputs : 83% (10TeV) vs 78.5% (7TeV).
  - Number of resolved ISR emissions comparable @ 10 TeV and 7 TeV.
  - Figures: jet # in single-lepton $t\bar{t}$ events, pt>20 GeV, $|\eta|$ <2.5,
  - events are requested to pass the baseline analysis cuts.



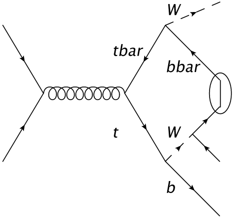


Nonperturbative Systematics: Colour Reconnections - 1

- CR non-perturbative modell. aspect implemented in PYTHIA,
- improve the description of the data for minbias/UE.



No Colour Reconnections
qqbar from
colour singlet W
hadronize separately
from the rest of
the event

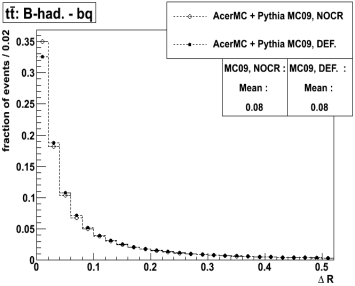


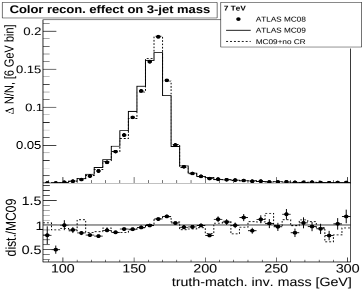
Colour Reconnections :
q from
colour singlet W
hadronizes with
the rest of the
event, for example
bbar from tbar

- MC09: $P_{reco} \sim 75\%$ - per string piece, affects the MC Event Record
- MC09 rates appropriate for the use in high-pt events cf. PYTHIA authors.

Nonperturbative Systematics: Colour Reconnections - 2

- MC09: high CR rates,
- MC08: different CR model, much lower CR activity.
- Left Figure: δR between b-quark entering the and B-had. leaving the string.
- Right Figure: turn CR off in MC09, no UE retuning $\Rightarrow$ reco. t-mass effects.





Nonperturbative Systematics < ISR/FSR Systematics

- Have estimated the Multiple Interactions/UE effects and b-fragm. effects,
- see Additional slides, showing the numbers here.

- Only an order of magnitude estimation;
- simple parameter variations and no UE/FSR retuning!

% events passing baseline analysis cuts					
MC09	MI$\uparrow$	MI$\downarrow$	b-frag: softer FF	b-frag: harder FF	**no CR**
36.5 %	37.1 %	36.2 %	36.6 %	36.9 %	**37.3 %**

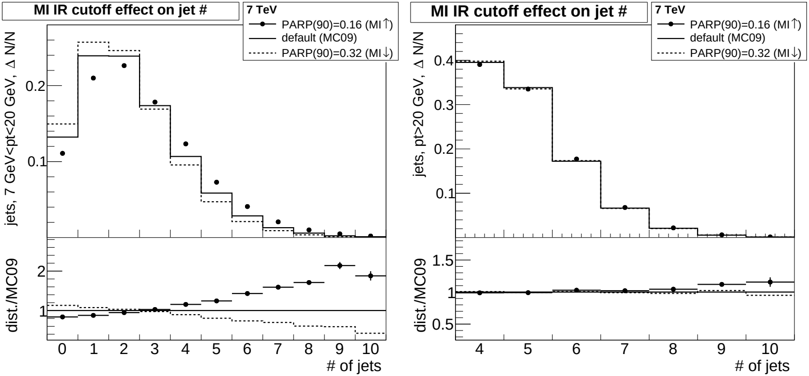
- Nonperturbative activity changes JES, but can also have less trivial effects (CR),
- nonperturbative systematics < ISR/FSR systematics, at least for the early data.

Nonperturbative Systematics: Underlying event

- Estimate order of magnitude by using the MI cutoff:

$$\mathrm{pt}_{\mathrm{min}}(\mathrm{E}_{\mathrm{CM}}) \propto (\mathrm{E}_{\mathrm{CM}})^{\mathrm{PARP(90)}},$$

- PARP(90) value = ? (important for UE extrapolation from Tevatront to LHC).
  - MC08, diffractive phenomenology: PARP(90)=0.16 (MI$\uparrow$),
  - MC09, experimental data: PARP(90)=0.25,
  - MC09, recent tunes have values up to: PARP(90)=0.32 (MI$\downarrow$).



- More UE increases soft jets #, JES, reconstructed top quark mass value . . . , but has only moderate effect on the hard jets.

Nonperturbative Systematics: b-fragmentation

- MC08: Peterson/SLAC (softer than MC09),
- MC09: Lund-Bowler, improve LEP description
  MC tuning PUB note: ATL-PHYS-PUB-2010-002,
- a harder fragm. function than MC09, but still ok with the LEP:
  PROFESSOR colab. tune: arXiv:0907.2973 [hep-ph],
- for the table in slide 11 I am using MC09 and changing the fragm. function only,
- no FSR etc. retuning done, so only useful for an order of magnitude.