Status and plans in Udine/ICTP

Kerim Suruliz (INFN Gruppo Collegato di Udine & ICTP, Trieste)

Bologna, November 24, 2009

The Udine/ICTP top physics group

The group consists of:

- Bobby Acharya (staff, INFN & ICTP, Trieste)
- Marina Cobal (staff, INFN & Università degli Studi di Udine)
- Michele Pinamonti (PhD, University of Trieste)
- Umberto de Sanctis (postdoc, SISSA)
- Kerim Suruliz (postdoc, INFN & ICTP, Trieste)

Our recent activities

Major activity: top pair production cross-section analysis in single lepton channel.

Important role in INT/PUB note at 10TeV.

Marina is the top cross-section group co-convener; Bobby was a co-editor of the PUB note.

Recent work on various topics - Michele on gg vs $q\bar{q}$ contributions to $t\bar{t}$ production, Also, W+jets data-driven background estimation.

Recent activities II

Umberto has previously been working on WW resonances. He's now going to focus on top. In particular, he's interested in working on $t\bar{t}$ resonances and high P_T tops. Umberto has been in touch with the relevant exotics subgroup - very positive response.

Our group has some previous experience with $Z' \rightarrow t\bar{t}$ - initial stages of collaboration with Univ. of Michigan.

Lots of other side activities - work in progress on triggers for $t\bar{t}$ in single lepton channel, QCD background estimation. These are done in collaboration with various people.

Participated in UAT stress test recently.

Plans for the near future

Writing a note on W+jets data-driven background estimation using W^+/W^- ratio. Hope to finish this before Christmas holidays.

Michele's work on gg vs $\bar{q}q$ nearly done.

Participating in the single lepton walkthrough in CERN on the 14th Dec (now postponed!). The aim of this is to present results updated to 7TeV, as well as think about early data ($\sim 10\mathrm{pb}^{-1}$ at 7TeV). Recall that the PUB note was for 200 pb^{-1} at 10TeV.

The walkthrough is meant to go beyond the PUB note in terms of level of realism: what data to we need? What if parts of the detector are not working? Focus on data-driven methods.

Plans for the near future II

Assuming everything goes OK with first collisions, start looking at 900GeV data in a few weeks' time.

Hope to collaborate with Milano and others from Italy on this topic.

We should look at 'real data' (i.e. cosmics) before then to make sure we're able to run on ESDs/partial AODs and face realistic detector conditions. tests in progress!

Once have a reasonable amount of data at 7TeV, start looking at

- fakes
- measuring W+jets background
- reconstructing W/Z
- and tops.

Plans for the near future III

Apart from real data: look at the re-reprocessed mc08; once they are finished, start looking at any mc09 samples at 7TeV.

Novel features (since reconstruction with r15): new jet algorithms (AntiKt), bugs fixed with respect to mc08 (e.g. such as working PDFInfo - making things like PDF error evaluation easier), etc.

There will be lots to do with mc09. Need to reconsider and improve our standard cross-section analysis (a cut reoptimisation will need to be redone).

Walkthrough for the baseline x-sec analysis

Things to be addressed:

- Results should be for 7TeV, at benchmark luminosities of $10/30/100\mathrm{pb}^{-1}$.
- Variation analyses should be presented which are designed for the case of detector problems.
- Data-driven strategies for background determination.
- Which datasets and data formats will be used?
- Which supporting measurements do we need?

Baseline analysis - results at 7TeV

- $t\bar{t}$ cross-section decreases by about 2.5
- W+jets cross-section decreases by about 1.5 - 1.7
- Single top decreases by about 2.0
- QCD decreases by about 1.5
- Assuming no major kinematical difference between events at 7 and 10 TeV for ttbar analyses, though events are more central at 7 TeV.

Now simply scale all the numbers from 10TeV, $200\mathrm{pb}^{-1}$ to 7TeV and the appropriate luminosity.

Use a reweighting procedure to perform this rescaling more accurately (e.g. using the PDFReweight tool). This is till far from perfect! Have to improvise due to lack of MC samples. This will be an issue when the real data comes, e.g. what if the beam energy is 3TeV but are samples are at 3.5TeV?

10 TeV Baseline Analysis in Single Lepton Channel at 200 pb^{-1}

Numbers of Selected Events				
	Electron Analysis		Muon Analysis	
Sample	**default**	**+M_W-cut**	**default**	**+M_W-cut**
$t\bar{t}$	2600	1286	3144	1584
W+jets	1305	448	1766	628
single top	210	81	227	98
Other	199	67	206	73
Signal	2600	1286	3144	1584
Background	1715	598	2199	799
S/B	1.5	2.1	1.4	2.0

This is what we had with 10TeV and 200pb^{-1}...

7 TeV Baseline Analysis in Single Lepton Channel at 60 pb^{-1}

Numbers of Selected Events				
	Electron Analysis		Muon Analysis	
Sample	**default**	**+M_W-cut**	**default**	**+M_W-cut**
$t\bar{t}$	312	154	377	190
W+jets	245	84	331	118
single top	32	12	34	15
Other	60	20	62	22
Signal	312	154	377	190
Background	337	116	427	155
S/B	0.9	1.3	0.9	1.2
$S/\sqrt{S+B}$	12	9	13	12

And this is what we have with 7TeV and 60 pb^{-1}. Have assumed that the other background cross-sections are unchanged from 10 to 7 TeV.

Other luminosities

Keeping the cuts fixed, we have for the standard selection in the electron and muon channels:

luminosity (pb^{-1})	S	B	$S/\sqrt{(S+B)}$
Electron channel			
10	52	56	5
30	156	169	8.7
100	520	562	15.8
Muon channel			
10	62	71	5.4
30	189	214	9.4
100	628	712	16.2

Uncertainty in $\sigma_{t\bar{t}}$ with 7TeV

The dominant sources of error are expected to be statistics (at low luminosity), the luminosity uncertainty and the JES. The latter two worsen slightly with respect to 10TeV, due to S/B being worse.

Luminosity (pb^{-1})	10	30	100
Statistical	20	12	6.3
Luminosity (20%)	25.9		
Luminosity (40%)	51.8		
JES 5%	+15.9-14.9		
JES 10%	+33.1-28.4		

Scenario of small integrated luminosity

If the integrated luminosity of the 7TeV data is low, we are severely limited by statistics. Might need to combine the cross section measurements in e and μ channels. Issues arise here: correlation of systematics.

Errors can't just be added in quadrature. Dilepton group has developed tools for taking this into account.

Also, consider loosening the cuts (3 jets > 40GeV fairly tight!) to improve statistics. This worsens the top reconstruction though - more radiation jets get through.

Another important question to be addressed: what is the smallest amount of data necessary to identify a top signal? When can we turn limits on the production cross-section into a measurement?

Data-driven background estimates

Data-driven background estimation: W+jets and QCD.

Milano group using method relying on W/Z ratio.

Recently we have been developing a method which relies of W^+/W^- asymmetry at LHC.

The basic idea is that p-p collisions will produce more W^+ than W^-.

However, for leptons coming from $t\bar{t}$, Z+jets, QCD, and most other backgrounds we expect equal numbers with positive and negative charge.

$\Rightarrow$ use this to measure W+jets background to $t\bar{t}$.

The ratio of W^+ and W^- production cross-sections is dependent on the com energy, s, via the pdfs. Roughly speaking

$$r = \sigma(W^+)/\sigma(W^-) = \frac{u(x_1)\bar{d}(x_2) + c(x_1)\bar{s}(x_2)}{\bar{u}(x_1)d(x_2) + \bar{c}(x_1)s(x_2)}$$

Setting $x_1 = x_2 = M_W/\sqrt{s}$ it can be seen that the ratio tends to 1 as $s \to \infty$ since for small x, $u \sim \bar{u}$ whereas for larger x the difference between valence and quarks is large and so $r > 1$. r is known to NNLO precision, the dominant uncertainty coming from PDF uncertainty. See 0905.3531 (Stirling et al). Ratio of total cross-sections of W^+ and W^- is predicted to be 1.373 ± 0.011 at 10TeV and 1.328 ± 0.010 at 14TeV.

l^+/l^- RATIOS FOR VARIOUS PROCESSES

Normalised results for W^+/W^- from ALPGEN samples.

Num jets	muon channel	electron channel
0	1.52 ± 0.02	1.49 ± 0.02
1	1.38 ± 0.03	1.32 ± 0.02
2	1.35 ± 0.09	1.39 ± 0.04
3	1.51 ± 0.11	1.43 ± 0.06
4	1.45 ± 0.05	1.49 ± 0.10

Results for the ttbar sample for electrons:

number of jets	+ events	- events	W^+/W^- ratio
0	119	163	0.73 ± 0.13
1	1772	1710	1.04 ± 0.05
2	6831	6990	0.98 ± 0.02
3	12277	12488	0.98 ± 0.02
4+	23383	23975	0.98 ± 0.01

Ratio compatible with 1, as expected for $t\bar{t}$.

Data-driven W+jets background determination

Now estimate the total W background to $t\bar{t}$ in the single lepton channel via

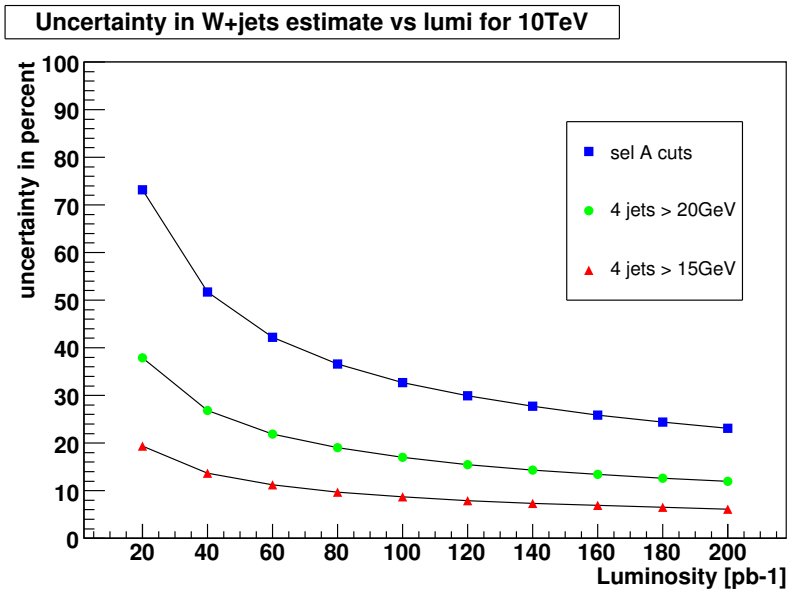
$$N_W = R(D^+ - D^-).$$

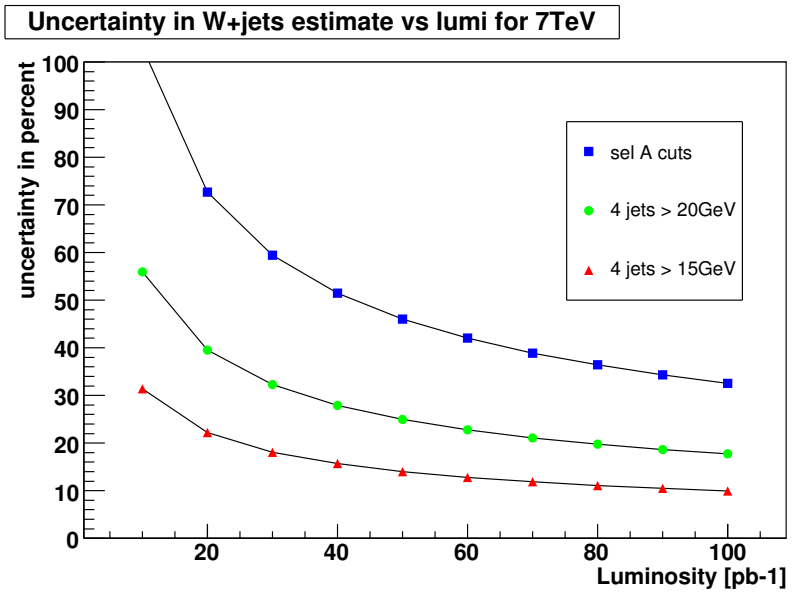
where $D^+(D^-)$ is the number of events passing the selection cuts, containing a positively (negatively) charged lepton, and R is a calculated from W+jets MC samples.

This works because everything else except W events will cancel out in $D^+ - D^-$, on average.

Statistical uncertainty in W+jets measurement

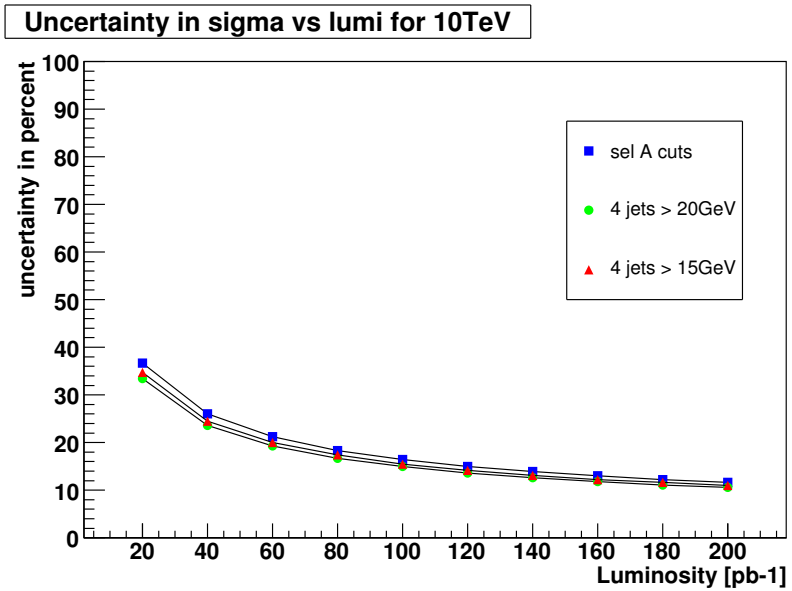
The dominant uncertainty here is statistical (as with the W/Z method). We can attempt to decrease this by loosening the cuts. What matters is the resulting error in σ.

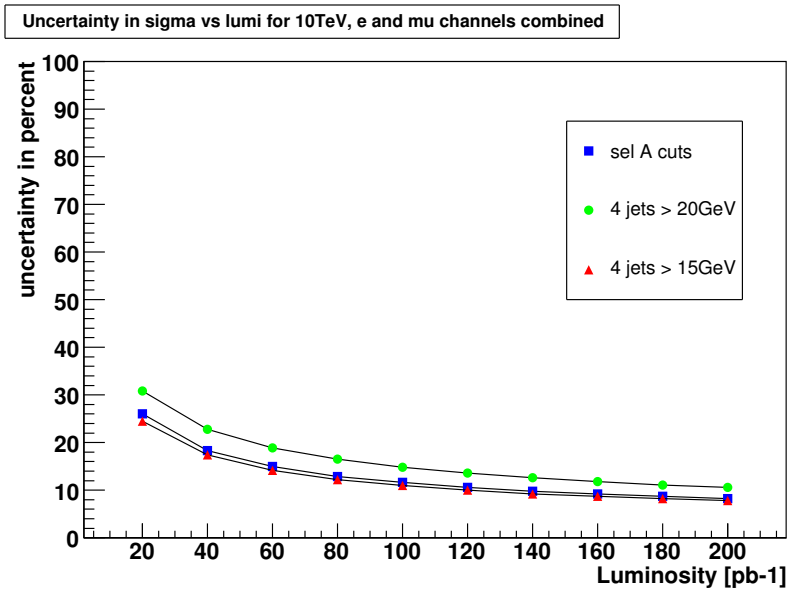




Statistical uncertainty in $\sigma_{t\bar{t}}$ @ 10TeV

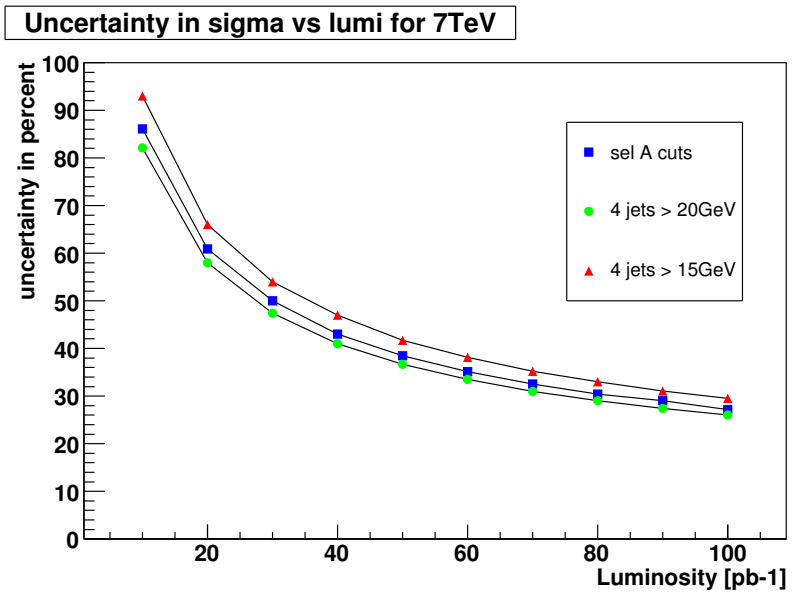
Uncertainty in $\sigma_{t\bar{t}}$ due to the statistical uncertainty in W+jets measurement, as a function of luminosity, at 10TeV. Left: electron channel; right; e and μ channels combined. Error seems fairly indepedent of choice of cuts.

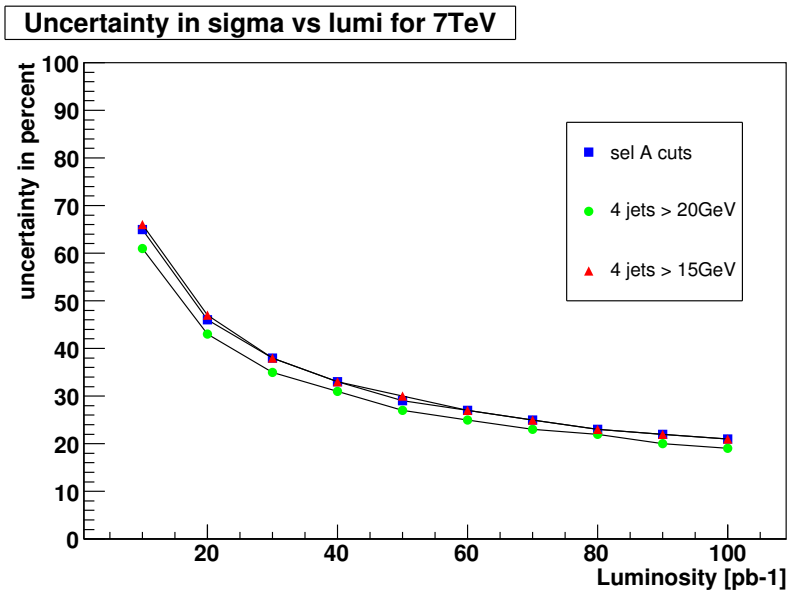




Statistical uncertainty in $\sigma_{t\bar{t}}$ @ 7TeV

Uncertainty in $\sigma_{t\bar{t}}$ due to the statistical uncertainty in W+jets measurement, as a function of luminosity, at 7TeV. Left: electron channel; right; e and μ channels combined.





Summary of data-driven W+jets estimate

Can use the W^+/W^- ratio to calculate W background to top pair production.

The method is very simple and straightforward to apply.

In order to minimise the statistical uncertainty on the size of W+jets background, the cuts should be relaxed. However, this has no major effect on the uncertainty in $\sigma_{t\bar{t}}$.

We need larger W+4,5 parton ALPGEN samples to calculate precisely the ratio W^+/W^- for the cuts using in the top x-sec group. There is a large uncertainty arising from limited MC statistics.

The charge asymmetry of t-channel single top is an important effect for tighter cuts. Will have to rely on MC to account for this.

Conclusions

Getting to the right frame of mind for early data.

Lots of effort going into data-driven background determination methods.

Need to think carefully about what we can do with 900GeV data. Basic checks that things are working well - which plots?

Approach to MC - how to rescale results from a difference centre of mass energy if samples not available. Also, how to combine results with data at different energies.

Technical issues: data formats! Common code, ntuples/DPDs? The possibility of using TAG?