FSP 101 ATLAS

Top pair production cross section in the met+jets channel at CDF

Giorgio Cortiana

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Top Physics Workshop, Grenoble, Oct 23rd-25th, 2008

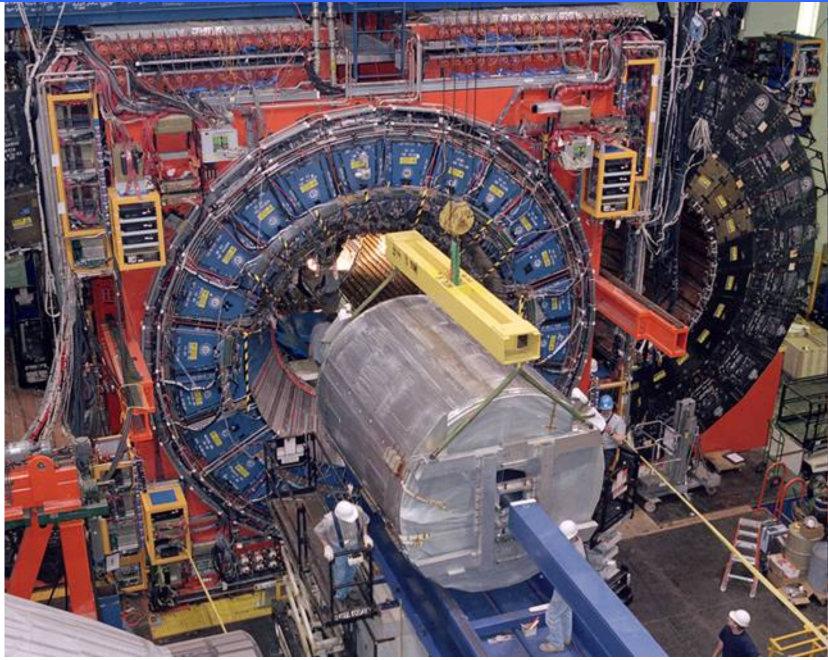
topics

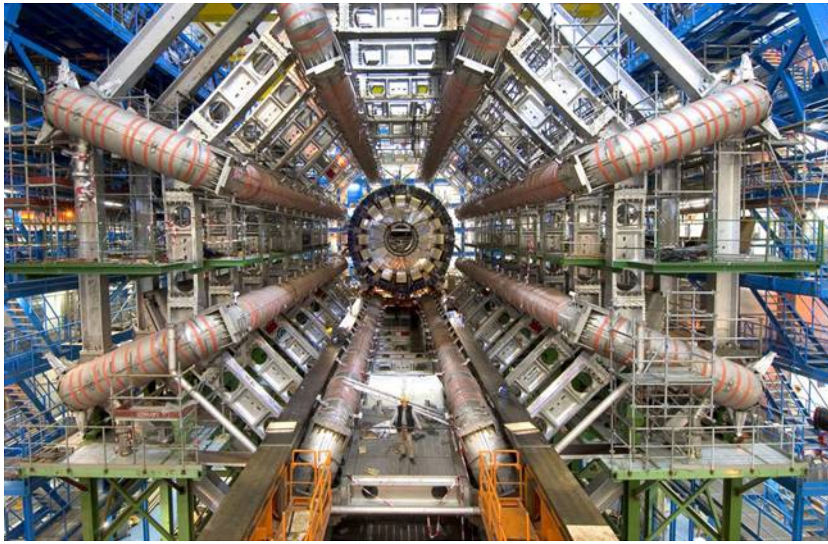
- Introduction CDF/Tevatron

- Top pair production cross section at Tevatron and LHC

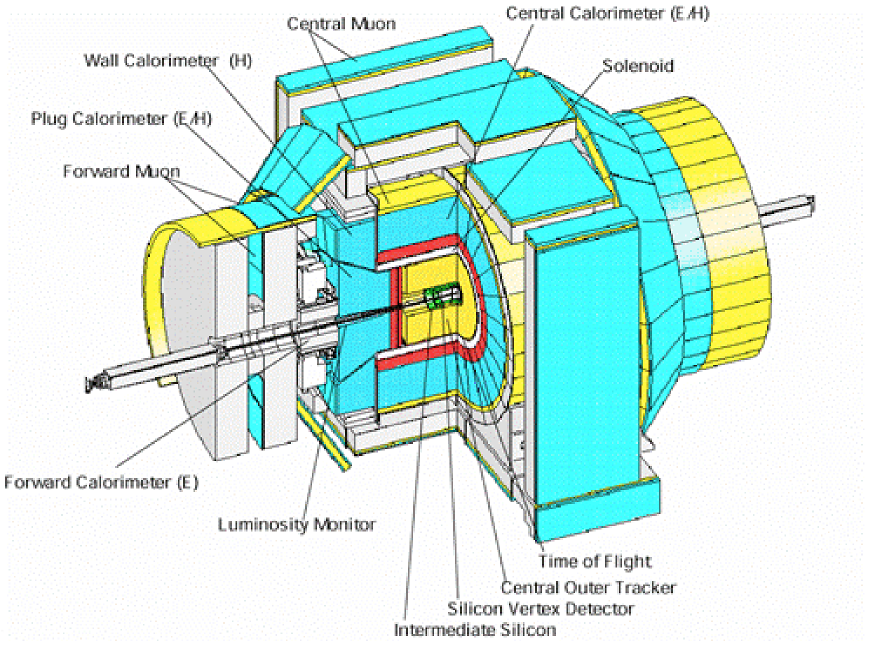
- The met+jets analysis at CDF:
  - introduction
  - kinematical selection
  - data-driven background estimate
  - background cross-checks
  - results

- Concluding remarks about the feasibility of the analysis in ATLAS



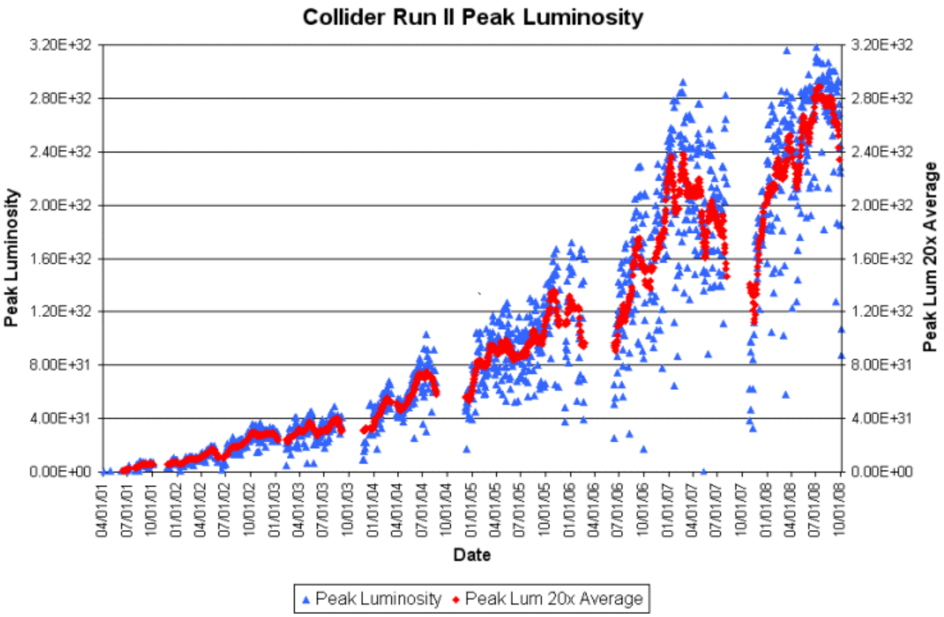


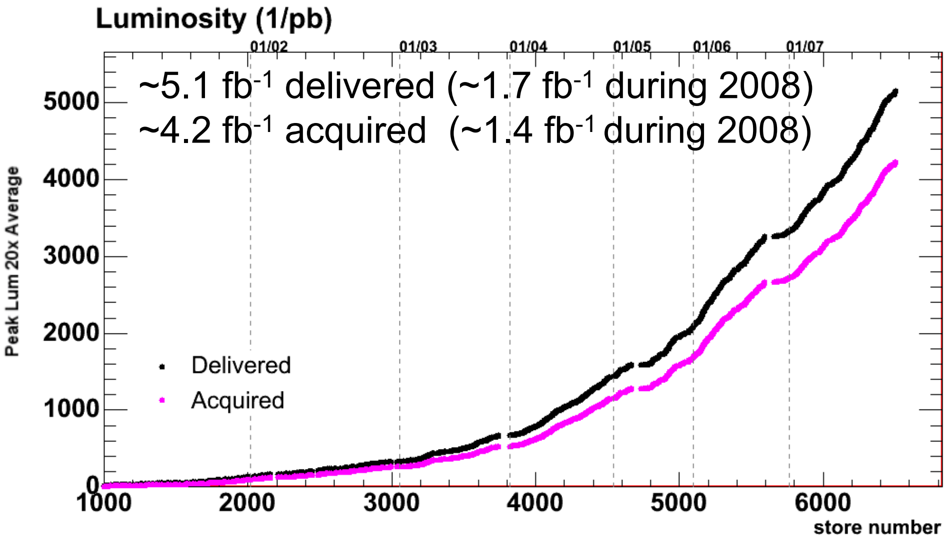
CDF/Tevatron overview



CDF is a multi purpose detector

- The tracking system is composed by multi wire drift tracking chamber surrounding a 7-layers silicon vertex detector, crucial to detect secondary vertices from heavy flavor decays.
- Energy measurements are provided by electromagnetic and hadronic calorimeters.
- Finally drift-tube chambers and scintillators used for muon detection cover the outer detector region.

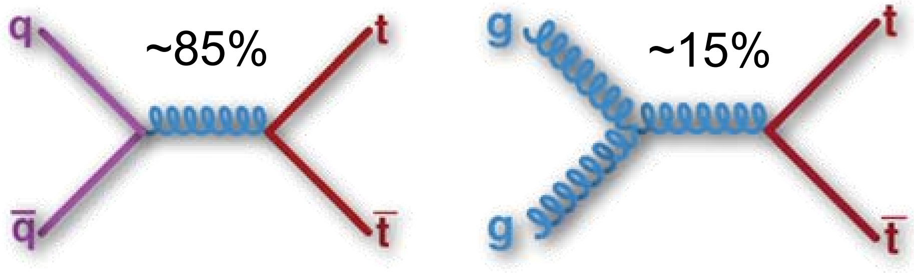


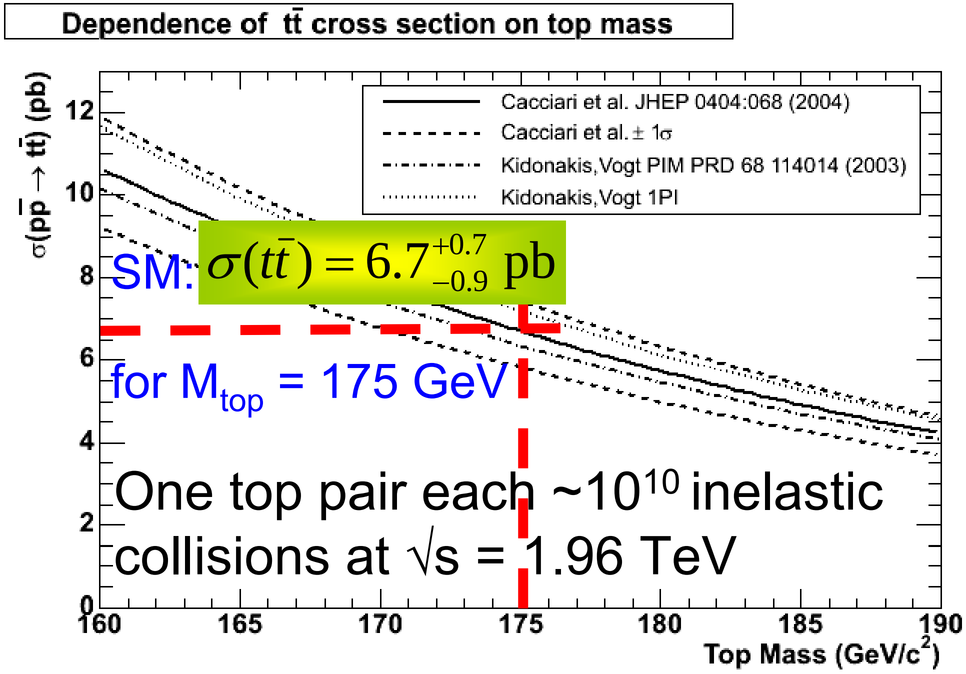


~5.1 fb^{-1} delivered (~1.7 fb^{-1} during 2008)
~4.2 fb^{-1} acquired (~1.4 fb^{-1} during 2008)

ttbar production @ √1.96 TeV

Top pair production through strong interactions

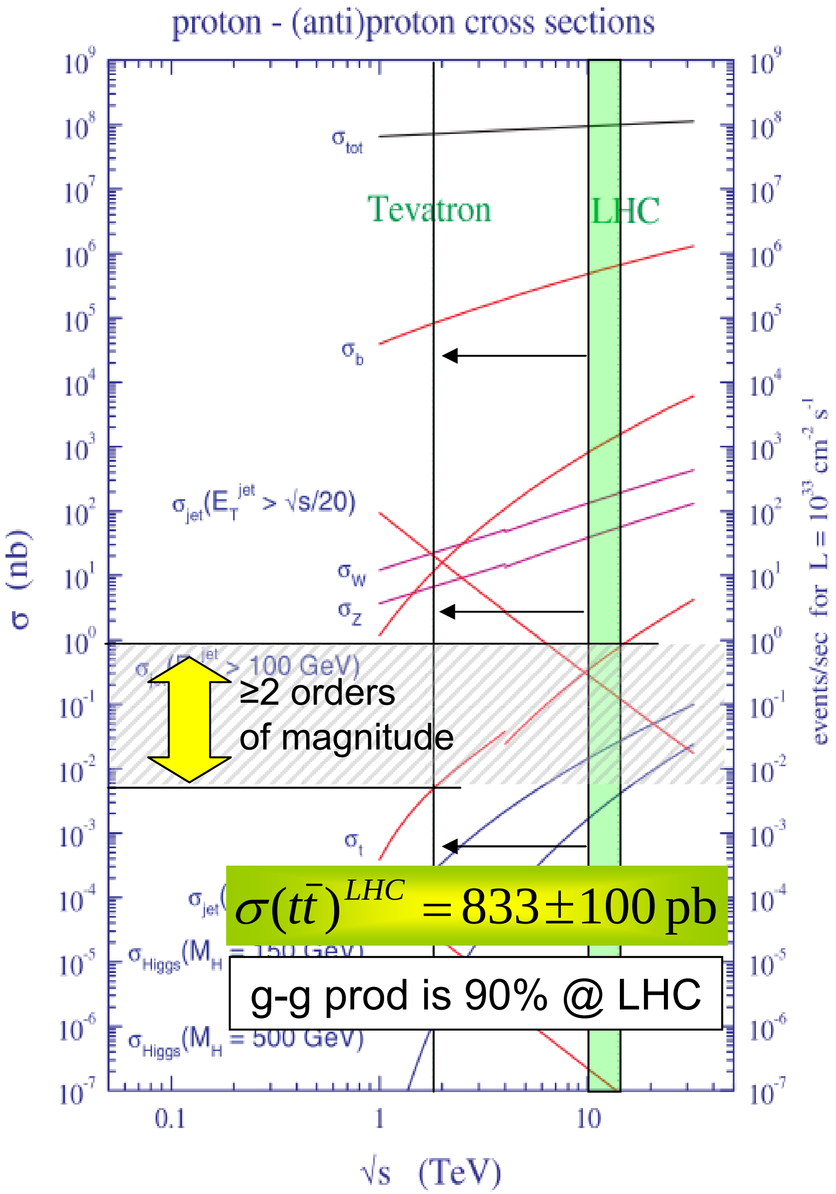




SM: $\sigma(t\bar{t}) = 6.7^{+0.7}_{-0.9}$ pb

for M_{top} = 175 GeV

One top pair each ~10^{10} inelastic collisions at √s = 1.96 TeV



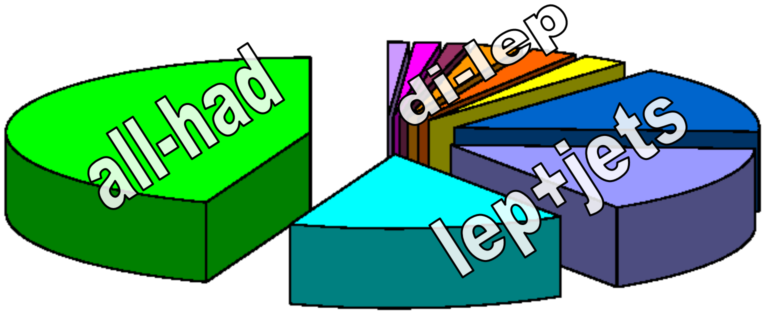
≥2 orders of magnitude

$\sigma(t\bar{t})^{LHC} = 833 \pm 100$ pb

g-g prod is 90% @ LHC

ttbar decay modes

At Tevatron top pairs are produced rarely, but in general provide clean handles allowing their separation from backgrounds

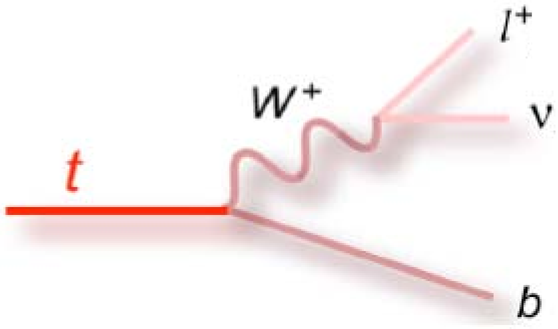


- events with W→lν decay contain high-p_T lepton/ high MET
  - possibility of triggering and/or selecting high purity samples
- each top pair decay mode contains nominally 2 b-quark jets
  - b-quark jets can be identified on the basis of b-properties: decays/lifetimes

Final states	**BR$_{tot}$**	**BR$_{used}$**	**Characteristics**
dilepton	10.3%	~5%	clean sample, low stat
lepton+jets	43.5%	~30%	golden channel, good S/B ratio
all-hadronic	46.2%	46.2%	challenging, huge background

met+jets analysis

…a new search channel



focus on missing ET from ν
rather than on lepton identification

- it is sensitive to leptonic W decays regardless of the lepton type
- has large acceptance with respect to W→τν decays.

✓ MET+jets new search channel in the top sector

✓ complementary and independent results w.r.t lepton-based and all-had measurements

✗ Large bkg: QCD, EWK+HF

- need an optimized kinematical + topological selection
- need b-jet identification to increase S/N ratio.

dataset and method

primary dataset shared with the all-hadronic analysis

- multi-jet trigger (**≥4 jets, ΔR=0.4, E_T ≥ 10 GeV, ΣE_T>125 GeV**)* dataset 311pb^{-1}

* before modifications for high luminosity scenario

MET+jets analysis
ttbar→blν bbarjj

yes

High missing E_T Selection?

no

All-Had analysis
ttbar→bjj bbarjj

[…]
see G.Compostella's talk

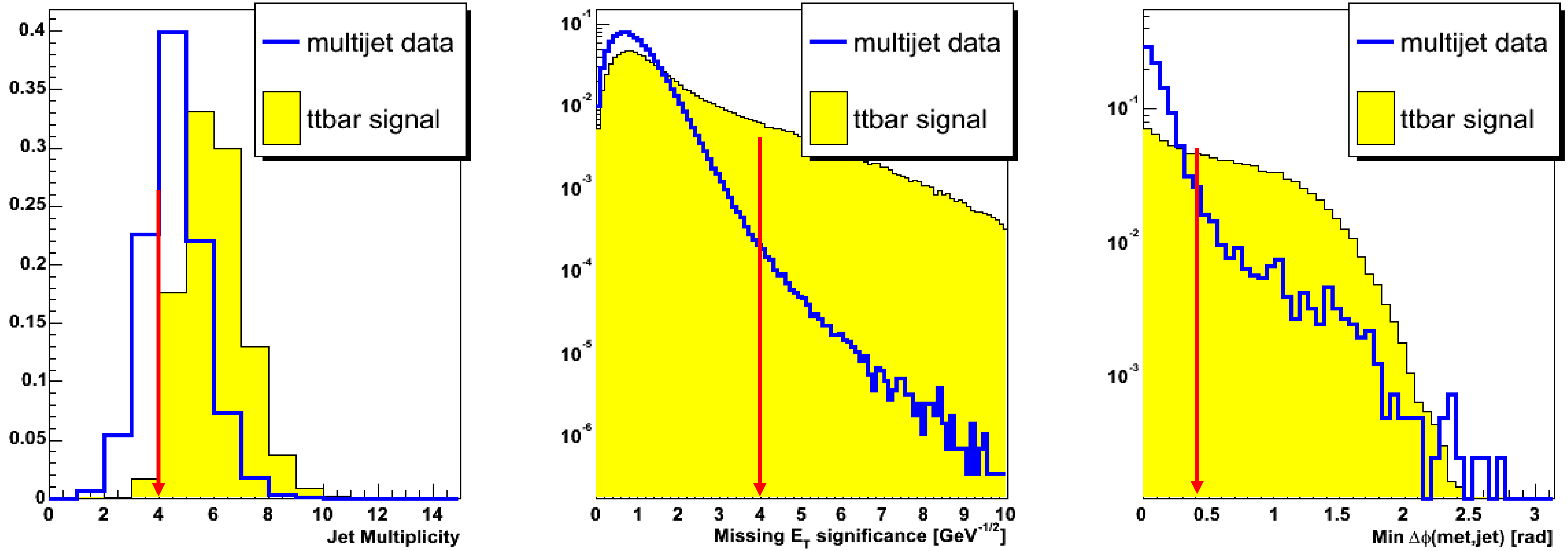
Clean up selection:

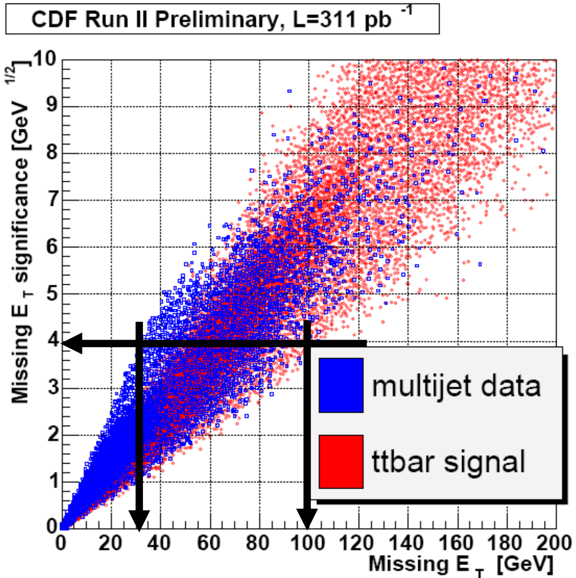
- **Tight leptons (e/μ) veto**
  - **no overlap w/ other L+J top analyses**
  - **increased relative contribution from W→τν +jets channel**

Kinematical selection + data driven background:

- **based on optimization procedure aimed at minimizing the stat. uncert. in the x-sec measurement (MC: pythia, 167fb^{-1}, m_{top} = 178 GeV)**
- **b-tagging matrix approach to predict the absolute amount of background**

kinematical selection





The cuts

- **N jets($E_T \geq 15$ GeV; $|\eta|<2.0$) ≥ 4**
- $\not{E}_T / \sqrt{\sum E_T} \geq 4.00$ **$GeV^{1/2}$**
- **min Δφ(met,jet) ≥ 0.4 rad**

promise a relative statistical uncertainty of 17.5%

kinematical selection - results

- **N jets($E_T \geq 15$ GeV; $|\eta|<2.0$) ≥ 4**
- $\not{E}_T / \sqrt{\sum E_T} \geq 4.00$ **GeV$^{1/2}$**
- **min $\Delta\phi$(met,jet) ≥ 0.4 rad**

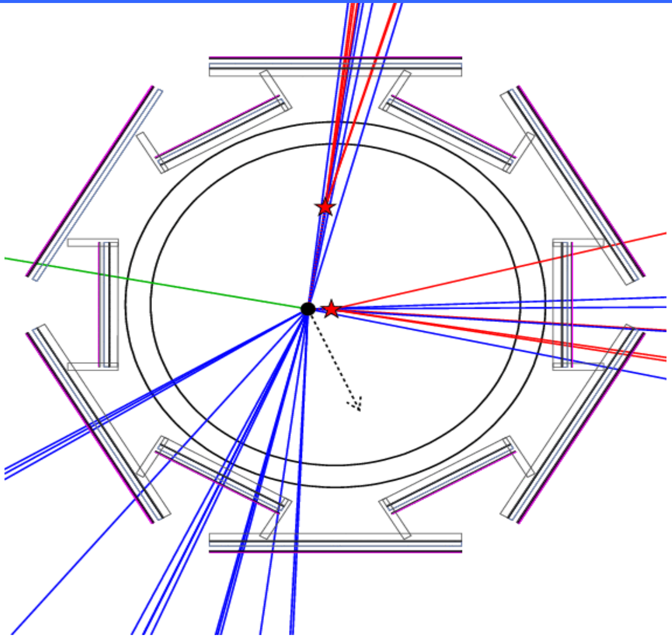
N events	**$MC_{\tau\text{+jets}}$**	**$MC_{\text{incl.}}$**	**Data**
total	149,323	1,021,024	4,249,644
$N_{\text{jets}} \geq 4$	72,708	549,138	2,781788
$MET_{\text{sig}} \geq 4$	29,830	78,145	3,996
min$\Delta\phi$(met,jets)≥0.4	19,079	49,848	597
≥1 b-tag	11,666	30,410	106
in 311 pb^{-1}	***21.7***	***56.5***	***106***

	S/N
w/o b-tagging	0.18
w/ b-tagging	1.14

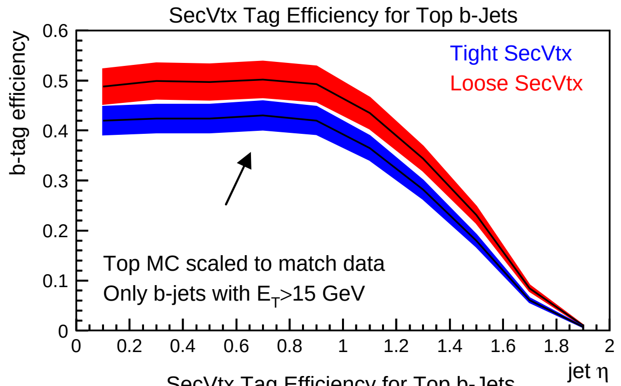
~40% of the total acceptance is provided by the tau+jets decay channel;

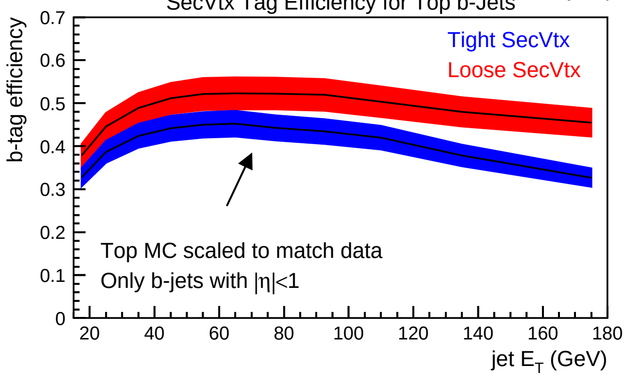
the remaining ~60% come from e/μ + jets ttbar events failing the tight lepton identification requirements

b-jets identification at CDF

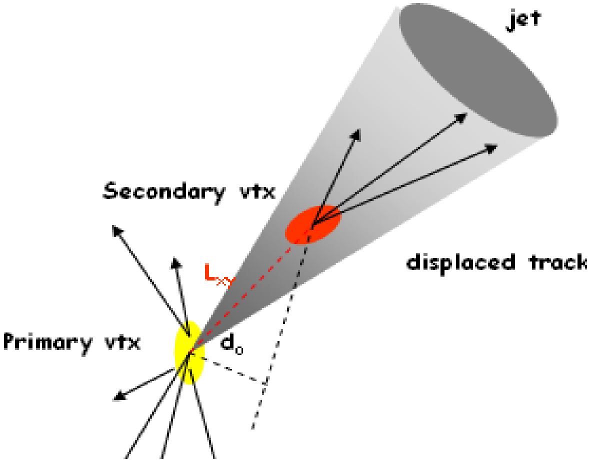


Tight and Loose tagging options to retain signal in double tag searches

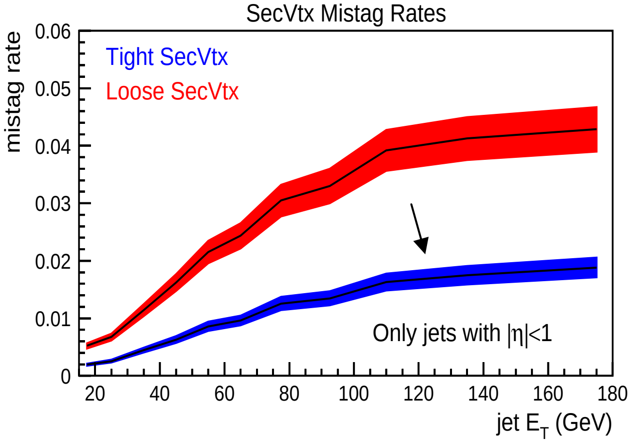




SECondary VerTeX tagger



- tracks with significant IP are used in a iterative fit to identify the secondary vertex inside the jet.



Efficiency drops at low jet E_T and high rapidity but is 45-50% for ttbar central b-jets

Mistag rates are kept typically at 1-2% (tight SecVtx)

data-driven background estimate

b-jets are copious in ttbar decays, rare in background processes

b-tag rates can distinguish the two components

Method implementation:

- Look at the b-tag rates, *P*, in multi-jet triggered data $\longrightarrow$ $P = \frac{N^{b-tagged}_{jets}}{N^{good}_{jets}}$

  using 3-jet ($E_T^{corr} \geq 15$ GeV, $|\eta| \leq 2.0$) events: $F_{top} = 2\times10^{-5}$

	3-jet events*	n_{AveTag}
data	879,187	0.065
exp ttbar (MC)	16.88	0.582

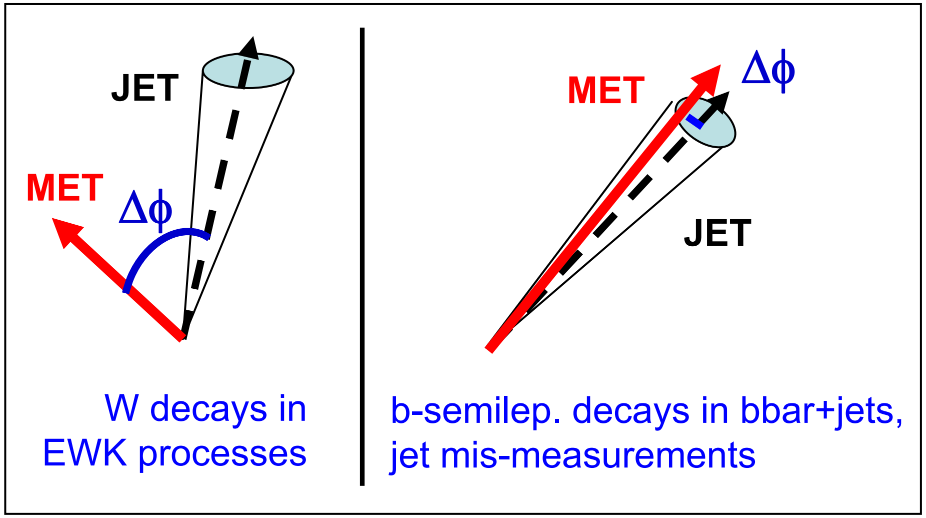
*the trigger requires 4 jets with raw E_T>10 GeV

- Use the variables on which the tag rate depends to construct a b-tagging matrix
- Check its predictions

background b-tagging rates

- A b-tag matrix is constructed as:

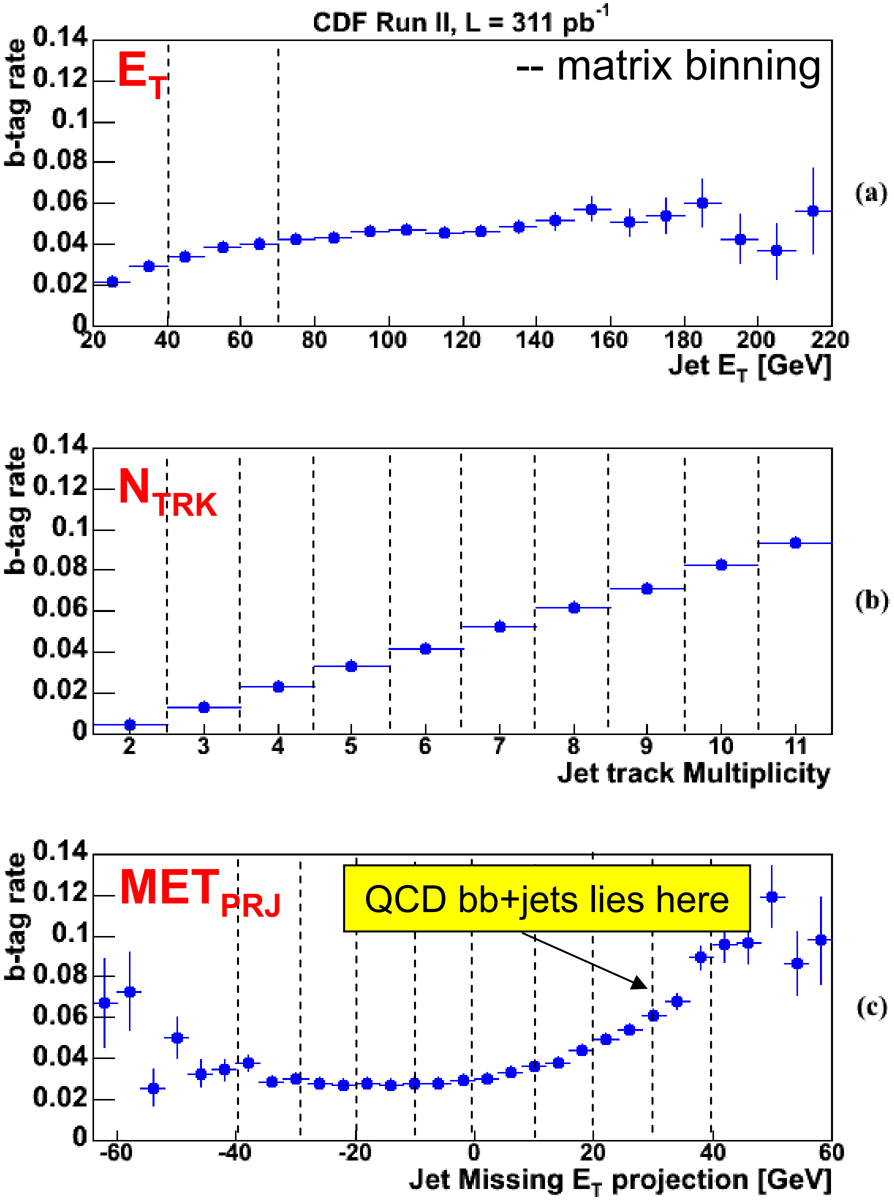
$$P\left(E_T, N_{trk}, MET_{PRJ}\right) = \frac{N_{jets}^{b-tagged}\left(E_T, N_{trk}, MET_{PRJ}\right)}{N_{jets}^{good}\left(E_T, N_{trk}, MET_{PRJ}\right)}$$



W decays in EWK processes

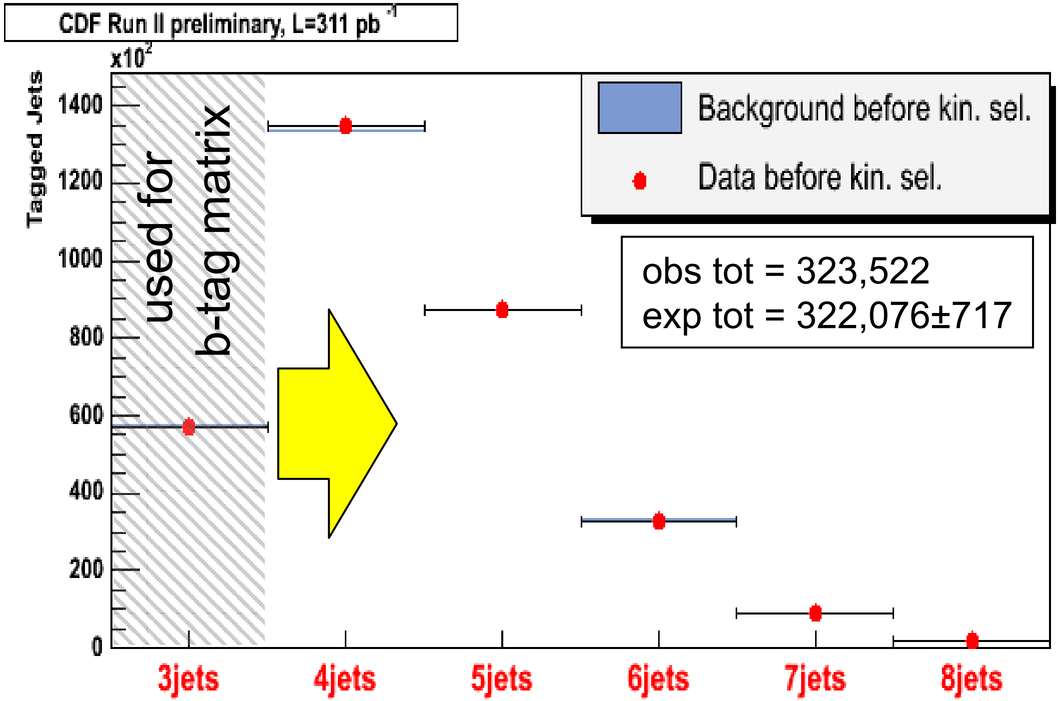
b-semilep. decays in bbar+jets, jet mis-measurements

- MET_{PRJ} has a consistent correlation with the heavy flavor component of the sample and allows to distinguish MET origins in relation to geometrical properties



b-tag matrix checks - 1

Extrapolate the tag rate from 3 jet events (where the matrix is calculated) to higher jet multiplicity events, before kinematical selection



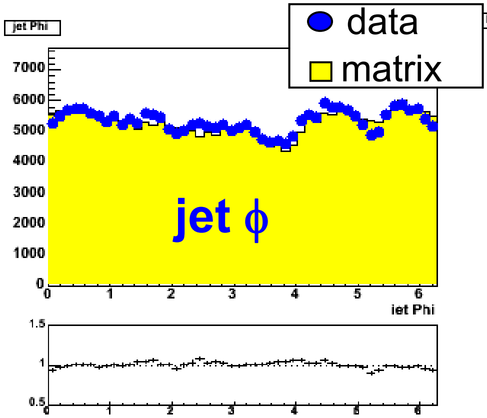
	3-jets	4-jets	5-jets	6-jets	7-jets	8-jets
obs	57,314	135,056	87,332	32,914	8,992	1,914
exp	57,314±233	133,275±546	87,156±370	33,184±149	9,147±43	2,000±10

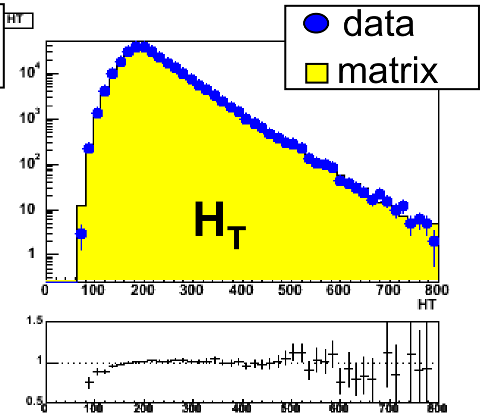
…stat only errors

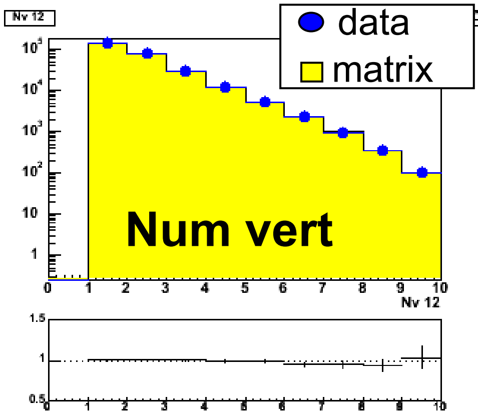
The agreement between the number of observed and matrix-predicted tagged jets is good for all jet multiplicities

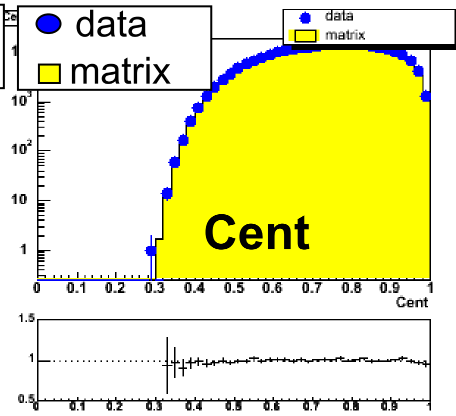
…but can we correctly predict kinematical distributions in the data? see next slide!

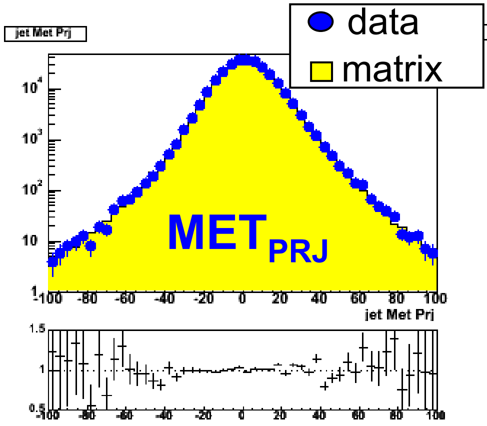
b-tag matrix checks - 2

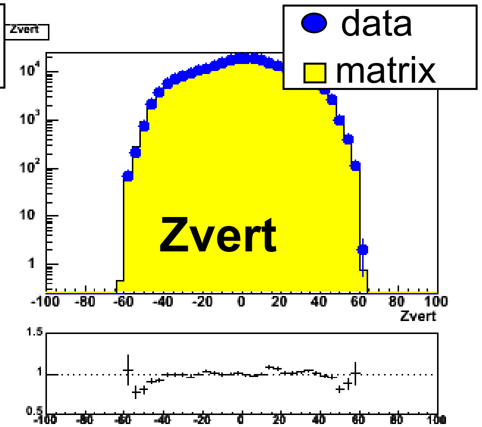


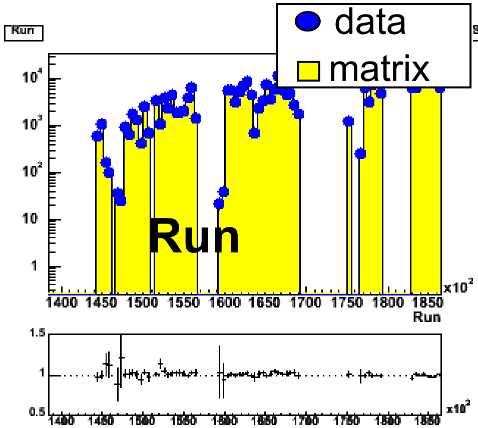


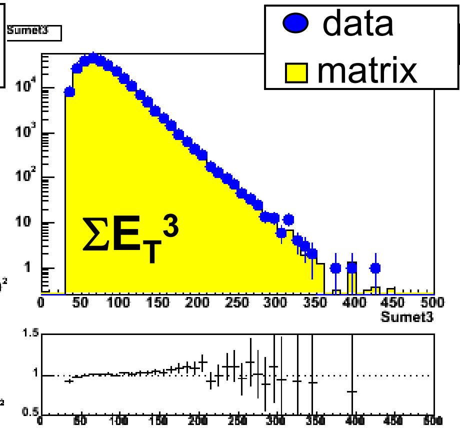






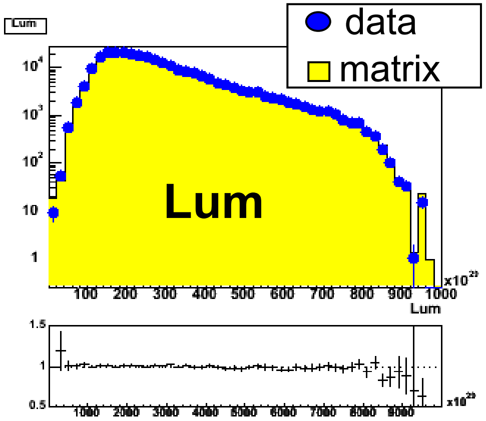


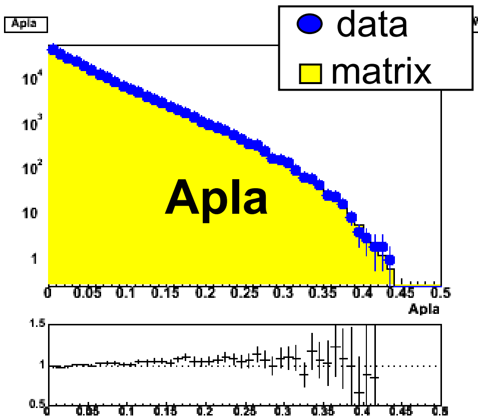


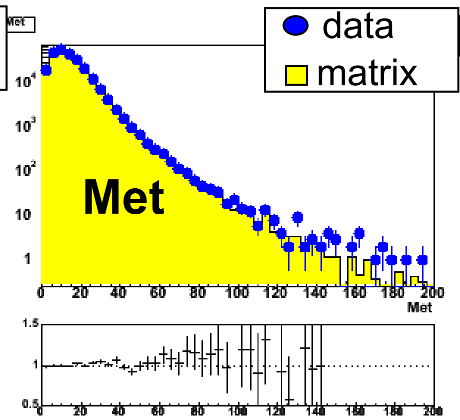


Jet Variables

Event variables

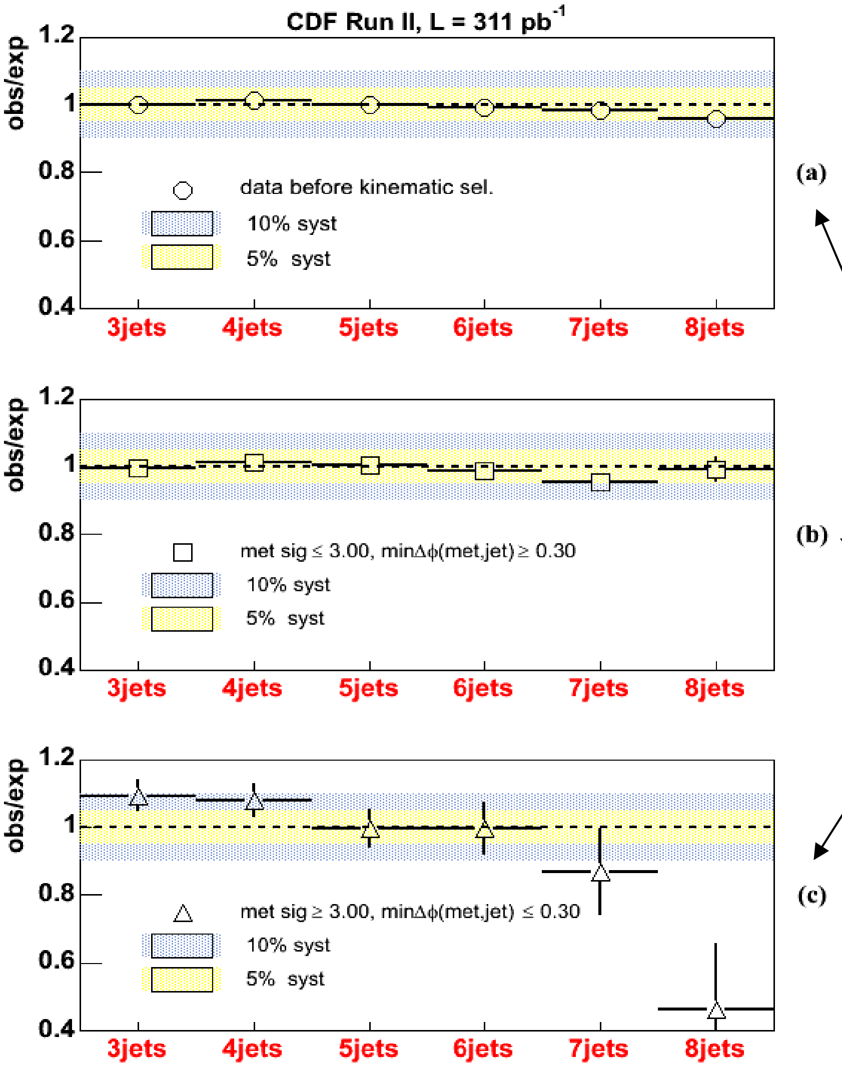


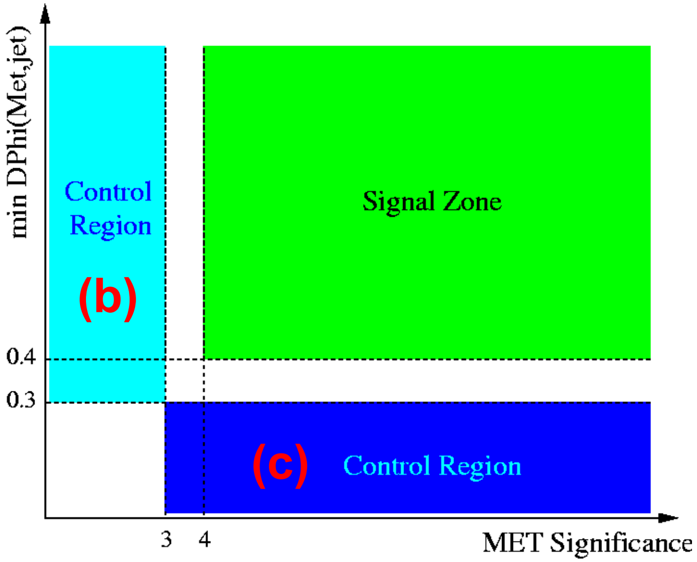




b-tag matrix checks - 3

The b-tag matrix background predictions are further checked in control samples, depleted of signal, obtained from data





(a) data before kinematical selection

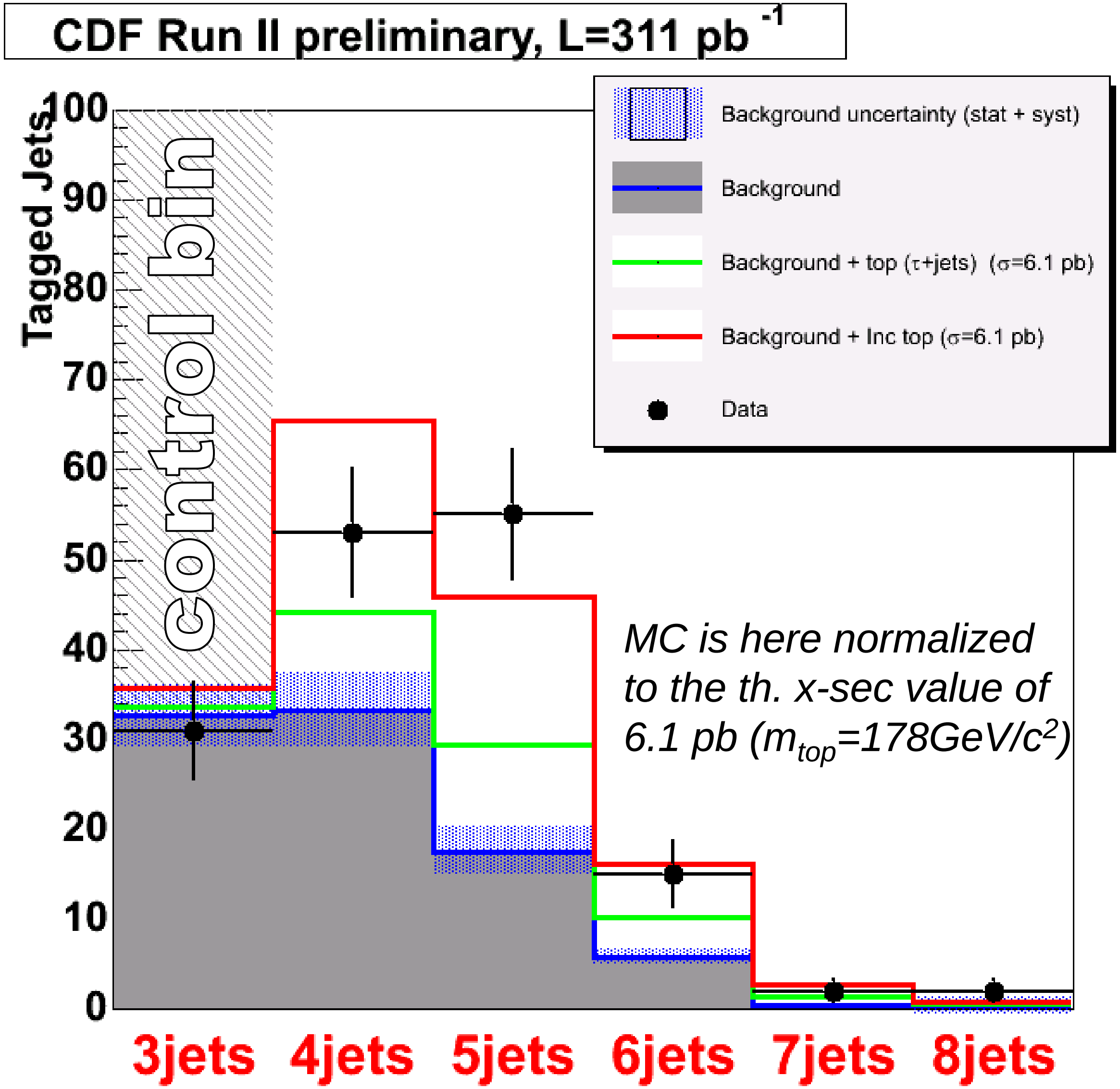
(b) data w/ met sig < 3 and minΔϕ > 0.3

(c) data w/ met sig > 3 and minΔϕ < 0.3

The matrix performs well in the control samples: the discrepancies in terms of the ratio obs/exp b-tags are limited at 10%

let's look at the selected sample

- N jets(E_T^{corr} ≥15; |η|<2.0) ≥ 3
- $\not{E}_T / \sqrt{\sum E_T} \geq 4.00$ GeV$^{1/2}$
- min Δϕ(met,jet) ≥ 0.4 rad
- ≥ 1 b-tagged jet



MC is here normalized to the th. x-sec value of 6.1 pb (m_{top}=178GeV/c^2)

- matrix-based background prediction is corrected with an iterative procedure to account for the ttbar presence in the pre-tag sample

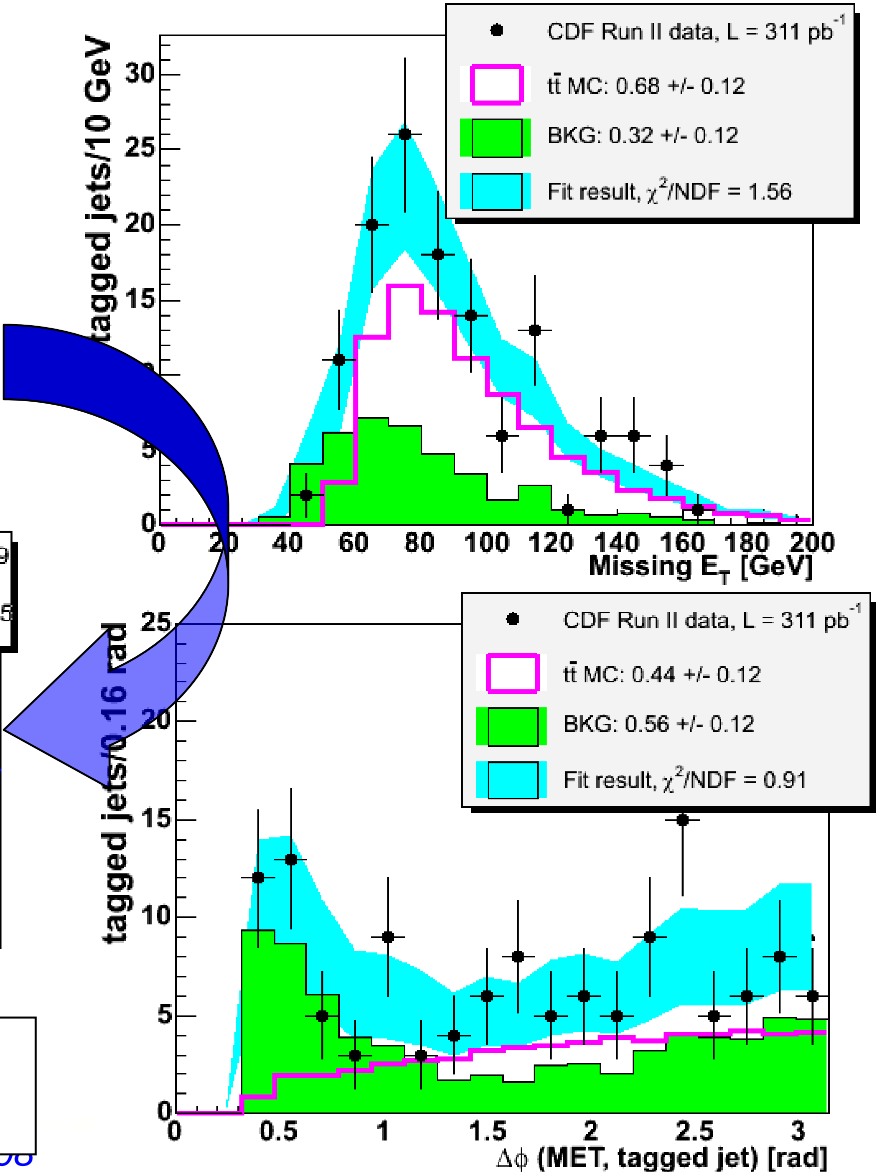
The excess is well consistent w.r.t. MC+BKG expectations in all jet bins

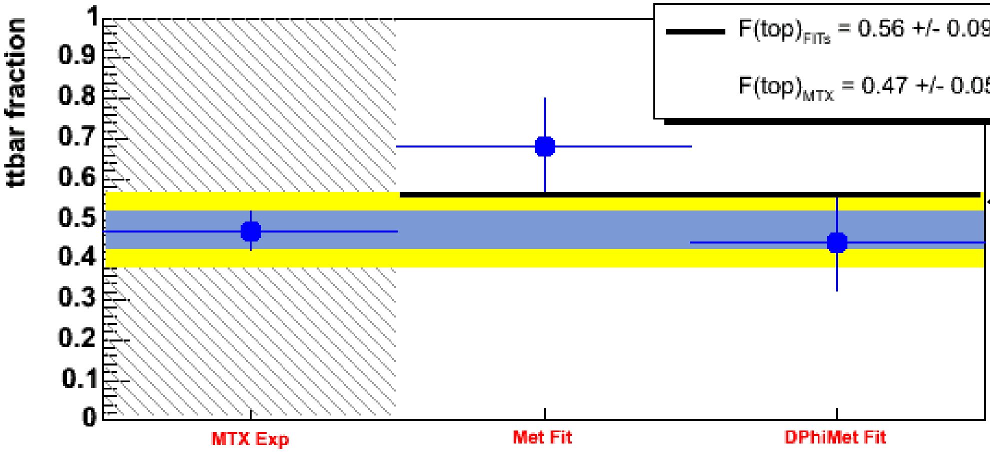
additional cross-checks

The excess attributed to top pair production is checked by looking to kinematical variables

Data distributions after kin sel + ≥1 tag are fit using a binned likelihood technique to the sum of:

- Inclusive $t\bar{t}$ template
- Matrix extracted bkg template





Fitted signal fractions are consistent with that calculated by the tag counting method

cross section measurement

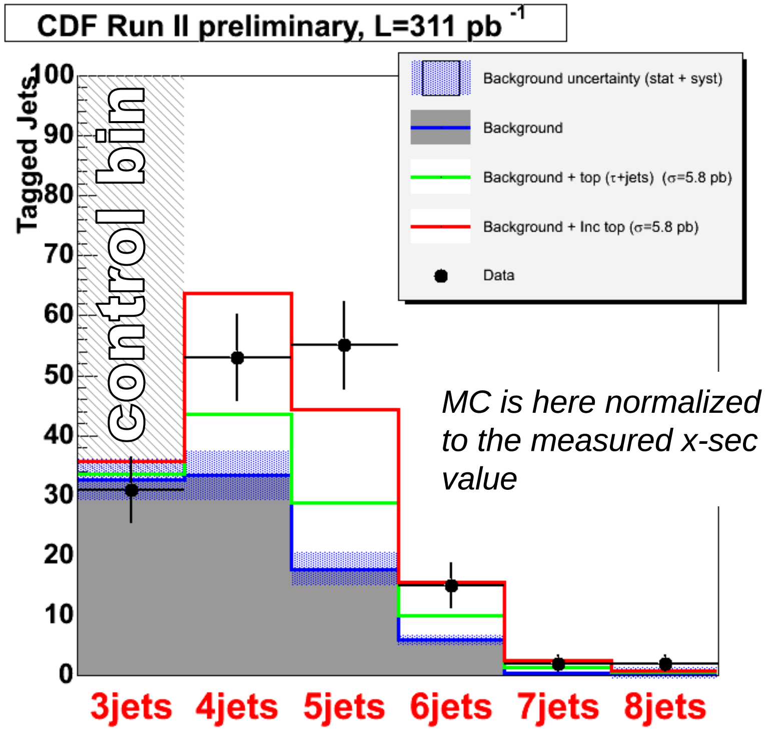
The cross section is calculated using a Poisson likelihood in which input parameters are subject to gaussian constraints.

for mtop = 178 GeV

$$\sigma_{ttbar} = 5.8 \pm 1.2\,(\text{stat})\ {}^{+0.9}_{-0.7}\,(\text{syst})\ \text{pb}$$

$$= 5.8\,{}^{+1.5}_{-1.4}\ \text{pb.}$$

σ_{stat}~20%
σ_{tot} ~25%



MC is here normalized to the measured x-sec value

- The cross section measurement changes by ± 0.05 pb for each -/+ 1 GeV/c^2 change in the m_{top} from the initial value of 178 GeV/c^2.

TABLE II: Relevant sources of systematic uncertainty.

Source	relative error
ϵ_{kin} systematics	
Generator dependence	8.2 %
Trigger acceptance	2.0 %
ISR/FSR	2.0 %
PDFs	1.6 %
Jet Energy Scale	1.5 %
Others	
Background prediction method	10.0 %
Luminosity measurement	6.0 %
ϵ_{tag}^{ave} (SECVTX scale factor)	5.8 %

PRL 96, 202002 (2006)

top mass measurement

MET + jets ttbar decays are also used to perform a top quark mass measurement

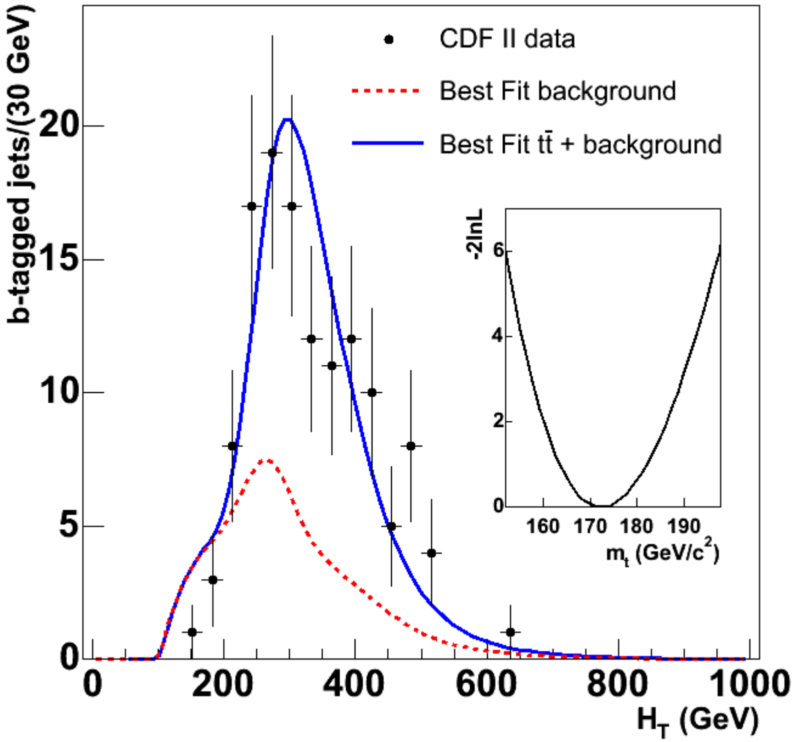
by adopting the template (TMT) method:

$$H_T = \Sigma E_T + MET$$

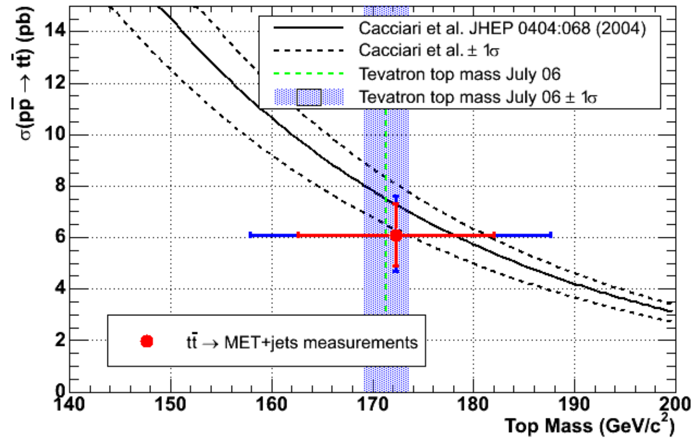
- data-driven background
- 21 signal templates parameterized linearly vs m_{top}
- likelihood fit to determine the best top mass

$$\boxed{M_{top} = 172.3\,^{+10.8}_{-9.6} \pm 10.8\ \mathrm{GeV}/c^2}$$

- systematic uncertainty dominated by JES

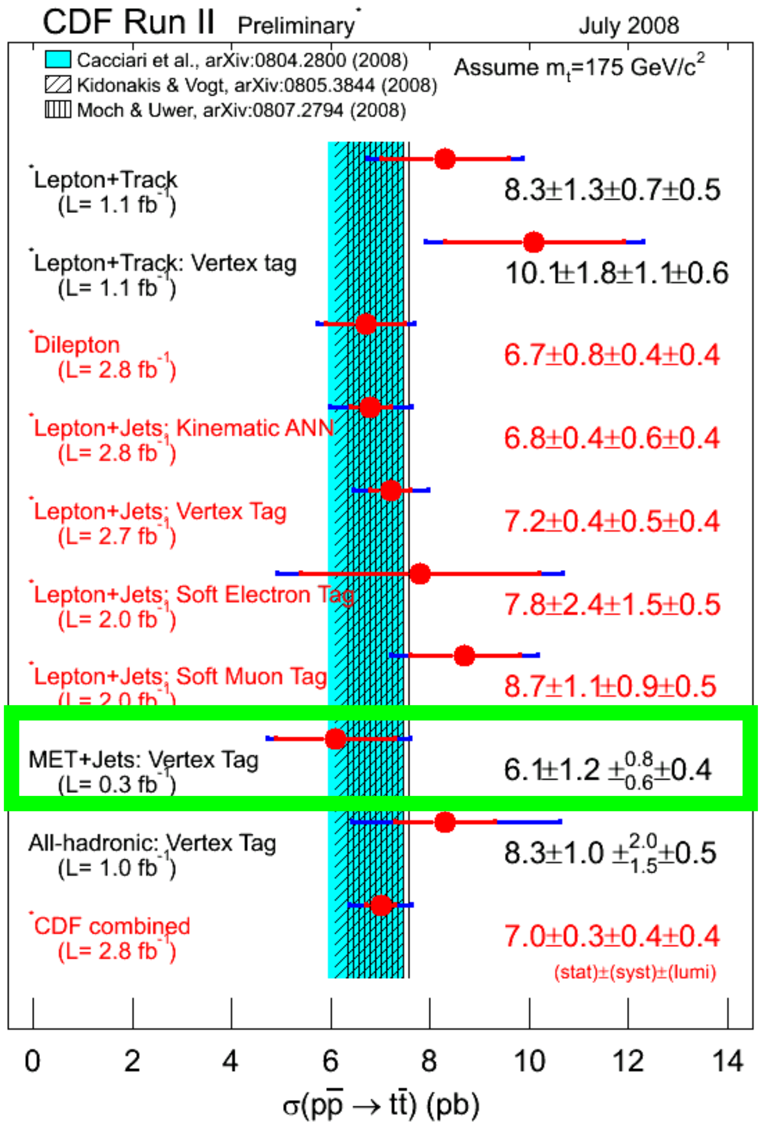


CDF Run II Preliminary, L=311 pb^{-1}



PRD 75 111103 2007

conclusions



- The MET+jets analysis explores high missing E_T sub sample of multi-jet data
- provided the first cross section measurements in the MET+jets channel.

- The measurement is sensitive to leptonic W decays regardless of lepton type and has large acceptance with respect to $W \rightarrow \tau\nu$ decays.
- The x-sec result is competitive (carried the ~30% of the weight in the previous combination of 0.7 fb^{-1}).

The same channel can be used for top mass measurements via template method.

For far H_T templates were used, but more sophisticated hadronic top mass reconstruction (max-P_T, minΔR etc..) as in ATLAS could be used

a refined NN-based analysis on is ongoing with ~2fb^{-1} (G. Compostella)

remarks

... feasibility of the analysis in ATLAS

The met+jets analysis needs:

- good understanding of jet/MET resolutions
- good understanding of b-tagging algorithms

⇒ not suitable for very first data, maybe with O(100 pb^{-1})

….see Ariel Schwartzman’s, and Richard Hawkings’ talks

remarks

... feasibility of the analysis in ATLAS

The met+jets analysis needs:

- good understanding of jet/MET resolutions
- good understanding of b-tagging algorithms

Trigger issues:

Need to study carefully which ATLAS dataset to use:

- multi-jet vs missing E_T/+jets triggers
- understanding pre-scales in trigger menus for low and high instantaneous luminosities

…and lot of work would need to be done… in any case it will be interesting to try it!

EF chain	P	L2 chain	P	L1 item	P
xe15	1	xe15	1	XE15	30000
xe20	1	xe20	1	XE20	7000
xe25	1	xe25	1	XE25	1500
xe30	1	xe30	1	XE30	200
xe40	1	xe40	1	XE40	20
xe70	1	xe70	1	XE70	1
xe80	1	xe80	1	XE80	1
j70_xe30	1	j70_xe30	1	J70_XE30	1
J10	1	J10	1	J10	42000
J50	1	J23	1	J10	42000
J115	1	J60	1	J23	2000
J140	1	J80	1	J35	500
J180	1	J110	1	J42	100
3J25	1	3J15	1	3J10	20000
3J60	1	3J35	1	3J18	1000
3J180	1	3J120	1	3J70	1
4J45	1	4J15	1	4J10	4000
4J80	1	4J35	1	4J18	100
4J95	1	4J50	1	4J23	20
4J125	1	4J80	1	4J35	1

information from 10^{31} trigger menu*

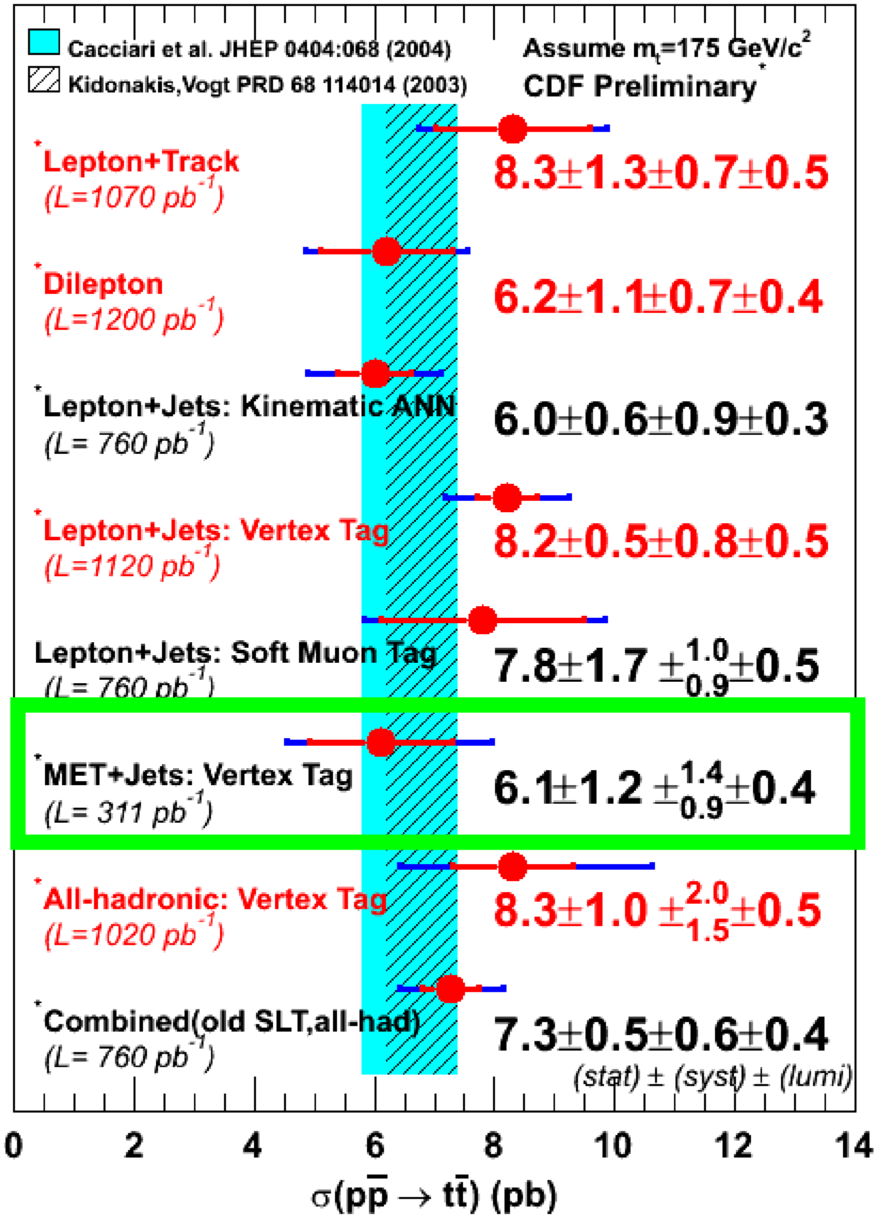
*information on trigger menu for L=10^{31}cm^{-2}s^{-1} from: http://tbold.web.cern.ch/tbold/view_menu.php?name=lumi1E31_14.2.20

- the end -

… thank you!

- backup slides -

previous CDF-combination results



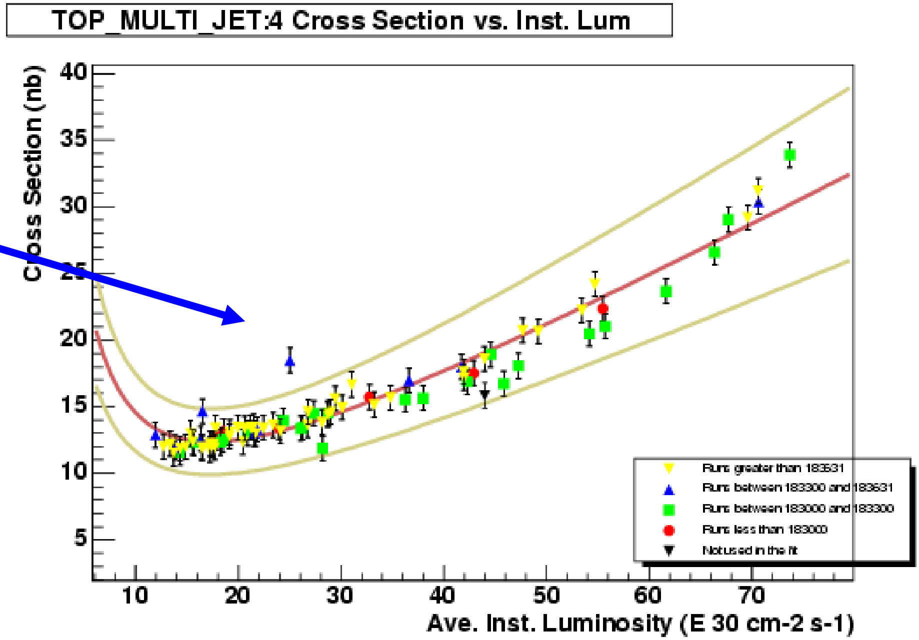
- The measurement is **sensitive to leptonic W decays** regardless of lepton type and has large **acceptance with respect to W→τν decays.**
- Recent CDF top pair **x-sec combination** demonstrated the **competitiveness of the measurement** which carried the ~30% of the weight in the previous combined result.

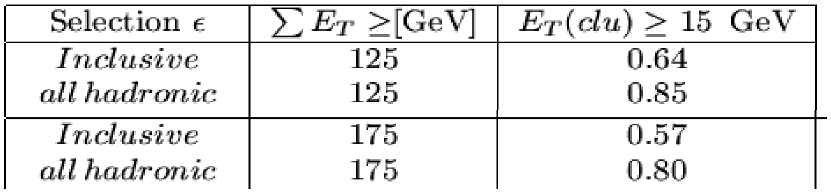
CDF multi-jet trigger at low lumi

The multijet trigger used for the MET+jets analysis required modification in order to cope with increasing instantaneous luminosity provided by the Tevatron.

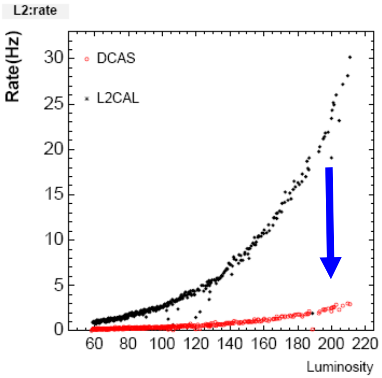
As a results the ΣE_T cut accompanying the jet selection was increased from 125 GeV to 175 GeV.

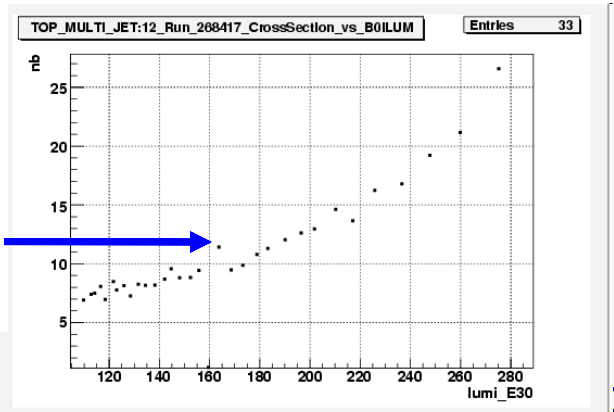
Later (2007), all CDF calorimeter trigger system was upgraded in order to use full calorimeter granularity, and adopt fixed cone algorithms already at the second trigger level.

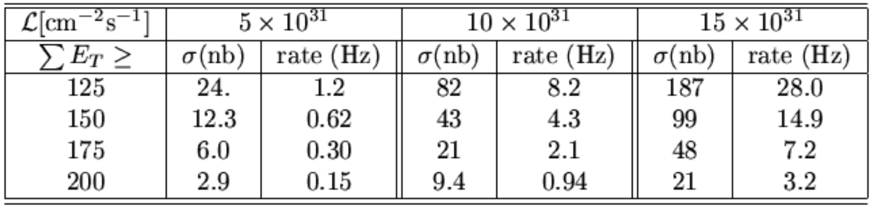




Selection ϵ	$\sum E_T \geq$[GeV]	$E_T(clu) \geq 15$ GeV
Inclusive	125	0.64
all hadronic	125	0.85
Inclusive	175	0.57
all hadronic	175	0.80

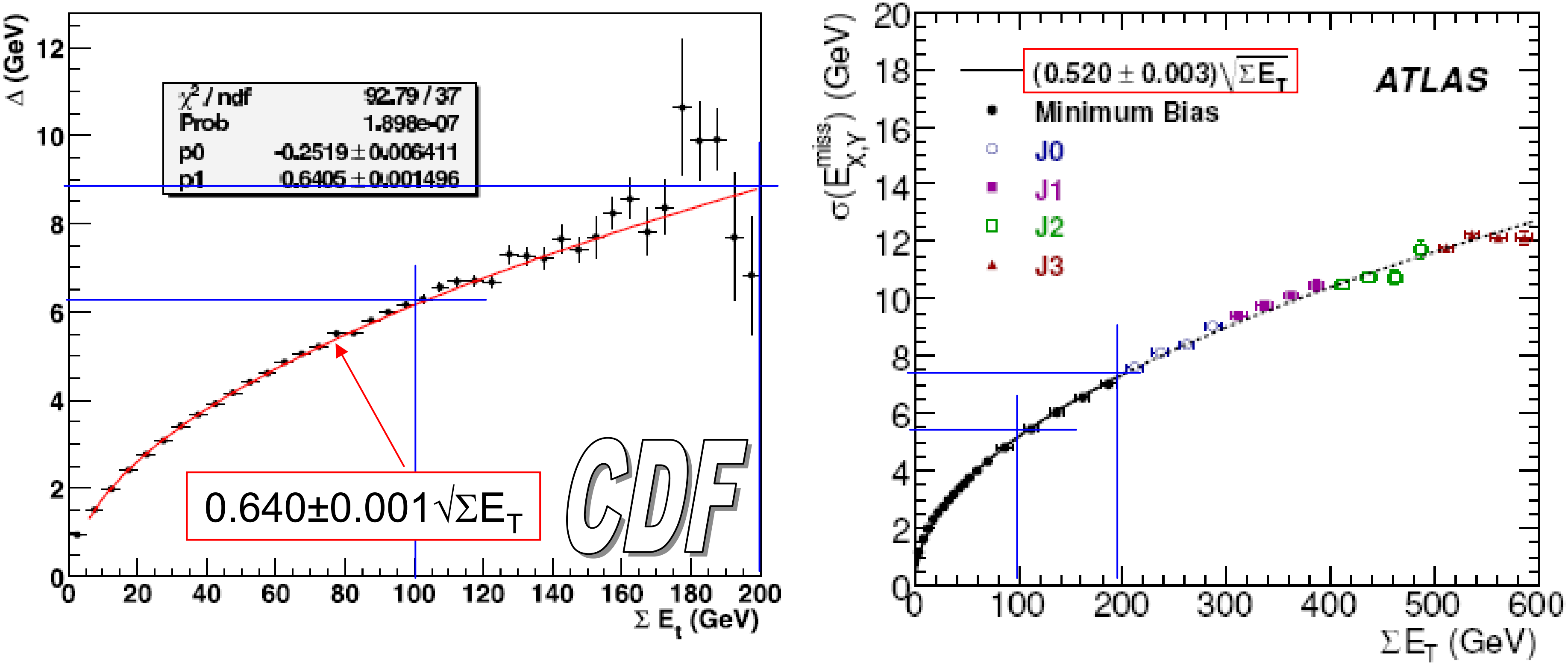






$\mathcal{L}[\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	5×10^{31}		10×10^{31}		15×10^{31}	
$\sum E_T \geq$	σ(nb)	rate (Hz)	σ(nb)	rate (Hz)	σ(nb)	rate (Hz)
125	24.	1.2	82	8.2	187	28.0
150	12.3	0.62	43	4.3	99	14.9
175	6.0	0.30	21	2.1	48	7.2
200	2.9	0.15	9.4	0.94	21	3.2

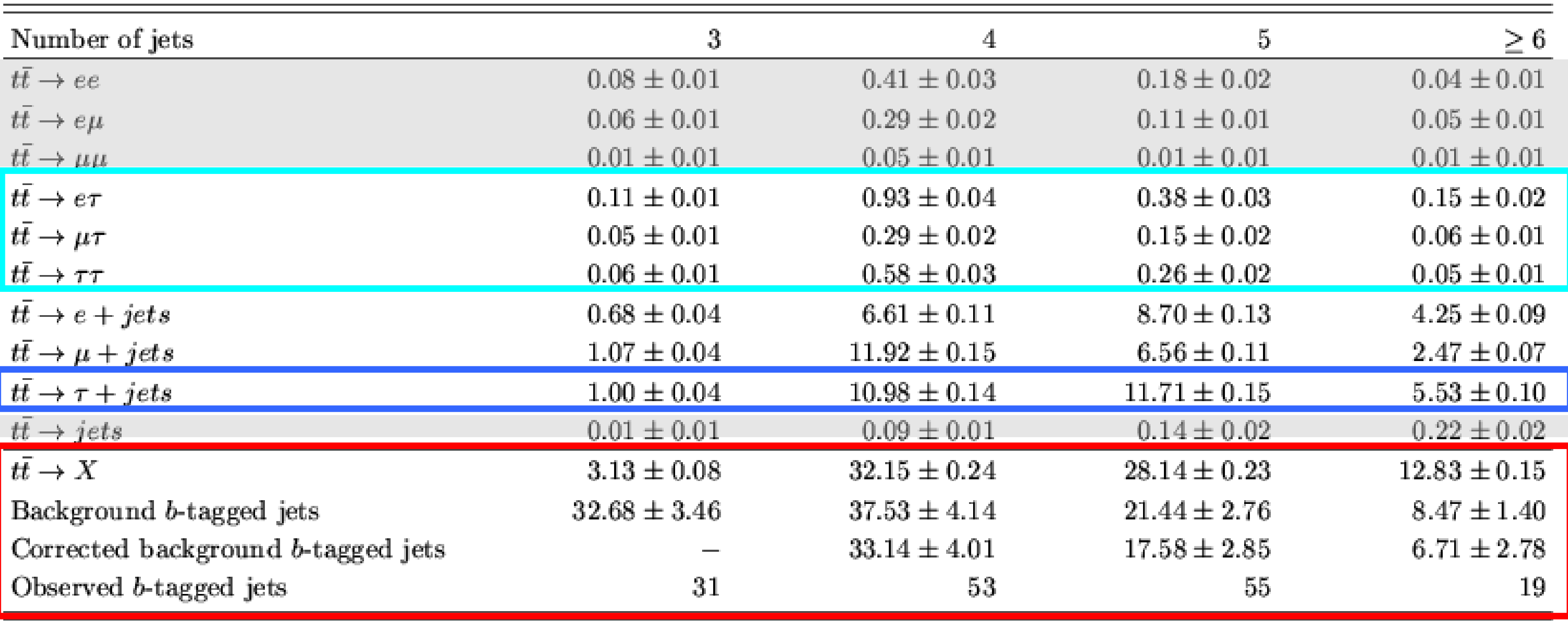
met resolution CDF vs ATLAS



Missing E_T resolution in min bias events

let’s look at the selected sample

…which ttbar decays are entering our selection?



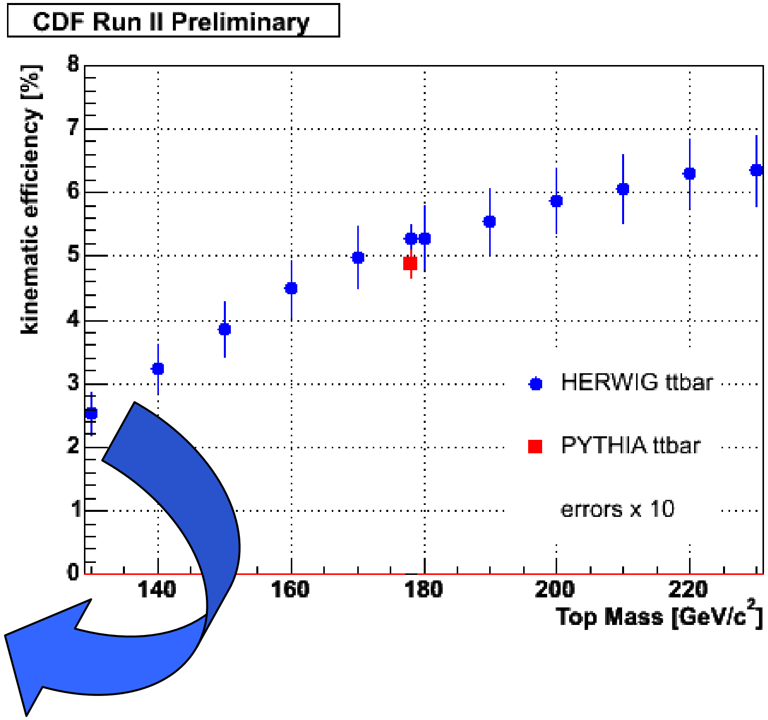
Number of jets	3	4	5	≥ 6
$t\bar{t} \to ee$	0.08 ± 0.01	0.41 ± 0.03	0.18 ± 0.02	0.04 ± 0.01
$t\bar{t} \to e\mu$	0.06 ± 0.01	0.29 ± 0.02	0.11 ± 0.01	0.05 ± 0.01
$t\bar{t} \to \mu\mu$	0.01 ± 0.01	0.05 ± 0.01	0.01 ± 0.01	0.01 ± 0.01
$t\bar{t} \to e\tau$	0.11 ± 0.01	0.93 ± 0.04	0.38 ± 0.03	0.15 ± 0.02
$t\bar{t} \to \mu\tau$	0.05 ± 0.01	0.29 ± 0.02	0.15 ± 0.02	0.06 ± 0.01
$t\bar{t} \to \tau\tau$	0.06 ± 0.01	0.58 ± 0.03	0.26 ± 0.02	0.05 ± 0.01
$t\bar{t} \to e + jets$	0.68 ± 0.04	6.61 ± 0.11	8.70 ± 0.13	4.25 ± 0.09
$t\bar{t} \to \mu + jets$	1.07 ± 0.04	11.92 ± 0.15	6.56 ± 0.11	2.47 ± 0.07
$t\bar{t} \to \tau + jets$	1.00 ± 0.04	10.98 ± 0.14	11.71 ± 0.15	5.53 ± 0.10
$t\bar{t} \to jets$	0.01 ± 0.01	0.09 ± 0.01	0.14 ± 0.02	0.22 ± 0.02
$t\bar{t} \to X$	3.13 ± 0.08	32.15 ± 0.24	28.14 ± 0.23	12.83 ± 0.15
Background b-tagged jets	32.68 ± 3.46	37.53 ± 4.14	21.44 ± 2.76	8.47 ± 1.40
Corrected background b-tagged jets	–	33.14 ± 4.01	17.58 ± 2.85	6.71 ± 2.78
Observed b-tagged jets	31	53	55	19

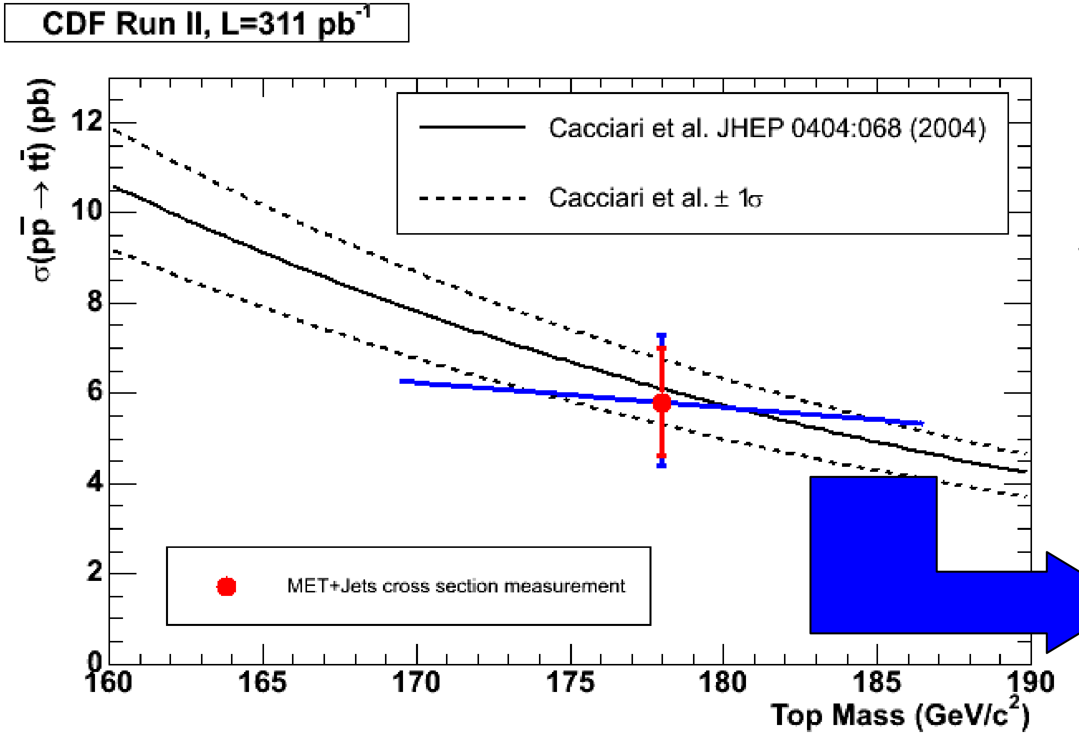
~44% of the total acceptance is provided by the W→τν ttbar decays

cross section vs top quark mass

m_{top} in the base Monte Carlo sample was set to 178 GeV/c^2.

We used different other MC signal samples to evaluate the kinematical selection and cross section dependence on m_{top}





- The cross section measurement changes by ± 0.05 pb for each -/+ 1 GeV/c^2 change in the m_{top} from the initial value of 178 GeV/c^2.
- We mesure:

σ(ttbar) = $6.0 \pm 1.2\ (\text{stat})^{+0.9}_{-0.7}\ (\text{syst})$ pb

for m_{top} = 175 GeV/c^2.

iterative top subtraction procedure

- The final sample kin sel + ≥1 tag consists of 106 events for a total of $N_{obs} = 127$ positive tagged jets.
- From tagging matrix prediction we expect $N_{exp} = 67.4 \pm 7.2$ tags
- We need to correct the tagging matrix prediction in order to account for the ttbar presence in the pre-tagging sample by using an iterative method:

$$N'_{\exp} = N^{fix}_{\exp}\frac{N_{evt} - N^{ttbar}_{evt}}{N_{evt}} = N^{fix}_{\exp}\frac{N_{evt} - \dfrac{N_{obs} - N_{\exp}}{\varepsilon^{ave}_{tag}}}{N_{evt}}$$

- The procedure stops when $|N_{exp}' - N_{exp}| < 1\%$.
- 10.0 tags out of 67.4 are attributed in this way to the ttbar presence in the pre-tagging sample.
- $N_{exp}' = 57.4 \pm 8.1$ is the corrected background amount to be used for a cross section measurement.